

数据复制服务

实时同步

文档版本 57
发布日期 2024-02-28



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编： 518129

网址： <https://www.huawei.com>

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目录

1 同步方案概览	1
2 数据同步拓扑介绍	11
3 入云	14
3.1 将 MySQL 同步到 MySQL.....	14
3.2 将 MySQL 同步到 PostgreSQL.....	32
3.3 将 MySQL 同步到 GaussDB 分布式版.....	47
3.4 将 MySQL 同步到 GaussDB 主备版.....	61
3.5 将 MySQL 同步到 GaussDB(DWS).....	75
3.6 将 MySQL 同步到 GaussDB(for MySQL).....	93
3.7 将 MySQL 同步到 MariaDB.....	110
3.8 将 PostgreSQL 同步到 PostgreSQL.....	125
3.9 将 PostgreSQL 同步到 GaussDB(DWS).....	144
3.10 将 PostgreSQL 同步到 GaussDB 主备版.....	160
3.11 将 PostgreSQL 同步到 GaussDB 分布式版.....	177
3.12 将 Oracle 同步到 MySQL.....	194
3.13 将 Oracle 同步到 GaussDB(for MySQL).....	214
3.14 将 Oracle 同步到 GaussDB 主备版.....	230
3.15 将 Oracle 同步到 GaussDB 分布式版.....	248
3.16 将 Oracle 同步到 DDM.....	266
3.17 将 Oracle 同步到 GaussDB(DWS).....	282
3.18 将 Oracle 同步到 PostgreSQL.....	300
3.19 将 DDM 同步到 MySQL.....	315
3.20 将 DDM 同步到 GaussDB(DWS).....	328
3.21 将 DDM 同步到 DDM.....	339
3.22 将 DB2 for LUW 同步到 GaussDB 主备版.....	350
3.23 将 DB2 for LUW 同步到 GaussDB 分布式版.....	362
3.24 将 DB2 for LUW 同步到 GaussDB(DWS).....	373
3.25 将 TiDB 同步到 GaussDB(for MySQL).....	385
3.26 将 Microsoft SQL Server 同步到 GaussDB(DWS).....	399
3.27 将 Microsoft SQL Server 同步到 GaussDB 主备版.....	414
3.28 将 Microsoft SQL Server 同步到 GaussDB 分布式版.....	429
3.29 将 Microsoft SQL Server 同步到 Microsoft SQL Server.....	444

3.30 将 MongoDB 同步到 DDS.....	457
3.31 将 MariaDB 同步到 MariaDB.....	470
3.32 将 MariaDB 同步到 MySQL.....	485
3.33 将 MariaDB 同步到 GaussDB(for MySQL).....	502
3.34 将 GaussDB(for MySQL)同步到 GaussDB(for MySQL).....	519
4 出云.....	536
4.1 将 MySQL 同步到 MySQL.....	536
4.2 将 MySQL 同步到 Kafka.....	553
4.3 将 MySQL 同步到 CSS/ES.....	567
4.4 将 MySQL 同步到 Oracle.....	581
4.5 将 MySQL 同步到 MariaDB.....	595
4.6 将 DDM 同步到 MySQL.....	612
4.7 将 DDM 同步到 Oracle.....	624
4.8 将 DDM 同步到 Kafka.....	636
4.9 将 DDS 同步到 MongoDB.....	647
4.10 将 DDS 同步到 Kafka.....	660
4.11 将 PostgreSQL 同步到 PostgreSQL.....	670
4.12 将 PostgreSQL 同步到 Kafka.....	688
4.13 将 GaussDB 分布式版同步到 MySQL.....	700
4.14 将 GaussDB 分布式版同步到 Oracle.....	716
4.15 将 GaussDB 分布式版同步到 GaussDB(DWS).....	734
4.16 将 GaussDB 分布式版同步到 Kafka.....	751
4.17 将 GaussDB 分布式版同步到 GaussDB 分布式版.....	764
4.18 将 GaussDB 分布式版同步到 GaussDB 主备版.....	782
4.19 将 GaussDB 主备版同步到 MySQL.....	799
4.20 将 GaussDB 主备版同步到 Oracle.....	815
4.21 将 GaussDB 主备版同步到 GaussDB(DWS).....	833
4.22 将 GaussDB 主备版同步到 Kafka.....	849
4.23 将 GaussDB 主备版同步到 GaussDB 分布式版.....	863
4.24 将 GaussDB 主备版同步到 GaussDB 主备版.....	879
4.25 将 GaussDB(for MySQL)同步到 MySQL.....	895
4.26 将 GaussDB(for MySQL)同步到 GaussDB(DWS).....	909
4.27 将 GaussDB(for MySQL)同步到 Kafka.....	923
4.28 将 GaussDB(for MySQL)同步到 CSS/ES.....	936
4.29 将 GaussDB(for MySQL)同步到 Oracle.....	950
4.30 将 MariaDB 同步到 MariaDB.....	962
4.31 将 Microsoft SQL Server 同步到 Kafka.....	977
5 自建到自建.....	988
5.1 将 MySQL 同步到 Kafka.....	988
5.2 将 MySQL 同步到 CSS/ES.....	1002
5.3 将 MySQL 同步到 GaussDB 主备版.....	1015
5.4 将 MySQL 同步到 GaussDB 分布式版.....	1030

5.5 将 Oracle 同步到 Kafka.....	1045
5.6 将 Oracle 同步到 GaussDB 主备版.....	1058
5.7 将 Oracle 同步到 GaussDB 分布式版.....	1076
5.8 将 PostgreSQL 同步到 Kafka.....	1094
5.9 将 GaussDB 主备版同步到 MySQL.....	1106
5.10 将 GaussDB 主备版同步到 Oracle.....	1122
5.11 将 GaussDB 主备版同步到 Kafka.....	1140
5.12 将 GaussDB 主备版同步到 GaussDB 分布式版.....	1154
5.13 将 GaussDB 主备版同步到 GaussDB 主备版.....	1170
5.14 将 GaussDB 分布式版同步到 MySQL.....	1186
5.15 将 GaussDB 分布式版同步到 Oracle.....	1203
5.16 将 GaussDB 分布式版同步到 Kafka.....	1221
5.17 将 GaussDB 分布式版同步到 GaussDB 分布式版.....	1234
5.18 将 GaussDB 分布式版同步到 GaussDB 主备版.....	1253
5.19 将 DB2 for LUW 同步到 GaussDB 主备版.....	1270
5.20 将 DB2 for LUW 同步到 GaussDB 分布式版.....	1281
5.21 将 Microsoft SQL Server 同步到 Kafka.....	1292
6 任务管理.....	1303
6.1 创建同步任务.....	1303
6.2 查询同步进度.....	1317
6.3 查看同步日志.....	1319
6.4 对比同步项.....	1320
6.5 对象管理.....	1332
6.5.1 编辑同步对象.....	1332
6.5.2 导入同步对象.....	1335
6.5.3 对象名映射.....	1336
6.5.4 查看同步映射.....	1342
6.5.5 数据加工.....	1343
6.6 参数管理.....	1351
6.6.1 修改任务参数.....	1351
6.6.2 查看参数修改历史.....	1354
6.7 任务生命周期.....	1355
6.7.1 查看任务详情.....	1355
6.7.2 修改任务信息.....	1355
6.7.3 修改连接信息.....	1357
6.7.4 修改流速模式.....	1357
6.7.5 编辑同步任务.....	1361
6.7.6 续传同步任务.....	1361
6.7.7 跳跃续传同步任务.....	1362
6.7.8 暂停同步任务.....	1363
6.7.9 重置同步任务.....	1366
6.7.10 重置同步位点.....	1369

6.7.11 重启同步任务.....	1370
6.7.12 跳过 DDL.....	1370
6.7.13 双 AZ 任务倒换.....	1371
6.7.14 克隆同步任务.....	1373
6.7.15 同步规格变更.....	1374
6.7.16 退订包周期任务.....	1375
6.7.17 结束同步任务.....	1377
6.7.18 删除同步任务.....	1378
6.7.19 批量导入任务.....	1379
6.7.20 任务状态说明.....	1380
7 标签管理.....	1382
8 连接诊断.....	1384
9 查看异常数据.....	1386
10 对接云审计服务.....	1387
10.1 支持审计操作的关键列表.....	1387
10.2 如何查看审计日志.....	1387
11 对接云监控服务.....	1389
11.1 支持的监控指标.....	1389
11.2 设置告警规则.....	1393
11.3 查看监控指标.....	1394
12 对接云日志服务.....	1396
12.1 日志配置管理.....	1396
12.2 查看或下载日志.....	1397
13 同步场景操作参考.....	1399
13.1 Kafka 消息格式.....	1399
13.2 Kafka 认证方式.....	1407
13.3 GaussDB 分布式版为源强制结束任务.....	1410
13.4 GaussDB 主备版为源强制结束任务.....	1411
13.5 PostgreSQL 为源强制结束任务.....	1412
13.6 通过创建事件触发器和函数实现 PostgreSQL 增量 DDL 同步.....	1414
14 附录.....	1417
14.1 通用使用建议.....	1417
14.2 DRS 预检查项一览表.....	1418
A 修订记录.....	1429

1 同步方案概览

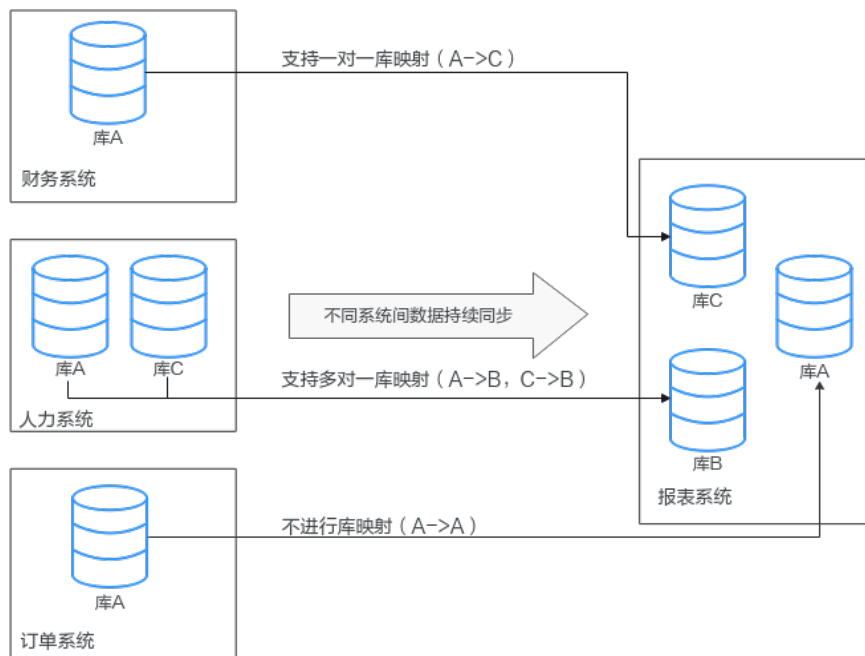
实时同步是指在不同的系统之间，将数据通过同步技术从一个数据源拷贝到其他数据库，并保持一致，实现关键业务的数据实时流动。

实时同步不同于迁移，迁移是以整体数据库搬迁为目的，而实时同步是维持不同业务之间的数据持续性流动。

常用场景：实时分析，报表系统，数仓环境。

同步特点：实时同步功能聚焦于表和数据，并满足多种灵活性的需求，例如多对一、一对多，动态增减同步表，不同表名之间同步数据等。

图 1-1 多对一实时同步



支持的数据库类型

目前，DRS实时同步支持的源数据库和目标数据库类型，如下表所示。

表 1-1 同步方案

源数据库类型	目标数据库类型	同步模式	相关文档
<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 RDS for MySQL 	RDS for MySQL	增量数据同步 全量数据同步 全量+增量数据同步	将MySQL同步到MySQL（入云）
	GaussDB(for MySQL)	增量数据同步 全量+增量数据同步	将MySQL同步到GaussDB(for MySQL)
	RDS for PostgreSQL	全量数据同步 全量+增量数据同步	将MySQL同步到PostgreSQL
	GaussDB分布式版	增量数据同步 全量数据同步 全量+增量数据同步	将MySQL同步到GaussDB分布式版（入云） 将MySQL同步到GaussDB分布式版（自建）
	GaussDB主备版	增量数据同步 全量数据同步 全量+增量数据同步	将MySQL同步到GaussDB主备版（入云） 将MySQL同步到GaussDB主备版（自建）
	GaussDB(DWS)	增量数据同步 全量数据同步 全量+增量数据同步	将MySQL同步到GaussDB(DWS)
	RDS for MariaDB 说明 目前仅支持白名单用户使用。	增量数据同步 全量数据同步 全量+增量数据同步	将MySQL同步到MariaDB（入云）
RDS for MySQL	<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 	增量数据同步 全量+增量数据同步	将MySQL同步到MySQL（出云）
	Kafka	增量数据同步 全量+增量数据同步	将MySQL同步到Kafka（出云）

源数据库类型	目标数据库类型	同步模式	相关文档
	CSS/ES	全量+增量数据同步	将MySQL同步到CSS/ES (出云)
	<ul style="list-style-type: none"> 本地自建Oracle数据库 ECS自建Oracle数据库 	全量+增量数据同步	将MySQL同步到Oracle
	<ul style="list-style-type: none"> 本地自建MariaDB数据库 ECS自建MariaDB数据库 其他云上MariaDB数据库 <p>说明 目前仅支持白名单用户使用。</p>	增量数据同步 全量数据同步 全量+增量数据同步	将MySQL同步到MariaDB (出云)
<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 	Kafka	增量数据同步 全量+增量数据同步	将MySQL同步到Kafka (自建-自建)
	CSS/ES	全量+增量数据同步	将MySQL同步到CSS/ES (自建-自建)
<ul style="list-style-type: none"> 本地自建PostgreSQL数据库 ECS自建PostgreSQL数据库 其他云上PostgreSQL数据库 RDS for PostgreSQL 	RDS for PostgreSQL	增量数据同步 全量数据同步 全量+增量数据同步	将PostgreSQL同步到PostgreSQL
	GaussDB(DWS)	全量+增量数据同步	将PostgreSQL同步到GaussDB(DWS)
	GaussDB主备版	增量数据同步 全量数据同步 全量+增量数据同步 <p>说明 目前单增量数据同步仅支持白名单用户使用。</p>	将PostgreSQL同步到GaussDB主备版

源数据库类型	目标数据库类型	同步模式	相关文档
	GaussDB分布式版	增量数据同步 全量数据同步 全量+增量数据同步 说明 目前单增量数据同步仅支持白名单用户使用。	将PostgreSQL同步到GaussDB分布式版
<ul style="list-style-type: none"> 本地自建PostgreSQL数据库 ECS自建PostgreSQL数据库 	Kafka	增量数据同步	将PostgreSQL同步到Kafka（自建-自建）
RDS for PostgreSQL	<ul style="list-style-type: none"> 本地自建PostgreSQL数据库 ECS自建PostgreSQL数据库 	增量数据同步 全量数据同步 全量+增量数据同步	将PostgreSQL同步到PostgreSQL（出云）
	Kafka	增量数据同步	将PostgreSQL同步到Kafka（出云）
DDM	RDS for MySQL	增量数据同步 全量数据同步 全量+增量数据同步	将DDM同步到MySQL（入云）
	<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 	增量数据同步 全量数据同步 全量+增量数据同步	将DDM同步到MySQL（出云）
	DDM	全量+增量数据同步	将DDM同步到DDM
	GaussDB(DWS)	全量+增量数据同步	将DDM同步到GaussDB(DWS)
	<ul style="list-style-type: none"> 本地自建Oracle数据库 ECS自建Oracle数据库 	增量数据同步 全量数据同步 全量+增量数据同步	将DDM同步到Oracle

源数据库类型	目标数据库类型	同步模式	相关文档
	Kafka	增量数据同步	将DDM同步到Kafka
<ul style="list-style-type: none"> 本地自建Oracle数据库 ECS自建Oracle数据库 	RDS for MySQL	增量数据同步 全量数据同步 全量+增量数据同步	将Oracle同步到MySQL
	GaussDB(for MySQL)	全量数据同步 全量+增量数据同步	将Oracle同步到GaussDB(for MySQL)
	RDS for PostgreSQL	全量数据同步 全量+增量数据同步	将Oracle同步到PostgreSQL
	GaussDB主备版	增量数据同步 全量数据同步 全量+增量数据同步	将Oracle同步到GaussDB主备版(入云) 将Oracle同步到GaussDB主备版(自建-自建)
	GaussDB分布式版	增量数据同步 全量数据同步 全量+增量数据同步	将Oracle同步到GaussDB分布式版(入云) 将Oracle同步到GaussDB分布式版(自建-自建)
	DDM	全量数据同步 全量+增量数据同步	将Oracle同步到DDM
	GaussDB(DWS)	增量数据同步 全量数据同步 全量+增量数据同步	将Oracle同步到GaussDB(DWS)
	Kafka	增量数据同步	将Oracle同步到Kafka
GaussDB分布式版	<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 RDS for MySQL 	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB分布式版同步到MySQL(出云) 将GaussDB分布式版同步到MySQL(自建-自建)

源数据库类型	目标数据库类型	同步模式	相关文档
	<ul style="list-style-type: none"> 本地自建Oracle数据库 ECS自建Oracle数据库 	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB分布式版同步到Oracle (出云) 将GaussDB分布式版同步到Oracle (自建-自建)
	GaussDB(DWS)	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB分布式版同步到GaussDB(DWS)
	Kafka	增量数据同步	将GaussDB分布式版同步到Kafka (出云) 将GaussDB分布式版同步到Kafka (自建-自建)
	GaussDB分布式版	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB分布式版同步到GaussDB分布式版 (出云) 将GaussDB分布式版同步到GaussDB分布式版 (自建-自建)
	GaussDB主备版	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB分布式版同步到GaussDB主备版 (出云) 将GaussDB分布式版同步到GaussDB主备版 (自建-自建)
GaussDB主备版	<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 RDS for MySQL 	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB主备版同步到MySQL (出云) 将GaussDB主备版同步到MySQL (自建-自建)
	<ul style="list-style-type: none"> 本地自建Oracle数据库 ECS自建Oracle数据库 	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB主备版同步到Oracle (出云) 将GaussDB主备版同步到Oracle (自建-自建)

源数据库类型	目标数据库类型	同步模式	相关文档
	Kafka	增量数据同步	将GaussDB主备版同步到Kafka (出云) 将GaussDB主备版同步到Kafka (自建-自建)
	GaussDB(DWS)	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB主备版同步到GaussDB(DWS)
	GaussDB分布式版	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB主备版同步到GaussDB分布式版 (出云) 将GaussDB主备版同步到GaussDB分布式版 (自建-自建)
	GaussDB主备版	全量数据同步 增量数据同步 全量+增量数据同步	将GaussDB主备版同步到GaussDB主备版 (出云) 将GaussDB主备版同步到GaussDB主备版 (自建-自建)
GaussDB(for MySQL)	GaussDB(for MySQL)	增量数据同步 全量+增量数据同步	将GaussDB(for MySQL)同步到GaussDB(for MySQL)
	<ul style="list-style-type: none"> • 本地自建MySQL数据库 • ECS自建MySQL数据库 • 其他云上MySQL数据库 • RDS for MySQL 	增量数据同步 全量+增量数据同步	将GaussDB(for MySQL)同步到MySQL
	GaussDB(DWS)	全量+增量数据同步	将GaussDB(for MySQL)同步到GaussDB(DWS)
	Kafka	增量数据同步 全量+增量数据同步	将GaussDB(for MySQL)同步到Kafka

源数据库类型	目标数据库类型	同步模式	相关文档
	CSS/ES	全量+增量数据同步	将GaussDB(for MySQL)同步到CSS/ES
	<ul style="list-style-type: none"> 本地自建Oracle数据库 ECS自建Oracle数据库 	全量+增量数据同步	将GaussDB(for MySQL)同步到Oracle
DDS	<ul style="list-style-type: none"> 自建MongoDB 其他云上MongoDB 	增量数据同步支持以下方式： 副本集到副本集 集群到集群（源数据库集群版本不小于4.0）	将DDS同步到MongoDB
	Kafka	增量数据同步	将DDS同步到Kafka
DB2 for LUW	GaussDB分布式版	全量数据同步 全量+增量数据同步	将DB2 for LUW同步到GaussDB主备版（入云） 将DB2 for LUW同步到GaussDB主备版（自建）
	GaussDB主备版	全量数据同步 全量+增量数据同步	将DB2 for LUW同步到GaussDB分布式版（入云） 将DB2 for LUW同步到GaussDB分布式版（自建）
	GaussDB(DWS)	全量数据同步 全量+增量数据同步	将DB2 for LUW同步到GaussDB(DWS)
TiDB	GaussDB(for MySQL)	全量+增量数据同步	将TiDB同步到GaussDB(for MySQL)

源数据库类型	目标数据库类型	同步模式	相关文档
<ul style="list-style-type: none"> 本地自建 Microsoft SQL Server 数据库 ECS 自建 Microsoft SQL Server 数据库 其他云上 Microsoft SQL Server 数据库 RDS for SQL Server 	GaussDB(DWS)	全量+增量数据同步	将Microsoft SQL Server同步到 GaussDB(DWS)
	GaussDB主备版	增量数据同步 全量数据同步 全量+增量数据同步	将Microsoft SQL Server同步到 GaussDB主备版
	GaussDB分布式版	增量数据同步 全量数据同步 全量+增量数据同步	将Microsoft SQL Server同步到 GaussDB分布式版
	Microsoft SQL Server 说明 目前仅支持白名单用户使用。	全量+增量数据同步	将Microsoft SQL Server同步到 Microsoft SQL Server
RDS for SQL Server	Kafka 说明 目前仅支持白名单用户使用。	增量数据同步	将Microsoft SQL Server同步到 Kafka (出云)
<ul style="list-style-type: none"> 本地自建 Microsoft SQL Server 数据库 ECS 自建 Microsoft SQL Server 数据库 其他云上 Microsoft SQL Server 数据库 	Kafka 说明 目前仅支持白名单用户使用。	增量数据同步	将Microsoft SQL Server同步到 Kafka (自建-自建)
<ul style="list-style-type: none"> 本地自建 Mongo 数据库 ECS 自建 Mongo 数据库 其他云上 Mongo 数据库 DDS 	DDS	全量+增量数据同步支持以下方式： 副本集到副本集	将MongoDB同步到DDS
<ul style="list-style-type: none"> 本地自建 MariaDB 数据库 ECS 自建 MariaDB 数据库 其他云上 MariaDB 数据库 	RDS for MariaDB	全量+增量数据同步	将MariaDB同步到 MariaDB

源数据库类型	目标数据库类型	同步模式	相关文档
<ul style="list-style-type: none"> 本地自建 MariaDB数据库 ECS自建 MariaDB数据库 其他云上 MariaDB数据库 	RDS for MySQL 说明 目前仅支持白名单用户使用。	增量数据同步 全量+增量数据同步	将MariaDB同步到MySQL
<ul style="list-style-type: none"> 其他云上 MariaDB数据库 	GaussDB(for MySQL) 说明 目前仅支持白名单用户使用。	增量数据同步 全量+增量数据同步	将MariaDB同步到GaussDB(for MySQL)
RDS for MariaDB	<ul style="list-style-type: none"> 本地自建 MariaDB数据库 ECS自建 MariaDB数据库 其他云上 MariaDB数据库 	全量+增量数据同步	将MariaDB同步到MariaDB

2 数据同步拓扑介绍

DRS实时同步功能目前支持多种拓扑类型，用户可根据自己的需求进行规划，详细说明可参考以下内容。

📖 说明

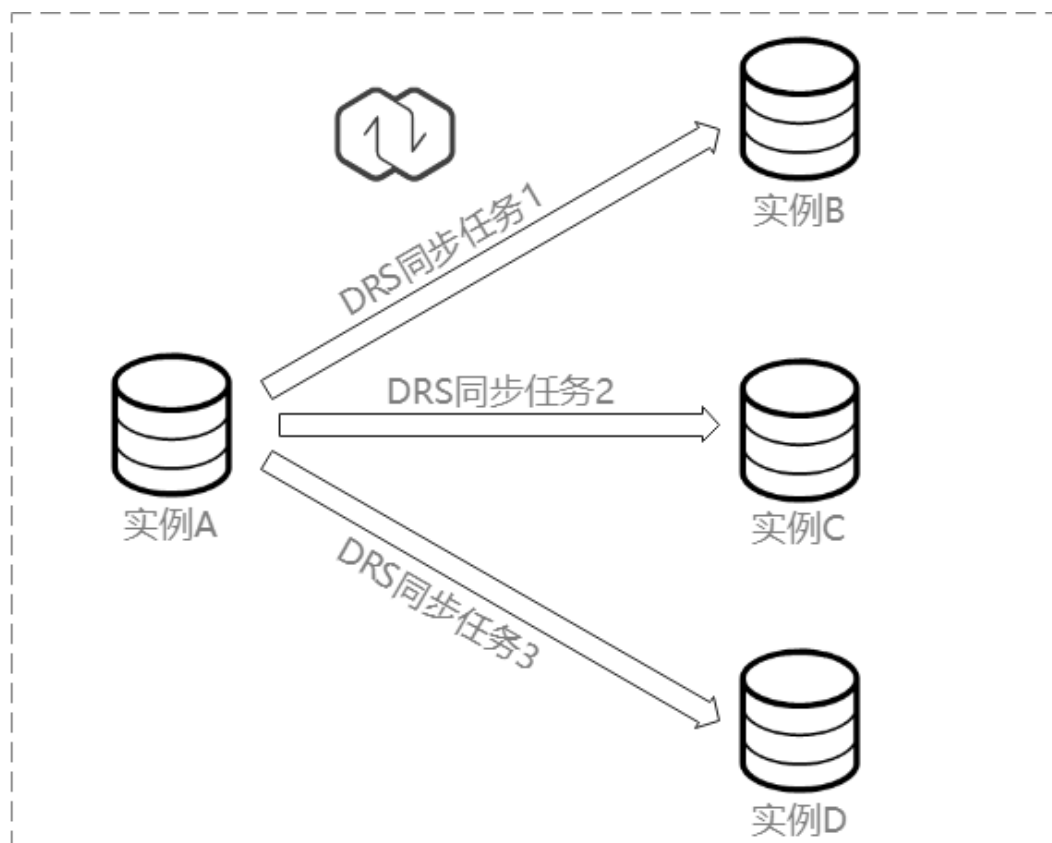
为保证同步数据的一致性，请勿对目标数据库中的同步对象进行修改操作。

一对一实时同步



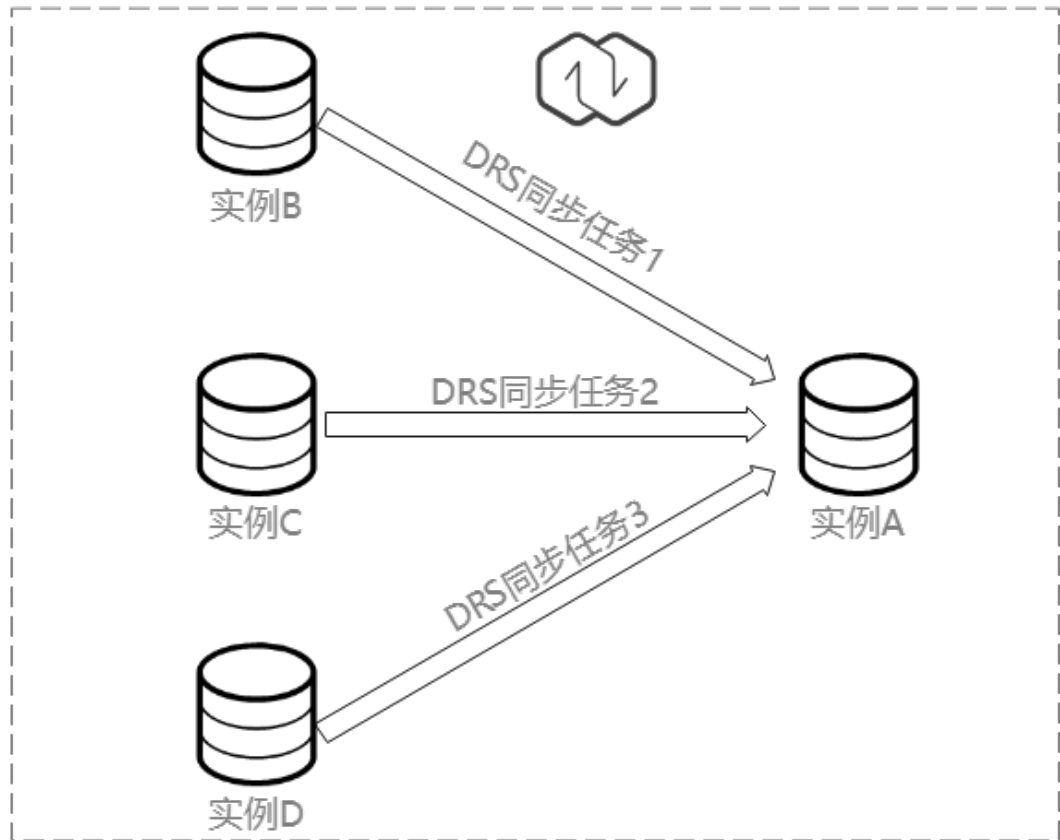
用户创建一个同步任务来实现一对一的实时同步。

一对多实时同步



用户需要用户创建多个同步任务来实现一对多的实时同步。例如从实例A同步到实例B、实例C和实例D，需要创建三个同步任务。

多对一实时同步



用户需要用户创建多个同步任务来实现多对一的实时同步。例如从实例B、实例C和实例D同步到实例A，需要创建三个同步任务。

多对一场景约束及操作建议，可参考[常见问题](#)。

3 入云

3.1 将 MySQL 同步到 MySQL

支持的源和目标数据库

表 3-1 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本）ECS自建MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本）其他云上MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本）RDS for MySQL（5.5、5.6、5.7、8.0版本）	RDS for MySQL（5.5、5.6、5.7、8.0版本） 说明 仅支持目标库版本等于或高于源库版本。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-2](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

表 3-2 数据库账号权限

类型名称	全量同步	全量+增量同步、增量同步
源数据库连接账号	需要具备如下最小权限： SELECT、SHOW VIEW、EVENT。	需要具备如下最小权限： SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。
目标数据库连接账号	需要具备如下最小权限： SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES。 RDS for MySQL实例的root账户默认已具备上述权限。当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。	

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-3](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-3 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none">• 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。• 仅支持MyISAM和InnoDB表的同步。• 不支持事件（EVENT）和触发器（TRIGGER）的同步。• 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如：<pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

注意事项

为保证任务的正常运行，DRS提供了自动化的检查方法，在您启动DRS任务前，会针对源数据库和目标数据库的相关配置、条件进行检查。主要的检查项和对应的处理建议请参考[预检查项](#)。除了预检查项目之外，您还需要注意的事情参见[表3-4](#)。

表 3-4 注意事项

类型名称	使用和操作限制
源数据库限制	<ul style="list-style-type: none">不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。源数据库不允许存在与目标库同名的无主键表。如果源数据库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。

类型名称	使用和操作限制
使用限制	<p>通用：</p> <ul style="list-style-type: none"> ● 不建议修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。 ● 不允许源库执行清除Binlog的操作。 ● 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。 <p>全量同步阶段：</p> <ul style="list-style-type: none"> ● 任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。 ● 对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量数据同步阶段，目标数据库会产生大量Binlog引起磁盘空间占用过大，因此在全量数据同步阶段目标数据库Binlog本地默认调整为只保留最新5个，全量同步结束后目标数据库Binlog本地保留时长恢复为用户当前配置。如果因业务需要，目标数据库本地Binlog保留时长需要保持用户当前配置，需提交工单申请开启。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 <p>增量同步阶段：</p> <ul style="list-style-type: none"> ● 如果设置session变量character_set_client为binary，可能导致乱码。 ● 表级同步时，增量同步过程支持使用Online DDL，可参考DRS实时同步支持使用Online DDL工具吗。 ● 支持任务再编辑追加同步对象。 ● 支持断点续传功能，在主机系统崩溃的情况下，对于非事务性的无主键的表可能会出现重复插入数据的情况。 ● 支持部分DDL语句。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY，客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 增量同步支持表的重命名，源表和目標表必須都在对象选择里。 ● 不支持源数据库进行恢复到某个备份点的操作（PITR）。 <p>结束任务：</p> <ul style="list-style-type: none"> ● 任务正常结束：

类型名称	使用和操作限制
	<p>请确保任务结束前，不要断开源和目标数据库的网络连通性，并在结束任务时关注同步日志上报的状态，达到数据库完整同步效果。</p> <ul style="list-style-type: none">● 任务强制结束： 强制结束将直接释放DRS资源。 <p>常见故障排查：</p> <ul style="list-style-type: none">● 在任务创建、启动、全量同步、增量同步、结束等过程中，如有遇到问题，可先参考“故障排查”章节进行排查。

类型名称	使用和操作限制
其他限制	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 对于源数据库和目标数据库的版本有差异的同步链路，由于不同版本之间的部分特性差异，可能会出现语法兼容性问题，详情可参考MySQL不同版本之间的语法差异。 ● 表名映射时，不支持同步视图、存储过程、函数依赖的表，不支持同步表的外键约束。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 进行表级多对一同步时，源库不允许存在无主键表。 ● 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 源数据库中的库名、表名如果存在非成对的{}，可能导致表结构同步失败。进行数据映射后，映射到目标库中的库名、表名如果存在非成对{}，可能导致表结构同步失败。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持分区表的分区字段进行列映射。 ● 源库和目标库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 当目标数据库为5.7版本时，因为版本限制，json类型中的浮点会丢失小数点后末位的0，存在精度损失导致内容对比不一致。 ● 在创建DRS任务之前，如果源或者目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 同步任务创建后，目标数据库不能设置为“只读”。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标数据库多的列要求非空且没有默认值，源库insert数据，同步到目标数据库后多的列为null，不符合目标数据库要求。 - 目标数据库多的列设置固定默认值，且有唯一约束。源库insert多条数据后，同步到目标数据库后多的列为固定默认值，不符合目标数据库要求。 ● 如果涉及多对一场景的同步任务，可参考多对一的场景约束及操作建议。 ● 如果涉及表级汇集的多对一同步任务，则不支持DDL，否则会导致同步全部失败。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename

类型名称	使用和操作限制
	<p>操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤。</p> <ul style="list-style-type: none"> - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。

前提条件

- [已登录数据复制服务控制台](#)。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。

操作步骤

本小节以MySQL->RDS for MySQL的实时同步为示例，介绍如何使用数据复制服务创建两个数据库实例之间的实时同步任务，其他存储引擎的配置流程类似。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-1 同步任务信息

⚠️ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 3-5 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 3-2 同步实例信息



表 3-6 同步实例信息

参数	描述
数据流动方向	选择“入云”，即目标端数据库为本云数据库。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例。可根据业务场景选择“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。

参数	描述
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	<p>用户所创建的关系型数据库实例。</p> <p>说明</p> <ul style="list-style-type: none"> - 目标数据库实例不支持选择只读实例。 - 目标数据库实例可以和源数据库选择同一个实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
内网安全组	<p>请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 3-3 任务类型



表 3-7 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-4 可用区</p>

- 企业项目和标签

图 3-5 企业项目和标签



表 3-8 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 同步实例创建成功后，在“源库及目标库”页面，填选源库信息和目标库信息后，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过，勾选协议，单击“下一步”。

入云和出云场景中源库和目标库需要填选的信息不同，具体以界面为准。

- 源库信息

图 3-6 源库信息

源库信息

IP地址或域名 请确保本云安全组允许访问源数据库公网地址，[点击这里](#)

端口

数据库用户名

不支持数据库参数和系统数据库同步，源数据库参数设置和用户将不会同步至目标数据库中，请在目标数据库中使用参数组修改参数，手工创建用户。

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

● 测试成功

表 3-9 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库用户名所对应的密码。支持在任务创建后修改密码。 任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面的，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。
SSL安全连接	如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

- 目标库信息

图 3-7 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码 

SSL安全连接

✔ 测试成功 

表 3-10 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	目标数据库对应的用户名。

参数	描述
数据库密码	目标数据库用户名对应的密码。支持在任务创建后修改密码。 任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。
SSL安全连接	如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择数据冲突策略和同步对象，单击“下一步”。

图 3-8 同步模式

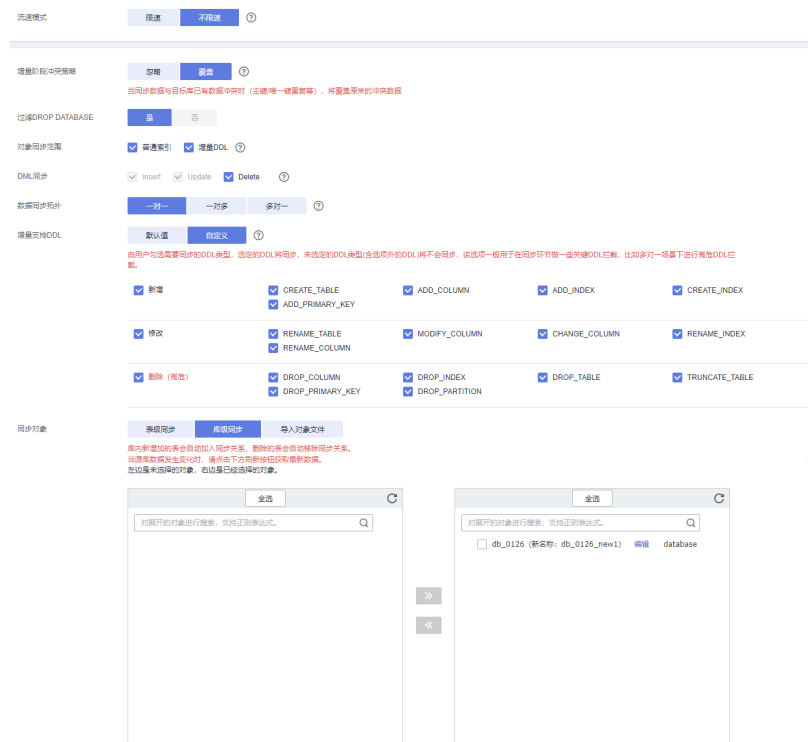




表 3-11 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-9 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 如果同步的对象在源数据库和目标数据库中的命名不同，可在右侧已选对象框中修改目标库对象映射名称，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 3-10 数据加工



源数据库	新名称	列名	操作类型	数据类型	操作
		a1	以serverName@database@table为列	varchar(256)	
		a2	默认值 123456789	long	
db_has_pk_test1_table16 db_has_pk_test1_table12 db_has_pk_test1_table14 db_has_pk_test1_table10		a3	以create_time为列	datetime	
		a4	以update_time为列	timestamp	
		a5	表达式 concat('current_database_', '_current_table')	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-11 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-12 任务启动设置



表 3-12 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.2 将 MySQL 同步到 PostgreSQL

支持的源和目标数据库

表 3-13 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本） ECS自建MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本） 其他云上MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本） RDS for MySQL（5.5、5.6、5.7、8.0版本） 	RDS for PostgreSQL（9.5、9.6、10、11、12、13、14版本）

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-14](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

表 3-14 数据库账号权限

类型名称	全量同步	全量+增量同步
源数据库连接账号	需要具备如下最小权限： SELECT	需要具备如下最小权限： SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。
目标数据库连接账号	RDS for PostgreSQL实例的账号默认已具有权限。	

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-15](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-15 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> 仅支持同步表结构、表数据、索引信息。 仅支持同步MyISAM和InnoDB表。 不支持同步存储过程等其他数据库对象。 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

注意事项

为保证任务的正常运行，DRS提供了自动化的检查方法，在您启动DRS任务前，会针对源数据库和目标数据库的相关配置、条件进行检查。主要的检查项和对应的处理建议请参考[预检查项](#)。除了预检查项目之外，您还需要注意的事情参见[表3-16](#)。

表 3-16 注意事项

类型名称	使用和操作限制
源数据库限制	<ul style="list-style-type: none"> 不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 如果源数据库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。

类型名称	使用和操作限制
使用限制	<p>通用：</p> <ul style="list-style-type: none"> ● 如果源数据库为RDS for MySQL，支持修改端口，修改之后同步任务失败，需要通过重试后继续进行同步。 ● 如果源数据库为非RDS for MySQL，不支持修改端口。 ● 不支持修改源和目标数据库的IP、账号和密码。 ● 不允许源库执行清除Binlog的操作。 ● 不允许源库进行恢复操作。 <p>全量同步阶段：</p> <ul style="list-style-type: none"> ● 任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。 ● DRS会向目标库PostgreSQL写入大量数据，会导致PostgreSQL的wal日志量急剧增长，PostgreSQL的磁盘有被写满的风险。可以通过在全量同步前关闭PostgreSQL的日志备份功能，减少wal日志的生产，同步完成后再将其打开的方式进行规避（具体操作方法可参考设置自动备份策略）。 <p>注意 关闭日志备份会影响数据库的灾备恢复，请根据实际情况谨慎选择。</p> <p>增量同步阶段：</p> <ul style="list-style-type: none"> ● 如果设置session变量character_set_client为binary，可能导致乱码。 ● 支持部分DDL语句。 <ul style="list-style-type: none"> - 支持表字段的增删改，例如： <pre>alter table `ddl_test` add column `c2` varchar(25); alter table `ddl_test` modify column `c1` varchar(50); alter table `ddl_test` alter c1 set default 'xxx';</pre> - 支持修改表索引，例如： <pre>alter table `ddl_test` drop primary key; alter table `ddl_test` add primary key(id); alter table `ddl_test` add index `ddl_test_uk`(id); alter table `ddl_test` drop index `ddl_test_uk`;</pre> - 表级同步支持增加列、修改列、增加主键和普通索引。 - 如果目标表不存在，则该DDL会被忽略掉。 - 不支持DROP_DATABASE、DROP_TABLE、TRUNCATE_TABLE、CREATE_VIEW、DROP_VIEW。 - 不支持使用Online DDL。 <p>结束任务：</p> <ul style="list-style-type: none"> ● 任务正常结束： 请确保任务结束前，不要断开源和目标数据库的网络连通性，并在结束任务时关注同步日志上报的状态，达到数据库完整同步效果。 ● 任务强制结束： 强制结束将直接释放DRS资源。 <p>常见故障排查：</p>

类型名称	使用和操作限制
	<ul style="list-style-type: none">在任务创建、启动、全量同步、增量同步、结束等过程中，如有遇到问题，可先参考“故障排查”章节进行排查。

类型名称	使用和操作限制
其他限制	<ul style="list-style-type: none"> ● 由于MySQL中视图支持as select ... from a join b where ...等语法，PostgreSQL不支持，可能会导致视图同步失败。 ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标数据库PostgreSQL中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，在全量同步中，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由"原库名_原表名_原索引名"计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图等。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 在创建DRS任务之前，如果源或者目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 网络中断在30秒内恢复的，不影响实时同步，如果超过30秒，则会导致同步任务失败。 ● 支持通过映射方式实现多个库对一个库的实时同步，且映射库之间不允许存在同名表，该功能需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 ● 索引同步不区分索引类型，同步到目标数据库都是btree索引。 ● 在磁盘空间允许的情况下，建议源数据库Binlog保存时间越长越好，建议为3天，设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置Binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的Binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置Binlog保留时间可参考设置RDS for MySQL本地Binlog日志清理。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标数据库多的列要求非空且没有默认值，源库insert数据，同步到目标数据库后多的列为null，不符合目标数据库要求。 - 目标数据库多的列设置固定默认值，且有唯一约束。源库insert多条数据后，同步到目标数据库后多的列为固定默认值，不符合目标数据库要求。 ● 仅支持记录违反非空约束的异常数据、char类型或varchar类型超出字段长度限制的异常数据。 ● 若无主键表中包含大字段（ blob、text、clob、nclob、bytea ），增量同步阶段存在大字段数据不一致的可能性。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> 源数据库同步的表如果有AUTO_INCREMENT属性，结束任务时，DRS会自动刷新同步表的序列整型列所对应的PostgreSQL自增列（sequence）起始值，刷新值为该列的最大值+10000。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-13 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目

* 任务名称 ?

描述 ?
0/256

表 3-17 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-14 同步实例信息



表 3-18 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“PostgreSQL”。
网络类型	<p>此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为关系型PostgreSQL数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量： 该模式为数据库一次性同步，适用于可中断业务的数据同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。

- 任务类型

图 3-15 任务类型



表 3-19 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-16 企业项目和标签



表 3-20 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过后，勾选协议，单击“下一步”。

图 3-17 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功

表 3-21 ECS 自建库场景源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

参数	描述
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-18 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码 

表 3-22 目标库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的关系型PostgreSQL数据库实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明

目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 3-19 同步模式

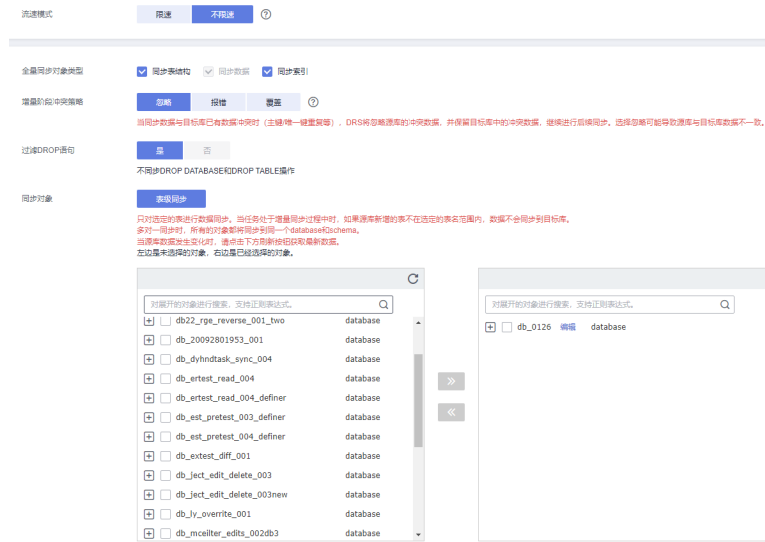



表 3-23 同步对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-20 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> ● 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 ● 否，则表示过程中将相关操作同步到目标库。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。</p> <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

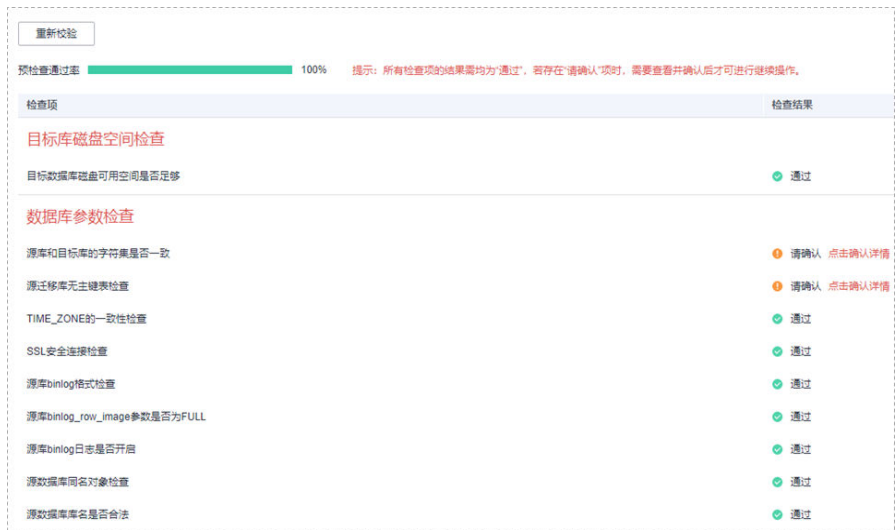
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-21 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。

图 3-22 任务启动设置




表 3-24 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。

参数	描述
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.3 将 MySQL 同步到 GaussDB 分布式版

支持的源和目标数据库

表 3-25 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 RDS for MySQL 	GaussDB分布式

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-26 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账户需要具备如下权限： <ul style="list-style-type: none"> - 全量同步：SELECT。 授权语句：GRANT SELECT ON <database>.<table> to <user>; - 全量+增量和增量同步：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 授权语句：GRANT SELECT, LOCK TABLES ON <database>.<table> to drsUser; GRANT REPLICATION SLAVE, REPLICATION CLIENT ON *.* to <user>; • 目标库GaussDB账户需要具备如下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>; - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> • 仅支持同步表结构、表数据、索引信息，不支持同步存储过程等其他数据库对象。 • 仅支持同步有主键表，不支持同步无主键表。 • 增量同步不支持同步DDL。 • 虚拟列会被同步为普通列，且不写入任何数据，也不做增量同步。 • 不支持的数据类型有：xml、包含bit类型的json、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon、interval。 • 仅支持MyISAM和InnoDB表的同步。 • 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 • 对象名同步到目标库后会转换成小写，因此选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 • 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 • 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 • 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 • 增量同步时，必须设置MySQL源数据库的server_id。MySQL5.7，server_id的取值范围在1 - 4294967296之间。 • 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 • 源数据库中的库名、表名不能包含：.<'>/\以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> • 目标数据库实例的运行状态必须正常。 • 目标数据库实例必须有足够的磁盘空间。 • 目标数据库的时区设置必须与源数据库一致。 • 任务配置的映射数据库必须在目标库已经存在。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，在全量同步中，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由"原库名_原表名_原索引名"计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key。 ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 网络中断在30秒内恢复的，不影响实时同步，如果超过30秒，则会导致同步任务失败。 ● 目标数据库与源数据库字符集不一致可能会导致同步后数据不一致或者同步失败。 ● 数据类型不兼容时，可能引起同步失败。 ● 源库表同步至目标库后分布方式为哈希分布，暂不支持复制分布。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 源库和目标库均相同的任务不允许出现重复同步的情况，如：A任务和B任务同时将源库的同一张表的实时同步到目标库的同一张表中，可能导致数据不一致和同步失败。 ● DRS全量同步表结构时，源库中char、varchar、nvarchar、enum、set字符类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），扩大倍数和源库中列的字符集有关。如源库字符集为UTF8的按照字节长扩大3倍，为UTF8MB4的按照字节长扩大4倍。 ● DRS全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 增量同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步过程中，不支持源库DDL的复制。 ● 增量同步场景下，不支持源数据库进行恢复操作。 ● 增量同步过程中，不支持delete和update目标端主键类型为tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary的表。 ● 任务启动、任务全量同步阶段，不建议做删除类型的DDL操作，可能会引起任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库进行恢复操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。 ● 如果无主键表中包含大字段（tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary），增量同步阶段存在大字段数据不一致的可能性。 ● 由于浮点型数据是采用近似值存储，而不是作为精确值存储，此外操作系统平台和底层实现也对浮点型数据有影响，因此FLOAT，DOUBLE类型的数据是不精确的，所以在MySQL和GaussDB同步浮点型数据，存在一定差异的可能性，可参考MySQL的官方文档。 ● 由于MySQL与GaussDB针对零时间（"0000-00-00 00:00:00"）内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。 ● 不支持两阶段事务。 ● 源端同步的表如果有AUTO_INCREMENT属性，结束任务时，DRS会自动刷新同步表的序列整型列所对应的GaussDB自增列（sequence）起始值，刷新值为该列的最大值+10000。

操作步骤

本小节以MySQL到GaussDB分布式版的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-23 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-27 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-24 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎操作，以免因为配置项错误，需要重新创建任务。

* 数据流动方向: [入云] [出云] [双向] ?

* 源数据库引擎: MySQL Oracle DB2 for LUW DBM MariaDB MongoDB PostgreSQL Microsoft SQL Server GaussDB(for MySQL) TDB

* 目标数据库引擎: MySQL GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版 PostgreSQL GaussDB(for MySQL)

* 网络类型: 公网网络 ?

* 目标数据库实例: [无可用实例] ? 查看数据库实例 查看不可用实例

* 同步实例所在子网: [选择子网] ? 查看子网 查看已占用的IP地址

* 同步模式: [全量] [增量] ?

表 3-28 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	可用的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 3-25 任务类型



表 3-29 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-26 企业项目和标签



表 3-30 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。 如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-27 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中使用时参数模板设定

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-31 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-28 目标库信息

目标库信息

数据库实例名称: [模糊处理]

数据库用户名: [输入框]

数据库密码: [输入框]

测试连接 ✔ 测试成功

表 3-32 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-29 同步模式

同步模式: 限速 不限速

全量同步对象类型: 同步表结构 同步数据 同步索引

增量同步对象选择: 忽略 按库 按表

同步对象: 全部同步 导入对象文件

对展开的库进行搜索，支持正则表达式。


- chen hao database
- db_a_double_az_001 database
- db_kernel001sync_kernel_001 database
- db_snarpe_inneed_001 database
- db_snarpe_inneed_001_three database
- db_snarpe_inneed_001_two database
- db_testless_compare_001 database
- db_test2ess_compare_001 database
- db_uxtargetcloud_003 database
- db_yssel_to_pg_at_001 database
- dengvuhuan1 database
- Auh1 database

对展开的库进行搜索，支持正则表达式。

- database
- _rdstords_fromrds_vpc_001 编辑
- db_1_e_az_reedit_001 编辑

表 3-33 同步模式和对象

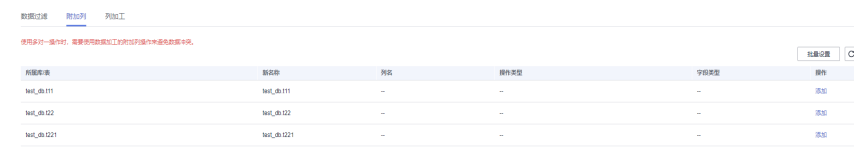
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-30 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 3-31 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-32 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-33 任务启动设置




表 3-34 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.4 将 MySQL 同步到 GaussDB 主备版

支持的源和目标数据库

表 3-35 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 RDS for MySQL 	GaussDB主备版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-36 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账户需要具备如下权限： <ul style="list-style-type: none"> - 全量同步：SELECT。 授权语句：GRANT SELECT ON <database>.<table> to <user>; - 全量+增量和增量同步：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 授权语句：GRANT SELECT, LOCK TABLES ON <database>.<table> to drsUser; GRANT REPLICATION SLAVE, REPLICATION CLIENT ON *.* to <user>; • 目标库GaussDB账户需要具备如下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>; - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> • 仅支持同步表结构、表数据、索引信息，不支持同步存储过程等其他数据库对象。 • 增量同步不支持同步DDL。 • 虚拟列会被同步为普通列，且不写入任何数据，也不做增量同步。 • 不支持的数据类型有：xml、包含bit类型的json、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon、interval。 • 仅支持MyISAM和InnoDB表的同步。 • 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 • 对象名同步到目标库后会转换成小写，因此选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 • 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 • 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 • 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 • 增量同步时，必须设置MySQL源数据库的server_id。MySQL5.7，server_id的取值范围在1 - 4294967296之间。 • 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 • 源数据库中的库名、表名不能包含：.<'>/\以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> • 目标数据库实例的运行状态必须正常。 • 目标数据库实例必须有足够的磁盘空间。 • 目标数据库的时区设置必须与源数据库一致。 • 任务配置的映射数据库必须在目标库已经存在。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，在全量同步中，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由"原库名_原表名_原索引名"计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key。 ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 网络中断在30秒内恢复的，不影响实时同步，如果超过30秒，则会导致同步任务失败。 ● 目标数据库与源数据库字符集不一致可能会导致同步后数据不一致或者同步失败。 ● 数据类型不兼容时，可能引起同步失败。 ● 源库表同步至目标库后分布方式为哈希分布，暂不支持复制分布。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 源库和目标库均相同的任务不允许出现重复同步的情况，如：A任务和B任务同时将源库的同一张表的实时同步到目标库的同一张表中，可能导致数据不一致和同步失败。 ● DRS全量同步表结构时，源库中char、varchar、nvarchar、enum、set字符类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），扩大倍数和源库中列的字符集有关。如源库字符集为UTF8的按照字节长扩大3倍，为UTF8MB4的按照字节长扩大4倍。 ● DRS全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 增量同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步过程中，不支持源库DDL的复制。 ● 增量同步场景下，不支持源数据库进行恢复操作。 ● 增量同步过程中，不支持delete和update目标端主键类型为tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary的表。 ● 任务启动、任务全量同步阶段，不建议做删除类型的DDL操作，可能会引起任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库进行恢复操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。 ● 如果无主键表中包含大字段（tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary），增量同步阶段存在大字段数据不一致的可能性。 ● 由于浮点型数据是采用近似值存储，而不是作为精确值存储，此外操作系统平台和底层实现也对浮点型数据有影响，因此FLOAT，DOUBLE类型的数据是不精确的，所以在MySQL和GaussDB同步浮点型数据，存在一定差异的可能性，可参考MySQL的官方文档。 ● 不支持两阶段事务。 ● 由于MySQL与GaussDB针对零时间（"0000-00-00 00:00:00"）内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。 ● 源端同步的表如果有AUTO_INCREMENT属性，结束任务时，DRS会自动刷新同步表的序列整型列所对应的GaussDB自增列（sequence）起始值，刷新值为该列的最大值+10000。

操作步骤

本小节以MySQL到GaussDB主备版的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-34 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-37 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-35 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免造成资源浪费，需要重新创建任务。

数据流动方向: [入云] [出云] [自建云] ?

源数据库引擎: MySQL, Oracle, DB2 for LUW, DDM, MariaDB, MongoDB, PostgreSQL, Microsoft SQL Server, GaussDB(for MySQL), TIB

目标数据库引擎: MySQL, GaussDB(DWS), GaussDB 分布式版, GaussDB 主备版, PostgreSQL, GaussDB(for MySQL)

网络类型: 公网网络 ?

目标数据库实例: [无可用实例] [查看数据库实例] [查看不可用实例]

同步实例所在子网: [选择子网] ? [查看子网] [查看已占用的IP地址]

同步模式: [全量] [增量] [全量] [增量] ?

表 3-38 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“GaussDB主备版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	<p>可用的GaussDB主备版实例。</p>
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 3-36 任务类型



表 3-39 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-37 企业项目和标签



表 3-40 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-38 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中使用时参数模板设定

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-41 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-39 目标库信息

目标库信息

数据库实例名称: [模糊]

数据库用户名:

数据库密码:

✔ 测试成功

表 3-42 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-40 同步模式

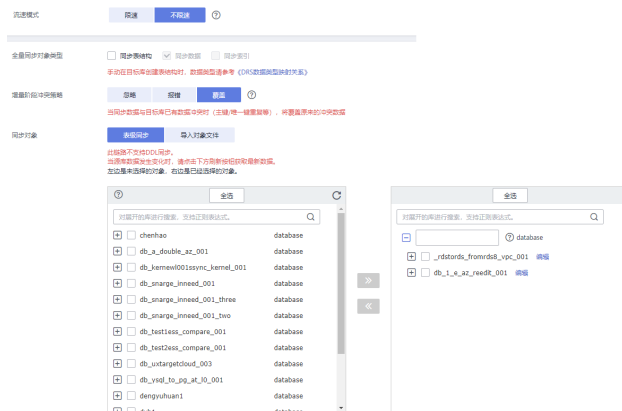



表 3-43 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-41 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 3-42 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-43 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-44 任务启动设置




表 3-44 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.5 将 MySQL 同步到 GaussDB(DWS)

支持的源和目标数据库

表 3-45 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MySQL数据库ECS自建MySQL数据库其他云上MySQL数据库RDS for MySQL	<ul style="list-style-type: none">GaussDB(DWS)集群 8.1.3、8.2.0版本

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-46 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账号需要具备如下权限：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 目标数据库账号必须具有每张表的如下权限：INSERT、SELECT、UPDATE、DELETE、CONNECT、CREATE。
同步对象约束	<ul style="list-style-type: none"> • 支持表、普通索引（B-Tree索引）、约束（主键、空、非空）的同步，不支持视图、外键、存储过程、触发器、函数、事件、虚拟列、唯一约束、唯一索引的同步。 • 全量同步支持comment，增量不支持。 • 不支持的数据类型有：xml、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon。 • 仅支持MyISAM和InnoDB表的同步。 • 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 • 对象名同步到目标库后会转换成小写，因此选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 • 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。 • 同步对象列名不能为CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID等GaussDB(DWS)禁止的字段，否则会导致任务失败。 • 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 必须设置MySQL源数据库的server - id，server - id的取值范围在2 - 4294967296之间。 ● 源数据库中的库名、表名不能包含：.<'>`\"以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 目标数据库的时区设置必须与源数据库一致。如果时区设置不一致，同步后会涉及时间数据类型的值显示不一致，修改为统一时区后数据显示一致。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：索引引用表等。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库进行恢复操作。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 同步无主键表时，如果存在唯一键，可能因数据冲突导致写入失败。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。 ● 如果无主键表中包含大字段（blob、text、clob、nclob、bytea），增量同步阶段存在大字段数据不一致的可能性。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 源库为RDS for MySQL实例时，支持源端多张表对 GaussDB(DWS)一张表的映射。详细操作可参考创建MySQL到GaussDB(DWS)同步任务。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● MySQL 5.7版会过滤check约束，如果源数据库为5.7版时，不建议使用check约束，可能会导致数据不一致。 ● 网络中断在30秒内恢复的，不影响实时同步，如果超过30秒，则会导致同步任务失败。 ● MySQL的唯一约束（unique key）在同步到目标GaussDB(DWS)后会改写成普通约束，主键（primary key）的同步保持不变。 ● 目标数据库与源数据库字符集不一致可能会导致同步后数据不一致或者同步失败。 ● 数据类型不兼容时，可能引起同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 实时同步过程中，若源库为RDS时，支持修改端口，修改之后同步任务失败，需要通过重试后继续进行同步。 ● 实时同步过程中，若源库为非RDS时，不支持修改端口。 ● 实时同步过程中，不支持IP、账号、密码修改。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB(DWS)中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名 (可能被截断)+_key。其中哈希值由“原库名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：哈希值+原约束名 (可能被截断)+_key，或表名_原索引名。 ● 全量阶段基于索引实现上差异，只保留普通B-Tree索引，其他索引默认不同步。GaussDB(DWS)索引过多时会影响存储空间和入库性能，建议用户可根据业务需求按需建索引。 ● DRS全量同步表结构时，源库中char、varchar、nvarchar、enum、set字符类型长度在目标库会按照字节长自动扩大 (因为目标库为字节长)。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。增量阶段同步分区表会依赖GaussDB(DWS)支持的类型创建分区表，可能会创建失败。 ● 全量同步timestamp类型时，默认值中的on update current_timestamp语法将不会同步到目标库GaussDB(DWS)中。 ● 增量同步过程中支持DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, ALTER_COLUMN, DROP_CONSTRAINT、ADD_CONSTRAINT、CREATE_INDEX, DROP_INDEX, RENAME_INDEX, RENAME_COLUMN, 客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 多对一场景下，执行column重命名操作，必须停业务操作，不然会有数据不一致的风险。 - 多对一场景下，推荐只同步加列DDL，其他的DDL同步可能会因为目标表被修改而导致任务失败或数据不一致。 - 多对一场景下，执行ADD_COLUMN时，需保证每张表加列的类型一致，否则有可能导致任务失败。 - 新增和修改表名、列名、索引名时不能超出63字符，否则任务会失败。 - 执行DDL语句创建索引时，语句中表名+索引名超出63字符可能出现重名情况，造成索引创建失败。 - 源库无主键表增加主键的时候，必须含有第一列，否则任务会失败。 - 增量阶段执行DDL操作时，如果目标表不存在，则该DDL会被忽略掉。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 增量阶段，源数据库执行CHANGE COLUMN修改列信息，如果该列在目标数据库GaussDB(DWS)中为分布列，则该语句会被忽略，因为GaussDB(DWS)不支持修改分布列。 - 增量阶段，由于源和目标数据库索引规则不一样，所以不支持RENAME INDEX的DDL操作。 - 增量阶段，不建议CHAR(0)与其他字符拼接进行同步，例如插入CONCAT('a',CHAR(0),'b')，可能造成数据不一致。 - 增量阶段，JSON数据类型里不建议CHAR(34)（双引号）、CHAR(92)（反斜杠）与其他字符拼接进行同步，例如插入JSON_OBJECT('\$0', CONCAT('a', CHAR(34), 'b'))，同步到目标端后会多出一个反斜杠，造成数据不一致。 <ul style="list-style-type: none"> ● 增量同步DDL中的SQL来源于用户执行的原生SQL，全量同步的表结构语法是经过SQL语法层处理之后的SQL语句，两者之前存在一些类型或者约束差异，可能会导致全量和增量创建的结构或者字段类型不同，详细说明可参考同步到GaussDB(DWS)的使用建议。 ● 增量同步过程中，不支持指定位置加列的DDL操作（例如ALTER TABLE ddl_test ADD COLUMN c2 AFTER/FIRST c1;），DRS会删除AFTER/FIRST属性，可能会导致列顺序不一致。 ● 增量同步过程中，库级同步不支持Online DDL，表级同步目前只支持阿里云DMS产生的Online DDL。 ● 增量同步过程中，库级同步新建的无主键表必须包含以下列中的其中之一，详细说明可参考GaussDB(DWS)参数说明。 INTEGER TYPES: TINYINT、SMALLINT、INT、BIGINT、NUMERIC/DECIMAL; CHARACTER TYPES: CHAR、BPCHAR、VARCHAR、VARCHAR2、NVARCHAR2; DATE/TIME TYPES: DATE、TIME、TIMETZ、TIMESTAMP、TIMESTAMPTZ、INTERVAL、SMALLDATETIME; 无主键表会指定可作为分布列的其中一列作为分布列，如果所有列类型都无法作为分布列，会导致建表失败，DRS同步中断。 ● 增量同步过程中，不支持delete和update目标端主键类型为binary, text, blob, clob的表。 ● 库级同步，在增量同步过程中，不能在源库创建名称字母相同但大小写不同的表，否则会导致其中一张表无法同步。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 源端同步的表如果有AUTO_INCREMENT属性，结束任务时，DRS会自动刷新同步表的序列整型列所对应的GaussDB(DWS)自增列（sequence）起始值，刷新值为该列的最大值+10000。

操作步骤

本小节以RDS for MySQL到GaussDB(DWS)多对一场景的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-45 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-47 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-46 同步实例信息

同步实例信息

以下信息确认后不可修改，请谨慎填写，以免因配置错误造成数据丢失或迁移失败。

* 数据源方向: [按钮: 入云, 出云, 数据同步] ?

DRS要求源数据库和目标数据库至少有一方为本云数据库实例。入云指目标数据库为本云数据库实例的场景，本云数据库实例之间的同步，请返回入云。

* 源数据库引擎: MySQL, Oracle, DB2 for LUW, DGM, GaussDB(for MySQL)分布式版, MongoDB, PostgreSQL, Microsoft SQL Server

* 目标数据库引擎: MySQL, GaussDB(DWS), GaussDB 分布式版, GaussDB 主备版, PostgreSQL, GaussDB(for MySQL)

MySQL、Oracle等数据库与GaussDB不同，存在DOL同步失败的可能性，建议先做实验验证GaussDB DWS支持DOL同步。

* 网络类型: VPC网络 ?

* DRS任务类型: 单AZ, 双AZ

采用单个节点部署策略，与多AZ部署方式相比，每个任务只创建一个节点，也具有高性价比，适用于业务数据量较小、周期较长，并对容灾要求不严格的场景。

* 同步数据库实例: [下拉菜单] ? 数据源数据库实例 数据不可用实例

同步实例所在子网: [下拉菜单] ? 数据子网 数据源已占用IP地址

* 同步模式: [按钮: 全量, 增量, 全量+增量] ?

该模式为数据库实时同步，通过全量过程完成目标数据库的初始化，增量同步的通过解析日志实现，增量同步的通过数据库日志实现，增量同步和数据库数据一致性。

表 3-48 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“GaussDB(DWS)”。
网络类型	<p>此处以“VPC网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	可用的GaussDB(DWS)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将用户选择的数据库对象和数据一次性同步至目标端数据库。 - 增量 该模式通过解析日志等技术，将源端产生的增量数据持续同步至目标端数据库。

● 任务类型

图 3-47 任务类型



表 3-49 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-48 可用区</p>

- 企业项目和标签

图 3-49 企业项目和标签



表 3-50 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-50 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中使用参数模板设定

数据库类型 自建库 RDS实例

数据库实例名称 [查看数据库实例](#) [查看不可选实例](#)

数据库用户名

数据库密码

SSL安全连接

测试成功 [?](#)

表 3-51 源库信息

参数	描述
数据库类型	源数据库类型，可选“自建库”和“RDS实例”，此处以“RDS实例”为示例。
数据库实例名称	选择待同步的RDS实例。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该任务后自动清除。

图 3-51 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

测试成功

表 3-52 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(DWS)实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象。单击“下一步”。

图 3-52 同步模式

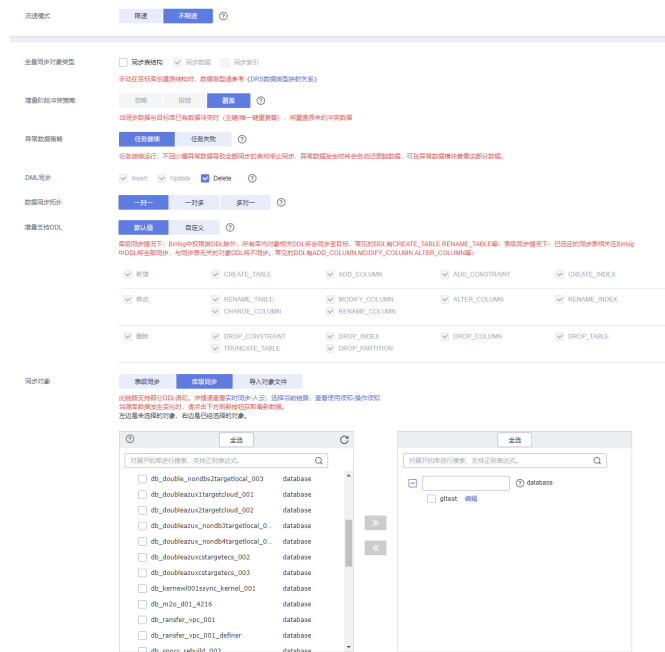



表 3-53 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-53 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。 <p>由于GaussDB(DWS)的表结构具有业务逻辑，建议结合业务情况在GaussDB(DWS)专家指导下自建表结构，并确保表名、列名、列类型的准确性，否则会出现数据同步失败、数据同步精度丢失等情况。</p>

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
异常数据策略	<p>源和目标数据库的客观差异差异，可能导致部分数据无法写入成功，出现异常数据，请选择出现异常数据时的任务策略。</p> <ul style="list-style-type: none"> ● 任务继续：不因少量异常数据导致全部数据均停止同步，异常数据发生时将会自动记录脏数据，可在“异常数据”模块查看这部分数据。 ● 任务失败：出现异常数据后，任务失败停止。
DML同步	选择需要同步的DML操作，默认都勾选。
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明 “数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>

参数	描述
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> ● “默认值”： <ul style="list-style-type: none"> - 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 - 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 ● “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 <p>说明</p> <ul style="list-style-type: none"> ● “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 ● 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 ● 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 如果同步的对象在源数据库和目标数据库中的命名不同，可在右侧已选对象框中修改目标库对象映射名称。也可通过对象名映射功能实现多张表对一张表的同步，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用多对一操作时，需要使用数据加工的附加列操作来避免数据冲突。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 源库和目标库多对一的表的结构要一致。 选择“导入对象文件”，具体步骤和说明可参考导入同步对象。文件导入对象也支持多张表对一张表的映射，由于消息体限制最多导入10000个表（表名长度过长或者规则过长也会影响导入数量）。 <p>说明</p> <ul style="list-style-type: none"> 选择数据的时候支持对展开的库搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以保持待选择的对象为最新源数据库对象。 对象名同步到目标库后会转换成小写，因此选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。 选择库级同步时，在增量同步过程中，不能在源库创建名称字母相同但大小写不同的表，否则会导致其中一张表无法同步。

步骤5 在“数据加工”页面，可对需要加工的表对象进行数据过滤或添加附加列，单击“下一步”。

- 如果需要设置数据过滤，选择“数据过滤”，设置相关过滤规则。
- 如果需要设置添加附加列，选择“附加列”，单击“操作”列的“添加”，选填需要添加的列名和操作类型信息。

相关操作可参考[数据加工](#)。

图 3-54 数据加工



源库/表	目标库/表	别名	操作类型	字段类型	操作
db1.jack_011	db1.jack_011	--	--	--	添加
db1.test2	db1.test2	--	--	--	添加
db1.test3	db1.test3	--	--	--	添加

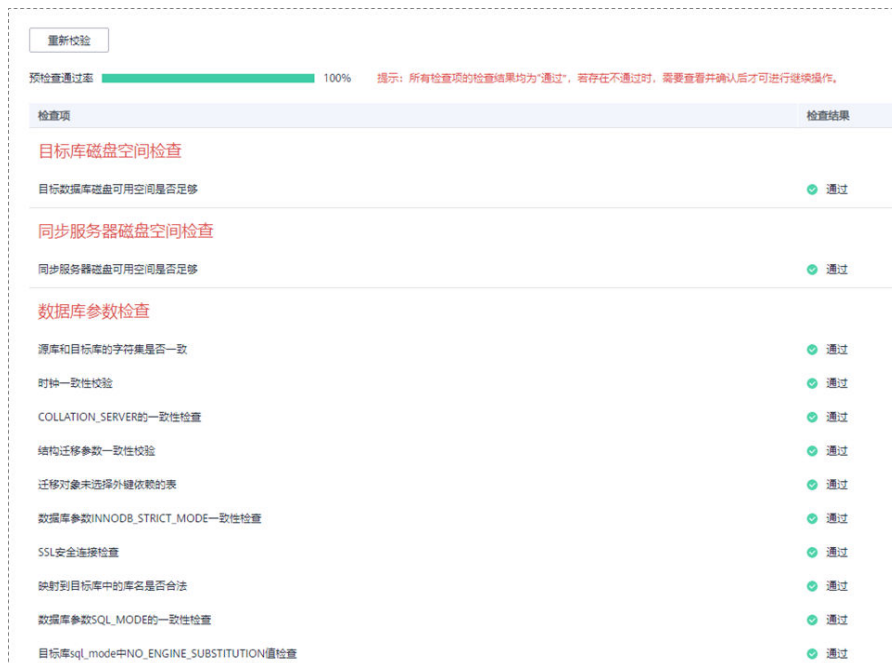
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-55 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 3-56 任务启动设置



表 3-54 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会出现一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.6 将 MySQL 同步到 GaussDB(for MySQL)

支持的源和目标数据库

表 3-55 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建MySQL数据库 ECS自建MySQL数据库 其他云上MySQL数据库 RDS for MySQL 	<ul style="list-style-type: none"> GaussDB(for MySQL)主备实例

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-56 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none">• 源数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。• 目标数据库账号必须拥有如下权限：SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES，GaussDB(for MySQL)实例的root账户默认已具备上述权限。 当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步。 ● 视图、存储过程、函数依赖的表不支持做表名映射。 ● 任务做表名映射时，不支持该表的外键约束同步。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 仅支持MyISAM和InnoDB表的同步。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 ● 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。 ● 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名、表名、视图名不能包含：'<'>\以及非ASCII字符。 ● 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● 不支持非MyISAM和非InnoDB的表同步到RDS。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibttmp1。 ● 数据库映射时，源库中存在视图、存储过程等对象，可能会导致失败。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例必须是GaussDB(for MySQL)主备实例。 ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 ● 源库不允许存在与目标库同名的无主键表。 ● 源库和目标库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源或者目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 任务创建后，目标库不能设置为只读。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 当在全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量同步过程中不支持DDL操作。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN,

类型名称	使用和操作限制
	<p>CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY, 客户可以根据自身需求, 在对象选择页面选择需要同步的DDL类型。</p> <ul style="list-style-type: none"> - 增量同步支持表的重命名, 源表和目标表必须都在对象选择里。 • 使用了附加列功能, 单表的列数超过500时, 对该表添加附加列可能会超过列数上限, 会导致任务失败。 • 建议将expire_log_day参数设置在合理的范围, 确保恢复时断点处的binlog尚未过期, 以保证服务中断后的顺利恢复。 • 需要DRS任务同步索引的情况下, 目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引, 增量阶段DDL根据索引名操作索引时会触发错位问题。

操作步骤

本小节以MySQL到GaussDB(for MySQL)的实时同步为示例, 介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面, 单击“创建同步任务”。

步骤2 在“同步实例”页面, 填选区域、项目、任务名称、描述、同步实例信息, 单击“开始创建”。

- 任务信息

图 3-57 同步任务信息

1、该页面仅有任务名称和描述可修改, 其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源, 在配置完源库及目标库基本信息后, 系统会去创建资源, 为后续数据迁移做准备, 虚拟资源一旦创建好后, 就不能修改。

区域: [Dropdown] ⓘ
不同区域的资源之间内网不互通, 请选择靠近您客户的区域, 可以降低网络时延, 提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ⓘ

描述: [Text Area] ⓘ
0/256

表 3-57 任务和描述

参数	描述
区域	当前所在区域, 可进行切换。

参数	描述
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 3-58 同步实例信息



表 3-58 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“GaussDB(for MySQL)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。

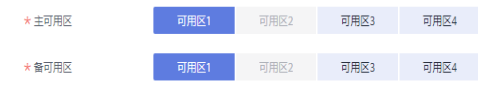
参数	描述
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	<p>可用的“GaussDB(for MySQL)”实例。</p>
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 <p>该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。</p> <ul style="list-style-type: none"> - 增量。 <p>增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。</p> <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 3-59 任务类型



表 3-59 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-60 可用区</p> 

- 企业项目和标签

图 3-61 企业项目和标签



表 3-60 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-62 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中[使用参数模板设置](#)

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-61 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-63 目标库信息

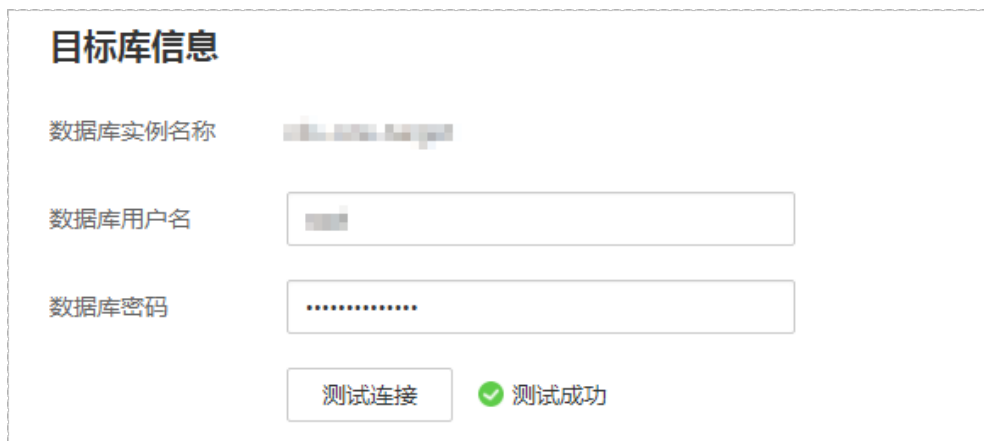


表 3-62 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-64 同步模式

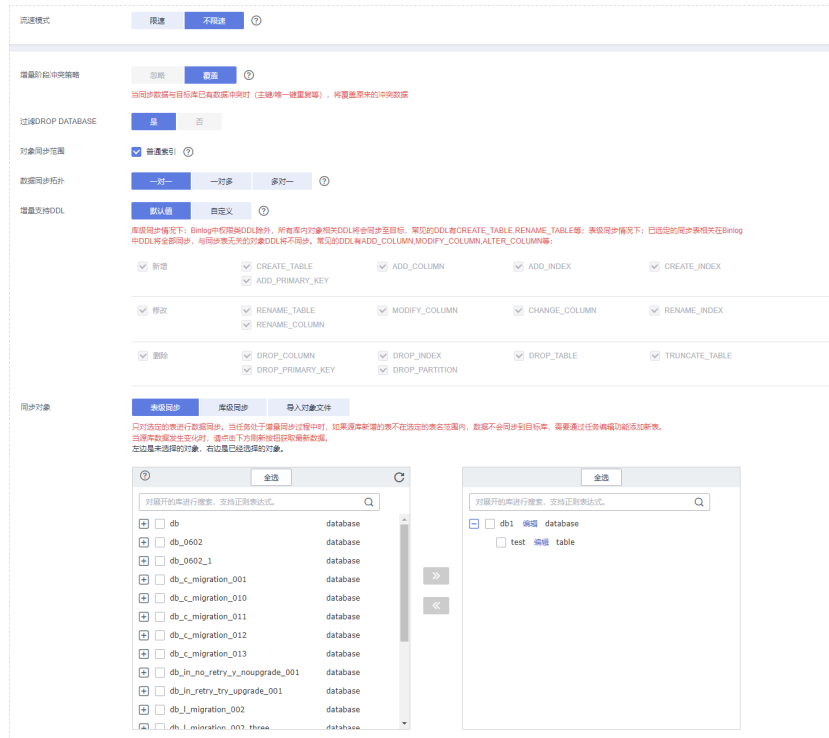



表 3-63 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-65 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>增量阶段冲突策略</p>	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>选择是否同步“普通索引”。</p> <p>DRS将默认同步主键/唯一索引，普通索引是指除主键/唯一索引以外的其他类型索引。勾选普通索引将会同步全部的索引，不勾选则仅同步主键/唯一索引，普通索引不会同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，可对需要加工的表对象进行数据过滤或添加附加列，完成后单击“下一步”。

- 如果需要设置数据过滤，选择“数据过滤”，设置相关过滤规则。
- 如果需要设置添加附加列，选择“附加列”，单击“操作”列的“添加”，选填需要添加的列名和操作类型信息。

相关操作可参考[数据加工](#)。

图 3-66 数据加工



表名	库名	操作类型	字符类型
db1.table_011	db1.table_011	添加	
db1.table_012	db1.table_012	添加	
db1.table_013	db1.table_013	添加	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-67 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。

图 3-68 任务启动设置




表 3-64 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.7 将 MySQL 同步到 MariaDB

支持的源和目标数据库

表 3-65 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本） ECS自建MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本） 其他云上MySQL数据库（MySQL 5.5、5.6、5.7、8.0版本） RDS for MySQL（5.5、5.6、5.7、8.0版本） 	RDS for MariaDB 10.5版本

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考表 3-66进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

表 3-66 数据库账号权限

类型名称	全量同步	全量+增量同步、增量同步
源数据库连接账号	需要具备如下最小权限： SELECT、SHOW VIEW、EVENT。	需要具备如下最小权限： SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。
目标数据库连接账号	需要具备如下最小权限： SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES，RDS for MariaDB实例的root账户默认已具备上述权限。	

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考表3-67。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-67 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none">支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。仅支持MyISAM和InnoDB表的同步。不支持事件（EVENT）和触发器（TRIGGER）的同步。不支持源库中同步对象的表结构中带加密表标识（ENCRYPTION）。视图、存储过程、函数依赖的表不支持做表名映射。任务做表名映射时，不支持该表的外键约束同步。任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如：<pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

注意事项

为保证任务的正常运行，DRS提供了自动化的检查方法，在您启动DRS任务前，会针对源数据库和目标数据库的相关配置、条件进行检查。主要的检查项和对应的处理建议请参考[预检查项](#)。除了预检查项目之外，您还需要注意的事情参见[表3-68](#)。

表 3-68 注意事项

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的表名、视图名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibttmp1。 ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● 除了MySQL和MariaDB系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● DRS同步时会有大量数据写入目标库，目标库max_allowed_packet参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 ● 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 对于源数据库和目标数据库的版本有差异的同步链路，由于不同版本之间的部分特性差异，可能会出现语法兼容性问题，详情可参考MySQL或MariaDB不同版本之间的语法差异。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 进行表级多对一同步时，源库不允许存在无主键表。 ● 源库不允许存在与目标库同名的无主键表。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 如果源库MySQL不支持TLS1.2协议，需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 不支持分区表的分区字段进行列映射。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 任务创建后，目标库不能设置为只读。 ● 实时同步过程中，如果修改了源库或者目标库的用户名、密码，会导致同步任务失败，需要在数据复制服务控制台将上述信息重新修改正确，然后重试任务可继续进行实时同步。一般情况下不建议在同步过程中修改上述信息。 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 当在全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量同步过程中不支持DDL操作。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY, 客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 增量同步支持表的重命名，源表和目標表必須都在对象选择里。 ● 表级同步时，增量同步过程支持使用Online DDL，可参考DRS实时同步支持使用Online DDL工具吗。 ● 增量同步支持任务再编辑追加同步对象。 ● 表级同步中，对于多对一场景下的表在数据加工页面设置了附加列为来源列，如果目标库中已经存在多对一映射之后的表，请在目标库中删除这张表或者清除这张表的数据，否则将不会通过附加列来创建联合主键（这在同步过程中会导致数据冲突并忽略，从而到导致数据不一致）。 ● 源库MySQL版本为8.0时，部分排序字符集（utf8mb4_0900_as_ci、utf8mb4_0900_as_cs、utf8mb4_0900_bin、utf8mb4_cs_0900_ai_ci等支持Unicode 9.0的排序字符集），在目标库不支持。 ● 由于引擎和版本差异，部分MySQL中的函数在MariaDB中不支持，不支持的函数有：MBRCOVEREDBY、ST_BUFFER_STRATEGY、ST_GeoHash、ST_IsValid、ST_LatFromGeoHash、ST_LongFromGeoHash、ST_PointFromGeoHash、ST_SIMPLIFY、ST_VALIDATE、(8.0)JSON_ARRAYAGG、JSON_OBJECTAGG、JSON_PRETTY、JSON_STORAGE_FREE、JSON_STORAGE_SIZE、JSON_TABLE。 ● 目标数据库MariaDB执行分区交换的时候不支持VALIDATION语法验证，在源数据库MySQL执行分区交换SQL的时候不要带校验语法，例如： ALTER TABLE t1 EXCHANGE PARTITION p0 WITH TABLE t2 WITH VALIDATION; ● 增量过程中请不要在源库MySQL的json类型中写入\n、\t、\r等字符，否则会导致数据不一致。 ● 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

操作步骤

本小节以MySQL->RDS for MariaDB的实时同步为示例，介绍如何使用数据复制服务创建两个数据库实例之间的实时同步任务，其他存储引擎的配置流程类似。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-69 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-69 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含!= < > & ' " \ 特殊字符。

- 同步实例信息

图 3-70 同步实例信息

同步实例信息

以下配置项均不可修改，请谨慎选择，以免因配置不当造成数据丢失或业务中断。

源实例信息: [MySQL, Oracle, DB2 LUW, DDM, MariaDB, MongoDB, PostgreSQL, Microsoft SQL Server, GaussDB for MySQL, TDS]

目标实例信息: [MySQL, GaussDB(DWS), GaussDB 分析型数据库, GaussDB 主备版, MariaDB, PostgreSQL, GaussDB for MySQL]

实例名称: [Text Field]

实例名称类型: [下拉菜单]

实例名称格式: [Text Field]

实例名称在子网: [Text Field]

实例模式: [下拉菜单]

表 3-70 同步实例信息

参数	描述
数据流动方向	选择“入云”，即目标端数据库为本云数据库。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“MariaDB”。
网络类型	<p>此处以“公网网络”为示例。可根据业务场景选择“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	用户所创建的RDS for MariaDB实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 3-71 可用区



表 3-71 任务可用区

参数	描述
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-72 可用区</p>

- 企业项目和标签

图 3-73 企业项目和标签



表 3-72 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 同步实例创建成功后，在“源库及目标库”页面，填选源库信息和目标库信息后，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 3-74 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

● 测试成功

表 3-73 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库用户名所对应的密码。
SSL安全连接	<p>如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。</p> <p>说明</p> <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

- 目标库信息

图 3-75 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

SSL安全连接

● 测试成功

表 3-74 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS for MariaDB实例，不可进行修改。
数据库用户名	目标数据库对应的用户名。
数据库密码	目标数据库用户名对应的密码。
SSL安全连接	如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择数据冲突策略和同步对象，单击“下一步”。

图 3-76 同步模式

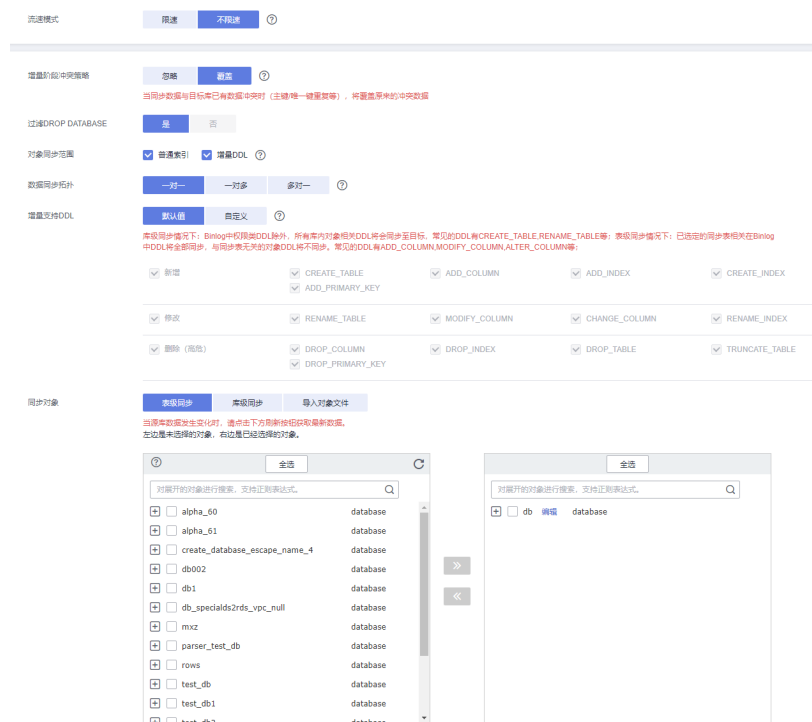




表 3-75 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-77 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>增量阶段冲突策略</p>	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 如果同步的对象在源数据库和目标数据库中的命名不同，可在右侧已选对象框中修改目标库对象映射名称，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 3-78 数据加工

源库对象	新名称	列名	操作类型	数据类型	操作
			以serverName@database@table为列	varchar(256)	操作
			默认值 123456789	long	
db_hai_pk_test1_table16 db_hai_pk_test1_table12 db_hai_pk_test1_table14 db_ha...			以create_time为列	datetime	
			以update_time为列	timestamp	
			表达式 concat(_current_database, @_current_table)	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

📖 说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-79 任务启动设置

The screenshot shows a configuration interface for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Latency Threshold (s)): A toggle switch is turned on.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned on.
- * 任务异常自动结束时间** (Task Abnormal Automatic End Time): An input field containing the number '14', with a help icon and a red note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (Task will automatically end after a period of abnormal state. Unit is days).

表 3-76 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.8 将 PostgreSQL 同步到 PostgreSQL

支持的源和目标数据库

表 3-77 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建数据库 (PostgreSQL 9.4、9.5、9.6、10、11、12、13、14、15 版本) • ECS自建数据库 (PostgreSQL 9.4、9.5、9.6、10、11、12、13、14、15 版本) • 其他云上数据库 (PostgreSQL 9.4、9.5、9.6、10、11、12、13、14、15 版本) • RDS for PostgreSQL (9.5、9.6、10、11、12、13、14、15版本) 	<p>RDS for PostgreSQL (9.5、9.6、10、11、12、13、14、15版本)</p> <p>说明 仅支持目标库大版本等于或高于源库大版本。</p>

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-78](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-78 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 不支持实例级同步：每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持的字段类型： 数字类型、货币类型、字符类型、二进制数据类型、日期/时间类型、布尔类型、枚举类型、几何类型、网络地址类型、位串类型、文本搜索类型、UUID类型、XML类型、JSON类型、数组、复合类型、范围类型。 ● 全量同步的范围： <ul style="list-style-type: none"> - 库级同步支持：模式、表、索引、约束、视图、物化视图、序列、存储过程、规则、触发器、外键、排序规则、插件、编码转换信息、聚合函数、操作符、统计扩展、转换信息、文本搜索配置、函数、数据类型、类型转换、用户、事件触发器、文本搜索解析器、文本搜索模板的同步。 表级同步时，仅支持表、视图、物化视图、序列、用户、普通索引的同步。 文件导入时，支持表的同步。 - 不支持系统模式（“pg_”开头的任何模式、“information_schema”、“sys”、“utl_raw”、“dbms_lob”、“dbms_output”和“dbms_random”）、系统表、系统用户、表空间、外部数据包装器、外部服务器、用户映射、发布、订阅等其他对象。 <p>说明 支持同步的对象有如下限制：</p> <ul style="list-style-type: none"> ● 对象名称：库名不可以包含+"%\<>”，模式名和表名不可以包含"!<>”，列名不可以包含"和"。 ● 表：不同步临时表。表级同步时，表的约束、索引及规则会一起同步，表的触发器不会同步。 ● 模式：不同步public模式的权限。表级同步时，会同步目标库已存在模式的权限。 ● 函数：不同步C语言函数，不同步带有leakproof属性或带有support属性的函数。 ● 插件：不同步插件的元数据。 ● 数据类型：不同步基本数据类型。 ● 类型转换：不同步二进制强制型的类型转换。 ● 事件触发器：受限于目标库版本，仅目标库为RDS for PostgreSQL 11.11及以上版本时支持。 ● 文本搜索解析器：受限于目标库版本，仅目标库为RDS for PostgreSQL 11.11及以上版本时支持。 ● 文本搜索模板：受限于目标库版本，仅目标库为RDS for PostgreSQL 11.11及以上版本时支持。 ● 用户：不同步目标库已存在的用户，不同步用户的superuser属性、replication属性、bypassrls属性，不同步superuser用户的成员关系。对象owner/grantor为superuser用户时，不同步它的owner/grantor。目标库为华为云RDS for PostgreSQL时，要同步用户的密码不能包含用户名。表级同步时，不同步源库用户的默认访问权限。superuser用户同步至目标库后为普通用户，默认授予root权限（仅当目标库版本为RDS for PostgreSQL 11.11及以上时支持）。

类型名称	使用须知
	<ul style="list-style-type: none"> ● 增量同步的范围: <ul style="list-style-type: none"> - 支持部分DML (Data Manipulation Language) : 包括 INSERT、UPDATE、DELETE。 - 支持部分DDL (Data Define Language) : 包括TRUNCATE (仅PostgreSQL 11及以上版本支持)、CREATE SCHEMA、CREATE TABLE、DROP TABLE、ALTER TABLE (包含ADD COLUMN、DROP COLUMN、ALTER COLUMN、RENAME COLUMN、ADD CONSTRAINT、DROP CONSTRAINT、RENAME)、CREATE SEQUENCE、DROP SEQUENCE、ALTER SEQUENCE、CREATE INDEX、ALTER INDEX、DROP INDEX、CREATE VIEW、ALTER VIEW、COMMENT ON COLUMN、COMMENT ON TABLE、COMMENT ON SCHEMA、COMMENT ON SEQUENCE、COMMENT ON INDEX、COMMENT ON VIEW。 表级同步时, DDL仅支持TRUNCATE (仅PostgreSQL 11及以上版本支持)、DROP TABLE、ALTER TABLE、COMMENT ON COLUMN、COMMENT ON TABLE, 其中ALTER TABLE支持: ADD COLUMN、DROP COLUMN、ALTER COLUMN、RENAME COLUMN、ADD CONSTRAINT、DROP CONSTRAINT、RENAME。 - 不支持: 无日志表 (UNLOGGED TABLE) 的DML, 临时表的DML。 <p>说明 同步DDL的实现原理: 在源库通过事件触发器捕获DDL语句, 并且记录在特定的表中, 因此需要提前在源库创建事件触发器、函数等。详细操作可参考通过创建事件触发器和函数实现PostgreSQL增量DDL同步。</p>

数据库账号权限要求

在使用DRS进行同步时, 连接源库和目标库的数据库账号需要满足以下权限要求, 才能启动实时同步任务。不同类型的同步任务, 需要的账号权限也不同, 详细可参考[表 3-79](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查, 并给出处理建议。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号, 避免因为数据库账号密码修改, 导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后, 请尽快[修改DRS任务中的连接信息](#), 避免任务连接失败后自动重试, 导致数据库账号被锁定影响使用。

表 3-79 数据库账号权限

类型名称	全量同步	全量+增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限，系统表pg_catalog.pg_authid的select权限（用于同步用户的密码）。	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限，系统表pg_catalog.pg_authid的SELECT权限（用于同步用户的密码），无主键表的UPDATE、DELETE和TRUNCATE权限，REPLICATION连接权限。 说明 <ul style="list-style-type: none"> 无主键表的UPDATE、DELETE和TRUNCATE权限，仅用于对无主键表短暂加锁，以确保迁移后，目标库中无主键表数据与源库保持一致。 REPLICATION连接权限的添加方法： <ol style="list-style-type: none"> 在源数据库的“pg_hba.conf”配置文件的所有配置前增加一行配置“host replication <src_user_name> <drs_instance_ip>/32 <认证方式>”；认证方式可参考PostgreSQL官方文档pg_hba.conf文件配置，常见的认证方式有scram-sha-256等。 在源库使用SUPERUSER用户执行语句“select pg_reload_conf();”生效，或重启数据库实例生效。
目标数据库连接账号	<ul style="list-style-type: none"> 库级同步： <ul style="list-style-type: none"> 如果目标库不是postgres，需要具有CREATEDB权限。 如果目标库是postgres，需要具有postgres库的CONNECT和CREATE权限、对模式public的USAGE和CREATE权限。 表级同步： <ul style="list-style-type: none"> 如果需要同步库，需要具有CREATEDB权限。 如果需要同步模式，需要具有模式所在库的CONNECT权限、模式所在库上的CREATE权限。 如果需要同步模式下的对象，需要具有模式所在库的CONNECT权限、对象所在模式的USAGE权限、对象所在模式上的CREATE权限。 同步用户：需要具有CREATEROLE权限。 同步用户权限：同步用户的default privilege需要为系统默认值，否则可能导致目标库与源库的对象权限不一致。 说明 <ul style="list-style-type: none"> 同步事件触发器、文本搜索解析器、文本搜索模版时，不仅要求目标库版本为RDS for PostgreSQL 11.11及以上，还要求目标库连接用户为root用户或root用户的成员。 目标库版本为PostgreSQL 15时，库级同步及表级同步目标库没有建库的场景下，全量或增量同步的过程中，可能会由于没有模式public的USAGE及CREATE权限而导致同步失败。 	

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单增量和单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-80 使用须知

类型名称	使用和操作约定
任务启动前使用须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库的分区表触发器不可以设置为disable。 - 全量同步支持源库备机状态，但需要设置 hot_standby_feedback为on；增量同步不支持源库备机状态。 - 若要做增量同步： 源数据库的“pg_hba.conf”文件中包含如下的配置： host replication all 0.0.0.0/0 md5 源数据库参数wal_level必须配置为logical； 源数据库需提前安装test_decoding插件； 源数据库中无主键表的replica_identity属性必须为full； 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量； 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica_identity属性必须为full。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不可以存在同名的触发器。 - 同步对象依赖和关联的对象也须一起同步，否则可能导致同步失败。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标数据库的block_size参数值必须大于或等于源库中的对应参数值。 - 目标数据库和源数据库的lc_monetary参数值一致。 - 若要做增量同步，且同步对象包含外键、触发器或事件触发器，则目标数据库的session_replication_role参数必须设置为replica，同步结束后，此参数需改为原来的值。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标库不可以包含，与待同步对象类型相同且名称相同的对象，包括库、模式、表等。系统库、系统模式、系统表等除外。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为格式：<code>i_+哈希值+原索引名（可能被截断）+_key</code>。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：<code>c_+哈希值+原约束名（可能被截断）+_key</code>。 - 对于全量+增量和增量任务，启动前请确保源库中未启动长事务，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 对于全量+增量和增量任务，如果预校验存在“内部错误”且在任务未启动时结束，请参考PostgreSQL为源强制结束任务章节检查并删除流复制槽，防止源库中流复制槽残留。 - 对于全量+增量和增量任务，任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - 若选择同步DDL，须注意源库执行DDL时，确保在目标库上是兼容的。 <p>说明 同步DDL的实现原理是：在源库通过事件触发器捕获DDL语句，并且记录在特定的表中，然后将DDL语句同步至目标库。因此启动任务前，需要提前在源库创建事件触发器、函数等。详细操作可参考通过创建事件触发器和函数实现PostgreSQL增量DDL同步。</p>
全量同步过程使用须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程使用须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源数据库表的主键或者唯一键（主键不存在时），否则可能导致增量数据不一致或任务失败。 ● 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 ● 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。

类型名称	使用和操作约定
同步对比使用须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。
结束任务使用须知	<ul style="list-style-type: none"> 正常结束任务： <ul style="list-style-type: none"> 将自动重置目标库序列值（自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000），若选择了同步用户，任务结束时将自动同步用户的成员关系。 全量实时同步任务结束时会对序列值进行比对，序列值采用逻辑一致性的判断原则，即自增序列当且仅当目标库对应序列值大于等于源库对应序列值，自减序列当且仅当目标库对应序列值小于等于源库对应序列值时，序列值判定为一致。 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 全量+增量实时同步任务正常结束时，如果目标库 session_replication_role 值为 replica，需手动将目标库 session_replication_role 参数值修改为原来的值。 强制结束任务： <ul style="list-style-type: none"> 需要手动刷新目标库中的序列值，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 强制结束全量+增量实时同步任务，如果目标库 session_replication_role 值为 replica，需手动将目标库 session_replication_role 参数值修改为原来的值。 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。
常见故障排查	<ul style="list-style-type: none"> 在任务创建、启动、全量同步、增量同步、结束等过程中，如有遇到问题，可先参考“故障排查”章节进行排查。

操作步骤

本小节以PostgreSQL->PostgreSQL的入云同步为示例，介绍如何使用数据复制服务配置VPC网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-80 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-81 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-81 同步实例信息

同步实例信息

数据流动方向: 入云

源数据库引擎: PostgreSQL

目标数据库引擎: PostgreSQL

表 3-82 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“PostgreSQL”。
目标数据库引擎	选择“PostgreSQL”。

参数	描述
网络类型	<p>此处以“VPC网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为RDS for PostgreSQL数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - 增量： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 3-82 任务类型



表 3-83 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-83 企业项目和标签



表 3-84 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过连接账号和密码访问。

📖 说明

此处源库类型分为ECS自建库和RDS实例，需要根据源数据库的实际来源选择相应的分类。两种场景下的参数配置不一样，需要根据具体场景进行配置。

- 场景一：ECS自建库源库信息配置

图 3-84 ECS 自建库场景源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设置和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

数据库类型 ECS自建库 RDS实例

VPC [查看虚拟私有云](#)

子网 [查看子网](#)

IP地址或域名

端口

指定数据库名称

数据库用户名

数据库密码

SSL安全连接

● 测试成功 [?](#)

表 3-85 ECS 自建库场景源库信息

参数	描述
数据库库类型	选择“ECS自建库”。
VPC	源数据库实例所在的虚拟专用网络，可以对不同业务进行网络隔离。您需要创建或选择所需的虚拟私有云。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用分区内才会有效，创建源数据库实例的子网需要开启DHCP功能，在创建过程中也不能关闭已选子网的DHCP功能。
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
指定数据库名称	选择是否指定数据库，开启后需手动输入数据库名称。

参数	描述
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。

📖 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

- 场景二：RDS实例源库信息配置

图 3-85 RDS 实例场景源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导入用户和作业。

数据库类型 ECS自建库 **RDS实例**

数据库实例名称 🔍 查看数据库实例 查看不可选实例

数据库用户名

数据库密码 👁️

🟢 测试成功 ?

表 3-86 RDS 实例场景源库信息

参数	描述
数据库类型	选择“RDS实例”。
数据库实例名称	选择待同步的关系型PostgreSQL数据库实例作为源数据库实例。
数据库用户名	源数据库实例的用户名。
数据库密码	源数据库的用户名所对应的密码。

图 3-86 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码 👁️

表 3-87 目标库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的关系型PostgreSQL数据库实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象和同步用户，单击“下一步”。

图 3-87 同步对象

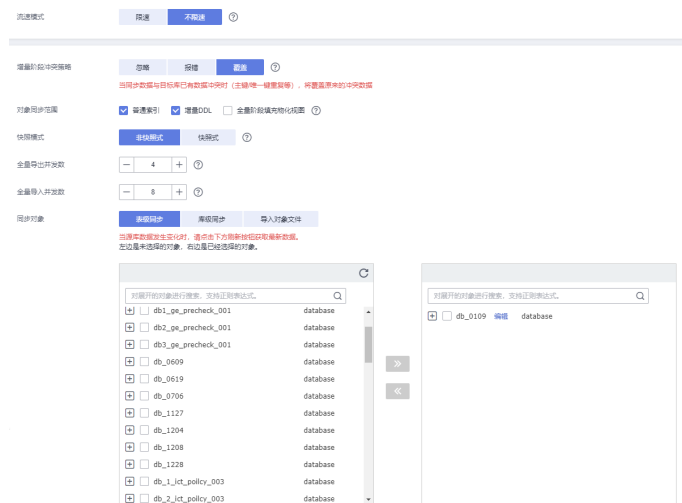


表 3-88 同步对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-88 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
对象同步范围	<p>对象同步范围支持选择普通索引、增量DDL同步和全量阶段填充物化视图，您可以根据业务需求选择是否进行同步。</p> <p>全量阶段填充物化视图：仅对源库已经填充的物化视图生效，该填充操作会影响全量同步性能，建议在全量同步完成后，手动在目标库填充。</p>
快照模式	<p>如果您选择的是全量+增量同步模式的任务，数据复制服务支持设置快照模式。</p> <p>全量阶段使用快照模式导出能够有效提升全量+增量场景下的数据同步效率，但PostgreSQL的快照机制会使导出期间数据库的历史数据不能被回收，可能有空间膨胀的现象。建议在全量或增量数据量大且源库磁盘空间充足的情况下使用该方式。</p>
全量导出并发数	全量导出对象的线程数量和导出数据的线程数量。
全量导入并发数	全量导入对象的线程数量和导入数据的线程数量。


参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 选择库级同步：全量同步时，同步对象为所选的库、库中对象的存量数据；增量同步时，同步除了无日志表、临时表外的所有表的DML及部分DDL。 选择表级同步：全量同步时，同步对象为所选的表、序列、视图或物化视图的存量数据；增量同步时，同步所选表的DML及部分DDL。 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为此格式：<code>i_+哈希值+原索引名</code>（可能被截断）+<code>_key</code>。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，同步后表上的原约束名将变为：<code>c_+哈希值+原约束名</code>（可能被截断）+<code>_key</code>。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。
同步用户	<p>数据库的同步过程中，同步用户需要进行单独处理。详细说明可参考表3-89。</p>

图 3-89 同步用户

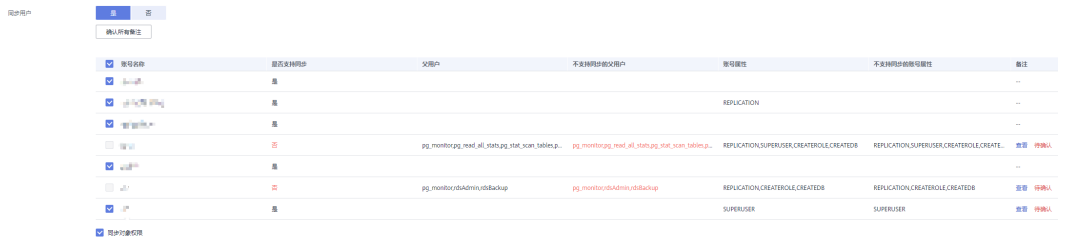


表 3-89 同步用户及权限

参数	描述
账号名称	源库对应的账号名称。

参数	描述
是否支持同步	账号是否支持同步。 一般分为两类：可同步的账号和不支持同步的账号。对于不支持同步的账号，在备注列的查看详情中会提示具体的原因。
父用户	账号所属父用户。
不支持同步的父用户	账号不支持同步的父用户。
账号属性	源库账号拥有的属性。
不支持同步的账号属性	目标库同步用户权限不足等原因导致的账号属性无法同步。
备注	不支持同步的父用户、不支持同步的账号属性的原因说明，所有备注项需确认后才可以进入下一步。
同步对象权限	是否同步账号对象对应的权限。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-90 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-91 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: "立即启动" (Immediately Start) and "稍后启动" (Start Later), with a help icon. Below are several settings:


- 任务异常通知设置**: A toggle switch is turned on.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: An input field containing the number "14", with a help icon and a red note: "任务处于异常状态一段时间后, 将会自动结束。单位为天。"

表 3-90 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.9 将 PostgreSQL 同步到 GaussDB(DWS)

支持的源和目标数据库

表 3-91 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建数据库 (PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本) • ECS自建数据库 (PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本) • 其他云上数据库 (PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本) • RDS for PostgreSQL (9.5、9.6、10、11、12、13、14版本) 	<ul style="list-style-type: none"> • GaussDB(DWS)集群

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-92](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-92 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 不支持实例级同步 每次至多同步一个库（database），同步多个库需要使用多个 DRS 任务。 ● 支持的字段类型 <ul style="list-style-type: none"> - 支持：tinyint、smallint、int、bigint、numeric、decimal、char、bpchar、varchar、text、date、time、timetz、timestamp、timestampz、interval 等 GaussDB(DWS)兼容 PostgreSQL 的数据类型。 - 不支持：xml、line、domain 类型同步。 ● 全量同步的范围 <ul style="list-style-type: none"> - 支持：模式、表、索引、约束、序列、自定义类型的同步。表级同步时，仅支持表、序列的同步。 - 不支持：系统模式（“pg_”开头的任何模式、“information_schema”、“sys”、“utl_raw”、“dbms_lob”、“dbms_output”和“dbms_random”）、系统表。 - 不支持：既是无主键表，又是分区表的自建表，可能会导致任务失败。 说明 支持同步的对象有如下限制： <ul style="list-style-type: none"> ● 对象名称：库名不可以包含“%”“<”，模式名和表名不可以包含“.”“<”，列名不可以包含“和”，列名不能为 CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID 等 GaussDB(DWS)禁止的字段。 ● 表：不同步临时表。表级同步时，表的约束、索引及规则会一起同步，表的触发器不会同步。 ● 模式：不同步 public 模式的权限。表级同步时，会同步目标库已存在模式的权限。 ● 增量同步的范围 <ul style="list-style-type: none"> - 支持：部分DML（Data Manipulation Language）：包括 INSERT、UPDATE、DELETE。 - 不支持：DDL（Data Define Language），无日志表（UNLOGGED TABLE）的 DML，临时表的 DML。 - 分区表的 DML 同步：若分区是新增分区，则该分区必须是新创建的分区，且该分区名不可以同已创建过的表名重复。该分区必须在所有 DML 数据都同步至目标库后才可以从源库删除。

数据库账号权限要求

在使用 DRS 进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考 [表 3-93](#) 进行赋权。DRS 会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-93 数据库账号权限

类型名称	全量同步	全量+增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限。	<p>数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限，无主键表的UPDATE、DELETE和TRUNCATE权限，REPLICATION连接权限。</p> <p>说明</p> <ul style="list-style-type: none"> • 无主键表的UPDATE、DELETE和TRUNCATE权限，仅用于对无主键表短暂加锁，以确保迁移后，目标库中无主键表数据与源库保持一致。 • REPLICATION连接权限的添加方法： <ol style="list-style-type: none"> 1. 在源数据库的“pg_hba.conf”配置文件的所有配置前增加一行配置“host replication <src_user_name> <drs_instance_ip>/32 <认证方式>”； 认证方式可参考PostgreSQL官方文档pg_hba.conf文件配置，常见的认证方式有scram-sha-256等。 2. 在源库使用SUPERUSER用户执行语句“select pg_reload_conf();”生效，或重启数据库实例生效。
目标数据库连接账号	<ul style="list-style-type: none"> • 库级同步：需要具有CREATEDB权限。 • 表级同步： <ul style="list-style-type: none"> - 如果需要同步库，需要具有CREATEDB权限。 - 如果需要同步模式，需要具有模式所在库的CONNECT权限、模式所在库上的CREATE权限。 - 如果需要同步模式下的对象，需要具有模式所在库的CONNECT权限、对象所在模式的USAGE权限、对象所在模式上的CREATE权限。 	

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-94 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库的分区表触发器不可以设置为disable。 - 全量同步支持源库备机状态，但需要设置 hot_standby_feedback为on；增量同步不支持源库备机状态。 - 若要做增量同步： 源数据库的“pg_hba.conf”文件中包含如下的配置： host replication all 0.0.0.0/0 md5 源数据库参数wal_level必须配置为logical。 源数据库需提前安装test_decoding插件。 源数据库中无主键表的replica identity属性必须为full。 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica identity属性必须为full。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 同步对象依赖和关联的对象也须一起同步，否则可能导致同步失败。 <p>说明 常见的关联关系：主外键关联表、表继承子表引用父表、表分区子分区表引用分区表、表自增列引用序列等。</p> ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标数据库的block_size参数值必须大于源库中的对应参数值。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标库不可以包含，与待同步对象类型相同且名称相同的对象，包括库、模式、表等。系统库、系统模式、系统表等除外。 - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知：

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 分区表只同步主表，转为普通表同步到目标库，分区表的数据将集中写入主表。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原约束名”计算得到：c_+哈希值+原约束名（可能被截断）+_key。 - 对于全量+增量和增量任务，启动任务前，请确保源库中未启动长事务，源库启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 对于全量+增量和增量任务，任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - DRS全量同步表结构时，源库中的character、character varying类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），默认扩大4倍，最大长度为10485760。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源数据库表的主键或者唯一键（主键不存在时），否则可能导致增量数据不一致或任务失败。 ● 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 ● 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。

类型名称	使用和操作约定
同步对比使用须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。
结束任务须知	<ul style="list-style-type: none"> 正常结束任务： <ul style="list-style-type: none"> 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 强制结束任务： <ul style="list-style-type: none"> 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。

操作步骤

本小节以PostgreSQL->GaussDB(DWS)的入云同步为示例，介绍如何使用数据复制服务配置VPC网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-92 同步任务信息

⚠ 1、该页面**仅有任务名称和描述可修改**，其他在点击开始创建后均不可修改。
 2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?	不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。
项目	<input type="text" value=""/>	
★ 任务名称	<input type="text" value="DRS-6435"/> ?	
描述	<div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> ?	

0/256

表 3-95 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 3-93 同步实例信息



表 3-96 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“PostgreSQL”。
目标数据库引擎	选择“GaussDB(DWS)”。
网络类型	<p>此处以“VPC网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为GaussDB(DWS)实例。

参数	描述
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 可用区

图 3-94 可用区



表 3-97 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-95 企业项目和标签



表 3-98 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none">- 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。- 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。- 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务同源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过连接账号和密码访问。

📖 说明

此处源库类型分为ECS自建库和RDS实例，需要根据源数据库的实际来源选择相应的分类。两种场景下的参数配置不一样，需要根据具体场景进行配置。

- **场景一**：ECS自建库源库信息配置

图 3-96 ECS 自建库场景源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中手动修改参数，手工导入用户和作业。

源库类型: ECS自建库 RDS实例

VPC: 查看虚拟私有云

子网: 查看子网

IP地址或域名:

端口:

指定数据库名称:

数据库用户名:

数据库密码: 显示/隐藏

SSL安全连接:

表 3-99 ECS 自建库场景源库信息

参数	描述
源库类型	选择“ECS自建库”。
VPC	源数据库实例所在的虚拟专用网络，可以对不同业务进行网络隔离。您需要创建或选择所需的虚拟私有云。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用分区内才会有效，创建源数据库实例的子网需要开启DHCP功能，在创建过程中也不能关闭已选子网的DHCP功能。
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
指定数据库名称	选择是否指定数据库，开启后需手动输入数据库名称。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

- 场景二：RDS实例源库信息配置

图 3-97 RDS 实例场景源库信息

表 3-100 RDS 实例场景源库信息

参数	描述
源库类型	选择“RDS实例”。
数据库实例名称	选择待同步的关系型PostgreSQL数据库实例作为源数据库实例。
数据库用户名	源数据库实例的用户名。
数据库密码	源数据库的用户名所对应的密码。

图 3-98 目标库信息

表 3-101 目标库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的GaussDB(DWS)实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象和同步用户，单击“下一步”。

图 3-99 同步模式

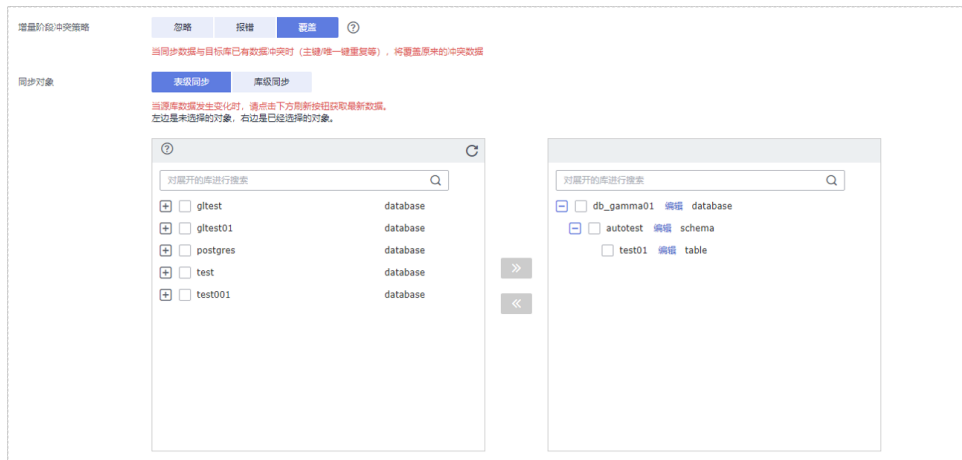



表 3-102 同步对象

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

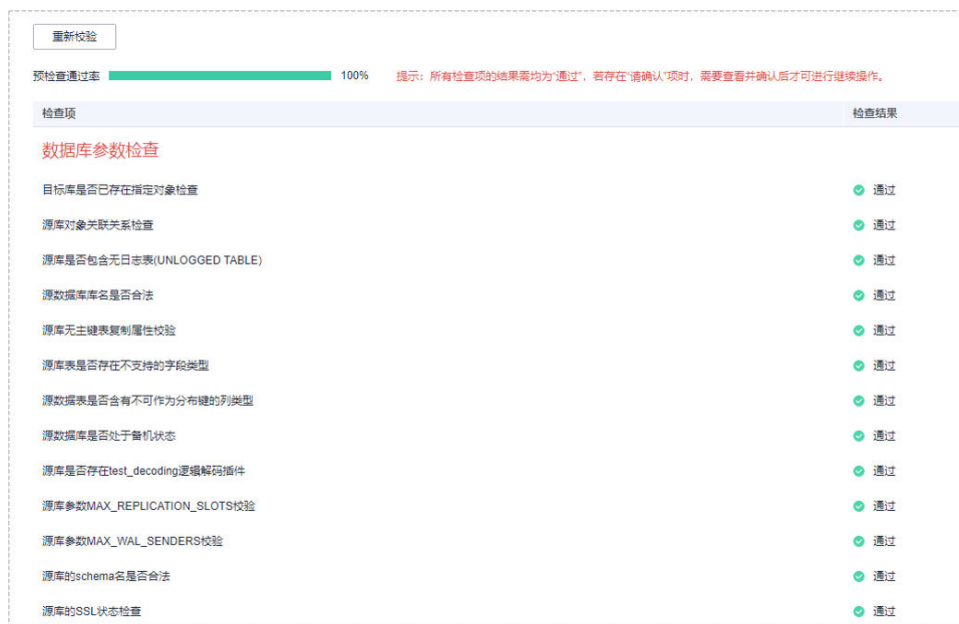
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-100 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-101 任务启动设置

The screenshot shows the 'Task Start Settings' configuration page. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Latency Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto End Time): An input field containing the value '14', with a help icon and a note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (Task will automatically end after a period of time in an abnormal state. Unit is days).

表 3-103 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.10 将 PostgreSQL 同步到 GaussDB 主备版

支持的源和目标数据库

表 3-104 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本）ECS自建数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本）其他云上数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本）RDS for PostgreSQL（9.5、9.6、10、11、12、13、14版本）	GaussDB主备版

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-105](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查

表 3-105 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 不支持实例级同步 每次至多同步一个库 (database) ， 同步多个库需要使用多个DRS任务。 ● 支持的字段类型 <ul style="list-style-type: none"> - 支持: tinyint、smallint、int、bigint、numeric、decimal、char、bpchar、varchar、text、date、time、timetz、timestamp、timestampz、interval等GaussDB兼容PostgreSQL的数据类型。 - 不支持: xml、line、domain、自建数据类型同步。 ● 全量同步的范围 <ul style="list-style-type: none"> - 支持: 模式、表、主键与唯一约束、表数据、序列的同步。 - 不支持: 系统模式 (以下模式在目标库为系统模式, 不支持同步: "pg_toast", "cstore", "snapshot", "sys", "dbms_job", "dbms_perf", "pg_catalog", "information_schema", "utl_file", "dbms_output", "dbms_random", "utl_raw", "dbms_sql", "dbms_lob", "dbe_perf", "pkg_service", "pkg_util", "dbe_file", "dbe_random", "dbe_output", "dbe_raw", "dbe_sql", "dbe_lob", "dbe_task", "blockchain", "db4ai", "dbe_pldebugger", "sqladvisor", "dbe_application_info", "dbe_match", "dbe_pldeveloper", "dbe_scheduler", "dbe_session", "dbe_utility", "dbe_sql_util") 、系统表。 - 不支持: 既是无主键表, 又是分区表的自建表, 可能会导致任务失败。 <p>说明 支持同步的对象有如下限制:</p> <ul style="list-style-type: none"> ● 对象名称: 库名不可以包含+"%\<>, 模式名和表名不可以包含".'\<>, 列名不可以包含"和'。 ● 表: 不同步临时表。目前仅支持表级同步, 表的约束会一起同步, 表的普通索引、规则和触发器不会同步。分区表目前只同步主表对象, 分区表的所有数据将集中写入主表。 ● 模式: 不同步public模式的权限。表级同步时, 会同步目标库已存在模式的权限。 <ul style="list-style-type: none"> ● 增量同步的范围 <ul style="list-style-type: none"> - 支持: 部分DML (Data Manipulation Language) : 包括 INSERT、UPDATE、DELETE。 - 不支持: DDL (Data Define Language) , 无日志表 (UNLOGGED TABLE) 的DML, 临时表的DML。 - 分区表的DML同步: 若分区是新增分区, 则该分区必须是新创建的分区, 且该分区名不可以同已创建过的表名重复。该分区必须在所有DML数据都同步至目标库后才可以从源库删除。

数据库账号权限要求

在使用DRS进行同步时, 连接源库和目标库的数据库账号需要满足以下权限要求, 才能启动实时同步任务。不同类型的同步任务, 需要的账号权限也不同, 详细可参考[表](#)

3-106进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-106 数据库账号权限

类型名称	全量同步	全量+增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限。	<p>数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限，无主键表的UPDATE、DELETE和TRUNCATE权限，REPLICATION连接权限。</p> <p>说明</p> <ul style="list-style-type: none"> • 无主键表的UPDATE、DELETE和TRUNCATE权限，仅用于对无主键表短暂加锁，以确保迁移后，目标库中无主键表数据与源库保持一致。 • REPLICATION连接权限的添加方法： <ol style="list-style-type: none"> 1. 在源数据库的“pg_hba.conf”配置文件的所有配置前增加一行配置“host replication <src_user_name> <drs_instance_ip>/32 <认证方式>”；认证方式可参考PostgreSQL官方文档pg_hba.conf文件配置，常见的认证方式有scram-sha-256等。 2. 在源库使用SUPERUSER用户执行语句“select pg_reload_conf();”生效，或重启数据库实例生效。
目标数据库连接账号	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。 • 目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。 	

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-107 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库的分区表触发器不可以设置为disable。 - 全量同步支持源库备机状态，但需要设置 hot_standby_feedback为on；增量同步不支持源库备机状态。 - 若要做增量同步： 源数据库的“pg_hba.conf”文件中包含如下的配置： host replication all 0.0.0.0/0 md5 源数据库参数wal_level必须配置为logical。 源数据库需提前安装test_decoding插件。 源数据库中无主键表的replica identity属性必须为full。 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica identity属性必须为full。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 同步对象依赖和关联的对象也须一起同步，否则可能导致同步失败。 <p>说明 常见的关联关系：视图引用表、视图引用视图、主外键关联表、表继承子表引用父表、表分区子分区表引用分区表、表自增列引用序列等。</p> ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标数据库参数max_prepared_transactions须大于源数据库。 - 目标数据库参数max_worker_processes须大于源数据库。 - 目标数据库参数max_locks_per_transaction须大于源数据库。 - 目标数据库参数max_connections须大于源数据库。 - 目标数据库和源数据库的lc_monetary参数值必须一致。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。

类型名称	使用和操作约定
	<p>目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。</p> <ul style="list-style-type: none"> ● 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 分区表只同步主表，转为普通表同步到目标库，分区表的数据将集中写入主表。 - 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。 - 对于全量+增量和增量任务，启动任务前，请确保源库中未启动长事务，源库启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 对于全量+增量和增量任务，任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - DRS全量同步表结构时，源库中的character、character varying类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），默认扩大4倍，最大长度为10485760。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源数据库表的主键或者唯一键（主键不存在时），否则可能导致增量数据不一致或任务失败。 ● 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 ● 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。

类型名称	使用和操作约定
同步对比使用须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。
结束任务须知	<ul style="list-style-type: none"> 正常结束任务： <ul style="list-style-type: none"> 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 强制结束任务： <ul style="list-style-type: none"> 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-102 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ⓘ
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ⓘ

描述: [Text Area] ⓘ
0/256

表 3-108 任务和描述

参数	描述
区域	当前所在区域，可进行切换。

参数	描述
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 3-103 同步实例信息



表 3-109 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“PostgreSQL”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	<p>此处以“VPC网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库的GaussDB主备版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库，包括：表、视图、存储过程等。 - 增量 该模式通过解析日志等技术，将源端产生的增量数据持续同步至目标端数据库。

- 可用区

图 3-104 可用区



表 3-110 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-105 企业项目和标签



表 3-111 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务同源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过连接账号和密码访问。

📖 说明

此处源库类型分为ECS自建库和RDS实例，需要根据源数据库的实际来源选择相应的分类。两种场景下的参数配置不一样，需要根据具体场景进行配置。

- **场景一：** ECS自建库源库信息配置

图 3-106 ECS 自建库场景源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

数据库类型 ECS自建库 RDS实例

VPC C 查看虚拟私有云

子网 ② 查看子网

IP地址或域名

端口

指定数据库名称

数据库用户名

数据库密码

SSL安全连接

● 测试成功 ?

表 3-112 ECS 自建库场景源库信息

参数	描述
数据库库类型	选择“ECS自建库”。
VPC	源数据库实例所在的虚拟专用网络，可以对不同业务进行网络隔离。您需要创建或选择所需的虚拟私有云。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用分区内才会有效，创建源数据库实例的子网需要开启DHCP功能，在创建过程中也不能关闭已选子网的DHCP功能。
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
指定数据库名称	选择是否指定数据库，开启后需手动输入数据库名称。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

- 场景二：RDS实例源库信息配置

图 3-107 RDS 实例场景源库信息

表 3-113 RDS 实例场景源库信息

参数	描述
源库类型	选择“RDS实例”。
数据库实例名称	选择待同步的关系型PostgreSQL数据库实例作为源数据库实例。
数据库用户名	源数据库实例的用户名。
数据库密码	源数据库的用户名所对应的密码。

图 3-108 目标库信息

表 3-114 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 3-109 同步模式

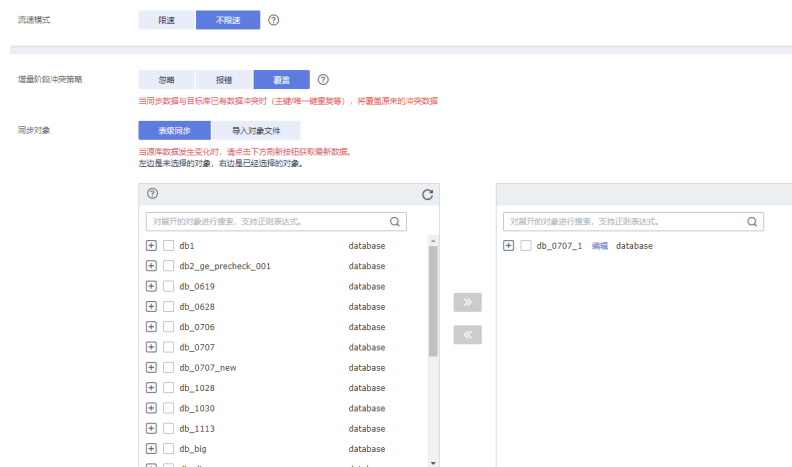



表 3-115 同步对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-110 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-111 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-112 任务启动设置



表 3-116 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.11 将 PostgreSQL 同步到 GaussDB 分布式版

支持的源和目标数据库

表 3-117 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本）ECS自建数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本）其他云上数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本）RDS for PostgreSQL（9.5、9.6、10、11、12、13、14版本）	GaussDB分布式版

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考表3-118。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-118 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 不支持实例级同步 每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持的字段类型 <ul style="list-style-type: none"> - 支持：tinyint、smallint、int、bigint、numeric、decimal、char、bpchar、varchar、text、date、time、timetz、timestamp、timestampz、interval等GaussDB兼容PostgreSQL的数据类型。 - 不支持：xml、line、domain、自建数据类型同步。 ● 全量同步的范围 <ul style="list-style-type: none"> - 支持：模式、表、主键与唯一约束、表数据、序列的同步。 - 不支持：系统模式（以下模式在目标库为系统模式，不支持同步："pg_toast", "cstore", "snapshot", "sys", "dbms_job", "dbms_perf", "pg_catalog", "information_schema", "utl_file", "dbms_output", "dbms_random", "utl_raw", "dbms_sql", "dbms_lob", "dbe_perf", "pkg_service", "pkg_util", "dbe_file", "dbe_random", "dbe_output", "dbe_raw", "dbe_sql", "dbe_lob", "dbe_task", "blockchain", "db4ai", "dbe_pldebugger", "sqladvisor", "dbe_application_info", "dbe_match", "dbe_pldeveloper", "dbe_scheduler", "dbe_session", "dbe_utility", "dbe_sql_util"）、系统表。 - 不支持：既是无主键表，又是分区表的自建表，可能会导致任务失败。 <p>说明 支持同步的对象有如下限制：</p> <ul style="list-style-type: none"> ● 对象名称：库名不可以包含+"%\<>，模式名和表名不可以包含".\<>，列名不可以包含"和"。 ● 表：不同步临时表。目前仅支持表级同步，表的约束会一起同步，表的普通索引、规则和触发器不会同步。分区表目前只同步主表对象，分区表的所有数据将集中写入主表。 ● 模式：不同步public模式的权限。表级同步时，会同步目标库已存在模式的权限。 <ul style="list-style-type: none"> ● 增量同步的范围 <ul style="list-style-type: none"> - 支持：部分DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 - 不支持：DDL（Data Define Language），无日志表（UNLOGGED TABLE）的DML，临时表的DML。 - 分区表的DML同步：若分区是新增分区，则该分区必须是新创建的分区，且该分区名不可以同已创建过的表名重复。该分区必须在所有DML数据都同步至目标库后才可以从源库删除。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-119](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-119 数据库账号权限

类型名称	全量同步	全量+增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限。	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限，无主键表的UPDATE、DELETE和TRUNCATE权限，REPLICATION连接权限。 说明 <ul style="list-style-type: none"> • 无主键表的UPDATE、DELETE和TRUNCATE权限，仅用于对无主键表短暂加锁，以确保迁移后，目标库中无主键表数据与源库保持一致。 • REPLICATION连接权限的添加方法： <ol style="list-style-type: none"> 1. 在源数据库的“pg_hba.conf”配置文件的所有配置前增加一行配置“host replication <src_user_name> <drs_instance_ip>/32 <认证方式>”；认证方式可参考PostgreSQL官方文档pg_hba.conf文件配置，常见的认证方式有scram-sha-256等。 2. 在源库使用SUPERUSER用户执行语句“select pg_reload_conf();”生效，或重启数据库实例生效。
目标数据库连接账号	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none"> • 需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。 • 目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。 	

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-120 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库的分区表触发器不可以设置为disable。 - 全量同步支持源库备机状态，但需要设置 hot_standby_feedback为on；增量同步不支持源库备机状态。 - 若要做增量同步： 源数据库的“pg_hba.conf”文件中包含如下的配置： host replication all 0.0.0.0/0 md5 源数据库参数wal_level必须配置为logical。 源数据库需提前安装test_decoding插件。 源数据库中无主键表的replica identity属性必须为full。 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica identity属性必须为full。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 同步对象依赖和关联的对象也须一起同步，否则可能导致同步失败。 <p>说明 常见的关联关系：主外键关联表、表继承子表引用父表、表分区子分区表引用分区表、表自增列引用序列等。</p> ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标数据库参数max_prepared_transactions须大于源数据库。 - 目标数据库参数max_worker_processes须大于源数据库。 - 目标数据库参数max_locks_per_transaction须大于源数据库。 - 目标数据库参数max_connections须大于源数据库。 - 目标数据库和源数据库的lc_monetary参数值必须一致。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。

类型名称	使用和操作约定
	<p>目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。</p> <ul style="list-style-type: none"> ● 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 分区表只同步主表，转为普通表同步到目标库，分区表的数据将集中写入主表。 - 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为此格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。 - 对于全量+增量和增量任务，启动任务前，请确保源库中未启动长事务，源库启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 对于全量+增量和增量任务，任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - DRS全量同步表结构时，源库中的character、character varying类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），默认扩大4倍，最大长度为10485760。 - 源数据库的唯一索引是表达式索引时，全量同步阶段目标数据库GaussDB分布式可能会因为无法创建该索引而导致任务失败。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿修改源数据库表的主键或者唯一键（主键不存在时），否则可能导致增量数据不一致或任务失败。 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据不一致或任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。
同步对比使用须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。
结束任务须知	<ul style="list-style-type: none"> 正常结束任务： <ul style="list-style-type: none"> 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 强制结束任务： <ul style="list-style-type: none"> 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-113 同步任务信息

⚠️ 1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述
0/256

表 3-121 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-114 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建迁移任务。

* 数据流动方向 入云 出云 自建-自建 ?

DRS要求源数据库或目标数据库中至少有一方为本云数据库实例，入云指目标数据库为本云数据库实例的场景，本云数据库实例之间的同步，请选择入云。

* 源数据库引擎 MySQL Oracle DB2 for LUW DDM MongoDB PostgreSQL Microsoft SQL Server TiDB

* 目标数据库引擎 GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版 PostgreSQL

* 网络类型 ?

* 目标数据库实例 ? 查看数据库实例 查看不可选实例

同步实例所在子网 ? 查看子网

* 同步模式 全量+增量 全量 增量 ?

该模式为数据持续实时同步，通过全量过程完成目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据一致。

表 3-122 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“PostgreSQL”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	<p>此处以“VPC网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库，包括：表、视图、存储过程等。 - 增量 该模式通过解析日志等技术，将源端产生的增量数据持续同步至目标端数据库。

- 可用区

图 3-115 可用区



表 3-123 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-116 企业项目和标签



表 3-124 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务同源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过连接账号和密码访问。

说明

此处源库类型分为ECS自建库和RDS实例，需要根据源数据库的实际来源选择相应的分类。两种场景下的参数配置不一样，需要根据具体场景进行配置。

- 场景一：ECS自建库源库信息配置

图 3-117 ECS 自建库场景源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导入用户和作业。

数据库类型 ECS自建库 RDS实例

VPC C 查看虚拟私有云

子网 ? 查看子网

IP地址或域名

端口

指定数据库名称

数据库用户名

数据库密码

SSL安全连接

● 测试成功 ? 帮助

表 3-125 ECS 自建库场景源库信息

参数	描述
数据库库类型	选择“ECS自建库”。
VPC	源数据库实例所在的虚拟专用网络，可以对不同业务进行网络隔离。您需要创建或选择所需的虚拟私有云。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用分区内才会有效，创建源数据库实例的子网需要开启DHCP功能，在创建过程中也不能关闭已选子网的DHCP功能。
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
指定数据库名称	选择是否指定数据库，开启后需手动输入数据库名称。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

- 场景二：RDS实例源库信息配置

图 3-118 RDS 实例场景源库信息

表 3-126 RDS 实例场景源库信息

参数	描述
源库类型	选择“RDS实例”。
数据库实例名称	选择待同步的关系型PostgreSQL数据库实例作为源数据库实例。
数据库用户名	源数据库实例的用户名。
数据库密码	源数据库的用户名所对应的密码。

图 3-119 目标库信息

表 3-127 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 3-120 同步模式

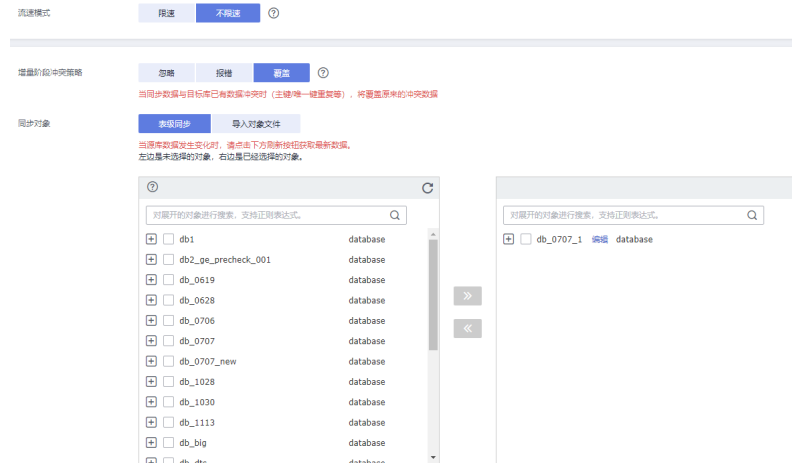



表 3-128 同步对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-121 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

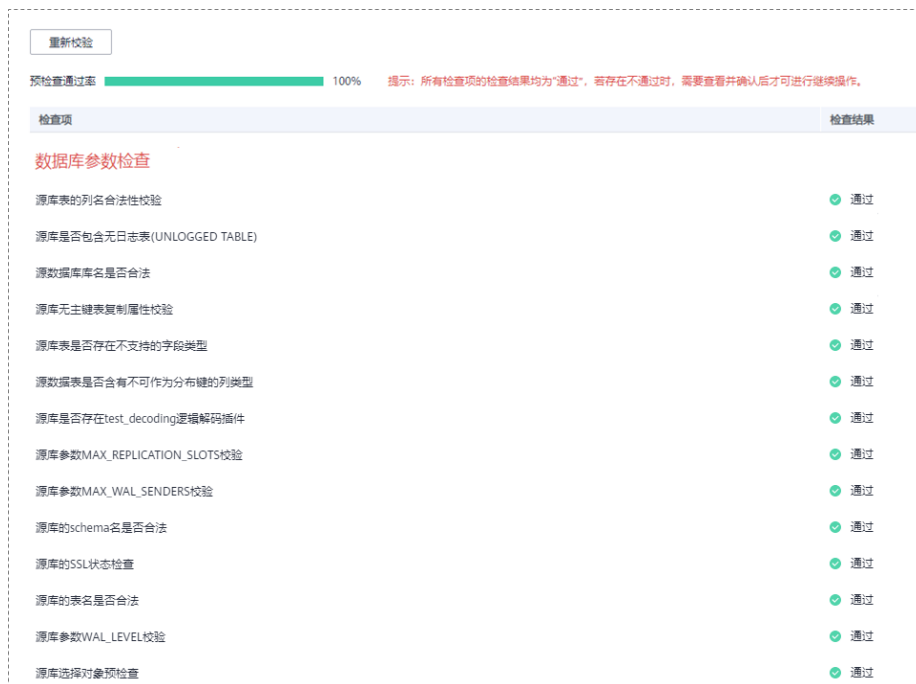
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-122 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-123 任务启动设置



表 3-129 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.12 将 Oracle 同步到 MySQL

支持的源和目标数据库

表 3-130 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本）ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本）	<ul style="list-style-type: none">RDS for MySQL

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-131](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-131 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 同步对象选择维度：支持表级同步，或者通过文件导入对象。 ● 支持的同步对象： <ul style="list-style-type: none"> - 支持库、表结构、主键、唯一键、普通索引、表数据的同步，其他数据库对象暂不支持，如存储过程、触发器、函数、序列、包、同义词、用户等。 - 全量阶段不支持bfile, xml、sdo_geometry、urowid、interval（精度大于6位）和自定义类型。 - 增量阶段不支持bfile, xml、interval、sdo_geometry、urowid、timestamp（精度大于6位）和自定义类型。 - 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 - 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 - 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：主外键关联表等。 - 不支持同步表结构中的partition，分区表在目的库同步为非分区表。 - 不支持默认值含有表达式的函数的表的同步。 - 不支持同步源库中的临时表。 - 不支持选择源数据库的空库进行同步。 - 不支持同步源库中有虚拟列的表。 - 如果表中只有LOB字段，可能出现数据不一致性情况。 - 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 - 针对无主键且无索引的表，非大字段的列必须大于3列，否则会因无法全列匹配导致增量异常。 <p>说明</p> <ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“`”、“ ”、“”、“?”、“!”、“”和“”等字符。 ● 对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-132](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-132 数据库账号权限

类型名称	全量同步	增量、全量+增量同步
源数据库连接账号	需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;)。	<ul style="list-style-type: none"> ● 12c及以上版本租户模式： <ul style="list-style-type: none"> - 12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;)。 - 12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;)，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。 ● 12c及以上版本非租户模式：需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;)。 ● 11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;)。 ● 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，

类型名称	全量同步	增量、全量+增量同步
		<p>若只开启表级补充日志，重建或者 RENAME 表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。</p> <ul style="list-style-type: none"> 12c及以上版本不支持使用 ORACLE_MAINTAINED=Y 的用户账号进行增量同步（system/sys除外），因为该属性的账号无日志解析权限。
目标数据库连接账号	必须拥有如下权限：SELECT、INSERT、CREATE、DROP、UPDATE、ALTER、DELETE、INDEX。	

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-133 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibttmp1。 - 源数据库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果：默认值函数可能会被置空。 创建表失败，导致对象对比不一致或者任务失败。 - Oracle单行记录不能超过8K（lob、long类型除外，会自动转换成MySQL的text、blob类型），原因是MySQL innodb引擎限制单行大小不能超过8K（text、blob类型除外）。 - 不建议以字符串类型作为主键或唯一键，因为Oracle的字符串作为主键、唯一键时区分空格，而MySQL不区分，可能导致数据不一致和死锁问题。 - 对于Oracle的binary_float或者binary_double类型，MySQL中不支持设置Nan、Inf、-Inf三种值，DRS默认会将这三种值转为0保存。 - Oracle的check约束同步到MySQL会失效，原因是MySQL不支持check约束。 - Oracle中建议列名不要取名AUTO_PK_ROW_ID，原因是这个列名在MySQL5.7中是保留列名，无法创建出来。 - Oracle中number(p, s)字段的精度不要超过p: [1, 38], s:[p-65, min(p, 30)]的精度表示范围。其中，s取值依赖于p的取值变化，即下限为p-65, 上限为p或30中取最小值。例如：当p=1, s的取值范围是[-64, 1]。当p=38, s取值范围是[-27, 30]。int字段的值不要超过（65，0）的精度表示范围。原因是MySQL数字的表示范围比Oracle小。 - Oracle中float(p)字段的精度不要超过p: [1, 32]的精度表示范围。原因是MySQL float类型的表示范围比Oracle小。 - Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 - 源数据库不允许存在索引列的长度之和超过目标库索引列长度限制的索引，具体长度要求请参见索引长度说明。 - 默认值不支持default user，MySQL没有对应的语法。 ● 目标库参数要求：

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - DRS同步时会有大量数据写入目标库，目标库 max_allowed_packet 参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 ● 目标库对象要求： <ul style="list-style-type: none"> - 源库和目标库时区设置必须一致。 - 选择同步表结构时，目标数据库不能存在待同步数据库。 - 目标数据库需要有足够的磁盘空间，约为源库空间大小的1.5倍。 - 目标数据库版本小于5.7.7时，源库单个索引的全部列的长度不得超过767，反之则不得超过3072。 - 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 数据类型不兼容时，可能引起同步失败。 - 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 - 如有中文、日文等特殊字符，业务连接Oracle数据库使用的编码需和Oracle数据库服务端编码一致，否则目标库会出现乱码。 - 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 - Oracle中表结构长度（所有列长字节数之和，char、varchar2等类型字节长度和编码有关）超过65535时，可能导致同步失败。 - Oracle中表结构同步到MySQL后表的字符集为UTF8MB4。 - 当Oracle字符集是WE8MSWIN1252时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 - 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 - Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 - 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 - 由于Oracle与MySQL的部分语法有明显区别，结构同步无法完全保证支持全部语法的转换，包括但不限于函数，表达式，依赖的系统表等。所以在同步过程中，会有在Oracle上存在，在MySQL中没有直接对应的语法，或者MySQL中有对应的语法，但当前还未适配转换的情况，这样会导致结构同步失败。这时，需要手工在目标数据库创建表结构。 - 同步对象支持设置事务的强一致性（事务同步到目标库的提交顺序和原子性与源库保持一致），性能相比默认模式有较大幅度降低。 - 打开事务强一致性开关，如果一次提交事务过大（大于256M），可能会导致内存溢出。 - 全量+增量或单增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 - lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。

类型名称	使用和操作限制
<p>全量同步阶段须知</p>	<ul style="list-style-type: none"> ● 任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。 ● 全量同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 全量同步过程中，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。 ● 全量同步过程中，不允许对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 全量同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 全量同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 全量数据同步阶段，目标数据库会产生大量Binlog引起磁盘空间占用过大，因此在全量数据同步阶段目标数据库Binlog本地默认调整为只保留最新5个，全量同步结束后目标数据库Binlog本地保留时长恢复为用户当前配置。如果因业务需要，目标数据库本地Binlog保留时长需要保持用户当前配置，需提交工单申请开启。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。

类型名称	使用和操作限制
增量同步阶段须知	<ul style="list-style-type: none"> ● 增量同步过程中，不允许修改、删除连接源和目标数据库的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步过程中，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。 ● 增量同步过程中，不允许对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 增量同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 增量同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 增量同步过程中，请勿修改源库或目标库表中的char类型字段为varchar，由于Oracle和MySQL表现不同，产生的空格可能会导致数据对比不一致。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果出现其余情况的DDL同步导致的任务异常，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify column以及truncate table的基本DDL，不支持默认值等的修改。 - 库级同步支持create table（表内定义不能包含函数）。 - DDL中的对象不能是目标库的关键字，比如index，where等。目标库MySQL关键字可参考MySQL官网文档。 - 目标库为8.0以下版本时，不支持alter table rename column。 - 增量DDL不支持全角、中文等特殊字符。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 表的对象名同步到目标库后会转换成小写，如ABC和abc。因此增量同步阶段，选择的源库的表中不能存在仅大小写不同的表，否则会导致同步失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。

操作步骤

本小节以Oracle到RDS for MySQL的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-124 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?
0/256

表 3-134 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-125 同步实例信息

同步实例信息

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建迁移任务。

- 数据流动方向**: 入云 (选中), 出云, 自建-自建
- 源数据库引擎**: MySQL, Oracle (选中), DB2 for LUW, DDM, MongoDB, PostgreSQL, Microsoft SQL Server, TiDB
- 目标数据库引擎**: MySQL (选中), DDM, GaussDB (DWS), GaussDB 分布式版, GaussDB 主备版, PostgreSQL, GaussDB(for MySQL)
- 网络类型**: 公网网络 (选中)
- 目标数据库实例**: 无可用实例 (选中)
- 同步实例所在子网**: 选择子网
- 同步模式**: 全量 + 增量 (选中), 全量, 增量

表 3-135 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	创建好的RDS for MySQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将用户选择的数据库对象和数据一次性同步至目标端数据库。 - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 3-126 任务类型



表 3-136 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p>

- 企业项目和标签

图 3-127 企业项目和标签



表 3-137 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-128 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码 👁

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-138 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-129 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功 [?](#)

表 3-139 目标库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的RDS for MySQL实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。支持在任务创建后修改密码。
SSL安全连接	<p>如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确后上传SSL证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-130 同步模式

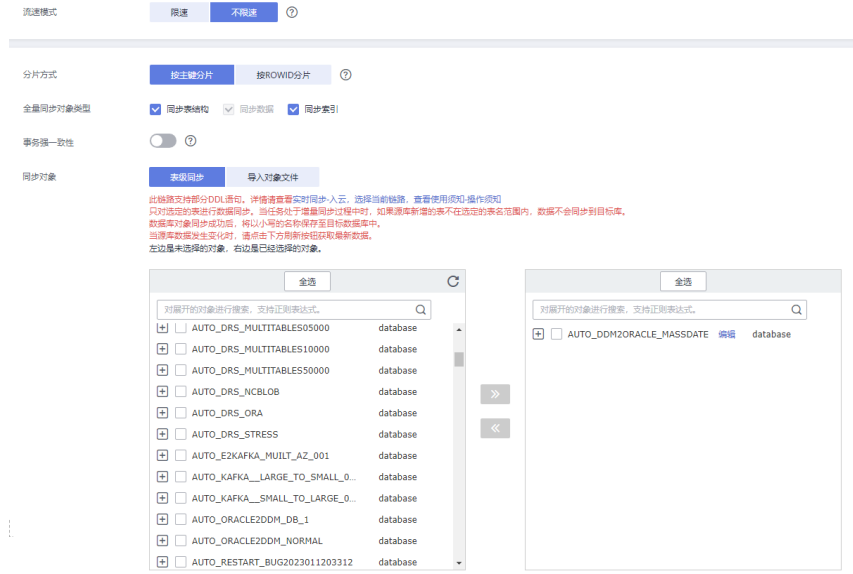



表 3-140 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-131 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
事务强一致性	<p>默认为关闭，打开后，事务同步到目标库的提交顺序和原子性与源库保持一致，性能相比默认模式有较大幅度降低。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。选择数据的时候支持搜索，以便您快速选择需要的数据库对象。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页，可对同步数据进行过滤，完成后单击“下一步”，详细可参考[“数据加工”](#)。

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的[“预检查不通过项修复方法”](#)。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-132 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Start Immediately) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on.
- * SMN主题**: A dropdown menu is present, with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: A text input field contains the value '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。'

表 3-141 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.13 将 Oracle 同步到 GaussDB(for MySQL)

支持的源和目标数据库

表 3-142 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建数据库 (Oracle 10g、11g、12c、18c、19c、21c版本) • ECS自建数据库 (Oracle 10g、11g、12c、18c、19c、21c版本) 	<ul style="list-style-type: none"> • GaussDB(for MySQL)主备实例

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-143 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> 12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。 12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> 11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库端：提供的目标数据库账号必须拥有如下权限：SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、INDEX、EVENT、RELOAD、CREATE VIEW、CREATE ROUTINE、TRIGGER。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持库、表结构、主键、唯一键、普通索引、表数据的同步，其他数据库对象暂不支持，如存储过程、触发器、函数、序列、包、同义词、用户等。 ● 全量阶段不支持bfile, xml、sdo_geometry、urowid、interval（精度大于6位）和自定义类型。 ● 增量阶段不支持bfile, xml、interval、sdo_geometry、urowid、timestamp（精度大于6位）和自定义类型。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会因无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● Oracle单行记录不能超过8K（lob、long类型除外，会自动转换成MySQL的text、blob类型），原因是MySQL innodb引擎限制单行大小不能超过8K（text、blob类型除外）。 ● 不建议以字符串类型作为主键或唯一键，因为Oracle的字符串作为主键、唯一键时区分空格，而MySQL不区分，可能导致数据不一致和死锁问题。 ● 对于Oracle的binary_float或者binary_double类型，MySQL中不支持设置Nan、Inf、-Inf三种值，DRS默认会将这三种值转为0保存。 ● Oracle的check约束同步到MySQL会失效，原因是MySQL不支持check约束。 ● Oracle中建议列名不要取名AUTO_PK_ROW_ID，原因是这个列名在MySQL5.7中是保留列名，无法创建出来。 ● Oracle中number(p, s)字段的精度不要超过p: [1, 38], s:[p-65, min(p, 30)]的精度表示范围。其中，s取值依赖于p的取值变化，即下限为p-65, 上限为p或30中取最小值。例如：当p=1, s的取值范围是[-64, 1]。当p=38, s取值范围是[-27, 30]。 int字段的值不要超过（65, 0）的精度表示范围。原因是MySQL数字的表示范围比Oracle小。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 库名、表名不支持的字符有：非ASCII字符、“.”、“>”、“<”、“\”、“_”、“ ”、“,”、“?”、“!”、“”和“'”。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 不支持选择源数据库的空库进行同步。 ● 源数据库不允许存在索引列的长度之和超过目标库索引列长度限制的索引，具体长度要求请参见索引长度说明。 ● 默认值不支持default user，MySQL没有对应的语法。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例必须是GaussDB(for MySQL)主备实例。 ● 选择同步表结构时，目标数据库不能存在待同步数据库。 ● DRS同步时会有大量数据写入目标库，目标库 max_allowed_packet 参数过小会导致无法写入，建议将目标库 max_allowed_packet参数值设置为大于100MB。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：主外键关联表等。 ● 源库和目标库时区设置必须一致。 ● 数据类型不兼容时，可能引起同步失败。 ● 如有中文、日文等特殊字符，业务连接Oracle数据库使用的编码需和Oracle数据库服务端编码一致，否则目标库会出现乱码。 ● 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● Oracle中表结构长度（所有列长字节数之和，char、varchar2等类型字节长度和编码有关）超过65535时，可能导致同步失败。 ● Oracle中表结构同步到MySQL后表的字符集为UTF8MB4。 ● 当Oracle字符集是WE8MSWIN1252时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 表等对象名同步到目标库后会转换成小写，如ABC和abc。因此增量同步阶段，选择的源库的表中不能存在仅大小写不同的表，否则，会导致同步失败。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 对于同步中的数据库对象，在同步期间，目标库不能进行写入操作，否则会导致数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及

类型名称	使用和操作限制
	<p>11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。</p> <ul style="list-style-type: none"> ● 同步过程中，如果修改源库或目标库表中的char类型字段为varchar，由于Oracle和MySQL表现不同，产生的空格可能会导致数据对比不一致。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果出现其余情况的DDL同步导致的任务异常，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify column以及truncate table的基本DDL，不支持默认值等的修改。 - 库级同步支持create table（表内定义不能包含函数）。 - DDL中的对象不能是目标库的关键字，比如index，where等。 - 增量DDL不支持全角、中文等特殊字符。 ● 库级映射和表级映射均不区分大小写，例如映射为abc与映射为ABC，同步到目标库后均为abc。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。

操作步骤

本小节以Oracle到GaussDB(for MySQL)的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-133 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-144 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-134 同步实例信息

同步实例信息 ?

以下信息默认不可修改，请谨慎填写，以免影响配置准确性，需要重新创建迁移任务。

* 数据源方向: [单选: 入云] [单选: 出云] [单选: 同域] ?

DRS最多支持数据源和目标数据库至少有一方为本云数据库实例。入云指目标数据库为本云数据库实例的同步，本云数据库实例之间的同步，请勾选入云。

* 源数据库引擎: MySQL Oracle DB2 for LUW DDM MongoDB PostgreSQL Microsoft SQL Server TiDB

* 目标数据库引擎: MySQL DDM GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版 PostgreSQL GaussDB for MySQL

* 网络类型: [下拉: 公网网络] ?

DRS将自动创建DRS实例所属弹性公网IP，请在创建实例时选择弹性公网IP。

* 目标数据库实例: [下拉: 新建实例] [单选: 选择现有实例]

同步实例所在子网: [下拉: 选择子网] [单选: 新建子网] ?

* 同步模式: [单选: 全量] [单选: 增量] [单选: 全量]

创建方式: 创建方式: 通过全量过程创建和增量过程创建初始化，需要同时配置全量过程和增量过程，再选择和目标任务源和目标数据库一致。

表 3-145 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“GaussDB(for MySQL)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	创建好的GaussDB(for MySQL)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将用户选择的数据库对象和数据一次性同步至目标端数据库。

● 任务类型

图 3-135 任务类型



表 3-146 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-136 企业项目和标签



表 3-147 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-137 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-148 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-138 目标库信息

表 3-149 目标库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。支持在任务创建后修改密码。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-139 同步模式

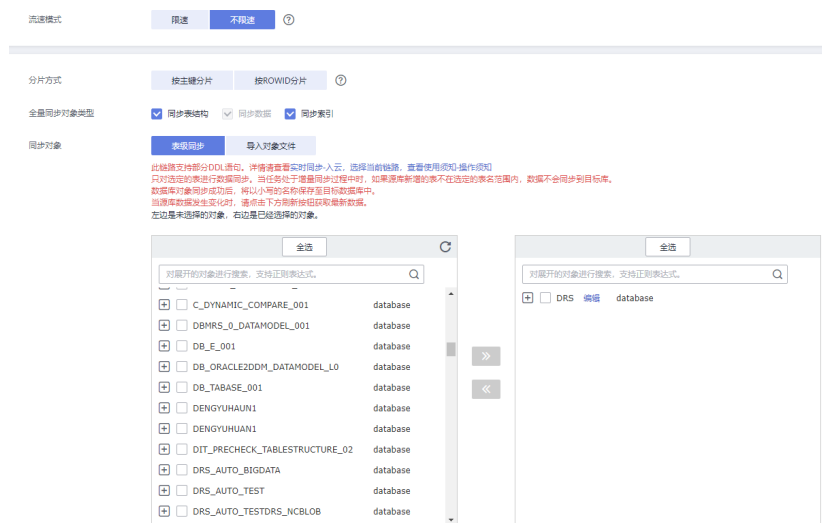



表 3-150 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-140 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。选择数据的时候支持搜索，以便您快速选择需要的数据库对象。</p> <ul style="list-style-type: none"> 选择导入对象文件，具体步骤和说明可参考导入同步对象。 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页，可对同步数据进行过滤，完成后单击“下一步”，详细可参考[“数据加工”](#)。

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的[“预检查不通过项修复方法”](#)。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-141 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 3-151 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.14 将 Oracle 同步到 GaussDB 主备版

支持的源和目标数据库

表 3-152 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） • ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） 	GaussDB主备版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-153 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量和增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> 12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。 12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> 11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>;

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 全量阶段不支持bfile、xml、sdo_geometry、urowid和自定义类型。增量阶段不支持bfile、xml、sdo_geometry、urowid、interval和自定义类型。 ● timestamp和interval day to second类型支持的最大精度是6。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 全量同步结构不支持位图索引、倒排索引、函数索引。 ● 增量同步LOB类型仅支持BasicFiles属性，不支持SecureFiles属性，支持的LOB类型大小限10M以内。 ● 对于TIMESTAMP WITH TIME ZONE类型，根据目标库时区做转换后不得大于“9999-12-31 23:59:59.999999”。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会因为无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。 ● 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 ● 选择手动创建表结构时，目标库中的时间类型是否带有时区需要与源库中保持一致，否则可能会因为时区转换导致时间数据不一致。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“`”、“ ”、“/”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是GaussDB主备版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。主键支持的类型可参考Oracle数据库->GaussDB数据库映射关系。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 当Oracle字符集是WE8MSWIN1252、WE8ISO8859P1时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 当每行的行内存储类型数据的总长度超过GaussDB的限制时（默认是8192字节），可能会导致任务失败。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 如果源库和目标库字符集不一致，如源库是ZHS16GBK，目标库是UTF8，由于ZHS16GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 使用DRS同步表结构的场景下，同一个schema中，同步的表中，表、约束、索引等不能有忽略大小写后的同名对象，比如表"A"中有索引名"inx1"，表B中有索引名"a"，表A和索引"a"忽略大小写重名了，会导致结构同步失败。多个schema映射到一个schema的场景，源库的多个schema中也不能包含同名但字母大小写不同的表、约束和索引。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步表结构时，如果目标库存在与源库同名约束会导致建表失败。 ● 全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量或单增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果转换失败或执行失败会忽略错误，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify

类型名称	使用和操作限制
	<p>column以及truncate table的基本DDL，不支持默认值等的修改。</p> <ul style="list-style-type: none"> - 增量DDL新增的字段不会自动扩展字符长度。 ● 增量同步阶段，修改抓取任务的启动位点主要用于重新同步数据。 <ul style="list-style-type: none"> - 修改抓取位点之后，上一次对象级对比结果不会展示出来。 - 单独修改抓取任务的启动位点，会把该位点同步到回放任务的启动位点上，即回放任务的启动位点和抓取任务的启动位点一致，不影响用户单独修改回放任务的启动位点。 ● 全量和增量同步不支持隐藏列（UNUSED, INVISIBLE）。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。 ● 当双AZ任务发生切换（例如主任务AZ级故障）时，如果主任务延迟较大，或源数据库有长事务未提交，任务倒换后可能因源数据库日志被清理导致任务无法启动，或着因为延迟、长事务期间的DDL变更导致同步异常。 ● 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-142 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。	
项目	<input type="text" value=""/>
* 任务名称	<input type="text" value="DRS-6435"/> ?
描述	<div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> <div style="text-align: right; font-size: x-small;">0/256</div> ?

表 3-154 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-143 同步实例信息



表 3-155 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的GaussDB主备版实例。

参数	描述
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。

- 任务类型

图 3-144 任务类型



表 3-156 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-145 企业项目和标签



表 3-157 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 3-146 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name

PDB名称 ⓘ

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-158 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

图 3-147 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

✔ 测试成功

表 3-159 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	目标数据库的用户名。

参数	描述
数据库密码	目标数据库的用户名所对应的密码。数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 3-148 同步模式

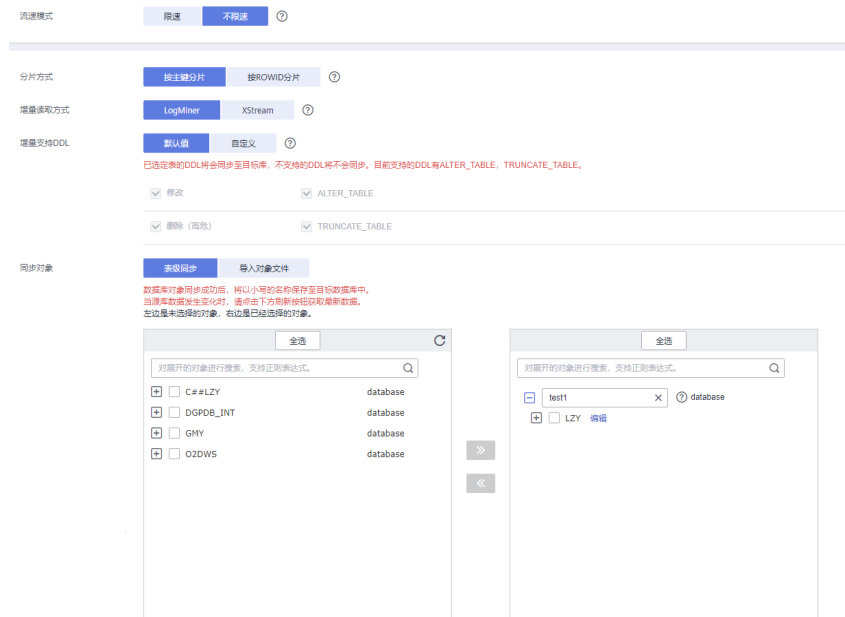



表 3-160 同步对象

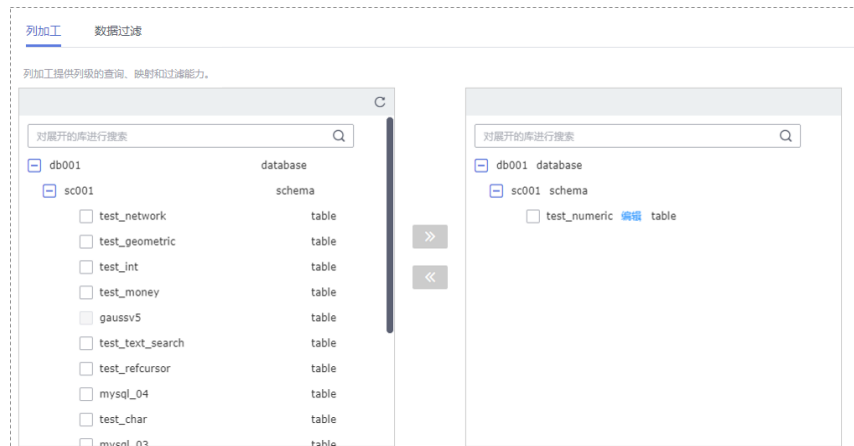
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-149 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
增量读取方式	<p>选择增量阶段读取日志方式。</p> <ul style="list-style-type: none"> LogMiner: 采用Oracle官方接口读取重做日志文件, 稳定性好。 Xstream: 需要开启Xstream接口。 <p>目前建议使用LogMiner。</p>
增量支持DDL	<p>选择增量同步的DDL类型, 用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”: 表级同步情况下, 已选定的同步表DDL将会同步至目标, 不支持的DDL将不同步。目前支持的DDL有CREATE_TABLE, ALTER_TABLE, TRUNCATE_TABLE。 “自定义”: 用户根据需求, 选择需要同步的DDL类型, 不同链路支持的DDL类型以界面显示为准。
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 可选表级同步或导入对象文件, 您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中, 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射, 具体操作可参考对象名映射。 选择导入对象文件, 具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面, 选择需要加工的数据, 进行列加工或者数据过滤。

- 如果不需要数据加工, 单击“下一步”。
- 如果需要加工列或者数据过滤, 参考[数据加工](#)中的列加工, 设置相关规则。

图 3-150 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-151 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 3-152 任务启动设置

表 3-161 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.15 将 Oracle 同步到 GaussDB 分布式版

支持的源和目标数据库

表 3-162 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） • ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） 	GaussDB分布式版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-163 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量和增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <ul style="list-style-type: none"> 12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。 12c及以上版本非租户模式： <ul style="list-style-type: none"> 需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>;

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 全量阶段不支持bfile, xmltype、sdo_geometry、urowid和自定义类型。增量阶段不支持XMLTYPE、bfile, xmltype、sdo_geometry、urowid、interval year to month、interval day to second和自定义类型。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 增量同步LOB类型仅支持BasicFiles属性，不支持SecureFiles属性，支持的LOB类型大小限10M以内。 ● 全量同步结构不支持位图索引、倒排索引、函数索引。 ● timestamp和interval day to second类型支持的最大精度是6。 ● 增量同步时，源端或者目标端数据库异常会触发任务失败，数据库恢复后重试启动任务时触发全局启动，此时会忽略原有抓取或回放组件的状态，同时回放也会按照抓取的中断点位再启动。 ● 对于TIMESTAMP WITH TIME ZONE类型，根据目标库时区做转换后不得大于“9999-12-31 23:59:59.999999”。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。 ● 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 ● 选择手动创建表结构时，目标库中的时间类型是否带有时区需要与源库中保持一致，否则可能会因为时区转换导致时间数据不一致。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“\`”、“ ”、“/”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是GaussDB分布式版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。主键支持的类型可参考Oracle数据库->GaussDB数据库映射关系。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 当Oracle字符集是WE8MSWIN1252、WE8ISO8859P1时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 当每行的行内存储类型数据的总长度超过GaussDB的限制时（默认是8192字节），可能会导致任务失败。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 无主键表缺乏行的唯一性标志，网络不稳定时涉及重试，表数据存在少量不一致的可能性。 ● 如果源库和目标库字符集不一致，如源库是ZHS16GBK，目标库是UTF8，由于ZHS16GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 使用DRS同步表结构的场景下，同一个schema中，同步的表中，表、约束、索引等不能有忽略大小写后的同名对象，比如表"A"中有索引名"inx1"，表B中有索引名"a"，表A和索引"a"忽略大小写重名了，会导致结构同步失败。多个schema映射到一个schema的场景，源库的多个schema中也不能包含同名但字母大小写不同的表、约束和索引。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致 ● 全量同步分为表结构同步（含索引）、实时同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入实时同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步表结构时，如果目标库存在与源库同名约束会导致建表失败。 ● 全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量或单增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果转换失败或执行失败会忽略错误，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify

类型名称	使用和操作限制
	<p>column以及truncate table的基本DDL，不支持默认值等的修改。</p> <ul style="list-style-type: none"> - 增量DDL新增的字段不会自动扩展字符长度。 ● 增量同步阶段，修改抓取任务的启动位点主要用于重新同步数据。 <ul style="list-style-type: none"> - 修改抓取位点之后，上一次对象级对比结果不会展示出来。 - 单独修改抓取任务的启动位点，会把该位点同步到回放任务的启动位点上，即回放任务的启动位点和抓取任务的启动位点一致，不影响用户单独修改回放任务的启动位点。 ● 全量和增量同步不支持隐藏列（UNUSED, INVISIBLE）。 ● 当目标库为GaussDB分布式版时，增量同步阶段源库Oracle对分布列的update操作会导致在GaussDB数据库执行失败，造成数据不一致，所以业务上要尽量规避对分布列的更新。 ● 建议目标库禁用全局二级索引（GSI），否则可能会导致增量同步失败。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。 ● 当双AZ任务发生切换（例如主任务AZ级故障）时，如果主任务延迟较大，或源数据库有长事务未提交，任务倒换后可能因源数据库日志被清理导致任务无法启动，或着因为延迟、长事务期间的DDL变更导致同步异常。 ● 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-153 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-164 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-154 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建迁移任务。

* 数据流动方向: [入云] [出云] [新建] [编辑] ?
DRS要求源数据库和目标数据库中至少有一方为云数据库实例，入云目标数据库为非云数据库实例的场景，非云数据库实例之间的同步，请选传入云。

* 源数据库引擎: MySQL Oracle DB2 for LUW DDM MongoDB PostgreSQL Microsoft SQL Server TiDB

* 目标数据库引擎: MySQL DDM GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版 PostgreSQL GaussDB(for MySQL)

* 网络类型: [公网网络] ?
 DRS将会自动为DRS实例绑定弹性公网IP，请在各地机房将自助删除弹性公网IP。

* 目标数据库实例: [选择实例] [重置数据库实例] [查看不可选实例]

保持实例所在子网: [选择子网] ? [查看子网]

* 同步模式: [全量 + 增量] [全量] [增量] ?
该模式为数据库实时同步，通过全量过程与目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标数据库数据保持一致。

表 3-165 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。

- 任务类型

图 3-155 任务类型



表 3-166 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-156 企业项目和标签



表 3-167 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 3-157 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-168 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选项项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

图 3-158 目标库信息

表 3-169 目标库信息


参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 3-159 同步模式

表 3-170 同步对象

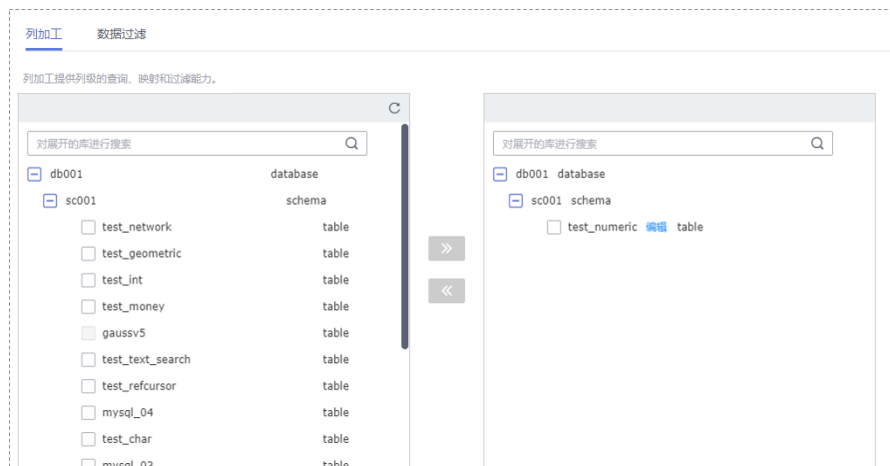
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-160 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
增量读取方式	<p>选择增量阶段读取日志方式。</p> <ul style="list-style-type: none"> LogMiner: 采用Oracle官方接口读取重做日志文件, 稳定性好。 Xstream: 需要开启Xstream接口。 <p>目前建议使用LogMiner。</p>
增量支持DDL	<p>选择增量同步的DDL类型, 用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”: 表级同步情况下, 已选定的同步表DDL将会同步至目标, 不支持的DDL将不同步。目前支持的DDL有CREATE_TABLE, ALTER_TABLE, TRUNCATE_TABLE。 “自定义”: 用户根据需求, 选择需要同步的DDL类型, 不同链路支持的DDL类型以界面显示为准。
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 可选表级同步或导入对象文件, 您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中, 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射, 具体操作可参考对象名映射。 选择导入对象文件, 具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面, 选择需要加工的数据, 进行列加工或者数据过滤。

- 如果不需要数据加工, 单击“下一步”。
- 如果需要加工列或者数据过滤, 参考[数据加工](#)中的列加工, 设置相关规则。

图 3-161 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-162 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-163 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 3-171 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.16 将 Oracle 同步到 DDM

支持的源和目标数据库

表 3-172 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建数据库 (Oracle 10g、11g、12c、18c、19c、21c版本) ECS自建数据库 (Oracle 10g、11g、12c、18c、19c、21c版本) 	<ul style="list-style-type: none"> DDM实例

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-173 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。</p> <p>12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> </p> - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库端：提供的目标数据库账号必须拥有如下权限：SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、INDEX、EVENT、RELOAD、CREATE VIEW。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 增量同步不支持DDL的同步。 ● 全量阶段不支持bfile, xml、sdo_geometry、urowid、interval（精度大于6位）和自定义类型。 ● 增量阶段不支持bfile, xml、interval、sdo_geometry、urowid、timestamp（精度大于6位）和自定义类型。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 目前只支持同步源库的数据，不支持同步源库表结构及其他数据库对象。 ● 用户需要在目标库根据源端逻辑库的表结构，自行在目标库创建对应的表结构及索引。未在目标库创建的对象，视为用户不选择这个对象进行同步。 ● 同步时需要在目标库创建表结构，目标库表结构要包含源库所有列，且主键要一致。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会因为无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● Oracle单行记录不能超过8K（lob、long类型除外，会自动转换成MySQL的text、blob类型），原因是MySQL innodb引擎限制单行大小不能超过8K（text、blob类型除外）。 ● 不建议以字符串类型作为主键或唯一键，因为Oracle的字符串作为主键、唯一键时区分空格，而MySQL不区分，可能导致数据不一致和死锁问题。 ● 对于Oracle的binary_float或者binary_double类型，MySQL中不支持设置Nan、Inf、-Inf三种值，DRS默认会将这三种值转为0保存。 ● Oracle中建议列名不要取名AUTO_PK_ROW_ID，原因是这个列名在MySQL5.7中是保留列名，无法创建出来。 ● Oracle中number(p, s)字段的精度不要超过p: [1, 38], s:[p-65, min(p, 30)]的精度表示范围。其中，s取值依赖于p的取值变化，即下限为p-65, 上限为p或30中取最小值。例如：当p=1, s的取值范围是[-64, 1]。当p=38, s取值范围是[-27, 30]。 int字段的值不要超过（65, 0）的精度表示范围。原因是MySQL数字的表示范围比Oracle小。 ● 不支持表名包含除下划线外的其他特殊字符的表的同步。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 不支持选择源数据库的空库进行同步。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> ● 同步前，目标数据库必须存在待同步数据库及表，且库名，表名，列名，索引名、约束名等必须为对应的小写名称。 ● DRS同步时会有大量数据写入目标库，目标库max_allowed_packet 参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作限制	<ul style="list-style-type: none"> ● 表对象名同步到目标库后会转换成小写，如ABC和abc。因此增量同步阶段，选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 如有中文、日文等特殊字符，业务连接Oracle数据库使用的编码需和Oracle数据库服务端编码一致，否则目标库会出现乱码。 ● Oracle中表结构同步到DDM后表的字符集为utf8mb4。 ● Oracle中表结构长度（所有列长字节数之和，char、varchar2等类型字节长度和编码有关）超过65535时，可能导致同步失败。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 数据类型不兼容时，可能引起同步失败。 ● 当Oracle字符集是WE8MSWIN1252时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，源库不能做DDL变更。 ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 对于同步中的数据库对象，在同步期间，目标库不能进行写入操作，否则会导致数据不一致。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 选择表级对象同步时，增量同步过程中不建议对表进行重命名操作。 ● 源库的用户对应目标库的数据库。 ● 源库用户、表结构信息同步至目标库后全部转换为小写。如表Ab及表AB同步至目标库为ab。 ● 不支持索引组织表的同步。 ● 同步任务全量阶段开始前，如有长时间未提交的事务，有可能丢失数据。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-164 同步任务信息

表 3-174 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-165 同步实例信息

表 3-175 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“DDM”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	可用的DDM实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将用户选择的数据库对象和数据一次性同步至目标端数据库。

- 任务类型

图 3-166 任务类型



表 3-176 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-167 企业项目和标签



表 3-177 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-168 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

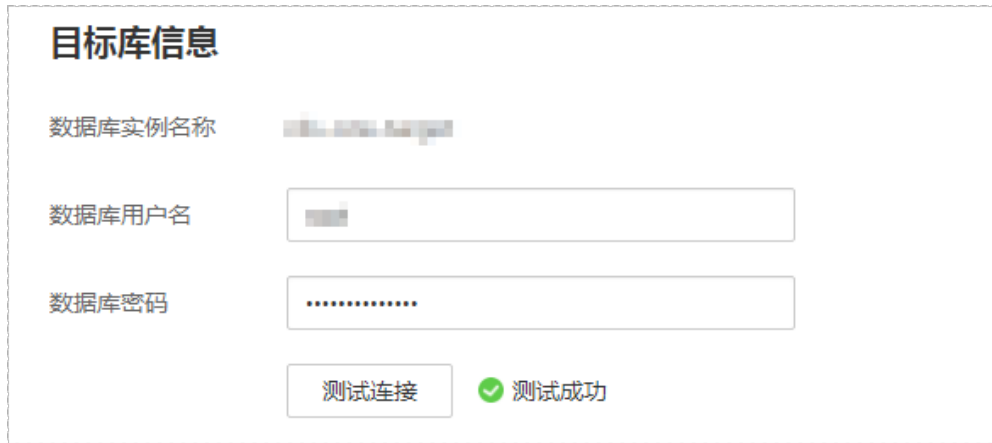
表 3-178 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该任务后自动清除。

图 3-169 目标库信息



目标库信息

数据库实例名称: [模糊处理]

数据库用户名: [输入框]

数据库密码: [输入框]

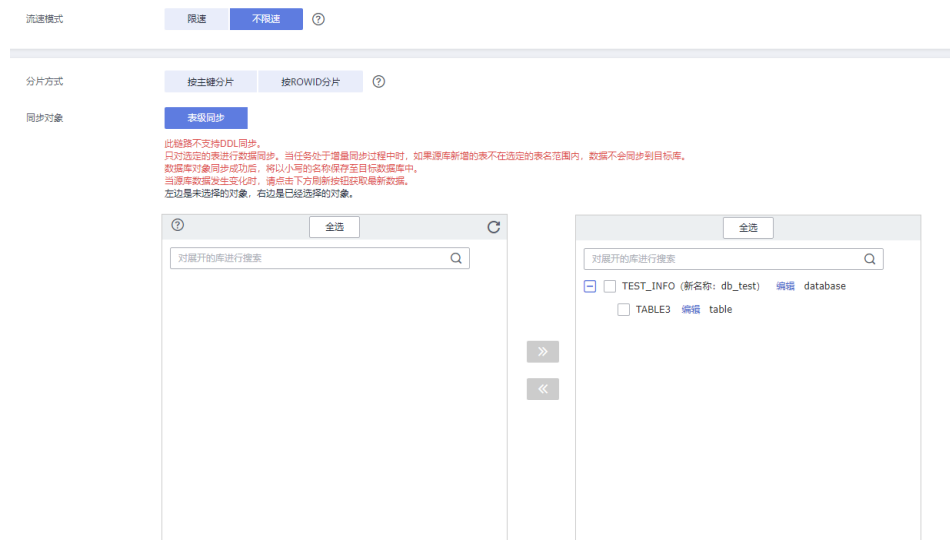
测试连接 ✔ 测试成功

表 3-179 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的DDM实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象。单击“下一步”。

图 3-170 同步模式



同步模式: 有限 不有限

分片方式: 按主键分片 按ROWID分片

同步对象: 表级同步

此策略不支持DDL同步。
只对选定的表进行数据同步，当任务处于增量同步过程中时，如果源库新增的表不在选定的表名范围内，数据不会同步到目标库。
数据库对象同步成功后，将以小写的名称保存至目标数据库中。
当源库数据发生变化时，请点击下方刷新按钮以获取最新数据。
左边是未选择的对象，右边是已经选择的对象。

对展开的库进行搜索 [搜索框]


对展开的库进行搜索 [搜索框]

TEST_INFO (新名称: db_test) [编辑] database

TABLE3 [编辑] table

表 3-180 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-171 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
分片方式	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-172 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-173 任务启动设置



表 3-181 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">当时延阈值设置为0时，不会发送通知给收件人。首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.17 将 Oracle 同步到 GaussDB(DWS)

支持的源和目标数据库

表 3-182 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） 	<ul style="list-style-type: none"> GaussDB(DWS)集群

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-183 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量和增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。</p> <p>12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> </p> - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库账号： <ul style="list-style-type: none"> - 目标数据库不存在DATABASE时需要有建库的权限；不存在SCHEMA时需要有在database中创建schema的权限；不存在表时需要有在对应schema中创建表的权限。 - 对于每张表需要有INSERT、SELECT、UPDATE、DELETE权限。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表、普通索引（B-Tree索引）、约束（主键、空、非空）的同步，不支持视图、外键、存储过程、触发器、函数、事件、虚拟列、唯一约束、唯一索引、外键索引、Check约束的同步。 ● 全量阶段不支持bfile, xml、sdo_geometry、urowid、interval（精度大于6位）和自定义类型。 ● 增量阶段不支持bfile, xml、interval、sdo_geometry、urowid、timestamp（精度大于6位）和自定义类型。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 对于TIMESTAMP WITH TIME ZONE类型，根据目标库时区做转换后不得大于“9999-12-31 23:59:59.999999”。 ● 仅支持to_date和sys_guid函数作为默认值，将其他函数作为默认值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，会出现以下情况： <ul style="list-style-type: none"> - 默认值函数会被置空。 - 创建表失败，导致任务失败。 - 增量同步的DDL如果包含目标数据库不存在的系统函数，会导致同步任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库表中的虚拟列。 ● 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 ● 选择手动创建表结构时，目标库中的时间类型是否带有时区需要与源库中保持一致，否则可能会因为时区转换导致时间数据不一致。 ● 源数据库Oracle中的物化视图同步到目标数据库会被转化为普通表。 ● 同步对象列名不能为CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID等GaussDB(DWS)禁止的字段，否则会导致任务失败。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> • 库名、表名不支持的字符有：非ASCII字符、“.”、“>”、“<”、“\”、“\`”、“ ”、“,”、“?”、“!”、“”和“'”。 • 不支持选择源数据库的空库进行同步。 • 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> • 目标数据库实例的运行状态必须正常。 • 目标数据库实例必须有足够的磁盘空间。 • 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。 ● 表等对象名同步到目标库后会转换成小写，如ABC会转换为abc。因此增量同步阶段，选择的源库的表中不能存在仅大小写不同的表，可能会导致同步失败。 ● 当Oracle字符集是WE8MSWIN1252时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 索引同步只同步普通索引，主键等约束在表结构中进行同步。 ● 同步表结构时，如果目标库存在与源库同名约束会导致建表失败。 ● 如有中文、日文等特殊字符，业务连接Oracle数据库使用的编码需和Oracle数据库服务端编码一致，否则目标库可能会出现乱码。 ● Oracle中表结构长度（所有列长字节数之和，char、varchar2等类型字节长度和编码有关）超过65535时，可能导致同步失败。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 使用DRS同步表结构的场景下，同一个schema中，同步的表中，表、约束、索引等不能有忽略大小写后的同名对象，比如表"A"中有索引名"inx1"，表B中有索引名"a"，表A和索引"a"忽略大小写重名了，会导致结构同步失败。多个schema映射到一个schema的场景，源库的多个schema中也不能包含同名但字母大小写不同的表、约束和索引。 ● 支持源端多张表对GaussDB(DWS)一张表的映射。详细操作可参考创建Oracle到GaussDB(DWS)同步任务

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 同步过程中，不允许删除连接源和目标数据库的用户的用户名、密码、权限，或修改目标数据库的端口号。 ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果转换失败或执行失败会忽略错误，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table modify column以及truncate table的基本DDL，不支持create index之类的语句。 ● 库级映射和表级映射均不区分大小写，例如映射为abc与映射为ABC，同步到目标库后均为abc。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-174 同步任务信息

表 3-184 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-175 同步实例信息

表 3-185 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“GaussDB(DWS)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	可用的GaussDB(DWS)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将用户选择的数据库对象和数据一次性同步至目标端数据库。 - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 3-176 任务类型



表 3-186 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-177 企业项目和标签



表 3-187 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-178 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-188 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。

参数	描述
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该任务后自动清除。

图 3-179 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码


 测试成功

表 3-189 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(DWS)实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象。单击“下一步”。

图 3-180 同步模式

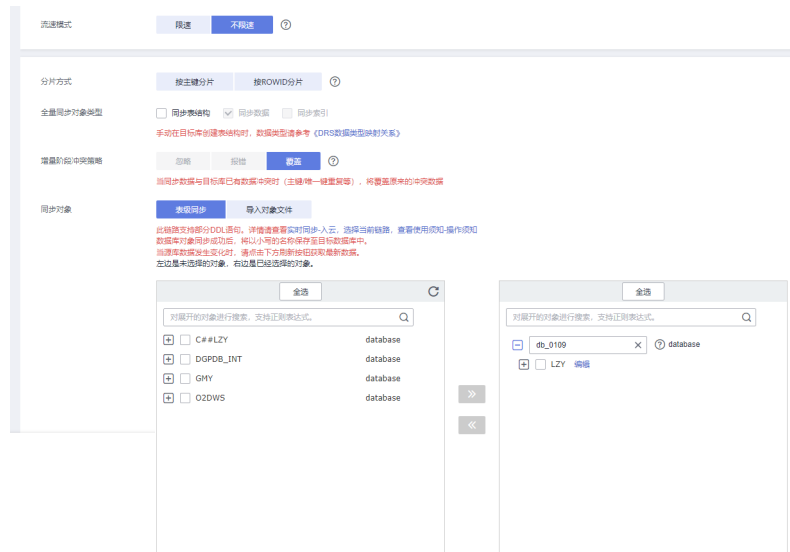



表 3-190 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-181 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
分片方式	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，可对需要加工的表对象进行数据过滤或添加附加列，检查无误后，单击“下一步”。

- 如果需要设置数据过滤，选择“数据过滤”，设置相关过滤规则。
- 如果需要设置添加附加列，选择“附加列”，单击“操作”列的“添加”，选填需要添加的列名和操作类型信息。

相关操作可参考[数据加工](#)。

图 3-182 数据加工



源数据库	表名称	列名	操作类型	字符类型	操作
CHENHAO_AB	CHENHAO_AB	--	--	--	添加
CHENHAO_ABC	CHENHAO_ABC	--	--	--	添加
CHENHAO_DIF	CHENHAO_DIF	--	--	--	添加
CHENHAO_IUFO_MEASPub_V\$FA	CHENHAO_IUFO_MEASPub_V\$FA	--	--	--	添加
CHENHAO_IUFO_MEASPub_X\$FA	CHENHAO_IUFO_MEASPub_X\$FA	--	--	--	添加
CHENHAO_NA	CHENHAO_NA	--	--	--	添加
CHENHAO_NB	CHENHAO_NB	--	--	--	添加
CHENHAO_ONE	CHENHAO_ONE	--	--	--	添加
CHENHAO_ONE1	CHENHAO_ONE1	--	--	--	添加
CHENHAO_ORACLE1	CHENHAO_ORACLE1	--	--	--	添加

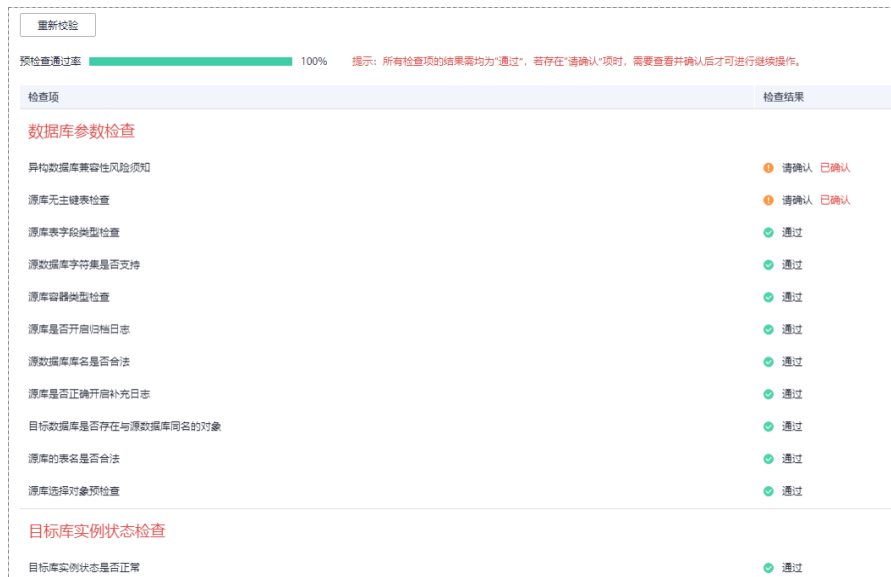
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-183 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-184 任务启动设置



表 3-191 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会出现一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.18 将 Oracle 同步到 PostgreSQL

支持的源和目标数据库

表 3-192 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本）ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本）	<ul style="list-style-type: none">RDS for PostgreSQL

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-193 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。</p> <p>12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> </p> - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库账号必须具有每张表的如下权限：INSERT、SELECT、UPDATE、DELETE、CONNECT、CREATE、REFERENCES。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表、索引、约束（主键、空、非空）的同步，不支持视图、外键、存储过程、触发器、函数、事件、虚拟列的同步。 ● 全量阶段不支持bfile, xml、sdo_geometry、urowid、interval（精度大于6位）和自定义类型。 ● 增量阶段不支持bfile, xml、interval、sdo_geometry、urowid、timestamp（精度大于6位）和自定义类型。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 对于TIMESTAMP WITH TIME ZONE类型，根据目标库时区做转换后不得大于“9999-12-31 23:59:59.999999”。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会因为无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。 ● 选择手动创建表结构时，目标库中的时间类型是否带有时区需要与源库中保持一致，否则可能会因为时区转换导致时间数据不一致。
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名不支持的字符有：非ASCII字符、“.”、“>”、“<”、“\”、“`”、“ ”、“,”、“?”、“!”、“””和“'”。 ● 不支持选择源数据库的空库进行同步。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。 ● 表等对象名同步到目标库后会转换成小写，如ABC会转换为abc。因此增量同步阶段，选择的源库的表中不能存在仅大小写不同的表，可能会导致同步失败。 ● 如有中文、日文等特殊字符，业务连接Oracle数据库使用的编码需和Oracle数据库服务端编码一致，否则目标库可能会出现乱码。 ● 不使用DRS同步表结构（自建表结构）的场景下，无主键表如果因表结构问题导致任务失败，修复表结构后续无法恢复，需要重置任务。 ● 同步表结构时，如果目标库存在与源库同名约束会导致建表失败。 ● 使用DRS同步表结构的场景下，同一个schema中，同步的表中，表、约束、索引等不能有忽略大小写后的同名对象，比如表"A"中有索引名"inx1"，表B中有索引名"a"，表A和索引"a"忽略大小写重名了，会导致结构同步失败。多个schema映射到一个schema的场景，源库的多个schema中也不能包含同名但字母大小写不同的表、约束和索引。 ● 当Oracle字符集是WE8MSWIN1252时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● Oracle中表结构长度（所有列长字节数之和，char、varchar2等类型字节长度和编码有关）超过65535时，可能导致同步失败。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 同步过程中，不允许删除连接源和目标数据库的用户的用户名、密码、权限，或修改目标数据库的端口号。 ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 索引同步只同步普通索引，主键等约束在表结构中进行同步。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果转换失败或执行失败会忽略错误，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify column以及truncate table的基本DDL，不支持默认值等的修改。 ● 表级映射不区分大小写，例如映射为abc与映射为ABC，同步到目标库后均为abc。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● 全量同步过程中，DRS会向目标库PostgreSQL写入大量数据，会导致PostgreSQL的wal日志量急剧增长，PostgreSQL的磁盘有被写满的风险。可以通过在全量同步前关闭PostgreSQL的日志备份功能，减少wal日志的生产，同步完成后再将其打开的方式进行规避（具体操作方法可参考设置自动备份策略）。 <p>注意 关闭日志备份会影响数据库的灾备恢复，请根据实际情况谨慎选择。</p>

操作步骤

本小节以Oracle到RDS for PostgreSQL的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-185 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会自动创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-194 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-186 同步实例信息

同步实例信息 ?

以下配置项均为必填项，配置完成后，请单击“开始创建”按钮，系统将自动创建资源。

源数据库实例: [Dropdown] ?

源数据库引擎: MySQL, Oracle, DB2 for LUW, DB2, Microsoft SQL Server, TDS

目标数据库引擎: MySQL, DB2, GaussDB(DWS), GaussDB (Doris), GaussDB 全量, PostgreSQL, GaussDB for PostgreSQL

同步策略: [Dropdown] ?

是否自动创建资源: 勾选后将自动创建资源，勾选后将自动创建资源，勾选后将自动创建资源。

同步实例名称: [Text Field] ?

同步实例名称: [Text Field] ?

同步策略: [Dropdown] ?

开始创建

表 3-195 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“PostgreSQL”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	创建好的RDS for PostgreSQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将用户选择的数据库对象和数据一次性同步至目标端数据库。

● 任务类型

图 3-187 任务类型



表 3-196 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-188 企业项目和标签



表 3-197 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-189 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-198 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-190 目标库信息

表 3-199 目标库信息


参数	描述
数据库实例名称	默认为创建任务时选择的RDS for PostgreSQL实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。支持在任务创建后修改密码。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-191 同步模式

表 3-200 同步模式和对象

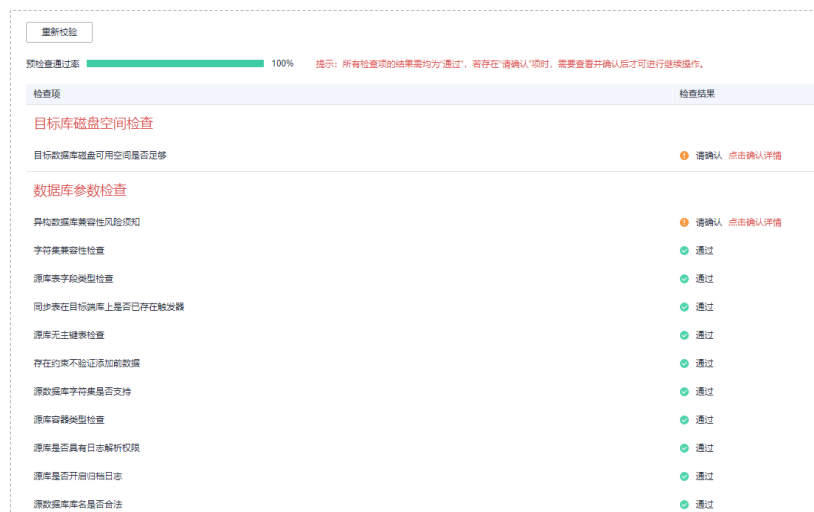
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-192 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。选择数据的时候支持搜索，以便您快速选择需要的数据库对象。</p> <ul style="list-style-type: none"> 选择导入对象文件，具体步骤和说明可参考导入同步对象。 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-193 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-194 任务启动设置

The screenshot shows a configuration interface for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Latency Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto End Time): An input field containing the number '14', with a help icon and a red note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 3-201 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.19 将 DDM 同步到 MySQL

支持的源和目标数据库

表 3-202 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • DDM实例 	<ul style="list-style-type: none"> • RDS for MySQL

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-203 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> 源数据库DDM账户至少需要具备一个权限，比如SELECT；DDM物理分片数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 提供的目标数据库账号必须拥有如下权限：SELECT、CREATE、ALTER、DROP、DELETE、INSERT、UPDATE。RDS for MySQL实例的root账户默认已具备上述权限。
同步对象约束	<ul style="list-style-type: none"> 全量同步支持数据、表结构和索引的同步。 源库不允许存在拆分键为timestamp类型的表。 源表的分库分表键要加到目标表的主键和唯一键中（也就是目标表的主键和唯一键中的列应该包含源表的分片列），避免数据冲突出现数据不一致问题。
源数据库要求	<ul style="list-style-type: none"> 增量同步时，源物理分片数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 源分库分表中间件中的库名、表名不能包含：'<>/\以及非ASCII字符。 MySQL源数据库建议开启skip-name-resolve，减少连接超时的可能性。 源物理分片数据库GTID状态建议为开启状态。
目标数据库要求	<ul style="list-style-type: none"> 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 目标数据库实例必须有足够的磁盘空间。 除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 目标数据库的字符集必须与源数据库一致。 目标数据库的时区设置必须与源数据库一致。 DRS同步时会有大量数据写入目标库，目标库max_allowed_packet参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 数据类型不兼容时，可能引起同步失败。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 支持断点续传功能，在主机系统崩溃的情况下，对于无主键的表可能会出现重复插入数据的情况。 ● 源数据库中存在主键或唯一键重复的数据时，直接同步将导致目标库数据比源库少，请务必检查并订正数据后启动同步。 ● 目标库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 任务创建后，目标库不能设置为只读。 ● 任务创建后，源数据库不支持增加逻辑库或修改旧逻辑库关联新的RDS，否则会导致数据无法正常同步或任务失败。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源端DDM正在同步的表做改变拆分键的操作；也不允许将单表/广播表改为拆分表，拆分表改为单表/广播表。 ● 全量数据同步阶段，目标数据库会产生大量Binlog引起磁盘空间占用过大，因此在全量数据同步阶段目标数据库Binlog本地默认调整为只保留最新5个，全量同步结束后目标数据库Binlog本地保留时长恢复为用户当前配置。如果因业务需要，目标数据库本地Binlog保留时长需要保持用户当前配置，需提交工单申请开启。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 ● 增量同步阶段，不支持源数据库进行恢复操作 ● 增量同步阶段，支持部分DDL操作。 <ul style="list-style-type: none"> - 不支持 DROP_DATABASE、DROP_TABLE、TRUNCATE_TABLE、CREATE_VIEW、DROP_VIEW。 - 不支持使用Online DDL。 - 支持创建表，例如： <pre>create table `ddl_test` (id int, c1 varchar(25), primary key(id)); create table `ddl_test_gho` like `ddl_test`;</pre> - 支持表的重命名，源表和目標表必須都在对象选择里面，例如：

类型名称	使用和操作限制
	<pre>rename table `ddl_test` to `ddl_test_new`;</pre> <ul style="list-style-type: none"> - 支持表字段的增和改，不支持删列，例如： <pre>alter table `ddl_test` add column `c2` varchar(25);</pre> <pre>alter table `ddl_test` modify column `c1` varchar(50);</pre> <pre>alter table `ddl_test` alter c1 set default '***';</pre> - 支持修改表索引，例如： <pre>alter table `ddl_test` drop primary key;</pre> <pre>alter table `ddl_test` add primary key(id);</pre> <pre>alter table `ddl_test` add index `ddl_test_uk`(id);</pre> <pre>alter table `ddl_test` drop index `ddl_test_uk`;</pre> - 表级同步支持增加列、修改列、增加主键和普通索引。 - 库级同步支持新建表、rename表、增加列、修改列、增加主键和普通索引。 - 新增和修改表名、列名、索引名时不能超出63字符，否则任务会失败。 - 源库无主键表增加主键的时候，必须含有第一列，否则任务会失败。 <ul style="list-style-type: none"> ● 增量同步阶段，对同一张表或列做DDL操作，需要在业务低峰期，并且时间间隔1分钟以上。 ● 任务再编辑时，对新添加的表，在再编辑变更过程中禁止大批量导出数据，建议在业务低峰期进行再编辑操作。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-195 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 3-204 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 3-196 同步实例信息

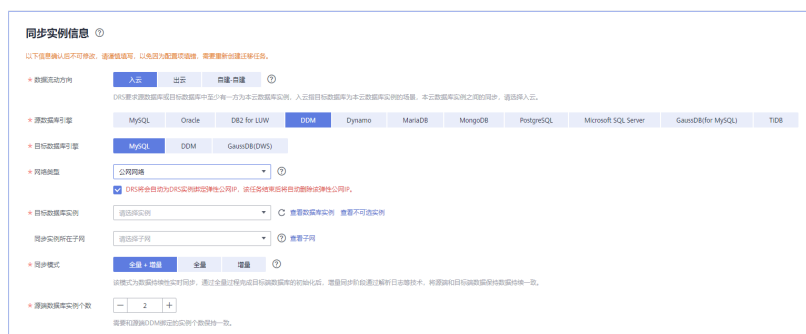


表 3-205 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“DDM”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的关系型数据库实例。

参数	描述
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。
源端数据库实例个数	源端DDM绑定的实例个数，默认值2，输入值在1到64之间，您需要根据源端实际情况设置该值大小。

● 任务类型

图 3-197 任务类型



表 3-206 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更佳。

● 企业项目和标签

图 3-198 企业项目和标签



表 3-207 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-199 源库信息

表 3-208 源库信息

参数	描述
中间件IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
中间件用户名	源数据库DDM中间件的用户名。
中间件密码	源数据库DDM中间件用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。
数据库实例	根据源库实际的分片数据库，填写对应的数据库信息。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-200 目标库信息

目标库信息

数据库实例名称

数据库用户名: root

数据库密码

SSL安全连接:

测试连接: ✔ 测试成功

表 3-209 目标库信息


参数	描述
数据库实例名称	默认为创建迁移任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。
SSL安全连接	如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 3-201 同步模式



表 3-210 同步模式和对象

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-202 预检查



说明

所有检查项结果均通过时, 若存在待确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-203 任务启动设置



表 3-211 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会出现一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.20 将 DDM 同步到 GaussDB(DWS)

支持的源和目标数据库

表 3-212 支持的数据库

源数据库	目标数据库
• DDM实例	• GaussDB(DWS)集群

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-213 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库DDM账户至少需要具备一个权限，比如SELECT；DDM物理分片数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 提供的目标数据库账号必须具有每张表的如下权限：INSERT、SELECT、UPDATE、DELETE、CONNECT、CREATE。
同步对象约束	<ul style="list-style-type: none"> • 全量同步支持数据、表结构和普通索引（B-Tree索引）的同步。 • 目标库不支持唯一键表，同步过程中将忽略源数据库中的唯一键表。 • 不支持同步无主键表，若选择同步的表中存在无主键表，则同步失败。 • 不支持唯一约束、唯一索引的同步。 • 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 • 同步对象列名不能为CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID等GaussDB(DWS)禁止的字段，否则会导致任务失败。
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源物理分片数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 • 在磁盘空间允许的情况下，binlog保存时间越长越好，建议为3天。 • 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 • 源分库分表中间件中的库名、表名不能包含：'<>/\以及非ASCII字符。 • 源物理分片数据库建议开启skip-name-resolve，减少连接超时的可能性。 • 源物理分片数据库的GTID状态建议为开启状态。 • 源物理分片数据库不支持枚举类型和set集合类型的。 • 源库的timestamp列的默认值，需要在目标库的合理取值内，否则会导致同步失败。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 目标数据库的时区设置必须与源数据库一致。
操作须知	<ul style="list-style-type: none"> ● 数据类型不兼容时，可能引起同步失败。 ● 源数据库中存在主键重复的数据时，直接同步将导致目标库数据比源库少，请务必检查并订正数据后启动同步。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 在源端，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB(DWS)中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由“原库名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key，或表名_原索引名。 ● 全量阶段基于索引实现上差异，只保留普通B-Tree索引，其他索引默认不同步。GaussDB(DWS)索引过多时会影响存储空间和入库性能，建议用户可根据业务需求按需建索引。 ● 任务创建后，源数据库不支持增加逻辑库或修改旧逻辑库关联新的RDS，否则会导致数据无法正常同步或任务失败。 ● 同步过程中，不允许源端DDM正在同步的表做改变拆分键的操作；也不允许将单表/广播表改为拆分表，拆分表改为单表/广播表。 ● 增量同步阶段，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步阶段，若需要对源库需要同步的表结构进行修改，则用户必须在目标库同步修改表结构。 ● 增量同步阶段，支持部分DDL语句，包括：ADD COLUMN、CREATE TABLE、MODIFY COLUMN、CREATE INDEX、DROP INDEX、RENAME INDEX。 ● 增量同步阶段，对同一张表或列做DDL操作，需要在业务低峰期，并且时间间隔1分钟以上。 ● 增量同步阶段，不支持源数据库进行恢复操作。 ● 增量同步阶段，不支持delete和update目标端主键类型为binary, text, blob, clob的表。 ● 任务再编辑时，对新添加的表，在再编辑变更过程中禁止大批量导出数据，建议在业务低峰期进行再编辑操作。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-204 同步任务信息

表 3-214 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-205 同步实例信息

表 3-215 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“DDM”。
目标数据库引擎	选择“GaussDB(DWS)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、“VPN和专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	可用的GaussDB(DWS)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>
源端数据库实例个数	源端DDM绑定的实例个数，默认值2，输入值在1到64之间，您需要根据源端实际情况设置该值大小。

- 任务类型

图 3-206 任务类型



表 3-216 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-207 企业项目和标签



表 3-217 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-208 源库信息

表 3-218 源库信息

参数	描述
中间件IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
中间件用户名	源数据库DDM中间件的用户名。
中间件密码	源数据库DDM中间件用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。
数据库实例	根据源库实际的分片数据库，填写对应的数据库信息。

说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 3-209 目标库信息

目标库信息

数据库实例名称: [实例名称]

数据库用户名: [用户名]

数据库密码: [密码]

测试连接 ✔ 测试成功

表 3-219 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(DWS)实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 3-210 同步模式

增量阶段冲突策略: 忽略 报错 覆盖

当同步数据与目标库已有数据冲突时（主键/唯一键重复等），将覆盖原来的冲突数据

过滤DROP DATABASE: 是 否

同步对象: 表级同步 库级同步 导入对象文件

此检验支持部分DDL语句，详情请查看《实时同步-使用须知-当前场操作须知》
当源库数据发生变化时，请点击下方刷新按钮获取最新数据。
左边是未选择的对象，右边是已经选择的对象。

对展开的库进行搜索


- autotest_3000table database
- autotest_dirtydata_100w database
- autotest_empty database
- ddm2dws_10 database
- dws_autotest_dirtydata database
- dws_autotest_normal database
- dws_moretoone database
- src_ddmtest database

对展开的库进行搜索

[] database

auto_test_with_all_data 编辑

表 3-220 同步模式和对象

参数	描述
增量阶段冲突策略	该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。
过滤DROP DATABASE	实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的表对象，填写需要添加的列名、类型、操作类型信息，检查无误后，单击“下一步”。可参考[数据加工](#)中的内容，设置相关规则，

图 3-211 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-212 预检查



说明

所有检查项结果均通过时, 若存在待确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-213 任务启动设置



表 3-221 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.21 将 DDM 同步到 DDM

支持的源和目标数据库

表 3-222 支持的数据库

源数据库	目标数据库
• DDM实例	• DDM实例

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-223 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库DDM账户至少需要具备一个权限，比如SELECT；DDM物理分片数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 目标中间件账户需要具备以下基本权限：CREATE、DROP、ALTER、INDEX、INSERT、DELETE、UPDATE、SELECT，同时必须具备扩展权限：全表SELECT权限。 • 目标中间件账户必须具备对所同步数据库的权限。
同步对象约束	<ul style="list-style-type: none"> • 全量同步支持数据、表结构和索引的同步。 • 源库不允许存在拆分键为timestamp类型的表。 • 仅支持MyISAM和InnoDB表的同步。
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源物理分片数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 • 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 • 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 • 源分库分表中间件中的库名、表名不能包含：'<>/\以及非ASCII字符。 • MySQL源数据库建议开启skip-name-resolve，减少连接超时的可能性。 • 源物理分片数据库GTID状态建议为开启状态。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> • 目标库需要提前创建逻辑库。 • 目标库若已存在数据，DRS在增量同步过程中源库相同主键的数据将覆盖目标库已存在的数据，因此在同步前需要用户自行判断数据是否需要清除，建议用户在同步前自行清空目标库。 • 目标实例及关联RDS实例的运行状态必须正常，若关联RDS实例是主备实例，复制状态也必须正常。 • 目标库关联RDS实例必须有足够的磁盘空间。 • 目标库关联RDS数据库的字符集必须与源数据库一致。 • 目标库实例若选择将时间戳类型（TIMESTAMP，DATETIME）的列作为分片键，则源库数据在同步到目标库之后，作为分片键的该时间戳类型列的秒精度将被丢弃。 • 目标数据库存在表的AUTO_INCREMENT值至少不能小于源库表的AUTO_INCREMENT值。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 数据类型不兼容时，可能引起同步失败。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 任务创建后，源数据库不支持增加逻辑库或修改旧逻辑库关联新的RDS，否则会导致数据无法正常同步或任务失败。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做删除类型的DDL操作，这样可能会引起任务同步失败。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许对源库需要同步的表结构进行修改。 ● 同步过程中，不允许源端DDM正在同步的表做改变拆分键的操作；也不允许将单表/广播表改为拆分表，拆分表改为单表/广播表。 ● 增量同步阶段，不支持源数据库进行恢复操作。 ● 增量同步阶段，支持部分DDL语句，包括：ADD COLUMN、CREATE TABLE、CREATE INDEX、DROP INDEX、RENAME INDEX、ADD INDEX。 ● 增量同步阶段，对同一张表或列做DDL操作，需要在业务低峰期，并且时间间隔1分钟以上。 ● 选择表级对象同步时，增量同步过程中不建议对表进行重命名操作。 ● 任务再编辑时，对新添加的表，在再编辑变更过程中禁止大批量导出数据，建议在业务低峰期进行再编辑操作。 ● 当目标DDM版本大于3.0.4.1版本，结束任务时，DRS会自动刷新DDM的自增列(sequence)起始值。 ● 建议将源中间件物理分片的expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-214 同步任务信息

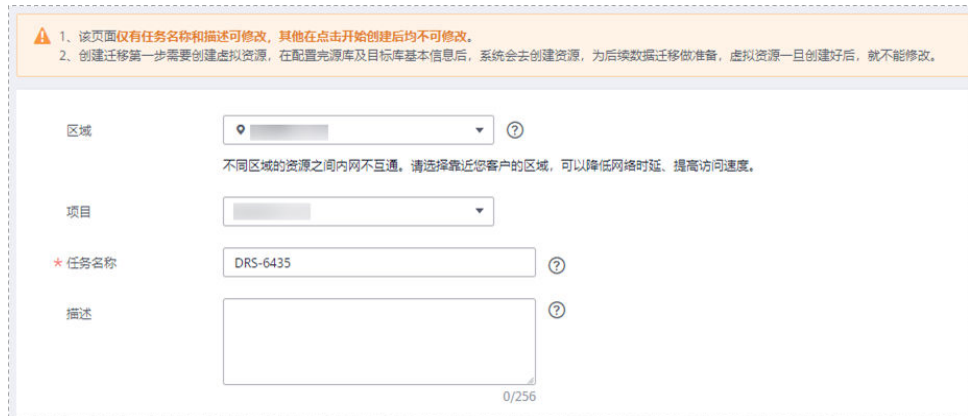


表 3-224 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-215 同步实例信息



表 3-225 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“DDM”。
目标数据库引擎	选择“DDM”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、“VPN和专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的DDM实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>
源端数据库实例个数	源端DDM绑定的实例个数，默认值2，输入值在1到64之间，您需要根据源端实际情况设置该值大小。

- 任务类型

图 3-216 任务类型



表 3-226 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-217 企业项目和标签



表 3-227 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-218 源库信息

表 3-228 源库信息

参数	描述
中间件IP地址或域名	源数据库DDM中间件的IP地址或域名。
端口	源数据库DDM中间件服务端口，可输入范围为1~65535间的整数。
中间件用户名	源数据库DDM中间件的用户名。
中间件密码	源数据库DDM中间件用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。
数据库实例	根据源库实际的分片数据库，填写对应的数据库信息。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-219 目标库信息

目标库信息

数据库实例名称: [模糊处理]

数据库用户名: [输入框]

数据库密码: [输入框]

测试连接 ✔ 测试成功

表 3-229 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的DDM实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 3-220 同步模式

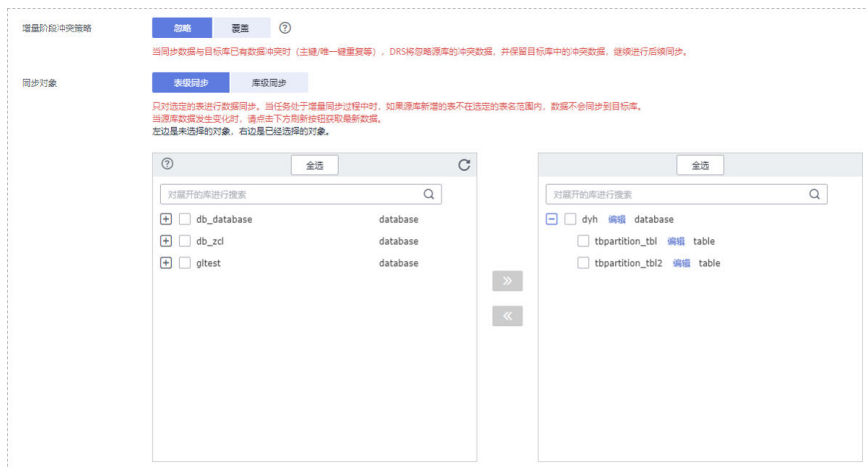



表 3-230 同步模式和对象

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。</p> <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-221 预检查



说明

所有检查项结果均通过时，若存在待确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-222 任务启动设置

The screenshot shows the 'Task Start Settings' configuration page. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Latency Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto End Time): An input field containing the value '14', with a help icon and a note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 3-231 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.22 将 DB2 for LUW 同步到 GaussDB 主备版

支持的源和目标数据库

表 3-232 支持的数据库

源数据库	目标数据库
DB2 for LUW 9.7、10.1、10.5、11.1、11.5版本	GaussDB主备版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-233 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步需要用户具有CONNECT、DATAACCESS权限。 - 全量+增量同步需要用户具有DBADM权限。 - 选择全量+增量同步时，需要确保开启源数据库的归档日志，确保需要同步的表的Datacapture属性为Y。 - 如果源数据库表结构存在DB2SECURITYLABEL字段类型，需要确保该用户具有该表全部数据的读权限。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 - SCHEMA级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 - 表级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下所有表的SELECT，UPDATE，INSERT和DELETE权限。 - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 不支持自定义类型。 ● timestamp类型支持的最大精度是6。 ● 全量同步结构不支持位图索引、倒排索引、函数索引、xml索引。 ● 增量同步支持的LOB类型大小限10M以内。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中执行的DDL语句。 ● 全量+增量任务不支持带XML字段的无主键表。 ● 浮点数类型作为主键时，增量同步可能会导致数据不一致。
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS不支持非ASCII字符、“.”、“>”、“<”、“\”、“.”、“ ”、“.”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：GBK、UTF8。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是本云GaussDB主备版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。
操作须知	<ul style="list-style-type: none"> ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步。 ● 不支持ARM虚拟机环境。 ● 如果无主键表中包含LOB、LONG等大字段，增量同步阶段存在大字段数据不一致的可能性。 ● 如果表结构存在DB2SECURITYLABEL字段类型，读取全量阶段数据可能会受DB2 LBAC（基于标签的访问控制）影响，导致源库和目标库数据不一致。 ● 如果源库和目标库字符集不一致，如源库是GBK，目标库是UTF8，由于GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 任务启动后，对于分区表的detach操作可能导致数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，源库不允许有未提交的DDL事务。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束，如有需要，用户必须手动在目标库创建对应表的默认值约束。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步阶段，支持任务再编辑追加同步对象。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-223 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [dropdown]

* 任务名称: DRS-6435 ?

描述: [text area] ?
0/256

表 3-234 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-224 同步实例信息

同步实例信息 ?

以下选项确认后不可修改，请谨慎填写，以免为后续操作带来麻烦，需要重新创建同步实例。

* 源实例所在区域: [dropdown] ?

* 源实例名称: MySQL Oracle DRS for IAM DDM MongoDB PostgreSQL Microsoft SQL Server TDS

* 目标实例所在区域: [dropdown] ?

* 目标实例名称: [dropdown] ?

* 同步模式: [radio buttons]

* 同步类型: [radio buttons]

* 同步方向: [radio buttons]

高级模式: [checkbox] ?

表 3-235 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“DB2 for LUW”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的GaussDB主备版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。

- 可用区

图 3-225 可用区



表 3-236 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-226 企业项目和标签



表 3-237 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 3-227 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码 👁

数据库名

SSL安全连接

同步驱动 已同步驱动: jjar

表 3-238 源库信息

参数	描述
IP地址或域名	源数据库DB2 for LUW的IP地址或域名。
端口	源数据库DB2 for LUW服务端口，可输入范围为1~65535间的整数。
数据库用户名	源DB2 for LUW的数据库账号。
数据库密码	源数据库账号所对应的密码。
数据库名	源DB2 for LUW中同步对象所属数据库的名称。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。目前仅DB2 for LUW 10.5及以上的版本支持SSL安全连接，DB2 for LUW 9.7和10.1版本暂不支持。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不使用SSL安全连接，请自行承担数据安全风险。 • DB2 for LUW SSL安全连接需要上传三个证书文件：jks证书、kdb证书和sth证书。 • kdb证书和sth证书测试连接时DRS不会进行校验，需要用户保证文件正确性，避免后续任务失败。
同步驱动	用户需要手动添加并上传源数据库DB2 for LUW对应的JDBC驱动包后，再进行测试连接。

说明

源数据库的数据库IP地址、用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-228 目标库信息

目标库信息

数据库实例名称: [模糊处理]

数据库用户名: [模糊处理]

数据库密码: [模糊处理]

测试连接 ✔ 测试成功

表 3-239 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	目标数据库的账号。
数据库密码	目标数据库账号所对应的密码。数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

说明


目标数据库的数据库的用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 3-229 同步模式



表 3-240 同步对象

参数	描述
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或者导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的数据过滤，设置相关规则。

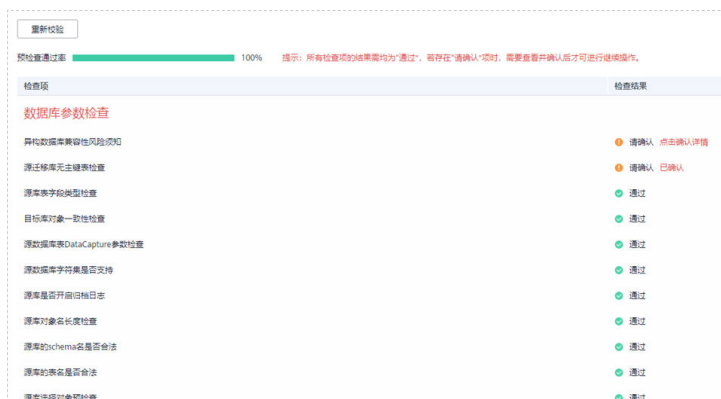
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-230 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-231 任务启动设置




表 3-241 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.23 将 DB2 for LUW 同步到 GaussDB 分布式版

支持的源和目标数据库

表 3-242 支持的数据库

源数据库	目标数据库
DB2 for LUW 9.7、10.1、10.5、11.1、11.5版本	GaussDB分布式版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-243 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库端： <ul style="list-style-type: none"> - 全量同步需要用户具有CONNECT、DATAACCESS权限。 - 全量+增量同步需要用户具有DBADM权限。 - 选择全量+增量同步时，需要确保开启源数据库的归档日志，确保需要同步的表的Datacapture属性为Y。 - 如果源数据库表结构存在DB2SECURITYLABEL字段类型，需要确保该用户具有该表全部数据的读权限。 • 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 - SCHEMA级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 - 表级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下所有表的SELECT，UPDATE，INSERT和DELETE权限。 - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有all privilege权限，详细信息可参考《GaussDB工具参考》。
同步对象约束	<ul style="list-style-type: none"> • 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 • 不支持自定义类型。 • timestamp类型支持的最大精度是6。 • 全量同步结构不支持位图索引、倒排索引、函数索引、xml索引。 • 增量同步支持的LOB类型大小限10M以内。 • 不支持默认值含有表达式的函数的表的同步。 • 不支持同步源库中的临时表。 • 不支持同步源库中执行的DDL语句。 • 全量+增量任务不支持带XML字段的无主键表。 • 浮点数类型作为主键时，增量同步可能会导致数据不一致。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“\`”、“ ”、“,”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：GBK、UTF8。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是本云GaussDB分布式版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步。 ● 不支持ARM虚拟机环境。 ● 如果无主键表中包含LOB、LONG等大字段，增量同步阶段存在大字段数据不一致的可能性。 ● 如果表结构存在DB2SECURITYLABEL字段类型，读取全量阶段数据可能会受DB2 LBAC（基于标签的访问控制）影响，导致源库和目标库数据不一致。 ● 如果源库和目标库字符集不一致，如源库是GBK，目标库是UTF8，由于GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 任务启动后，对于分区表的detach操作可能导致数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，源库不允许有未提交的DDL事务。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束，如有需要，用户必须手动在目标库创建对应表的默认值约束。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步阶段，支持任务再编辑追加同步对象。 ● 增量同步阶段，源库DB2 for LUW对主键列或无主键表的第一列的update操作会导致在GaussDB数据库更新分布列，有可能会造成数据不一致，所以业务上要尽量规避对上述列的更新。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-232 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-244 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-233 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向: 入云 | 出云 | 自建-自建 ?

DRS要求源数据库或目标数据库中至少有一方为本云数据库实例，入云指目标数据库为本云数据库实例的场景，本云数据库实例之间的同步，请选择入云。

* 源数据库引擎: MySQL | Oracle | DB2 for LUW | DDM | MongoDB | PostgreSQL | Microsoft SQL Server | TiDB

* 目标数据库引擎: GaussDB(DWS) | GaussDB 分布式版 | GaussDB 主备版

* 网络类型: 公网网络 ?
 DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除该弹性公网IP。

* 目标数据库实例: [下拉菜单] ? | 查看数据库实例 | 查看不可选实例

同步实例所在子网: [下拉菜单] ? | 查看子网

* 同步模式: 全量 + 增量 | 全量 ?

该模式为数据持续性实时同步，通过全量过程完成目标数据库的初始化后，增量同步和仅通过解析日志等技术，将源端和目标数据库数据保持一致。

表 3-245 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“DB2 for LUW”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。

- 可用区

图 3-234 可用区



表 3-246 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-235 企业项目和标签



表 3-247 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过，单击“下一步”。

图 3-236 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码 👁

数据库名

SSL安全连接

同步驱动 已同步驱动: jjar

表 3-248 源库信息

参数	描述
IP地址或域名	源数据库DB2 for LUW的IP地址或域名。

参数	描述
端口	源数据库DB2 for LUW服务端口，可输入范围为1~65535间的整数。
数据库用户名	源DB2 for LUW的数据库账号。
数据库密码	源数据库账号所对应的密码。
数据库名	源DB2 for LUW中同步对象所属数据库的名称。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。目前仅DB2 for LUW 10.5及以上的版本支持SSL安全连接，DB2 for LUW 9.7和10.1版本暂不支持。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不使用SSL安全连接，请自行承担数据安全风险。 • DB2 for LUW SSL安全连接需要上传三个证书文件：jks证书、kdb证书和sth证书。 • kdb证书和sth证书测试连接时DRS不会进行校验，需要用户保证文件正确性，避免后续任务失败。
同步驱动	用户需要手动添加并上传源数据库DB2 for LUW对应的JDBC驱动包后，再进行测试连接。

📖 说明

源数据库的数据库IP地址、用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-237 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

✔️ 测试成功

表 3-249 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	目标数据库的账号。
数据库密码	目标数据库账号所对应的密码。数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

说明

目标数据库的数据库的用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。


步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 3-238 同步模式



表 3-250 同步对象

参数	描述
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的数据过滤，设置相关规则。

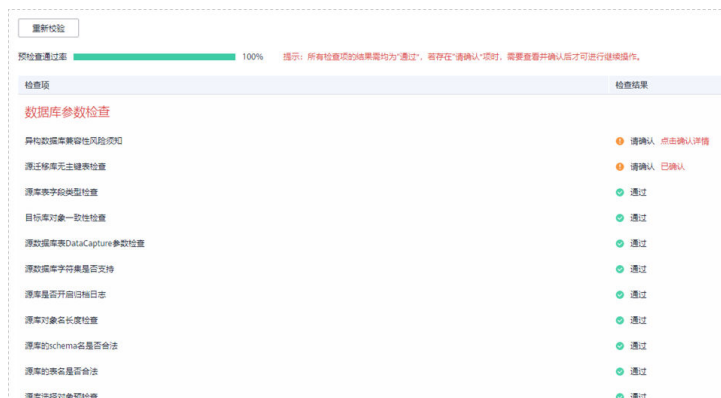
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-239 预检查



📖 说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-240 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: "立即启动" (Immediately Start) and "稍后启动" (Start Later), with a help icon. Below are several settings:


- 任务异常通知设置**: A toggle switch is turned on.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: An input field containing the number "14", with a help icon and a red note: "任务处于异常状态一段时间后, 将会自动结束。单位为天。"

表 3-251 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

3.24 将 DB2 for LUW 同步到 GaussDB(DWS)

支持的源和目标数据库

表 3-252 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • DB2 for LUW 9.7、10.1、10.5、11.1、11.5版本 	<ul style="list-style-type: none"> • GaussDB(DWS)

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-253 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步需要用户具有CONNECT、DATAACCESS权限。 - 全量+增量同步需要用户具有DBADM权限。 - 选择全量+增量同步时，需要确保开启源数据库的归档日志，确保需要同步的表的Datacapture属性为Y。 - 如果源数据库表结构存在DB2SECURITYLABEL字段类型，需要确保该用户具有该表全部数据的读权限。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 目标数据库不存在DATABASE时需要有建库的权限；不存在SCHEMA时需要有在database中创建schema的权限；不存在表时需要有在对应schema中创建表的权限。 - 对于每张表需要有INSERT、SELECT、UPDATE、DELETE权限。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 不支持自定义类型。 ● timestamp类型支持的最大精度是6。 ● 全量同步结构不支持位图索引、倒排索引、函数索引、xml索引。 ● 增量同步支持的LOB类型大小限10M以内。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中执行的DDL语句。 ● 全量+增量任务不支持带XML字段的无主键表。 ● 源库单个schema的同步对象数量不能超过32766。 ● 源库索引名超过63个字符时，目标数据库回放时将自动命名该索引。 ● 增量同步的表Datacapture属性需要为Y。 ● 浮点数类型作为主键时，增量同步可能会导致数据不一致。 ● 同步对象列名不能为CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID等GaussDB(DWS)禁止的字段，否则会导致任务失败。
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“.”、“ ”、“”、“?”、“!”、“”和“!”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 选择全量+增量同步时，需要确保开启源数据库的归档日志 ● 目前仅支持同步如下字符集：GBK、UTF8。
目标数据库要求	<ul style="list-style-type: none"> ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。 ● 不支持ARM虚拟机环境。 ● 如果无主键表中包含LOB、LONG等大字段，增量同步阶段存在大字段数据不一致的可能性。 ● 如果表结构存在DB2SECURITYLABEL字段类型，读取全量阶段数据可能会受DB2 LBAC（基于标签的访问控制）影响，导致源库和目标库数据不一致。 ● 如果源库和目标库字符集不一致，如源库是GBK，目标库是UTF8，由于GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 任务启动后，对于分区表的detach操作可能导致数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，源库不允许有未提交的DDL事务。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束，如有需要，用户必须手动在目标库创建对应表的默认值约束。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步阶段，支持任务再编辑追加同步对象。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-241 同步任务信息

⚠️ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述
0/256

表 3-254 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-242 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置错误导致数据迁移失败，需要重新创建迁移任务。

* 数据流动方向 入云 出云 ?

DRS要求源数据库和目标数据库中至少有一方为本云数据库实例。入云指目标数据库为本云数据库实例的场景，本云数据库实例之间的同步，请选择入云。

* 源数据库引擎 MySQL Oracle DB2 for LUW DDM MongoDB PostgreSQL Microsoft SQL Server TDB

* 目标数据库引擎 GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版

MySQL、Oracle等数据库多与DWS不同，存在DDL同步失败的可能性，建议方案实施前联系DWS专家详细DDL支持度。

* 网络类型 ?

DRS将会自动为DRS实例创建弹性公网IP，请在创建后自行调整弹性公网IP。

* 目标数据库实例 查看数据库实例 查看不可用实例

同步实例所在子网 ? 查看子网

* 同步模式 ?

该模式为数据持续性实时同步，通过全量过程内成员与源数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据保持一致。

表 3-255 同步实例信息

参数	描述
数据流动方向	选择“入云”，即目标端数据库为本云数据库。
源数据库引擎	选择“DB2 for LUW”。
目标数据库引擎	选择“GaussDB(DWS)”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。可根据业务场景选择“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的GaussDB(DWS)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 3-243 可用区



表 3-256 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-244 企业项目和标签



表 3-257 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 3-245 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

数据库名

SSL安全连接

同步驱动 已同步驱动: jjar

表 3-258 源库信息

参数	描述
IP地址或域名	源数据库DB2 for LUW的IP地址或域名。
端口	源数据库DB2 for LUW服务端口，可输入范围为1~65535间的整数。
数据库用户名	源DB2 for LUW的数据库账号。
数据库密码	源数据库账号所对应的密码。
数据库名	源DB2 for LUW中同步对象所属数据库的名称。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。目前仅DB2 for LUW 10.5及以上的版本支持SSL安全连接，DB2 for LUW 9.7和10.1版本暂不支持。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不使用SSL安全连接，请自行承担数据安全风险。 DB2 for LUW SSL安全连接需要上传三个证书文件：jks证书、kdb证书和sth证书。 kdb证书和sth证书测试连接时DRS不会进行校验，需要用户保证文件正确性，避免后续任务失败。
同步驱动	用户需要手动添加并上传源数据库DB2 for LUW对应的JDBC驱动包后，再进行测试连接。

说明

源数据库的数据库IP地址、用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 3-246 目标库信息

表 3-259 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(DWS)实例，不可进行修改。
数据库用户名	目标数据库的账号。
数据库密码	目标数据库账号所对应的密码。


说明

目标数据库的数据库的用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 3-247 同步模式

表 3-260 同步对象

参数	描述
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-248 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-249 任务启动设置



表 3-261 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.25 将 TiDB 同步到 GaussDB(for MySQL)

支持的源和目标数据库

表 3-262 支持的数据库

源数据库	目标数据库
TiDB (4.0.0及以上版本, 不包括dev版本)	GaussDB(for MySQL)主备实例

支持的同步对象范围

在使用DRS进行同步时, 不同类型的同步任务, 支持的同步对象范围不同, 详细情况可参考表3-263。DRS会根据用户选择, 在“预检查”阶段对同步对象进行自动检查。

表 3-263 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 支持的字段类型: BIGINT、BINARY、BIT、BLOB、BOOLEAN、CHAR、DATE、DATETIME、DECIMAL、DOUBLE、ENUM、FLOAT、INT、JSON、LONGBLOB、LONGTEXT、MEDIUMBLOB、MEDIUMINT、SET、SMALLINT、TEXT、TIME、TIMESTAMP、TINYBLOB、TINYINT、TINYTEXT、VARBINARY、VARCHAR、YEAR。 ● 全量同步的范围: <ul style="list-style-type: none"> - 支持: 选中表的表结构、数据、约束和索引。 - 同步对象的库名、表名不能包含: '<>\'以及非ASCII字符。 - 库名或映射后的名称不允许以ib_logfile开头, 也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 - 任务做表名映射时, 不支持该表的外键约束同步。 - 使用非字母、非数字、非下划线等字符作为库表名称, 或库表映射后名称包含中划线、井号时, 名称的长度请不要超过42个字符。 ● 增量同步的范围: <ul style="list-style-type: none"> - 支持部分DML (Data Manipulation Language): 包括INSERT、UPDATE、DELETE。 - 支持部分DDL (Data Definition Language): 包括CREATE TABLE、DROP TABLE、ALTER TABLE、RENAME TABLE、TRUNCATE TABLE。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-264](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-264 数据库账号权限

类型名称	全量+增量同步
源数据库连接账号	SELECT, CONFIG权限。
目标数据库连接账号	SELECT, CREATE, DROP, INSERT, DELETE, UPDATE, ALTER, REFERENCES, INDEX权限。

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-265 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源库需要部署Pump和Drainer组件，具体操作请参考TiDB Binlog集群部署。 - 修改Drainer组件的配置文件，设置输出类型为Kafka，具体操作请参考配置 Kafka Drainer。 注意 <ul style="list-style-type: none"> ● TiDB Binlog 与 TiDB v5.0 版本开始引入的以下特性不兼容： <ul style="list-style-type: none"> TiDB 聚簇索引特性： 开启 TiDB Binlog 后 TiDB 不允许创建非单个整数列作为主键的聚簇索引；已创建的聚簇索引表的数据插入、删除和更新动作不会通过 TiDB Binlog 同步到下游。如需同步聚簇索引表，请升级至 v5.1 版本。 TiDB 系统变量 tidb_enable_async_commit：启用 TiDB Binlog 后，开启该选项无法获得性能提升。 TiDB 系统变量 tidb_enable_1pc：启用 TiDB Binlog 后，开启该选项无法获得性能提升。 ● TiDB Binlog 与 TiDB v4.0.7 版本开始引入的以下特性不兼容： <ul style="list-style-type: none"> TiDB 系统变量 tidb_enable_amend_pessimistic_txn：两个功能存在兼容性问题，一起使用会造成 TiDB Binlog 复制数据不一致的正确性问题。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库待同步对象的库名、表名不能包含：'<>/\以及非ASCII字符。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标数据库参数lower_case_table_names等于1时，不允许同步包含大写字母的库或者表。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例必须是GaussDB(for MySQL)主备实例。 - 选择全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致或包含源端的列。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 - 目标库已启用的触发器不可以与同步表关联。 ● 其他使用须知： <ul style="list-style-type: none"> - 增量同步支持任务再编辑追加同步对象。 - 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。

类型名称	使用和操作约定
全量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。
同步对比须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。

操作步骤

本小节以TiDB->GaussDB(for MySQL)的入云同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-250 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-266 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-251 同步实例信息



表 3-267 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“TiDB”。
目标数据库引擎	选择“GaussDB(for MySQL)”。
网络类型	此处以“公网网络”为示例，可选“公网网络”和“VPN、专线网络”。 <ul style="list-style-type: none"> - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为GaussDB(for MySQL)实例。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。

参数	描述
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 可用区

图 3-252 可用区



表 3-268 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-253 企业项目和标签



表 3-269 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。 如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息、Kafka信息和目标库信息，单击“源库、Kafka和目标库”处的“测试连接”，分别测试并确定与源库、Kafka和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 3-254 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功

表 3-270 源库信息

参数	描述
IP地址或域名	源数据库TiDB的IP地址或域名。
端口	源数据库TiDB服务端口，可输入范围为1~65535间的整数，默认为4000。
数据库用户名	源数据库TiDB的用户名。

参数	描述
数据库密码	源数据库TiDB的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

图 3-255 Kafka 信息

Kafka 信息

IP地址 ?

Kafka地址只支持IP方式，不支持域名或主机名，且所有Kafka节点的listener配置也必须为IP方式

表 3-271 Kafka 信息

参数	描述
IP地址或域名	Kafka的IP地址或域名。DRS从Kafka中获取TiDB产生的Binlog数据，然后将对应的增量更新实时同步至到目标库中

图 3-256 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码 🔍

表 3-272 目标库信息

参数	描述
数据库实例名称	默认为创建任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明


源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。


步骤4 在“设置同步”页面，选择同步策略、Topic和同步对象，单击“下一步”。

图 3-257 同步模式



表 3-273 同步对象

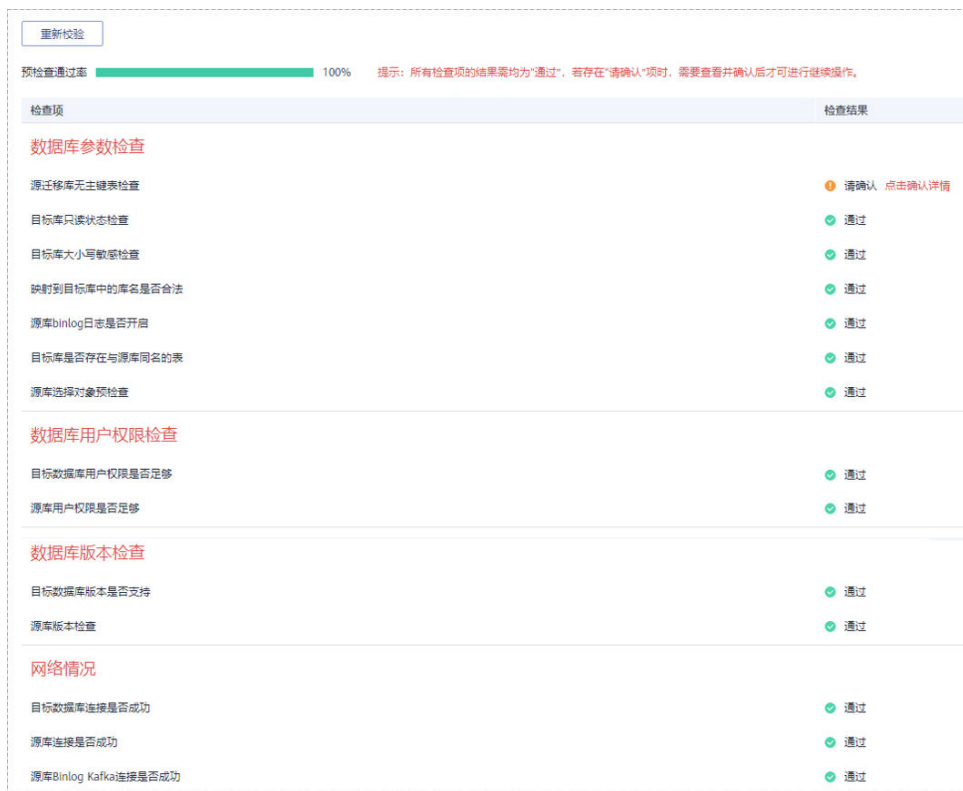
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-258 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
Topic	<p>请选择Kafka中存放TiDB binlog的Topic，否则将同步失败。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。</p> <p>使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-259 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 3-260 任务启动设置



表 3-274 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.26 将 Microsoft SQL Server 同步到 GaussDB(DWS)

支持的源和目标数据库

表 3-275 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建Microsoft SQL Server (企业版2012、2014、2016、2017、2019版本, 标准版2016 SP2及以上版本、2017、2019版本)ECS自建Microsoft SQL Server数据库 (企业版2012、2014、2016、2017、2019版本, 标准版2016 SP2及以上版本、2017、2019版本)其他云上Microsoft SQL Server数据库 (企业版2012、2014、2016、2017、2019版本, 标准版2016 SP2及以上版本、2017、2019版本)RDS for SQL Server (企业版2012、2014、2016、2017、2019版本, 标准版2016 SP2及以上版本、2017、2019版本)	<ul style="list-style-type: none">GaussDB(DWS)集群

支持的同步对象范围

在使用DRS进行同步时, 不同类型的同步任务, 支持的同步对象范围不同, 详细情况可参考[表3-276](#)。DRS会根据用户选择, 在“预检查”阶段对同步对象进行自动检查。

表 3-276 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none">● 支持的字段类型： TINYINT、SMALLINT、INT、BIGINT、DECIMAL、NUMERIC、FLOAT、REAL、SMALLMONEY、MONEY、BIT、DATE、DATETIME、DATETIME2、DATETIMEOFFSET、TIME、TIMESTAMP、XML、CHAR、VARCHAR、NCHAR、NVARCHAR、BINARY、VARBINARY、IMAGE、HIERARCHYID、NTEXT、TEXT、UNIQUEIDENTIFIER。● 不支持的字段类型： SQL_VARIANT、GEOMETRY、GEOGRAPHY。● 全量同步的范围：<ul style="list-style-type: none">- 支持选中表的表结构、数据和索引。- 同步对象的库名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。- 同步对象列名不能为CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID等GaussDB(DWS)禁止的字段，否则会导致任务失败。- 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。- 不支持同步源数据库中开启TDE（Transparent Data Encryption）加密的数据库。若不需要同步TDE加密的库，请在对象选择中去掉；若需要同步TDE加密的库，请先将TDE关闭。- 不支持列加密。- 不支持同步自增属性列。● 增量同步的范围：<ul style="list-style-type: none">- 支持DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。- 不支持DDL（Data Definition Language）。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-277](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-277 数据库账号权限

类型名称	全量+增量同步
源数据库连接账号	需要具备sysadmin权限，或者view server state权限以及待同步数据库的db_datareader或db_owner权限。
目标数据库连接账号	<ul style="list-style-type: none"> 目标库不存在待同步数据库时，需要CREATEDB权限。 目标库存在数据库但不存在SCHEMA时，需要数据库的CONNECT和CREATE权限。 目标库存在数据库和SCHEMA但不存在表时，需要数据库的CONNECT权限，SCHEMA的USAGE和CREATE权限。 目标库数据库、SCHEMA和表都存在时，需要数据库的CONNECT权限，SCHEMA的USAGE权限，表的INSERT、UPDATE、DELETE、SELECT、ALTER、REFERENCES和INDEX权限。

使用建议

注意

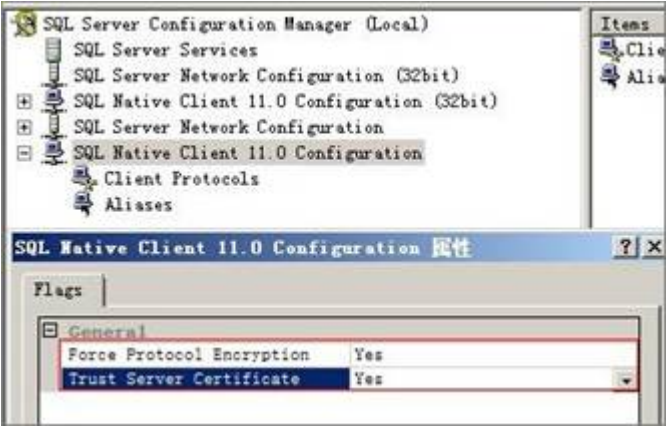
- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-278 使用须知

类型名称	使用和操作约定
<p>任务启动前须知</p>	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库模式必须设置为FULL模式。 - 源数据库必须开启SQL Server Agent代理服务。 - 源数据库若开启客户端配置中的“强制协议加密（ Force Protocol Encrypton ）”，必须同时开启“信任服务器证书（ trust server certificate ）”，如下图3-261所示： <p style="text-align: center;">图 3-261 查看客户端属性</p>  <ul style="list-style-type: none"> ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不允许存在被禁用的表的聚簇索引，否则会导致同步失败。 - 源库中不允许存在名称为cdc的用户名或schema。 - 源库待同步对象的库名、schema名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 源库待同步对象的表中列名不能包含如下字符：[]?。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例必须有足够的磁盘空间。 - 除了SQL Server系统数据库之外，目标数据库不能包含与源数据库同名的数据库。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 不支持源数据库主备切换，源数据库主备切换会导致同步任务失败。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 不支持源数据库Microsoft SQL Server为TLS 1.0、TLS 1.1协议的同步，如果需要同步，建议源库升级到TLS 1.2及以上版本。 - SQL Server为源的增量同步基于SQL Server提供的CDC能力，如果SQL Server源库的CDC功能出现异常（如被关闭、事务日志满等情况），则增量同步就会受到影响。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持DDL操作，源数据库进行的DDL操作不会同步到目标数据库。 ● 不支持大数据类型IMAGE、TEXT、NTEXT的删除操作。 ● 支持任务再编辑追加同步对象。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

操作步骤

本小节以Microsoft SQL Server->GaussDB(DWS)的入云同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-262 同步任务信息

⚠️ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述
0/256

表 3-279 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-263 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向 入云 出云 自建-自建 ?

DRS要求源数据库或目标数据库中至少有一方为本云数据库实例，入云指目标数据库为本云数据库实例的场景，本云数据库实例之间的同步，请忽略入云。

* 源数据库引擎 MySQL Oracle DB2 for LUW DDM MongoDB PostgreSQL Microsoft SQL Server TiDB

* 目标数据库引擎 GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版

* 网络类型 ?
 DRS将会自动为DRS实例绑定弹性公网IP，该任务完成后将自动删除该弹性公网IP。

* 目标数据库实例 查看数据库实例 查看不可选实例

同步实例所在子网 ? 查看子网

* 同步模式 全量 + 增量 ?

该模式为数据持续性实时同步，通过全量过程完成目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

表 3-280 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Microsoft SQL Server”。
目标数据库引擎	选择“GaussDB(DWS)”。

参数	描述
网络类型	<p>此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为GaussDB(DWS)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 可用区

图 3-264 可用区



表 3-281 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-265 企业项目和标签



表 3-282 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 3-266 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

数据库类型 ECS自建库 RDS实例

数据库实例名称 [查看数据库实例](#) [查看不可选实例](#)

数据库用户名

数据库密码

表 3-283 源库信息

参数	描述
数据库类型	选择“RDS实例”。
数据库实例名称	选择源数据库Microsoft SQL Server实例。
数据库用户名	源数据库Microsoft SQL Server的用户名。
数据库密码	源数据库Microsoft SQL Server的用户名所对应的密码。

图 3-267 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

表 3-284 目标库信息

参数	描述
数据库实例名称	默认为创建任务时选择的GaussDB(DWS)实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-268 同步模式

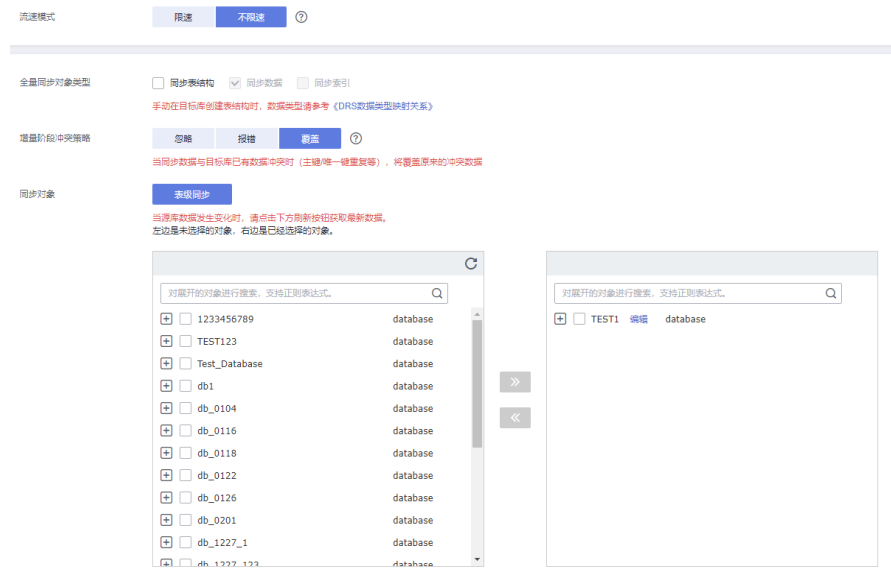



表 3-285 同步对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-269 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

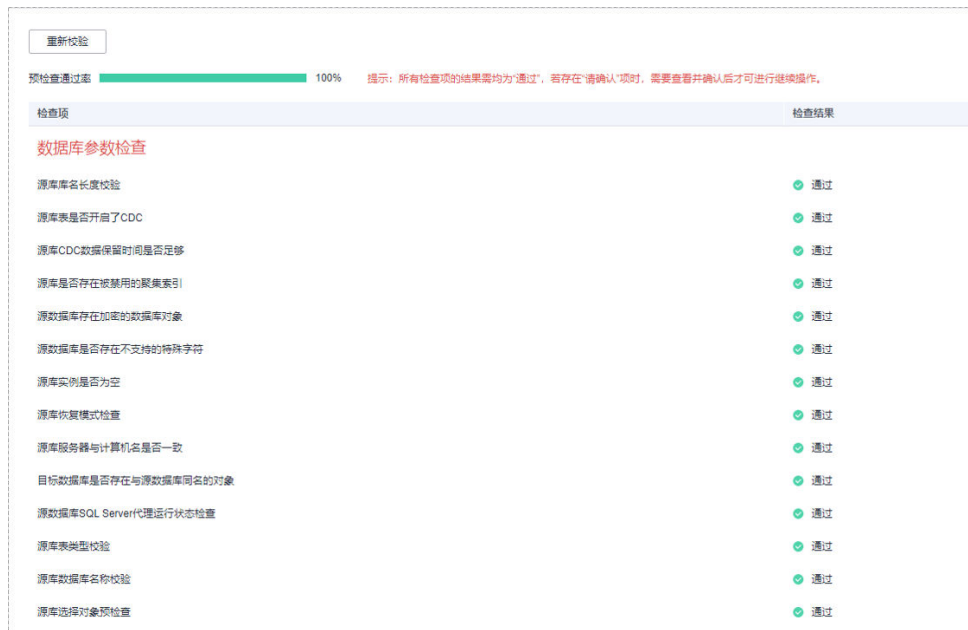
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-270 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-271 任务启动设置



表 3-286 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">当时延阈值设置为0时，不会发送通知给收件人。首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.27 将 Microsoft SQL Server 同步到 GaussDB 主备版

支持的源和目标数据库

表 3-287 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建Microsoft SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）ECS自建Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）其他云上Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）RDS for SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）	GaussDB主备版

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-288](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-288 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 支持的字段类型： TINYINT、SMALLINT、INT、BIGINT、DECIMAL、NUMERIC、FLOAT、REAL、SMALLMONEY、MONEY、BIT、DATE、SMALLDATETIME、DATETIME、DATETIME2、DATETIMEOFFSET、TIME、TIMESTAMP、XML、CHAR、VARCHAR、NCHAR、NVARCHAR、BINARY、VARBINARY、IMAGE、HIERARCHYID、NTEXT、TEXT、UNIQUEIDENTIFIER。 ● 不支持的字段类型：SQL_VARIANT、GEOMETRY、GEOGRAPHY。 ● 全量同步的范围： <ul style="list-style-type: none"> - 支持选中表的表结构、数据和索引。 - 同步对象的库名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 - 不支持同步源数据库中开启TDE（Transparent Data Encryption）加密的数据库。若不需要同步TDE加密的库，请在对象选择中去掉；若需要同步TDE加密的库，请先将TDE关闭。 - 不支持列加密。 - 不支持同步自增属性列。 ● 增量同步的范围： <ul style="list-style-type: none"> - 支持DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 - 不支持DDL（Data Definition Language）。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-289](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-289 数据库账号权限

类型名称	全量同步	增量、全量+增量同步
源数据库连接账号	需要具备sysadmin权限，或者待同步数据库的db_datareader或db_owner权限。	需要具备sysadmin权限，或者view server state权限以及待同步数据库的db_datareader或db_owner权限。
目标数据库连接账号	<ul style="list-style-type: none"> 库级权限：需要使用root或其他有sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 SCHEMA级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 表级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下所有表的SELECT，UPDATE，INSERT和DELETE权限。 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。 	

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

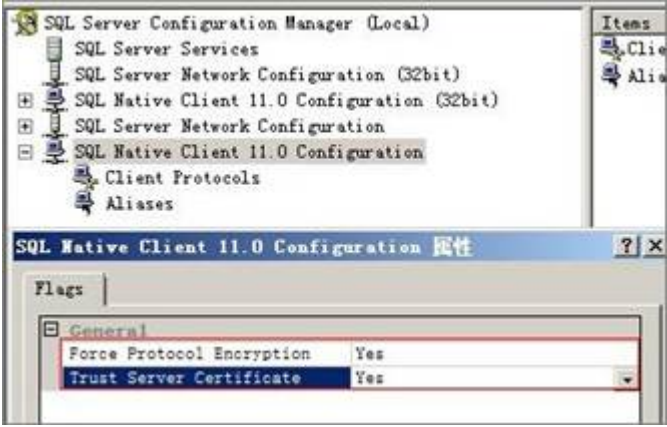
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-290 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库模式必须设置为FULL模式。 - 源数据库必须开启SQL Server Agent代理服务。 - 源数据库若开启客户端配置中的“强制协议加密（ Force Protocol Encrypton ）”，必须同时开启“信任服务器证书（ trust server certificate ）”，如下图3-272所示： <p style="text-align: center;">图 3-272 查看客户端属性</p>  <ul style="list-style-type: none"> ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不允许存在被禁用的表的聚簇索引，否则会导致同步失败。 - 源库中不允许存在名称为cdc的用户名或schema。 - 源库待同步对象的库名、schema名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 源库待同步对象的表中列名不能包含如下字符：[]?。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例必须有足够的磁盘空间。 - 除了SQL Server系统数据库之外，目标数据库不能包含与源数据库同名的数据库。 ● 其他使用须知： <ul style="list-style-type: none"> - 不支持源数据库主备切换，源数据库主备切换会导致同步任务失败。 - 不支持源数据库Microsoft SQL Server为TLS 1.0、TLS 1.1协议的同步，如果需要同步，建议源库升级到TLS 1.2及以上版本。 - SQL Server为源的增量同步基于SQL Server提供的CDC能力，如果SQL Server源库的CDC功能出现异常（如被关闭、事务日志满等情况），则增量同步就会受到影响。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 目标数据库为Oracle兼容模式时，会将写入的空字符串作为NULL处理，源数据库存在空字符串且具有非空约束时，会导致写入失败。
全量同步过程须知	<ul style="list-style-type: none"> • 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 • 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 • 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> • 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 • 请勿在目标库做写入操作，否则可能导致数据不一致。 • 不支持DDL操作，源数据库进行的DDL操作不会同步到目标数据库。 • 不支持大数据类型image、text、ntext的删除操作。 • 不支持delete和update目标端主键类型为blob、text、clob、nclob、bytea的表。 • 单增量任务再编辑增加同步对象时，在子任务变更结束前，不建议对新增同步的表进行DML操作，否则可能导致部分数据无法同步到目标数据库。
同步对比须知	<ul style="list-style-type: none"> • 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 • 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 • 不支持在全量同步过程中做对比。 • 不支持数据对比过程中做限速。

操作步骤

本小节以Microsoft SQL Server->GaussDB主备版的入云同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-273 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-291 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-274 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎选择，以免为后期配置带来麻烦，需要重新创建迁移任务。

* 数据流动方向: [In-Cloud] [Out-Cloud] [Bidirectional] ?

DRS要求源数据库和目标数据库中至少有一方为本地数据库实例，入云目标数据库为本地数据库实例的场景，本地数据库实例之间的同步，请勾选入云。

* 源数据库引擎: MySQL Oracle DB2 for LUW DDM MongoDB PostgreSQL Microsoft SQL Server TiDB

* 目标数据库引擎: GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版

* 网络类型: [Public Network] ?

DRS将会创建为DRS实例所属弹性公网IP，请任意选择并修改弹性公网IP。

* 目标数据库实例: [Select Instance] [Create New Instance] [Cannot Create Instance]

* 同步实例所在子网: [Select Subnet] ?

* 同步模式: [Full + Incremental] [Full] [Incremental] ?

该模式为典型库间实时同步，通过全量过程完成目标数据库的初始化后，增量同步时仅处理过解析日志等技术，将源端和目标数据库的增量数据保持一致。

表 3-292 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Microsoft SQL Server”。
目标数据库引擎	选择“GaussDB主备版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为GaussDB主备版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 3-275 可用区



表 3-293 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-276 企业项目和标签



表 3-294 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 3-277 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

数据库类型 ECS自建库 RDS实例

数据库实例名称 [查看数据库实例](#) [查看不可选实例](#)

数据库用户名

数据库密码

表 3-295 源库信息

参数	描述
数据库类型	选择“RDS实例”。
数据库实例名称	选择源数据库Microsoft SQL Server实例。
数据库用户名	源数据库Microsoft SQL Server的用户名。
数据库密码	源数据库Microsoft SQL Server的用户名所对应的密码。

图 3-278 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

表 3-296 目标库信息

参数	描述
数据库实例名称	默认为创建任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-279 同步模式

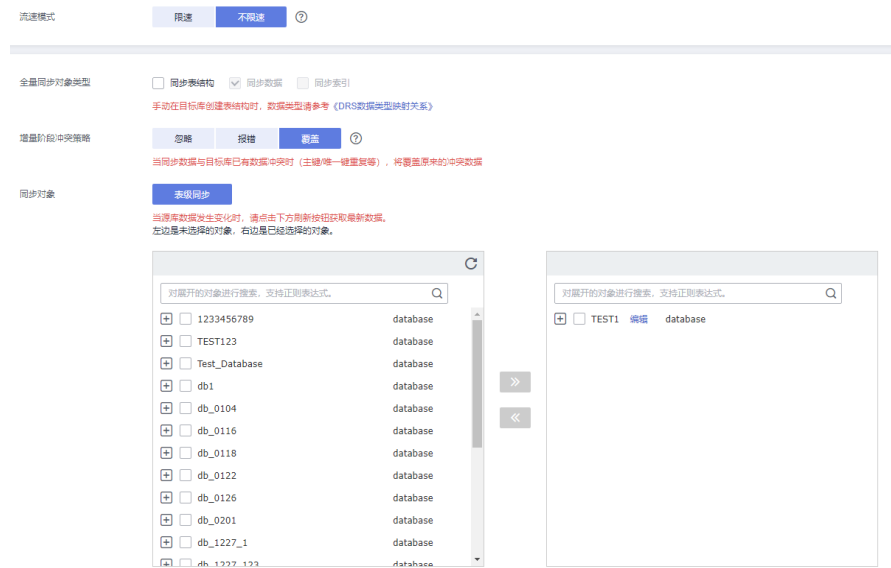



表 3-297 同步对象

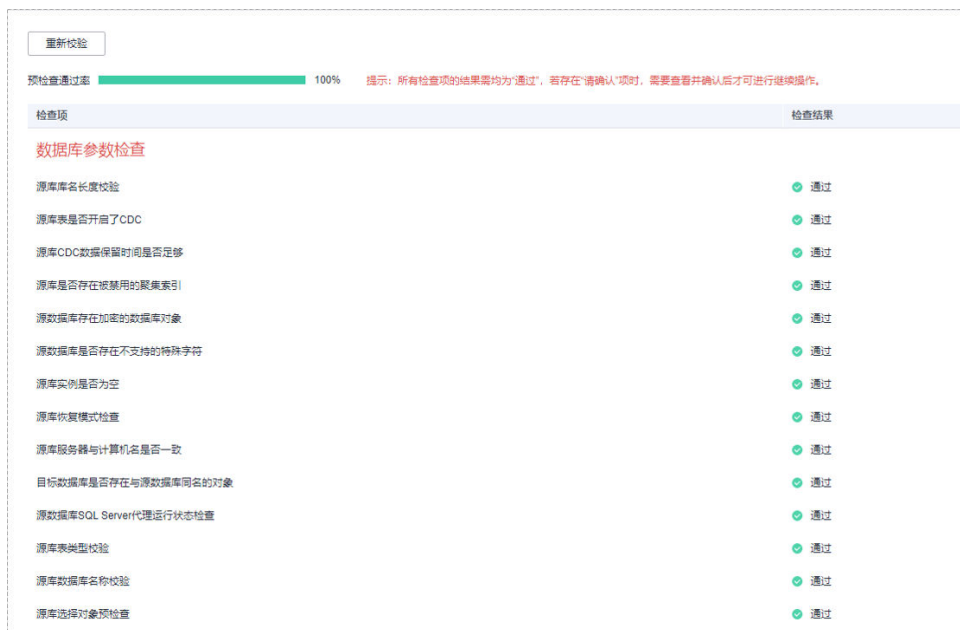
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-280 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保持选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-281 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-282 任务启动设置



表 3-298 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">当时延阈值设置为0时，不会发送通知给收件人。首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.28 将 Microsoft SQL Server 同步到 GaussDB 分布式版

支持的源和目标数据库

表 3-299 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建Microsoft SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）ECS自建Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）其他云上Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）RDS for SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）	GaussDB分布式版

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-300](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-300 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 支持的字段类型： TINYINT、SMALLINT、INT、BIGINT、DECIMAL、NUMERIC、FLOAT、REAL、SMALLMONEY、MONEY、BIT、DATE、SMALLDATETIME、DATETIME、DATETIME2、DATETIMEOFFSET、TIME、TIMESTAMP、XML、CHAR、VARCHAR、NCHAR、NVARCHAR、BINARY、VARBINARY、IMAGE、HIERARCHYID、NTEXT、TEXT、UNIQUEIDENTIFIER。 ● 不支持的字段类型：SQL_VARIANT、GEOMETRY、GEOGRAPHY。 ● 全量同步的范围： <ul style="list-style-type: none"> - 支持选中表的表结构、数据和索引。 - 同步对象的库名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 - 不支持同步源数据库中开启TDE（Transparent Data Encryption）加密的数据库。若不需要同步TDE加密的库，请在对象选择中去掉；若需要同步TDE加密的库，请先将TDE关闭。 - 不支持列加密。 - 不支持同步自增属性列。 ● 增量同步的范围： <ul style="list-style-type: none"> - 支持DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 - 不支持DDL（Data Definition Language）。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-301](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-301 数据库账号权限

类型名称	全量同步	增量、全量+增量同步
源数据库连接账号	需要具备sysadmin权限，或者待同步数据库的db_datareader或db_owner权限。	需要具备sysadmin权限，或者view server state权限以及待同步数据库的db_datareader或db_owner权限。
目标数据库连接账号	<ul style="list-style-type: none"> • 库级权限：需要使用root或其他有sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 • SCHEMA级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 • 表级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下所有表的SELECT，UPDATE，INSERT和DELETE权限。 • 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。 	

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

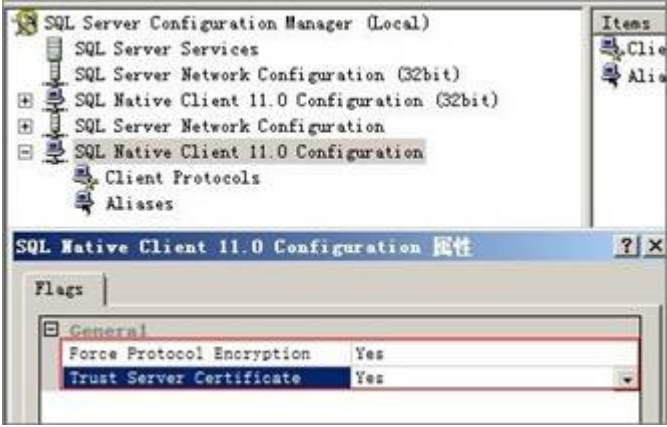
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-302 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库模式必须设置为FULL模式。 - 源数据库必须开启SQL Server Agent代理服务。 - 源数据库若开启客户端配置中的“强制协议加密（ Force Protocol Encrypton ）”，必须同时开启“信任服务器证书（ trust server certificate ）”，如下图3-283所示： <p style="text-align: center;">图 3-283 查看客户端属性</p>  <ul style="list-style-type: none"> ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不允许存在被禁用的表的聚簇索引，否则会导致同步失败。 - 源库中不允许存在名称为cdc的用户名或schema。 - 源库待同步对象的库名、schema名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 源库待同步对象的表中列名不能包含如下字符：[]?。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例必须有足够的磁盘空间。 - 除了SQL Server系统数据库之外，目标数据库不能包含与源数据库同名的数据库。 ● 其他使用须知： <ul style="list-style-type: none"> - 不支持源数据库主备切换，源数据库主备切换会导致同步任务失败。 - 不支持源数据库Microsoft SQL Server为TLS 1.0、TLS 1.1协议的同步，如果需要同步，建议源库升级到TLS 1.2及以上版本。 - SQL Server为源的增量同步基于SQL Server提供的CDC能力，如果SQL Server源库的CDC功能出现异常（如被关闭、事务日志满等情况），则增量同步就会受到影响。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 目标数据库为Oracle兼容模式时，会将写入的空字符串作为NULL处理，源数据库存在空字符串且具有非空约束时，会导致写入失败。
全量同步过程须知	<ul style="list-style-type: none"> • 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 • 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 • 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> • 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 • 请勿在目标库做写入操作，否则可能导致数据不一致。 • 不支持DDL操作，源数据库进行的DDL操作不会同步到目标数据库。 • 不支持大数据类型image、text、ntext的删除操作。 • 不支持delete和update目标端主键类型为blob、text、clob、nclob、bytea的表。 • 单增量任务再编辑增加同步对象时，在子任务变更结束前，不建议对新增同步的表进行DML操作，否则可能导致部分数据无法同步到目标数据库。
同步对比须知	<ul style="list-style-type: none"> • 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 • 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 • 不支持在全量同步过程中做对比。 • 不支持数据对比过程中做限速。

操作步骤

本小节以Microsoft SQL Server->GaussDB分布式版的入云同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-284 同步任务信息

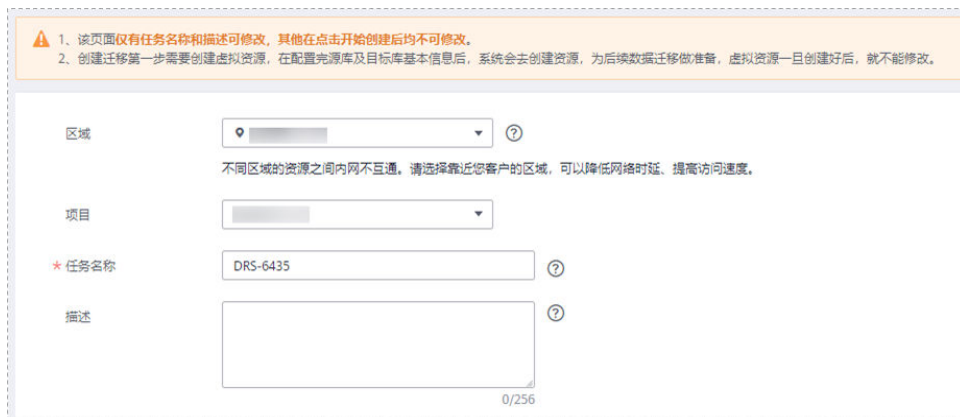


表 3-303 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-285 同步实例信息



表 3-304 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Microsoft SQL Server”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 3-286 可用区



表 3-305 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-287 企业项目和标签



表 3-306 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 3-288 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

数据库类型 ECS自建库 RDS实例

数据库实例名称 [查看数据库实例](#) [查看不可选实例](#)

数据库用户名

数据库密码

表 3-307 源库信息

参数	描述
数据库类型	选择“RDS实例”。
数据库实例名称	选择源数据库Microsoft SQL Server实例。
数据库用户名	源数据库Microsoft SQL Server的用户名。
数据库密码	源数据库Microsoft SQL Server的用户名所对应的密码。

图 3-289 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

表 3-308 目标库信息

参数	描述
数据库实例名称	默认为创建任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-290 同步模式

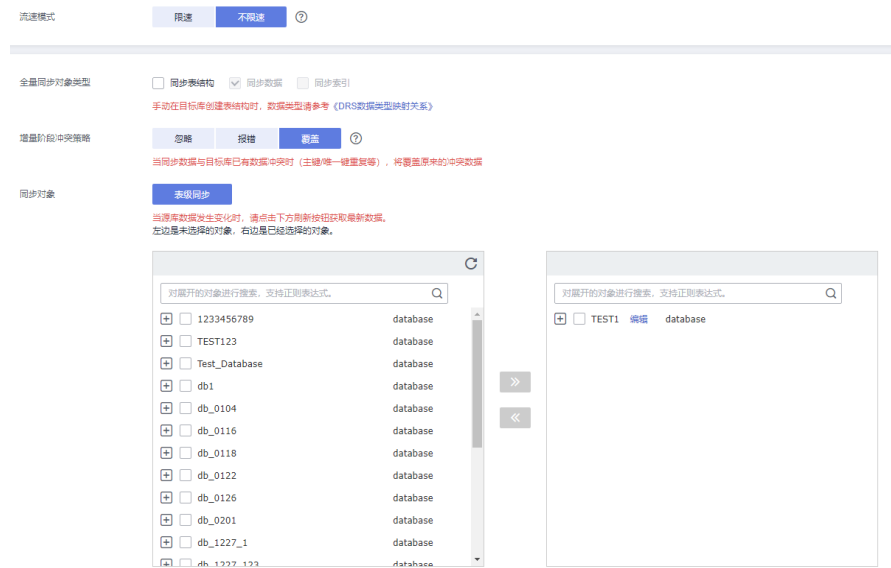



表 3-309 同步对象

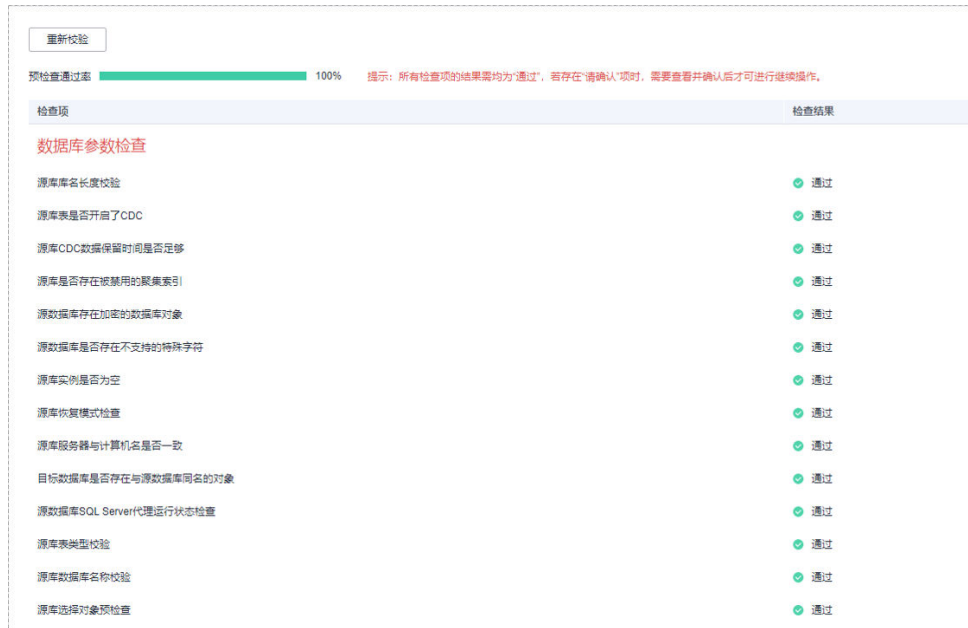
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-291 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-292 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 3-293 任务启动设置



表 3-310 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.29 将 Microsoft SQL Server 同步到 Microsoft SQL Server

支持的源和目标数据库

表 3-311 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建Microsoft SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）ECS自建Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）其他云上Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）RDS for SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本）	<ul style="list-style-type: none">RDS for SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本） <p>说明 仅支持目标库大版本等于或高于源库大版本。</p>

说明

目前仅支持白名单用户使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-312](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-312 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 支持的字段类型： TINYINT、SMALLINT、INT、BIGINT、DECIMAL、NUMERIC、FLOAT、REAL、SMALLMONEY、MONEY、BIT、DATE、SMALLDATETIME、DATETIME、DATETIME2、DATETIMEOFFSET、TIME、TIMESTAMP、XML、CHAR、VARCHAR、NCHAR、NVARCHAR、BINARY、VARBINARY、IMAGE、HIERARCHYID、NTEXT、TEXT、UNIQUEIDENTIFIER。 ● 不支持的字段类型：SQL_VARIANT、GEOMETRY、GEOGRAPHY。 ● 全量同步的范围： <ul style="list-style-type: none"> - 支持选中表的表结构、数据和索引。 - 同步对象的库名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 - 不支持同步源数据库中开启TDE（Transparent Data Encryption）加密的数据库。若不需要同步TDE加密的库，请在对象选择中去掉；若需要同步TDE加密的库，请先将TDE关闭。 - 不支持列加密。 - 不支持同步自增属性列。 ● 增量同步的范围： <ul style="list-style-type: none"> - 支持DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 - 不支持DDL（Data Definition Language）。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-313](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-313 数据库账号权限

类型名称	全量+增量同步
源数据库连接账号	最少需要具备sysadmin权限，或者view server state权限以及待同步数据库的db_datareader或db_owner权限。
目标数据库连接账号	如果库不存在，需要对库级别开启create any database权限； 如果库存在需要库级别开启connect， create table， alter any schema 和select权限。

使用建议

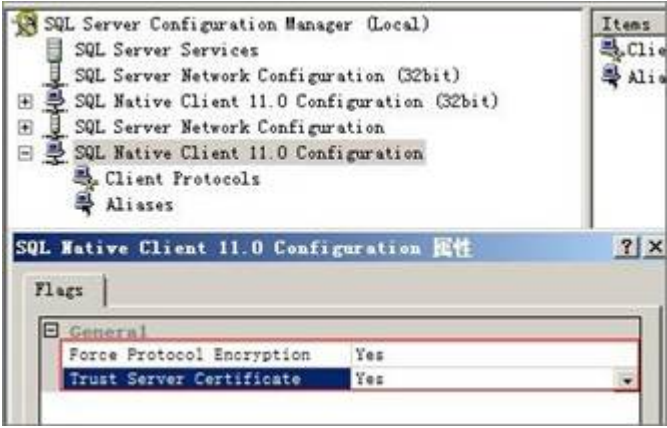
⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-314 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库模式必须设置为FULL模式。 - 源数据库必须开启SQL Server Agent代理服务。 - 源数据库若开启客户端配置中的“强制协议加密（ Force Protocol Encrypton ）”，必须同时开启“信任服务器证书（ trust server certificate ）”，如下图3-294所示： <p style="text-align: center;">图 3-294 查看客户端属性</p>  <ul style="list-style-type: none"> ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不允许存在被禁用的表的聚簇索引，否则会导致同步失败。 - 源库待同步对象的库名、schema名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 源库待同步对象的表中列名不能包含如下字符：[]?。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例必须有足够的磁盘空间。 ● 其他使用须知： <ul style="list-style-type: none"> - 不支持源数据库主备切换，源数据库主备切换会导致同步任务失败。 - 不支持源数据库Microsoft SQL Server为TLS 1.0、TLS 1.1协议的同步，如果需要同步，建议源库升级到TLS 1.2及以上版本。 - SQL Server为源的增量同步基于SQL Server提供的CDC能力，如果SQL Server源库的CDC功能出现异常（如被关闭、事务日志满等情况），则增量同步就会受到影响。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。 不支持DDL操作，源数据库进行的DDL操作不会同步到目标数据库。 不支持大数据类型IMAGE、TEXT、NTEXT的删除操作。
同步对比须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。

操作步骤

本小节以Microsoft SQL Server->Microsoft SQL Server的入云同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-295 同步任务信息

表 3-315 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。

参数	描述
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

• 同步实例信息

图 3-296 同步实例信息



表 3-316 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“Microsoft SQL Server”。
目标数据库引擎	选择“Microsoft SQL Server”。
网络类型	此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。 <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	目标数据库为RDS for SQL Server实例。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。

参数	描述
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 可用区

图 3-297 可用区



表 3-317 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-298 企业项目和标签



表 3-318 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。 如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 3-299 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

表 3-319 源库信息

参数	描述
数据库类型	选择“RDS实例”。
数据库实例名称	选择源数据库Microsoft SQL Server实例。
数据库用户名	源数据库Microsoft SQL Server的用户名。
数据库密码	源数据库Microsoft SQL Server的用户名所对应的密码。

图 3-300 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码 👁

表 3-320 目标库信息

参数	描述
数据库实例名称	默认为创建任务时选择的RDS for SQL Server实例，不可进行修改。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-301 同步模式

流速模式 ⓘ

全量同步对象类型 同步表结构 同步数据 同步索引

手动在目标库创建表结构时，数据类型请参考《DRS数据类型映射关系》

增量阶段冲突策略 ⓘ

当同步数据与目标库已有数据冲突时（主键/唯一键重复等），将覆盖原来的冲突数据

同步对象

当源库数据发生变化时，请点击下方刷新按钮获取最新数据。
左边是未选择的对象，右边是已经选择的对象。

对展开的对象进行搜索，支持正则表达式。


- 1233456789 database
- TEST123 database
- Test_Database database
- db1 database
- db_0104 database
- db_0116 database
- db_0118 database
- db_0122 database
- db_0126 database
- db_0201 database
- db_1227_1 database
- db_1227_123 database

对展开的对象进行搜索，支持正则表达式。

- TEST1 编辑 database

表 3-321 同步对象

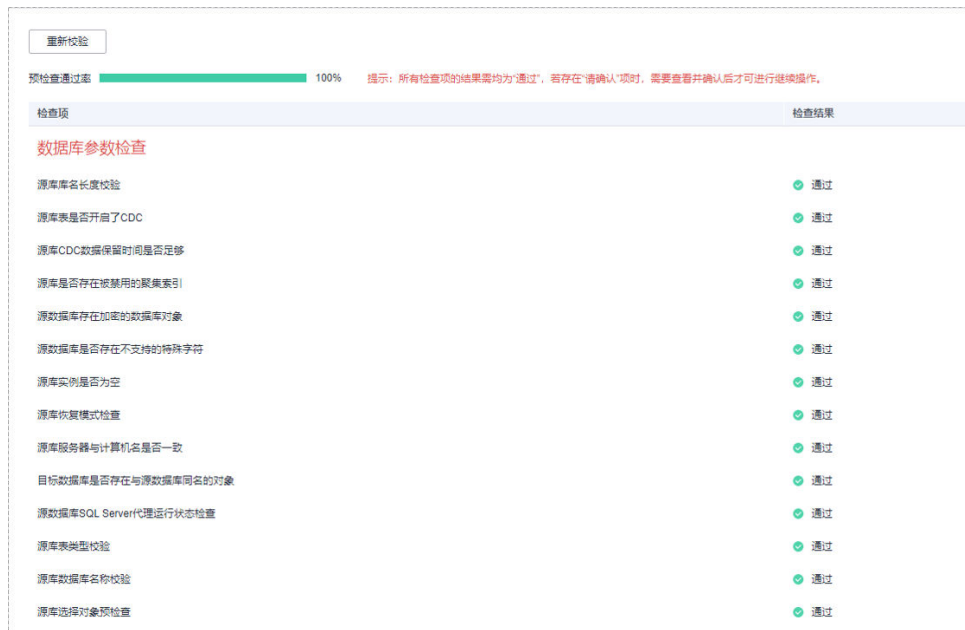
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-302 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步普通索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 同步对象选择的表个数不能超过1000。对于超过1000个表的情况，建议分批次同步（待上一个同步任务结束后，再重新创建任务）。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保持选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-303 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 3-304 任务启动设置



表 3-322 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.30 将 MongoDB 同步到 DDS

支持的源和目标数据库

表 3-323 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建Mongo数据库（3.2、3.4、3.6、4.0、4.2、4.4版本） ECS自建Mongo数据库（3.2、3.4、3.6、4.0、4.2、4.4版本） 其他云上Mongo数据库（3.2、3.4、3.6、4.0、4.2、4.4版本） DDS实例（3.2、3.4、4.0、4.2、4.4版本） 	<ul style="list-style-type: none"> DDS实例（3.4、4.0、4.2、4.4版本） <p>说明 仅支持目标库版本等于或高于源库版本。</p>

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 同步阶段，确保目标数据库无业务写入，保证同步前后数据一致。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约20MB/s的查询压力，以及占用2~4个CPU。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-324 使用须知

类型名称	使用和操作限制
数据库权限设置	<p>源数据库最小权限要求：</p> <ul style="list-style-type: none"> • 全量+增量同步权限要求： <ul style="list-style-type: none"> - 副本集：连接源数据库的用户需要有admin库的readAnyDatabase权限，有local库的read权限。 <p>目标数据库最小权限要求：连接目标数据库的用户需要有admin库的dbAdminAnyDatabase权限，有目标数据库的readWrite权限。</p>
同步对象约束	<ul style="list-style-type: none"> • 副本集：目前只支持集合(包括验证器，是否是固定集合)，索引和视图的同步。 • 不支持同步系统库（如local、admin、config等）和系统集合，如果业务数据在系统库下，则需要先将业务数据移动到用户数据库下，可以使用renameCollection命令进行移出。 • 不支持视图的创建语句中有正则表达式。 • 不支持_id字段没有索引的集合。 • 不支持BinData()的第一个参数为2。 • 源库为4.4版本集群时，不支持复合哈希索引，不支持复合哈希分片键。 • 源库为4.4版本副本集时，不支持复合哈希索引。 • 同步对象映射名称长度在1到63个字符之间，不支持以下字符： / \ . " \$ <>。 • 请勿在源库集合的String类型字段中存储非UTF-8类型字符串，否则会导致同步前后数据不一致。
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源数据库的Oplog日志必须打开。 • 在磁盘空间允许的情况下，建议源数据库Oplog保存时间越长越好，建议为3天。 • 不支持源数据库的库名包含\.\$和空格，集合名或视图名中不能包含\$或以system.开头。 • 源数据库不能是GeminiDB Mongo实例。
目标数据库要求	<ul style="list-style-type: none"> • 目标数据库实例的运行状态必须正常。 • 目标数据库实例必须有足够的磁盘空间。 • 目前不支持从高版本数据库到低版本数据库的同步。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 相互关联的数据对象要确保同时同步，避免同步因关联对象缺失，导致同步失败。常见的关联关系：视图引用集合、视图引用视图等。 ● 副本集：MongoDB数据库的副本集实例状态必须正常，要存在主节点。 ● 源数据库为副本集实例时，增量同步阶段支持如下操作： <ul style="list-style-type: none"> - 支持文档（document）新增、删除、更新。 - 支持集合（collection）新建、删除。 - 支持索引（index）新建、删除。 - 支持视图（view）新建，删除。 - 支持convertToCapped、collMod、renameCollection命令。 ● 对于在源数据库已经存在TTL索引的集合，或者在增量同步期间在源库数据创建了TTL索引的集合，由于源数据库和目标库数据库时区，时钟的不一致，不能保证同步完成之后数据的一致性。 ● 压缩参数block_compressor的取值，根据源库对应集合的stats().wiredTiger.creationString.block_compressor来决定，如果目标库已经存在对应的空集合，则不同步压缩参数；特殊的，对于不支持的压缩参数，按照目标库实例的net.compression.compressors参数决定。当目标库为DDS 4.2版本时，由于目标库不支持压缩参数的设置，DRS不同步压缩参数。 ● 如果源数据库的MongoDB服务不是单独部署的，而是和其他的服务部署在同一台机器，则必须要给源数据库的wiredTiger引擎加上cacheSizeGB的参数配置，建议值设为最小空闲内存的一半。 ● 选择集合同步时，增量同步过程中不建议对集合进行重命名操作。 ● 建议填写副本集所有的主节点和备节点信息，以防主备切换影响同步任务。如果填写的是主备多个节点的信息，注意所有的节点信息必须属于同一个副本集实例。 ● 为了保持数据一致性，在整个同步过程中，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做删除类型的DDL操作，比如删除数据库、集合、索引、文档、视图等，这样可能会引起任务同步失败。 ● 在整个同步过程中，不支持源数据库主备切换导致数据回滚的情况。 ● 不支持全量同步和增量同步阶段insert、update源库大于16MB的文档。 ● 任务增量阶段，为保证同步的性能，会集合级的进行并发回放。特殊地，以下情况只支持单线程写入，不支持并发回放： <ul style="list-style-type: none"> - 集合的索引中有unique key时；

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 集合属性的capped为true时； 这种情况可能会导致任务时延增高。 • 为了提高同步的速度，在开始同步之前，建议在源数据库删掉不需要的索引，只保留必须的索引。在同步过程中不建议对源库创建索引，如果必须要创建索引，请使用后台的方式创建索引。 • 为了防止回环，不支持在相同的源和目标库同时启动入云和出云任务。 <p>说明 同步任务支持源和目标为同一个数据库，为了避免回环，必须使用对象名映射功能进行重命名。</p>

操作步骤

本章节将以DDS副本集为示例，介绍在VPC网络场景下，通过数据复制服务配置DDS副本集同步任务的流程。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-305 同步任务信息

表 3-325 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-306 同步实例信息



表 3-326 同步实例信息

参数	描述
数据流动方向	选择入云。 入云指目标端数据库为本云数据库。
源数据库引擎	选择“MongoDB”。
目标数据库引擎	选择“DDS”。
网络类型	此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。 <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的目标数据库实例。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <p>全量+增量：该模式为数据库持续性同步，适用于对业务中断敏感的场景，通过全量同步过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据库保持数据持续一致。</p> <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>
源数据库实例类型	同步模式为“全量+增量”时，需要根据源数据库的具体来源进行设置，目前源库仅支持副本集，该项为非集群。

- 任务类型

图 3-307 任务类型

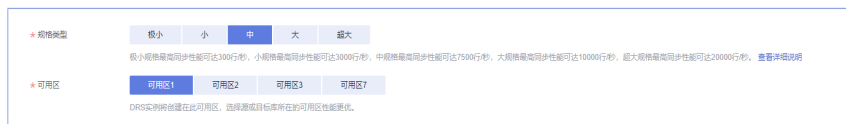


表 3-327 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-308 企业项目和标签



表 3-328 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 3-309 源库信息页面

源库信息

数据库类型 ECS自建库 DDS实例

数据库实例名称 查看数据库实例 查看不可选实例

账号认证数据库

数据库用户名

数据库密码 👁

测试连接
✔ 测试成功
?

表 3-329 源库信息

参数	描述
数据库类型	源数据库类型可以为“ECS自建库”或“DDS实例”，本示例选择“DDS实例”。
VPC	源数据库类型选择“ECS自建库”可见，指源数据库实例所在的虚拟专用网络，可以对不同业务进行网络隔离。
子网	源数据库类型选择“ECS自建库”可见，通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用分区内才会有效，创建源数据库实例的子网需要开启DHCP功能，在创建过程中也不能关闭已选子网的DHCP功能。
IP地址或域名	源数据库类型选择“ECS自建库”可见，源数据库的IP地址或域名。
数据库实例名称	源数据库类型选择“DDS实例”可见，选择作为源数据库的DDS实例。
账号认证数据库	填写的数据库账号所属的数据库名称。例如：华为云DDS实例默认的账号认证数据库为admin。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该同步任务后自动清除。

- 目标库信息配置

图 3-310 目标库信息

目标库信息

数据库实例名称

账号认证数据库

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-330 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的数据库实例，不可进行修改。
账号认证数据库	填写的数据库账号所属的数据库名称。例如：华为云DDS实例默认的账号认证数据库为admin。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的登录密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“同步设置”页面，设置同步对象，单击“下一步”。

图 3-311 设置同步对象

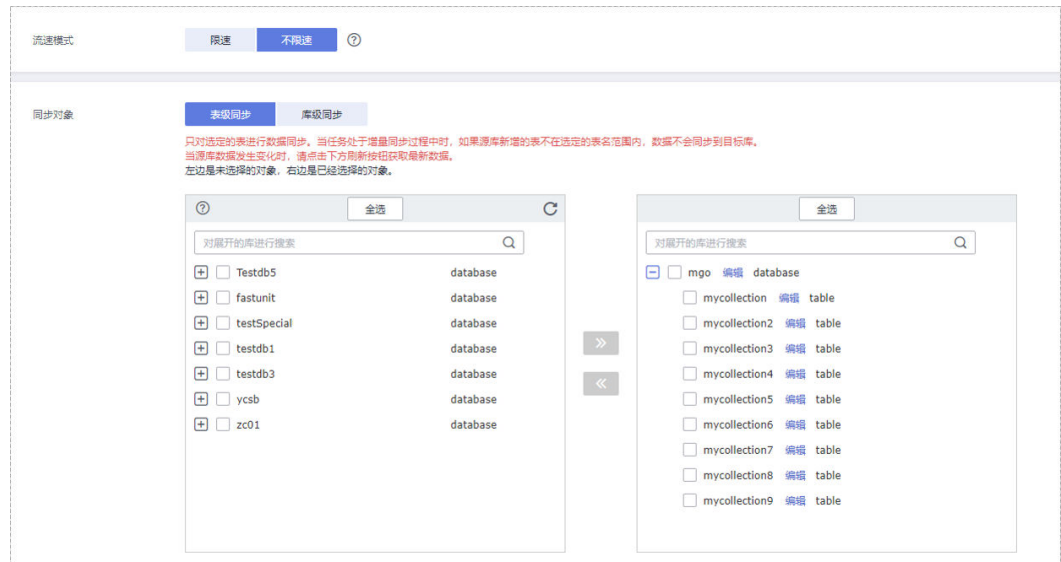




表 3-331 同步对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-312 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。映射后名称长度在1到63个字符之间，不支持以下字符：/ \ . " \$ < >。</p> <p>如果有切换源数据库的操作或源库同步对象变化的情况，请务必在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

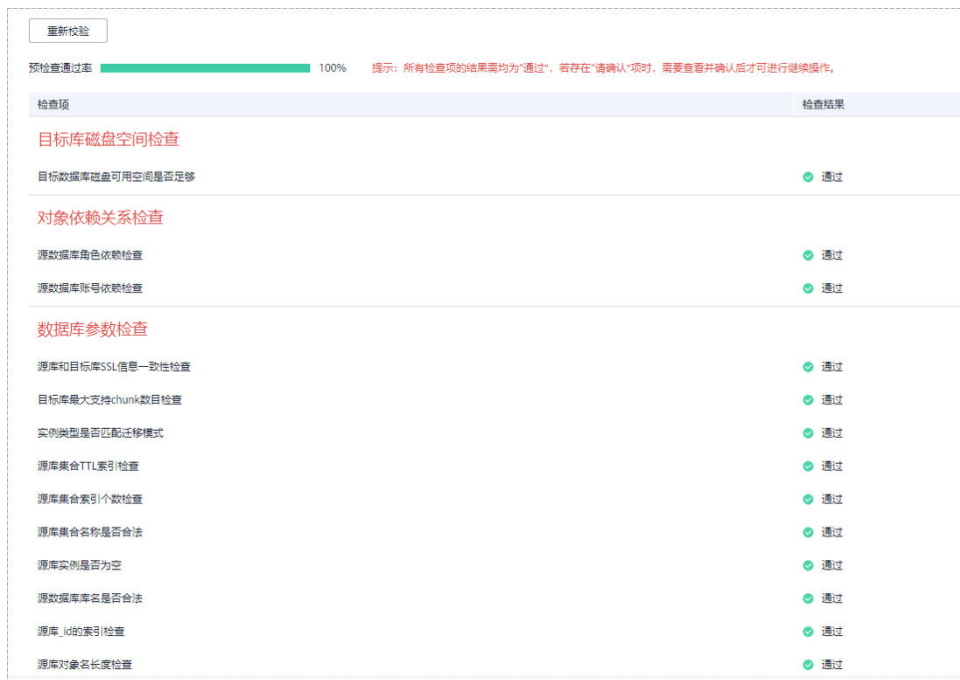
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且预检查通过率为100%时，单击“下一步”。

 **说明**

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

图 3-313 预检查



步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-314 任务启动设置




表 3-332 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。

参数	描述
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.31 将 MariaDB 同步到 MariaDB

支持的源和目标数据库

表 3-333 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MariaDB数据库 10.3、10.4、10.5版本ECS自建MariaDB数据库 10.3、10.4、10.5版本其他云上MariaDB数据库 10.3、10.4、10.5版本	RDS for MariaDB 10.5版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-334](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-334 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 仅支持MyISAM和InnoDB表的同步。 不支持事件、触发器的同步。 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 视图、存储过程、函数依赖的表不支持做表名映射。 任务做表名映射时，不支持该表的外键约束同步。 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-335](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-335 数据库账号权限

类型名称	全量+增量同步
源数据库连接账号	SELECT、SHOW VIEW、EVENT、LOCK TABLES、RELOAD、REPLICATION SLAVE、REPLICATION CLIENT（或BINLOG MONIOTOR，MariaDB 10.5及以上版本更改为此权限）。
目标数据库连接账号	SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES、INDEX，RDS for MariaDB实例的root账户默认已具备上述权限。

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-336 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 增量同步时，MariaDB源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 - 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 - 源数据库expire_logs_days参数值为0，可能会导致同步失败。 - 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 - 增量同步时，必须设置MariaDB源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源数据库中的库名不能包含：'<>\'以及非ASCII字符。 - 源数据库中的表名、视图名不能包含：'<>\'以及非ASCII字符。 - 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 不支持从高版本同步到低版本。 - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - DRS同步时会有大量数据写入目标库，目标库max_allowed_packet 参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 - 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 - 目标数据库实例必须有足够的磁盘空间。 - 除了系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 - 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“'”。 ● 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 - 源库不允许存在与目标库同名的无主键表。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 - 源库和目标库为RDS for MariaDB实例时，不支持带有TDE特性并建立具有加密功能表。 - 如果源库MariaDB不支持TLS1.2协议，需提交运维申请才能使用SSL测试连接。 - 如果源库版本低于10.4.3，选择库级同步，任务启动之后，请不要在对象选择的库里使用带json数据类型的建表语句创建表结构（如create table t1 (c1 json)），或者使用带json数据类型的加列语句（如alter table t1 add column c1 json），否则数据将不一致。 - 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 - 不支持源数据库恢复到之前时间点的操作(PITR)。 - 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 - 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 - 不支持强制清理binlog，否则会导致同步任务失败。 - 不支持分区表的分区字段进行列映射。 - 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 - 任务创建后，目标库不能设置为只读。 - 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

类型名称	使用和操作约定
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量数据同步阶段，目标数据库会产生大量Binlog引起磁盘空间占用过大，因此在全量数据同步阶段目标数据库Binlog本地默认调整为只保留最新5个，全量同步结束后目标数据库Binlog本地保留时长恢复为用户当前配置。如果因业务需要，目标数据库本地Binlog保留时长需要保持用户当前配置，需提交工单申请开启。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 增量同步支持DDL语句。 ● 增量同步支持表的重命名，源表和目标表必须都在对象选择里。 ● 增量同步支持任务再编辑追加同步对象。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

操作步骤

本章节介绍在公网网络场景下，通过数据复制服务配置MariaDB同步任务的流程。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-315 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-337 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-316 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎操作，以免因为配置失误，需要重新创建迁移任务。

* 数据流动方向: [入云] [出云] [自建-自建] ?
DRS要求源数据库和目标数据库中至少有一方为本云数据库实例。入云指目标数据库为本云数据库实例的场景，本云数据库实例之间的同步，请选择入云。

* 源数据库引擎: [MySQL] [Oracle] [DB2 for LUW] [DBM] [MariaDB] [MongoDB] [PostgreSQL] [Microsoft SQL Server] [TiDB]

* 目标数据库引擎: [MariaDB]

* 网络类型: [公网网络] ?
 DRS将会自动为DRS实例创建公网IP，该任务结束后将自动删除该弹性公网IP。

* 目标数据库实例: [无可用实例] ? [重置数据库实例] [重置不可用实例]

同步实例所在子网: [选择子网] ? [查看子网]

* 同步模式: [全库+增库] ?
该模式为数据库实例实时同步，通过全量过程或目标数据库的初始化，增量同步的数据通过解析日志等技术，将源端和目标数据库的数据保持一致。

表 3-338 同步实例信息

参数	描述
数据流动方向	选择入云。 入云指目标端数据库为本云数据库。
源数据库引擎	选择“MariaDB”。
目标数据库引擎	选择“MariaDB”。

参数	描述
网络类型	<p>此处以“公网网络”为示例，默认为公网网络类型，可按照需求选择“VPC网络”、“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
目标数据库实例	用户所创建的RDS for MariaDB数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <p>全量+增量：该模式为数据库持续性同步，适用于对业务中断敏感的场景，通过全量同步过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据库保持数据持续一致。</p> <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 3-317 任务类型



表 3-339 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明</p> <p>目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 3-318 企业项目和标签



表 3-340 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 3-319 源库信息页面

源库信息

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

● 测试成功 ?

表 3-341 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库用户名所对应的密码。支持在任务创建后修改密码。 任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该同步任务后自动清除。

- 目标库信息配置

图 3-320 目标库信息

目标库信息

数据库实例名称 xxxx-xxx-xxxx

* 数据库用户名

* 数据库密码

待实例创建成功后再进行测试连接

表 3-342 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS for MariaDB数据库实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的登录密码。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。


步骤4 在“同步设置”页面，设置同步对象，单击“下一步”。

图 3-321 设置同步对象



表 3-343 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-322 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”，参考[数据加工](#)章节，设置相关规则。

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且预检查通过率为100%时，单击“下一步”。

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

图 3-323 预检查



步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-324 任务启动设置




表 3-344 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.32 将 MariaDB 同步到 MySQL

支持的源和目标数据库

表 3-345 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本ECS自建MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本其他云上MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本	RDS for MySQL 5.5、5.6、5.7、8.0版本 说明 <ul style="list-style-type: none">源库为MariaDB 10.0、10.1版本时，建议同步到目标库MySQL 5.6及以上的版本。源库为MariaDB 10.2、10.3、10.4版本时，建议同步到目标库MySQL 5.7及以上的版本。源库为MariaDB 10.5版本时，建议同步到目标库MySQL 8.0版本。

说明

目前仅支持白名单用户使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-346](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-346 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 ● 仅支持MyISAM和InnoDB表的同步。 ● 不支持事件、触发器的同步。 ● 视图、存储过程、函数依赖的表不支持做表名映射。 ● 任务做表名映射时，不支持该表的外键约束同步。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 ● 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-347](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-347 数据库账号权限

类型名称	增量、全量+增量同步
源数据库连接账号	SELECT、SHOW VIEW、EVENT、LOCK TABLES、RELOAD、REPLICATION SLAVE、REPLICATION CLIENT（或BINLOG MONIOTOR，MariaDB 10.5及以上版本更改为此权限）。
目标数据库连接账号	SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES，RDS for MySQL实例的root账户默认已具备上述权限。当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-348 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 增量同步时，MariaDB源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 - 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 源数据库为自建MariaDB时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 源数据库为RDS for MariaDB时，设置binlog保留时间可参考《RDS用户指南》。 - 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 - 增量同步时，必须设置MariaDB源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源数据库中的库名不能包含：'<'>'以及非ASCII字符。 - 源数据库中的表名、视图名不能包含：'<'>'以及非ASCII字符。 - 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 不支持从高版本同步到低版本。 - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - DRS同步时会有大量数据写入目标库，目标库max_allowed_packet参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 - 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 - 目标数据库实例必须有足够的磁盘空间。 - 除了系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 - 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“'”。 ● 其他使用须知：

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 对于源数据库和目标数据库的版本有差异的同步链路，由于不同版本之间的部分特性差异，可能会出现语法兼容性问题，详情可参考MySQL或MariaDB不同版本之间的语法差异。 - 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 - 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 - 进行表级多对一同步时，源库不允许存在无主键表。 - 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 - 源库不允许存在与目标库同名的无主键表。 - 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 - 如果源库MariaDB不支持TLS1.2协议，需提交运维申请才能使用SSL测试连接。 - 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 - 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 - 不支持源数据库恢复到之前时间点的操作(PITR)。 - 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 - 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 - 不支持强制清理binlog，否则会导致同步任务失败。 - 不支持分区表的分区字段进行列映射。 - 任务创建后，目标库不能设置为只读。 - 多对一表级同步场景下，在“数据加工”页面设置了附加列为来源列，如果目标库中已经存在多对一映射之后的表，请在目标库中删除这张表或者清除表的数据，否则将不会通过附加列来创建联合主键。并且在同步过程中会因为数据冲突并忽略，而导致数据不一致。 - 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在

类型名称	使用和操作约定
	<p>的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。</p>
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量数据同步阶段，目标数据库会产生大量Binlog引起磁盘空间占用过大，因此在全量数据同步阶段目标数据库Binlog本地默认调整为只保留最新5个，全量同步结束后目标数据库Binlog本地保留时长恢复为用户当前配置。如果因业务需要，目标数据库本地Binlog保留时长需要保持用户当前配置，需提交工单申请开启。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY, 客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 增量同步支持表的重命名，源表和目标表必须都在对象选择里。 ● 增量同步支持任务再编辑追加同步对象。 ● 表级同步时，增量同步过程支持使用Online DDL，可参考DRS实时同步支持使用Online DDL工具吗。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

操作步骤

本章节介绍在公网网络场景下，通过数据复制服务配置MariaDB->RDS for MySQL同步任务的流程。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-325 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 3-349 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-326 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免影响数据迁移，请谨慎创建任务。

数据流向方向: [按钮: 入云] [按钮: 出云] [按钮: 自建自建] ?

源数据库引擎: MySQL, Oracle, DR2 for LUW, DDM, MariaDB, MongoDB, PostgreSQL, Microsoft SQL Server, GaussDB for MySQL, TD8

目标数据库引擎: MySQL, MariaDB, GaussDB for MySQL

网络类型: 公网网络 ?

DRS实例类型: WAL, AZ ?

目标数据库实例: [下拉菜单] ? [按钮: 复制数据库实例] [按钮: 删除不可用实例]

同步实例所在子网: [下拉菜单] ? [按钮: 查看子网] [按钮: 查看已占用的IP地址]

[按钮: 上一步] [按钮: 创建] [按钮: 取消] ?

表 3-350 同步实例信息

参数	描述
数据流动方向	选择入云。 入云指目标端数据库为本云数据库。
源数据库引擎	选择“MariaDB”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例，默认为公网网络类型，可按照需求选择“VPC网络”、“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	用户所创建的RDS for MySQL数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量：该模式为数据库持续性同步，适用于对业务中断敏感的场景，通过全量同步过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据库保持数据持续一致。 - “增量”：增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 3-327 规格类型



表 3-351 任务类型信息

参数	描述
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-328 可用区</p>

- 企业项目和标签

图 3-329 企业项目和标签



表 3-352 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 3-330 源库信息页面

源库信息

IP地址或域名

端口

数据库用户名

数据库密码 

SSL安全连接

● 测试成功 

表 3-353 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库用户名所对应的密码。支持在任务创建后修改密码。 任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该同步任务后自动清除。

- 目标库信息配置

图 3-331 目标库信息

目标库信息

数据库实例名称 RDS for MySQL

* 数据库用户名

* 数据库密码

待实例创建成功后再进行测试连接

表 3-354 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS for MySQL数据库实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的登录密码。

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“同步设置”页面，设置同步对象，单击“下一步”。

图 3-332 设置同步对象

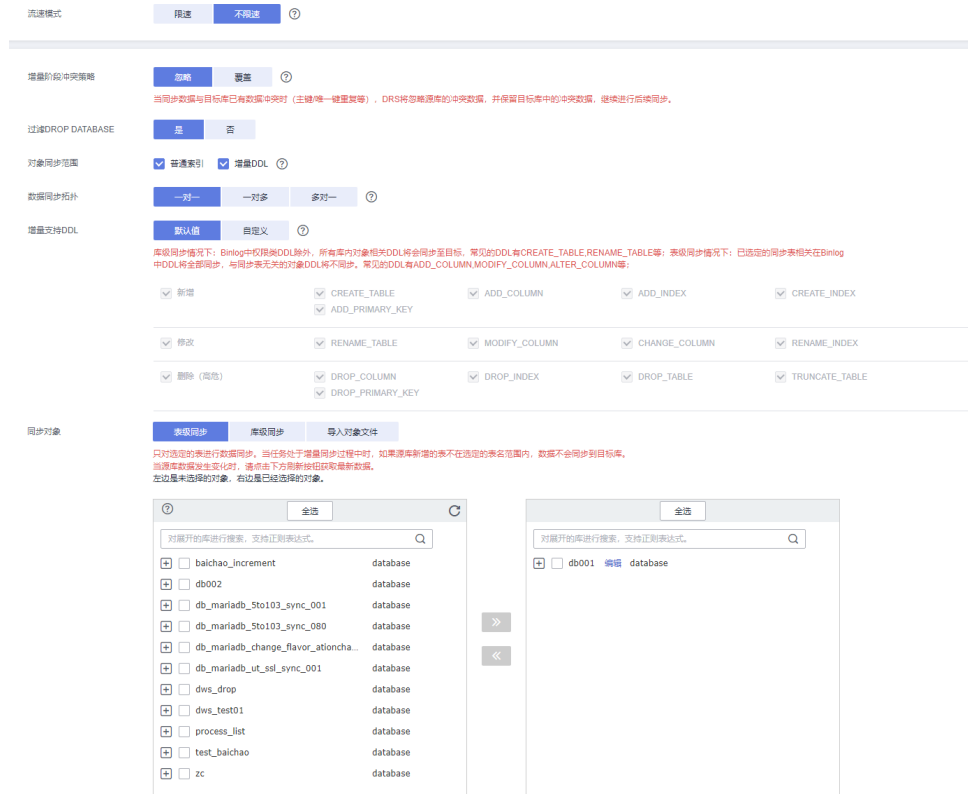



表 3-355 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-333 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明 “数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 3-334 数据加工



源对象/库	新名称	列名	操作类型	字段类型	操作
		a1	concat_name@database@table为列	varchar(256)	操作
		a2	取以值 123456789	long	
db_hai_pk_test1_table1@db_hai_pk_test1_table12 db_hai_pk_test1_table14 db_ha...		a3	concat_time为列	datetime	
		a4	concat_time为列	timestamp	
		a5	表达式 concat(_current_database, @_current_table)	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且预检查通过率为100%时，单击“下一步”。

📖 说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

图 3-335 预检查



步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-336 任务启动设置




表 3-356 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.33 将 MariaDB 同步到 GaussDB(for MySQL)

支持的源和目标数据库

表 3-357 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本ECS自建MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本其他云上MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本	GaussDB(for MySQL)主备实例

说明

目前仅支持白名单用户使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表3-358](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 3-358 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 ● 仅支持MyISAM和InnoDB表的同步。 ● 不支持事件、触发器的同步。 ● 视图、存储过程、函数依赖的表不支持做表名映射。 ● 任务做表名映射时，不支持该表的外键约束同步。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 ● 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 3-359](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-359 数据库账号权限

类型名称	增量、全量+增量同步
源数据库连接账号	SELECT、SHOW VIEW、EVENT、LOCK TABLES、RELOAD、REPLICATION SLAVE、REPLICATION CLIENT（或BINLOG MONIOTOR，MariaDB 10.5及以上版本更改为此权限）。
目标数据库连接账号	SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES，GaussDB(for MySQL)实例的root账户默认已具备上述权限。当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 3-360 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 增量同步时，MariaDB源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 - 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 源数据库为自建MariaDB时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 源数据库为RDS for MariaDB时，设置binlog保留时间可参考《RDS用户指南》。 - 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 - 增量同步时，必须设置MariaDB源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源数据库中的库名不能包含：'<>/'以及非ASCII字符。 - 源数据库中的表名、视图名不能包含：'<>/'以及非ASCII字符。 - 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 不支持从高版本同步到低版本。 - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - DRS同步时会有大量数据写入目标库，目标库max_allowed_packet参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 - 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例必须是GaussDB(for MySQL)主备实例。 - 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 - 目标数据库实例必须有足够的磁盘空间。 - 除了系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 - 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“'”。 ● 其他使用须知：

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 - 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 - 进行表级多对一同步时，源库不允许存在无主键表。 - 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 - 源库不允许存在与目标库同名的无主键表。 - 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 - 如果源库MariaDB不支持TLS1.2协议，需提交运维申请才能使用SSL测试连接。 - 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 - 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 - 不支持源数据库恢复到之前时间点的操作(PITR)。 - 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 - 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 - 不支持强制清理binlog，否则会导致同步任务失败。 - 不支持分区表的分区字段进行列映射。 - 任务创建后，目标库不能设置为只读。 - 多对一表级同步场景下，在“数据加工”页面设置了附加列为来源列，如果目标库中已经存在多对一映射之后的表，请在目标库中删除这张表或者清除表的数据，否则将不会通过附加列来创建联合主键。并且在同步过程中会因为数据冲突并忽略，而导致数据不一致。 - 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

类型名称	使用和操作约定
全量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。 全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY，客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 增量同步支持表的重命名，源表和目标表必须都在对象选择里。 增量同步支持任务再编辑追加同步对象。 表级同步时，增量同步过程支持使用Online DDL，可参考DRS实时同步支持使用Online DDL工具吗。
同步对比须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。

操作步骤

本章节介绍在公网网络场景下，通过数据复制服务配置MariaDB->GaussDB(for MySQL)同步任务的流程。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-337 同步任务信息

1、该页面**仅有任务名称和描述可修改**，其他在**点击开始创建后均不可修改**。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 3-361 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 3-338 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因配置错误，需要重新创建任务。

数据流动方向: [In-Cloud] [Out-Cloud] [Bidirectional] ?

源数据库引擎: MySQL, Oracle, DB2 for LUW, DBM, MariaDB, MongoDB, PostgreSQL, Microsoft SQL Server, GaussDB(for MySQL), TDB

目标数据库引擎: MySQL, MariaDB, GaussDB(for MySQL)

网络地址: [Public Network] ?

DRS任务类型: [Replica] [Reverse] ?

其他数据库实例: [Select Instance] ?

同步实例所在子网: [Select Subnet] ?

同步模式: [Full] [Incremental] ?

表 3-362 同步实例信息

参数	描述
数据流动方向	选择入云。 入云指目标端数据库为本云数据库。
源数据库引擎	选择“MariaDB”。
目标数据库引擎	选择“GaussDB(for MySQL)”。

参数	描述
网络类型	<p>此处以“公网网络”为示例，默认为公网网络类型，可按照需求选择“VPC网络”、“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	用户所创建的GaussDB(for MySQL)数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量：该模式为数据库持续性同步，适用于对业务中断敏感的场景，通过全量同步过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据库保持数据持续一致。 - “增量”：增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 3-339 规格类型



表 3-363 任务类型信息

参数	描述
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-340 可用区</p>

- 企业项目和标签

图 3-341 企业项目和标签



表 3-364 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 3-342 源库信息页面

源库信息

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功 ?

表 3-365 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	<p>源数据库用户名所对应的密码。支持在任务创建后修改密码。</p> <p>任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。</p>

参数	描述
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该同步任务后自动清除。

- 目标库信息配置

图 3-343 目标库信息

目标库信息

数据库实例名称 gaussdb-for-mysql

* 数据库用户名

* 数据库密码

待实例创建成功后再进行测试连接

表 3-366 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(for MySQL)数据库实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的登录密码。

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“同步设置”页面，设置同步对象，单击“下一步”。

图 3-344 设置同步对象

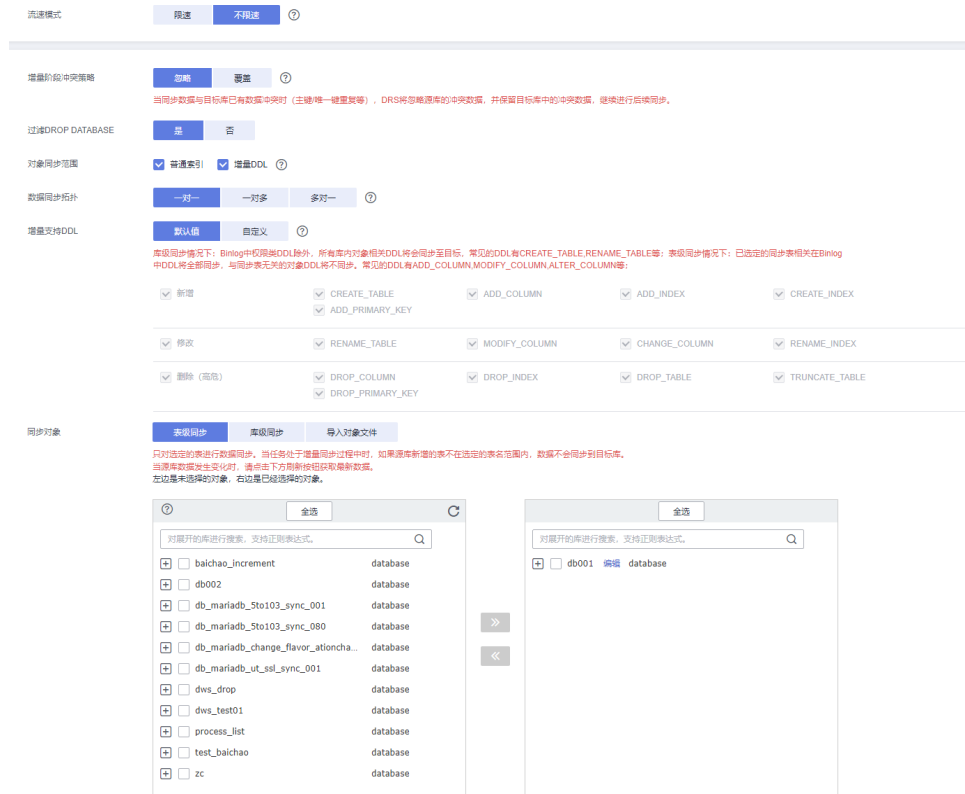




表 3-367 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-345 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明 “数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 3-346 数据加工



源对象/库	别名	列名	操作类型	数据类型	操作
		a1	concat_name@database@table为列	varchar(256)	操作
		a2	取以值 123456789	long	
db_hai_pk_test1_table1@db_hai_pk_test1_table12 db_hai_pk_test1_table14 db_ha...		a3	concat_time为列	datetime	
		a4	concat_time为列	timestamp	
		a5	表达式 concat(_current_database, @_current_table)	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

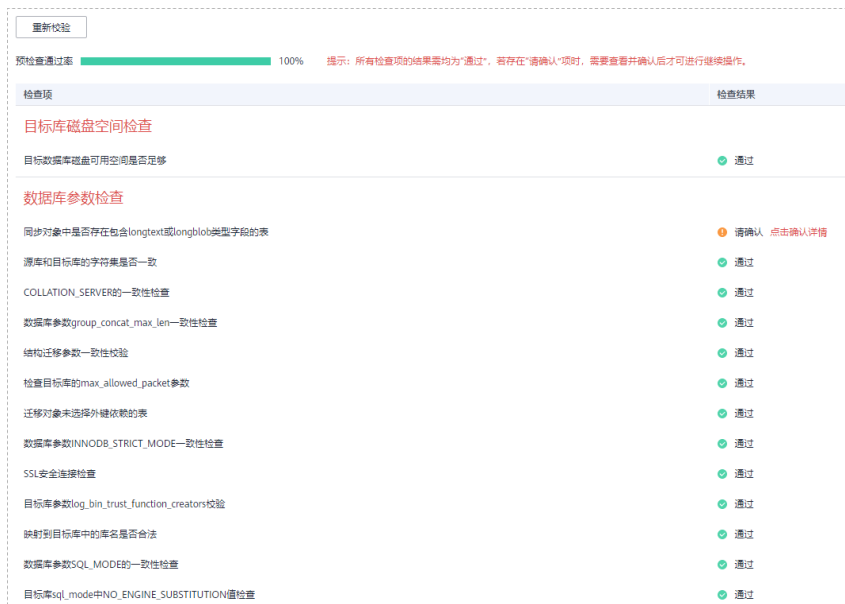
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且预检查通过率为100%时，单击“下一步”。

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

图 3-347 预检查



步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-348 任务启动设置




表 3-368 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

3.34 将 GaussDB(for MySQL)同步到 GaussDB(for MySQL)

支持的源和目标数据库

表 3-369 支持的数据库

源数据库	目标数据库
GaussDB(for MySQL)主备实例	GaussDB(for MySQL)主备实例

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 3-370 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none">• 源数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。• 目标数据库账号必须拥有如下权限：SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES，GaussDB(for MySQL)实例的root账户默认已具备上述权限。 当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步。 ● 视图、存储过程、函数依赖的表不支持做表名映射。 ● 任务做表名映射时，不支持该表的外键约束同步。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 仅支持MyISAM和InnoDB表的同步。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 ● 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。 ● 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名、表名、视图名不能包含：'<'>\以及非ASCII字符。 ● 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 数据库映射时，源库中存在视图、存储过程等对象，可能会导致同步失败。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例必须是GaussDB(for MySQL)主备实例。 ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 源库和目标库是相同的GaussDB(for MySQL)实例时，不支持没有库映射的实时同步。 ● 在创建DRS任务之前，如果源或者目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 源库不允许存在与目标库同名的无主键表。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 任务创建后，目标库不能设置为只读。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 当在全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量同步过程中不支持DDL操作。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY, 客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 增量同步支持表的重命名，源表和目標表必須都在对象选择里。 • 使用了附加列功能，单表的列数超过500时，对该表添加附加列可能会超过列数上限，会导致任务失败。 • 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 • 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

操作步骤

本小节以GaussDB(for MySQL)到GaussDB(for MySQL)的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 3-349 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 3-371 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 3-350 同步实例信息



表 3-372 同步实例信息

参数	描述
数据流动方向	选择“入云”。
源数据库引擎	选择“GaussDB(for MySQL)”。
目标数据库引擎	选择“GaussDB(for MySQL)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	可用的“GaussDB(for MySQL)”实例。

参数	描述
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 <p>该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。</p> <ul style="list-style-type: none"> - 增量。 <p>增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。</p> <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 3-351 任务类型



表 3-373 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明</p> <p>目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 3-352 可用区</p>

- 企业项目和标签

图 3-353 企业项目和标签



表 3-374 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 3-354 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中使用时参数模板设定

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 3-375 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该同步任务后自动清除。

图 3-355 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

✔ 测试成功

表 3-376 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库用户名对应的密码。 数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 3-356 同步模式

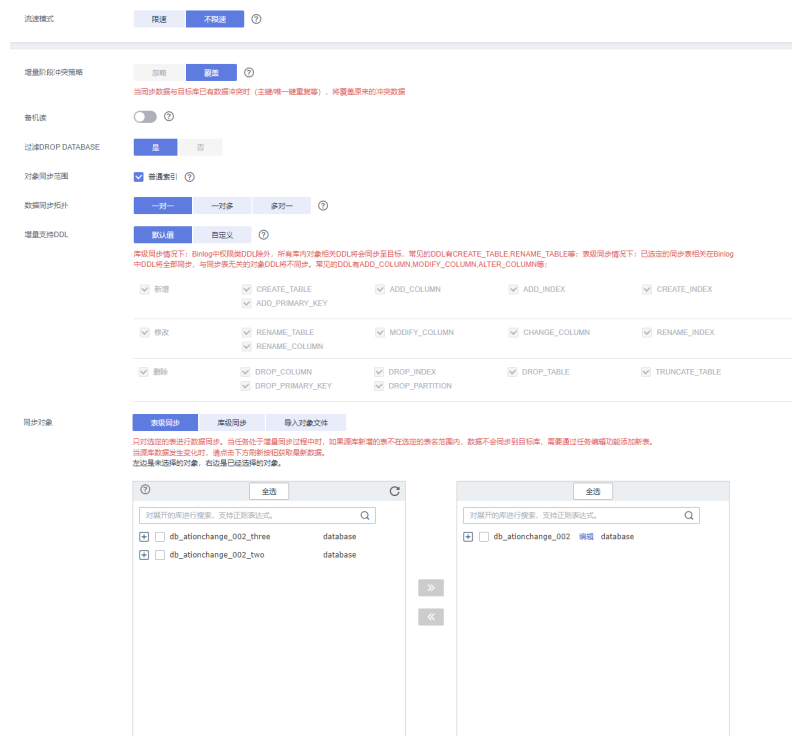



表 3-377 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 3-357 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>增量阶段冲突策略</p>	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
备机读	<p>步骤3中源数据库为可选实例时，支持备机读功能，可选打开或者关闭。</p> <ul style="list-style-type: none"> • 打开：全量数据将从备库读取，增量binlog从主库获取。 <ul style="list-style-type: none"> - 打开备机读取全量数据开关后，要求源端主节点与备节点实例状态正常，主备同步正常。同步对象中存在无主键表时，要求连接源端账户具有RELOAD权限。 - 源端在主节点执行DDL后，备节点会断开阻塞主备同步的连接，可能导致DRS任务全量同步失败。在未修改表列结构情况下，可通过续传尝试恢复任务。 - 源端节点发生主备角色变化时，可能会导致DRS任务全量同步失败，可通过续传尝试恢复任务。 - 仅源端为可选实例时支持备机读功能，不支持源端为自建库填写IP场景。 • 关闭：全量数据和增量binlog都将从主库获取。
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> • 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 • 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>选择是否同步“普通索引”。</p> <p>DRS将默认同步主键/唯一索引，普通索引是指除主键/唯一索引以外的其他类型索引。勾选普通索引将会同步全部的索引，不勾选则仅同步主键/唯一索引，普通索引不会同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>

参数	描述
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，可对需要加工的表对象进行数据过滤或添加附加列，完成后单击“下一步”。

- 如果需要设置数据过滤，选择“数据过滤”，设置相关过滤规则。
- 如果需要设置添加附加列，选择“附加列”，单击“操作”列的“添加”，选填需要添加的列名和操作类型信息。

相关操作可参考[数据加工](#)。

图 3-358 数据加工

源表名称	目标表	列名	操作类型	字符类型	操作
db1.jakl_011	db1.jakl_011	--	--	--	添加
db1.tak2	db1.tak2	--	--	--	添加
db1.tak3	db1.tak3	--	--	--	添加

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 3-359 预检查

检查项	检查结果
数据库参数检查	
结构迁移参数一致性检查	请确认, 点击确认详情
源迁移库无主键表检查	请确认, 点击确认详情
源库一对多映射检查	通过
源端存在不支持的外键引用操作	通过
源库选择的表中已包含附加列检查	通过
源库和目标库的字符集是否一致	通过
时区一致性检查	通过
COLLATION_SERVER的一致性检查	通过
数据库参数group_concat_max_len一致性检查	通过
TIME_ZONE的一致性检查	通过
检查目标库的max_allowed_packet参数	通过
迁移对象未选择外键依赖的表	通过

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 3-360 任务启动设置

启动时间 立即启动 稍后启动 ?

任务异常通知设置 ?

* SMN主题 ?

时延阈值(s) ?

数据异常通知


* 任务异常自动结束时间 ? 任务处于异常状态一段时间后，将会自动结束。单位为天。

表 3-378 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4 出云

4.1 将 MySQL 同步到 MySQL

支持的源和目标数据库

表 4-1 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">RDS for MySQL	<ul style="list-style-type: none">本地自建MySQL数据库ECS自建MySQL数据库其他云上MySQL数据库RDS for MySQL <p>说明 仅支持目标库版本等于或高于源库版本。</p>

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。如果同步不可避免业务高峰期，推荐使用同步限速功能，即“流速模式”选择“限速”。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。
- 如果涉及多对一场景的同步任务，可参考[多对一的场景约束及操作建议](#)。
- 如果涉及表级汇集的多对一同步任务，则不支持DDL，否则会导致同步全部失败。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-2 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账号需要具备如下权限： SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 提供的目标数据库账号必须拥有如下权限： SELECT、CREATE、INDEX、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES。当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步。 ● 视图、存储过程、函数依赖的表不支持做表名映射。 ● 任务做表名映射时，不支持该表的外键约束同步。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 仅支持MyISAM和InnoDB表的同步。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 ● 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。 ● 源数据库中的库名、表名如果存在非成对的{}，可能导致表结构同步失败。进行数据映射后，映射到目标库中的库名、表名如果存在非成对{}，可能导致表结构同步失败。 ● 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的表名、视图名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 RDS for MySQL设置binlog保留时间可参考《RDS用户指南》。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● 源数据库不支持选择只读实例。
目标数据库要求	<ul style="list-style-type: none"> ● 不支持从高版本同步到低版本。 ● 目标数据库必须有足够的磁盘空间。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● 除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● DRS同步时会有大量数据写入目标库，目标库max_allowed_packet参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 ● 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。 ● 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“'”。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 对于源数据库和目标数据库的版本有差异的同步链路，由于不同版本之间的部分特性差异，可能会出现语法兼容性问题，详情可参考MySQL不同版本之间的语法差异。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 进行表级多对一同步时，源库不允许存在无主键表。 ● 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 ● 源库不允许存在与目标库同名的无主键表。 ● 源库和目标库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果目标库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 当目标数据库为5.7版本时，因为版本限制，json类型中的浮点数会丢失小数点后末位的0，存在精度损失导致内容对比不一致。 ● 在创建DRS任务之前，如果源或者目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 任务创建后，目标库不能设置为只读。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 实时同步过程中，如果修改了源库或者目标库的用户名、密码，会导致同步任务失败，需要在数据复制服务控制台将上述信息重新修改正确，然后重试任务可继续进行实时同步。一般情况下不建议在同步过程中修改上述信息。 ● 实时同步过程中，如果修改了源库或者目标库端口，会导致同步任务失败。针对该情况，数据复制服务提供不同的处理机制。 <ul style="list-style-type: none"> - 对于源库端口，需要在数据复制服务控制台修改为正确的端口，然后重试任务可继续进行实时同步。 - 对于目标库端口，系统自动更新为正确的端口，需要重试任务即可进行同步。 一般情况下不建议在同步过程中修改端口。 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。 ● 全量同步过程中不支持DDL操作。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY，客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 增量同步支持表的重命名，源表和目標表必須都在对象选择里。 ● 增量同步支持任务再编辑追加同步对象。 ● 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

操作步骤

本小节以RDS for MySQL->MySQL的出云实时同步为示例，介绍如何使用数据复制服务创建两个数据库实例之间的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-1 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 4-3 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-2 同步实例信息

同步实例信息

以下选项均已不可修改，请谨慎填写，以免因配置错误，导致数据迁移任务失败。

数据流动方向: [入云] [出云] [新建] ?

源数据库引擎: MySQL, OceanBase 分布式数据库, OceanBase 主备库, MariaDB, DDS, PostgreSQL, OceanBase/MySQL

目标数据库引擎: MySQL, Oracle, CSES, Kafka

网络类型: [公网网络] ?

DRS将会自动为DRS实例创建网络IP，请勾选该选项并设置网络地址公网IP。

添加源库实例: [输入框] [重置和同步实例] [查看不可同步实例]

同步实例所在子网: [下拉菜单] [搜索子网] [查看已占用IP地址]

同步实例安全组: [选择安全组] [C]

同步模式: [全量+增量] [增量] ?

表 4-4 同步实例信息

参数	描述
数据流动方向	选择“出云”，即源端数据库为本云数据库。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“MySQL”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。可根据业务场景选择“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的RDS for MySQL数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-3 任务类型



表 4-5 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-4 企业项目和标签



表 4-6 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 同步实例创建成功后，在“源库及目标库”页面，填选源库信息和目标库信息后，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过，勾选协议，单击“下一步”。

入云和出云场景中源库和目标库需要填选的信息不同，具体以界面为准。

- 源库信息

图 4-5 源库信息

表 4-7 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	源数据库对应的用户名。
数据库密码	源数据库用户名对应的密码。
SSL安全连接	如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

源数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

- 目标库信息

图 4-6 目标库信息

目标库信息

IP地址或域名 请确保本云安全组允许访问源数据库公网地址，[点击这里](#)

端口

数据库用户名
不支持数据库参数和系统数据库同步，源数据库参数设定和用户将不会同步至目标数据库中，请在目标数据库中使用参数组修改参数，手工创建用户。

数据库密码

SSL安全连接
如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 4-8 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none">- 最大支持上传500KB的证书文件。- 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

目标数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择数据冲突策略和同步对象，单击“下一步”。

图 4-7 同步模式

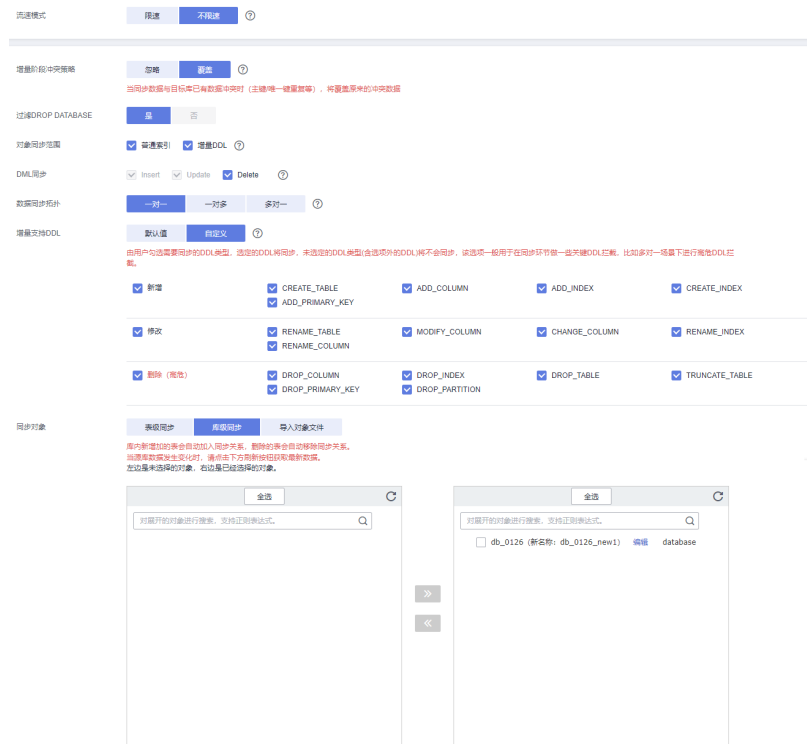




表 4-9 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-8 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 如果同步的对象在源数据库和目标数据库中的命名不同，可在右侧已选对象框中修改目标库对象映射名称，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 4-9 数据加工



源表/库	新名称	列名	操作类型	字符类型	操作
		a1	以serverName@database@table为列	varchar(256)	
		a2	默认值 123456789	long	
db_has_pk_test1_table16 db_has_pk_test1_table12 db_has_pk_test1_table14 db_has_pk_test1_table15		a3	以create_time为列	datetime	
		a4	以update_time为列	timestamp	
		a5	表达式 concat('current_database_', @@current_schema)	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-10 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 4-11 任务启动设置



表 4-10 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.2 将 MySQL 同步到 Kafka

支持的源和目标数据库

表 4-11 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">RDS for MySQL	<ul style="list-style-type: none">Kafka

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。如果同步不可避免业务高峰期，推荐使用同步限速功能，即“流速模式”选择“限速”。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，会存在3s以内的单表级锁定。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-12 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> 源数据库账号需要具备如下权限：SELECT、LOCK TABLES、SHOW VIEW、EVENT、REPLICATION SLAVE、REPLICATION CLIENT。
同步对象约束	<ul style="list-style-type: none"> 全量支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步，增量仅支持表数据和DDL同步。 仅支持MyISAM和InnoDB表的同步。 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。RDS for MySQL设置binlog保留时间可参考《RDS用户指南》。 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 源数据库中的库、表名不能包含：'<'>\以及非ASCII字符。 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。
目标数据库要求	<ul style="list-style-type: none"> 目标库为社区Kafka。 Kafka的auto.create.topics.enable参数建议为false。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 数据类型不兼容时，可能引起同步失败。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 实时同步过程中，如果修改了源库的用户名、密码，会导致同步任务失败，需要在数据复制服务控制台将上述信息重新修改正确，然后重试任务可继续进行实时同步。一般情况下不建议在同步过程中修改上述信息。 ● 实时同步过程中，如果修改了源库端口，会导致同步任务失败。针对该情况，系统自动更新为正确的端口，重试任务后即可进行同步。一般情况下不建议在同步过程中修改端口。 ● 实时同步过程中，对于因修改IP地址导致同步任务失败的情况，系统自动更新为正确的IP地址，需要重试任务可继续进行同步。一般情况下，不建议修改IP地址。 ● 全量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的identifier字段进行去重（同一个shardId）。 ● 增量同步过程中，请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 当在同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 选择表级对象同步时，同步过程中不建议对表进行重命名操作。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-12 同步任务信息

表 4-13 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-13 同步实例信息

表 4-14 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“MySQL”。

参数	描述
目标数据库引擎	选择“Kafka”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
源数据库实例	用户创建的RDS for MySQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： <ul style="list-style-type: none"> 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> - “增量”： <ul style="list-style-type: none"> 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-14 任务类型



表 4-15 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 4-15 可用区</p>

- 企业项目和标签

图 4-16 企业项目和标签



表 4-16 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-17 源库信息



表 4-17 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS实例，不可进行修改。

参数	描述
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-18 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

表 4-18 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

步骤**步骤2**中同步模式选择不同时，“设置同步”页面参数也不同，[图4-19](#)选择“全量+增量”为示例。

图 4-19 同步模式

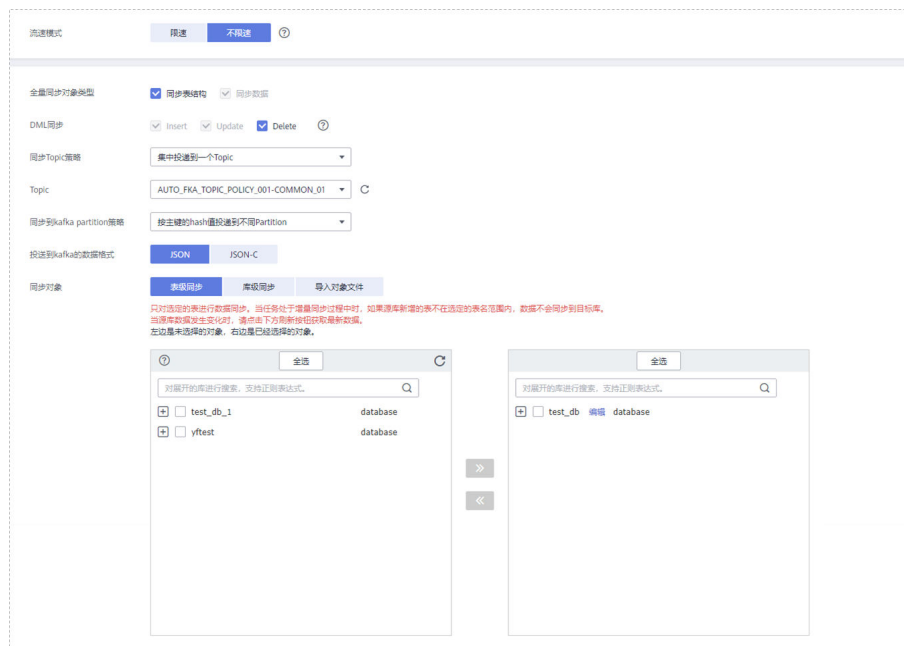



表 4-19 同步对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-20 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	全量同步对象类型，可选同步表结构、同步数据，根据实际需求进行选择要同步内容。
DML同步	选择需要同步的DML操作，默认都勾选。
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
同步Topic策略	同步Topic策略，可选择“集中投递到一个Topic”或者“自动生成Topic名字”。

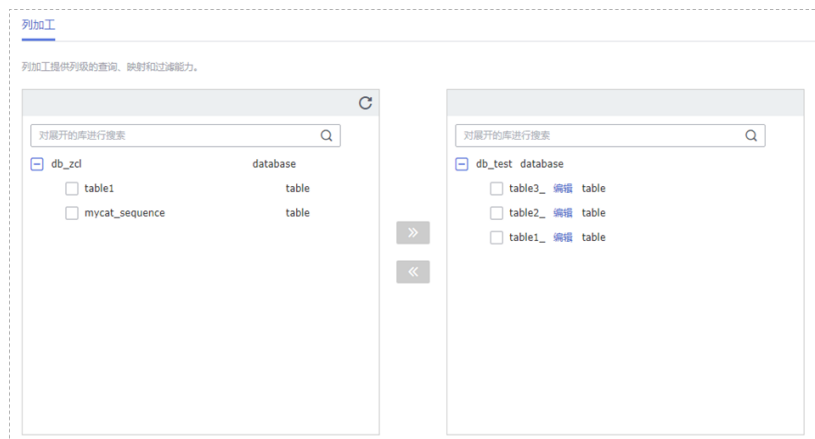
参数	描述
Topic	选择目标端需要同步到的Topic，同步Topic策略选择“集中投递到一个Topic”时可见。
Topic名字格式	<p>Topic名字格式，同步Topic策略选择“自动生成Topic名字”时可见。</p> <p>由于Kafka的机制限制，Topic名字只能包含ASCII字符、“.”、“_”及“-”，若超过该范围会导致创建Topic失败，任务异常。</p> <p>若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。</p> <p>Topic名字格式支持database和tablename两个变量，其他字符都当做常量。分别用\$database\$代替数据库名，\$tablename\$代替表名。</p> <p>例如：配置成\$database\$-\$tablename\$时，如果数据库名称为db1，表名为tab1，则Topic名字为db1-tab1。如果是DDL语句，\$tablename\$为空，则Topic名字为db1。</p>
Partition个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。</p>
副本个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。</p>
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名+表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 全部投递到Partition 0：适用于有事务要求的场景，事务保序，可以保证完全按照事务顺序消费，写入性能比较差，如果没有强事务要求，不推荐使用此选项。 按主键的hash值投递到不同的Partition：适用于一个表一个Topic的场景。
投送到kafka的数据格式	<p>选择MySQL投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。步骤步骤2的同步模式选择“增量”时可见，Avro格式目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。 JSON-C：一种能够兼容多个批量，流式计算框架的数据格式。 <p>详细格式可参考Kafka消息格式。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的列，进行列加工。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列，参考[数据加工](#)章节，设置相关规则。

图 4-21 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-22 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。

图 4-23 任务启动设置




表 4-20 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。

参数	描述
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.3 将 MySQL 同步到 CSS/ES

支持的源和目标数据库

表 4-21 支持的数据库

源数据库	目标数据库
RDS for MySQL 5.5、5.6、5.7、8.0版本	ElasticSearch 5.5、6.2、6.5、7.1、7.6、7.9、7.10版本

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-22 使用须知

类型名称	使用和操作限制
数据库权限设置	<p>全量+增量最小同步权限要求：</p> <ul style="list-style-type: none"> • 源数据库账号需要具备如下权限： SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 目标数据库账号必须拥有如下权限： READ、WRITE。
同步对象约束	<ul style="list-style-type: none"> • 支持表数据的同步。 • 不支持数据库、视图、索引、约束、函数、存储过程、触发器（TRIGGER）和事件（EVENT）的同步。 • 不支持系统库的同步以及事件状态的同步。 • 不支持主键为FLOAT类型的表的同步。 • 不支持无主键表的同步。 • 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的表名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 RDS for MySQL设置binlog保留时间可参考《RDS用户指南》。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● MySQL源数据库建议开启skip-name-resolve，减少连接超时的可能性。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 源库不支持mysql binlog dump命令。 ● 源数据库和目标数据库字符集需保持一致，否则同步失败。 ● 源数据库log_slave_updates参数需设置为开启状态，否则会导致同步失败。 ● 源数据库的binlog_row_image参数需设置为FULL，否则会导致同步失败。 ● 源数据库不支持选择只读实例。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库恢复到全量同步时间段范围内的PITR操作。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 源库时间字段目标库不支持的范围统一转成null。 ● 源库类型是binary类型会截断后面为0的字节，原因是源库是定长的自动补齐长度，目标库是变长类型。 ● binary的值会base64 加密后再写入目标库。 ● datetime类型源库无时区到目标库则是用户指定的。 ● 表字段名称全部转成小写。 ● 目标库数据_id如果指定源库多列生成需要用“:”分隔。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做DDL操作 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL操作）。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源库写入binlog格式为statement的数据。 ● 同步过程中，不允许源库执行清除binlog的操作。 ● 同步过程中，不允许在源库创建库名为ib_logfile的数据库。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 增量同步过程中，若源库存在分布式事务，可能会导致同步失败。 ● 增量同步会过滤所有的DDL操作。 ● 增量同步阶段，支持断点续传功能，在主机系统崩溃的情况下，对于非事务性的无主键的表可能会出现重复插入数据的情况。 ● 选择表级对象同步时，增量同步过程中不支持对表进行重命名操作。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 选择同步对象时，单次选择的库表名称和列名称的大小不能超过4M。如果超过该限制，可以通过再编辑同步对象功能，分批增加同步对象。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-24 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 4-23 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-25 同步实例信息

同步实例信息 ?

以下信息输入后不可修改，请谨慎填写，以免造成数据不连续，或者影响全量迁移任务。

数据流向: [In/Out] [Start] [Cancel] ?

源数据库引擎: MySQL, DDM, GaussDB(for MySQL) 分布式版, GaussDB 分布式版, GaussDB 主备版, DDS, PostgreSQL, GaussDB(for MySQL)

目标数据库引擎: MySQL, Oracle, CDB3, Kafka

同步类型: 公网同步 ?
 DRS同步类型为DRS实例创建公网同步, 该信息将同步到数据复制任务中。

同步实例名称: [Text Field] ? 设置实例名称, 设置不可见实例

同步实例所在子网: [Dropdown] ? 选择子网, 查看已占用的IP地址

同步模式: [Start] [Cancel] ?

请确认以上信息输入是否正确, 确认后系统会自动创建虚拟资源, 请留意网络和通过解析IP地址, 将源和目标数据库的数据保持一致。

表 4-24 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“CSS/ES”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户创建的RDS for MySQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

● 任务类型

图 4-26 任务类型



表 4-25 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-27 企业项目和标签



表 4-26 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。 如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-28 源库信息

表 4-27 源库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-29 目标库信息

目标库信息

IP地址或域名 ?
请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

SSL安全连接
如启用SSL安全连接, 请在目标库开启SSL, 并确保相关配置正确, 并上传SSL证书

加密证书

测试成功

表 4-28 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能, 用户可以选择是否开启对同步链路的加密。如果开启该功能, 需要用户上传SSL根证书。 说明 <ul style="list-style-type: none">• 最大支持上传500KB的证书文件。• 仅支持后缀为cer和pem的证书。• 如果不启用SSL安全连接, 请自行承担数据安全风险。

说明

目标库的数据库用户名和密码, 会被系统加密暂存, 直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面, 选择同步策略和同步对象, 单击“下一步”。

图 4-30 同步模式

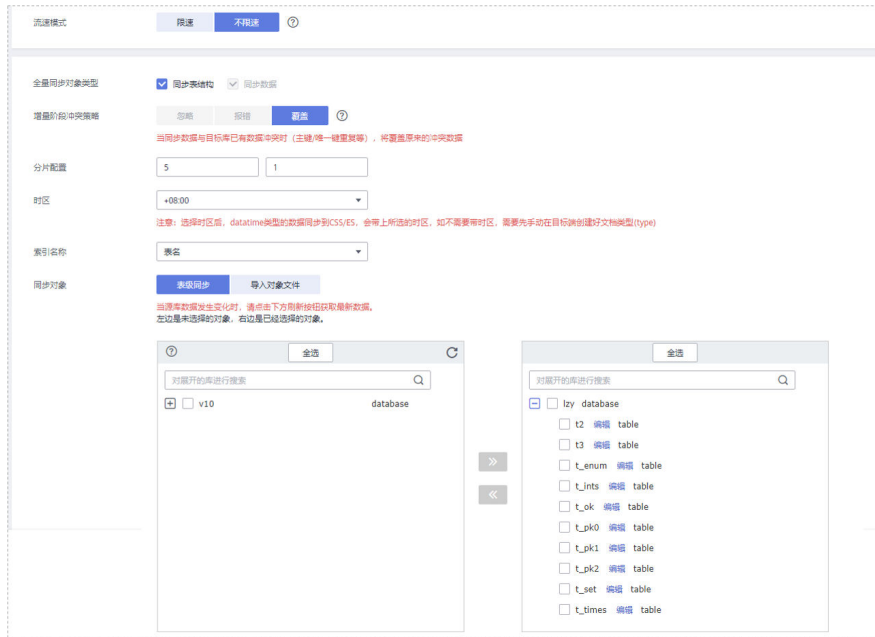



表 4-29 同步模式和对象

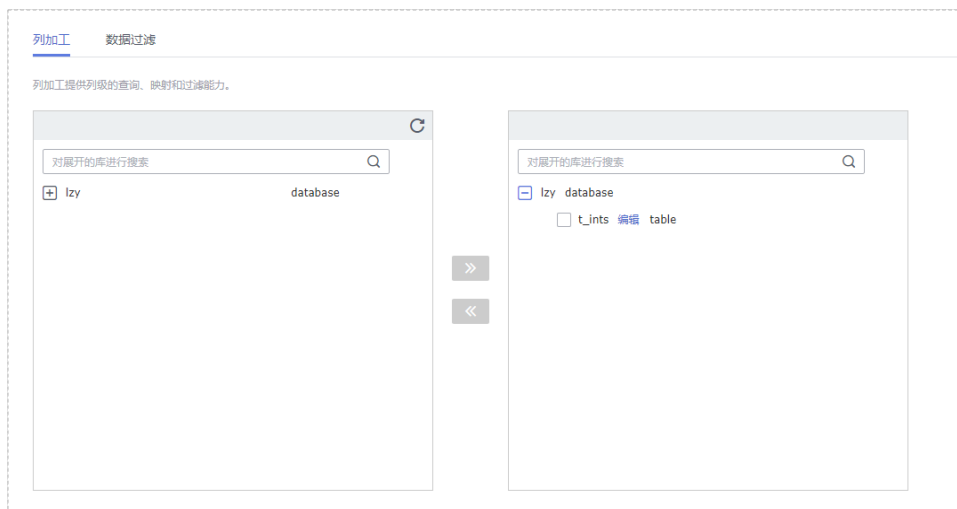
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-31 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> • 同步数据为必选项。 • 选择同步表结构的时候目标库不能有同名的表。 • 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，默认选择覆盖；全量阶段的冲突默认忽略。</p>
分片配置	<p>配置主分片数和分片副本数，主分片数默认为5，分片副本数默认为1。</p>

参数	描述
时区	选择时区后，datetime类型的数据同步到CSS/ES，会带上所选的时区。
索引名称	<ul style="list-style-type: none"> • 表名 选择为表名后，在目标Elasticsearch实例中创建的索引名称和表名一致。 • 库名+表名 选择为库名+表名后，在目标Elasticsearch实例中创建的索引名称为库名_表名。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> • 可以选择文件导入对象，具体步骤和说明可参考导入同步对象。 • 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 • 选择表级同步时，单次选择的库表名称和列名称的大小不能超过4M。如果超过该限制，可以通过再编辑同步对象功能，分批增加同步对象，参考编辑同步对象。 <p>说明</p> <ul style="list-style-type: none"> • 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 • 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 • 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 • 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 4-32 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-33 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-34 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-30 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时会有一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.4 将 MySQL 同步到 Oracle

支持的源和目标数据库

表 4-31 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • RDS for MySQL 	<ul style="list-style-type: none"> • 本地自建数据库 • ECS自建数据库

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-32 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账号需要具备如下权限：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 提供的目标数据库账号必须具有如下权限：ALTER ANY INDEX、ALTER ANY TABLE、ALTER SESSION、ANALYZE ANY、COMMENT ANY TABLE、CREATE ANY INDEX、CREATE ANY TABLE、CREATE SESSION、DELETE ANY TABLE、DROP ANY TABLE、INSERT ANY TABLE、SELECT ANY TABLE、SELECT ANY DICTIONARY、SELECT ANY TRANSACTION、UPDATE ANY TABLE、RESOURCE角色。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步支持数据、表结构和索引的同步。 ● 增量同步仅支持数据同步。 ● 不支持geometry, geometrycollection, linestring, multilinestring, multipoint, point, polygon等地理坐标类型 ● 不支持视图、约束、函数、存储过程、触发器（TRIGGER）和事件（EVENT）的同步。 ● 不支持系统库的同步以及事件状态的同步。 ● 目标库Oracle不支持空字符串，同步对象含有非空约束的字段中不能包含空字符串。 ● 源数据库MySQL支持的最大列数是1017，目标数据库Oracle支持的最大列数为1000，因为目标数据库限制，同步对象的列数不能大于1000。 ● 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的表名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 RDS for MySQL设置binlog保留时间可参考《RDS用户指南》。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● MySQL源数据库建议开启skip-name-resolve，减少连接超时的可能性。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 源库不支持mysql binlog dump命令。 ● 源数据库和目标数据库字符集需保持一致，否则同步失败。 ● 源数据库log_slave_updates参数需设置为开启状态，否则会导致同步失败。 ● 源数据库的binlog_row_image参数需设置为FULL，否则会导致同步失败。 ● 源数据库MySQL8.0目前不支持参数lower_case_table_names等于0的同步。 ● 源数据库不支持选择只读实例。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 由于数据库引擎支持不同，MySQL中的非空约束支持空字符串，Oracle中的非空约束不支持空字符串。数据同步过程中，如果非空约束的字段出现了空字符串，建议用户在目标端Oracle中将这个非空约束删除。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 源库时间字段默认值全0的统一转成1970-01-01 00:00:00。 ● 源库类型是decimal类型，精度超过38有效数字被截断补0，原因是目标库对应的number类型的最大精度是38。 ● 源库表中类型为varbinary、binary、timestamp的字段不能含有主键或唯一键约束。 ● 源库varchar类型的长度大于等于667时，该类型到Oracle中将会被转成clob类型。 ● 表字段名称全部转成大写。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做DDL操作 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL操作）。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源库写入binlog格式为statement的数据。 ● 同步过程中，不允许源库执行清除binlog的操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 同步过程中，不允许在源库创建库名为ib_logfile的数据库。 ● 增量同步场景下，不支持源数据库进行恢复到某个备份点的操作（PITR）。 ● 增量同步过程中，若源库存在分布式事务，可能会导致同步失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> 增量同步会过滤所有的DDL操作。 增量同步阶段，支持断点续传功能，在主机系统崩溃的情况下，对于非事务性的无主键的表可能会出现重复插入数据的情况。 选择表级对象同步时，增量同步过程中不支持对表进行重命名操作。 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-35 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown menu] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown menu]

* 任务名称: DRS-6435 ?

描述: [Text area] ?
0/256

表 4-33 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-36 同步实例信息

同步实例信息

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建迁移任务。

- 数据流动方向: 入云 | **出云** | 自建-自建
- 源数据库引擎: MySQL | DDM | GaussDB(for MySQL) 分布式版 | GaussDB 分布式版 | GaussDB 主备版 | DDS | PostgreSQL | GaussDB(for MySQL)
- 目标数据库引擎: MySQL | **Oracle** | CSS/ES | Kafka
- 网络类型: 公网网络
- 源数据库实例: 无可用实例 | 查看数据库实例 | 查看不可选实例
- 同步实例所在子网: 请选择子网 | 查看子网 | 查看已占用的IP地址
- 同步模式: 全量+增量

该模式为数据持续性实时同步，通过全量过程完成目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

表 4-34 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“Oracle”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	源数据库的RDS for MySQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 全量+增量。 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 任务类型

图 4-37 任务类型



表 4-35 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-38 企业项目和标签



表 4-36 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-39 源库信息

表 4-37 源库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-40 目标库信息

目标库信息

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ▼ ?

数据库用户名

数据库密码 👁

SSL安全连接

表 4-38 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群，建议使用scanip接入，提高访问性能。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

目标库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 4-41 同步模式

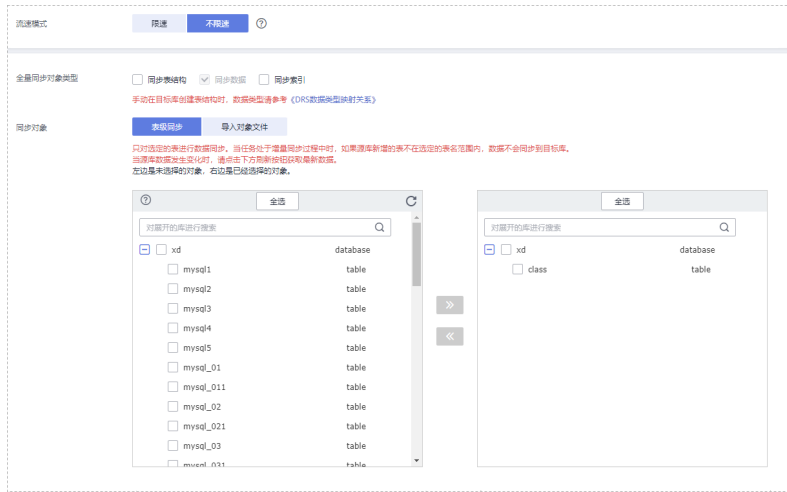



表 4-39 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-42 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> • 同步数据为必选项。 • 选择同步表结构的时候目标库不能有同名的表。 • 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。选择导入对象文件，具体步骤和说明可参考导入同步对象。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以保持待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-43 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 4-44 任务启动设置

表 4-40 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.5 将 MySQL 同步到 MariaDB

支持的源和目标数据库

表 4-41 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • RDS for MySQL (5.6、5.7、8.0版本) 	<ul style="list-style-type: none"> • 本地自建MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本 • ECS自建MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本 • 其他云上MariaDB数据库 10.0、10.1、10.2、10.3、10.4、10.5版本 <p>说明</p> <ul style="list-style-type: none"> • 源库为MySQL 5.6版本时，建议同步到目标库MariaDB 10.0、10.1及以上的版本。 • 源库为MySQL 5.7版本时，建议同步到目标库MariaDB 10.2、10.3、10.4及以上的版本。 • 源库为MySQL 8.0版本时，建议同步到目标库MariaDB 10.5版本。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表](#)

4-42进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

表 4-42 数据库账号权限

类型名称	全量同步	全量+增量同步、增量同步
源数据库连接账号	需要具备如下最小权限： SELECT、SHOW VIEW、EVENT。	需要具备如下最小权限： SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。
目标数据库连接账号	需要具备如下最小权限： SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES。	

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-43](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-43 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 仅支持MyISAM和InnoDB表的同步。 不支持事件（EVENT）和触发器（TRIGGER）的同步。 不支持源库中同步对象的表结构中带加密表标识（ENCRYPTION）。 视图、存储过程、函数依赖的表不支持做表名映射。 任务做表名映射时，不支持该表的外键约束同步。 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

注意事项

为保证任务的正常运行，DRS提供了自动化的检查方法，在您启动DRS任务前，会针对源数据库和目标数据库的相关配置、条件进行检查。主要的检查项和对应的处理建议请参考[预检查项](#)。除了预检查项目之外，您还需要注意的事情参见[表4-44](#)。

表 4-44 注意事项

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的表名、视图名不能包含：'<>\'"以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibttmp1。 ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● 除了MySQL和MariaDB系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● DRS同步时会有大量数据写入目标库，目标库max_allowed_packet参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 ● 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 对于源数据库和目标数据库的版本有差异的同步链路，由于不同版本之间的部分特性差异，可能会出现语法兼容性问题，详情可参考MySQL或MariaDB不同版本之间的语法差异。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 进行表级多对一同步时，源库不允许存在无主键表。 ● 源库不允许存在与目标库同名的无主键表。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 如果源库MySQL不支持TLS1.2协议，需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 不支持分区表的分区字段进行列映射。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 任务创建后，目标库不能设置为只读。 ● 实时同步过程中，如果修改了源库或者目标库的用户名、密码，会导致同步任务失败，需要在数据复制服务控制台将上述信息重新修改正确，然后重试任务可继续进行实时同步。一般情况下不建议在同步过程中修改上述信息。 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 当在全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量同步过程中不支持DDL操作。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY, 客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 增量同步支持表的重命名，源表和目標表必須都在对象选择里。 ● 表级同步时，增量同步过程支持使用Online DDL，可参考DRS实时同步支持使用Online DDL工具吗。 ● 增量同步支持任务再编辑追加同步对象。 ● 表级同步中，对于多对一场景下的表在数据加工页面设置了附加列为来源列，如果目标库中已经存在多对一映射之后的表，请在目标库中删除这张表或者清除这张表的数据，否则将不会通过附加列来创建联合主键（这在同步过程中会导致数据冲突并忽略，从而到导致数据不一致）。 ● 源库MySQL版本为8.0时，部分排序字符集（utf8mb4_0900_as_ci、utf8mb4_0900_as_cs、utf8mb4_0900_bin、utf8mb4_cs_0900_ai_ci等支持Unicode 9.0的排序字符集），在目标库不支持。 ● 由于引擎和版本差异，部分MySQL中的函数在MariaDB中不支持，不支持的函数有：MBRCOVEREDBY、ST_BUFFER_STRATEGY、ST_GeoHash、ST_IsValid、ST_LatFromGeoHash、ST_LongFromGeoHash、ST_PointFromGeoHash、ST_SIMPLIFY、ST_VALIDATE、(8.0)JSON_ARRAYAGG、JSON_OBJECTAGG、JSON_PRETTY、JSON_STORAGE_FREE、JSON_STORAGE_SIZE、JSON_TABLE。 ● 目标数据库MariaDB执行分区交换的时候不支持VALIDATION语法验证，在源数据库MySQL执行分区交换SQL的时候不要带校验语法，例如： ALTER TABLE t1 EXCHANGE PARTITION p0 WITH TABLE t2 WITH VALIDATION; ● 增量过程中请不要在源库MySQL的json类型中写入\n、\t、\r等字符，否则会导致数据不一致。 ● 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

操作步骤

本小节以RDS for MySQL->MariaDB的出云同步为示例，介绍如何使用数据复制服务创建两个数据库实例之间的实时同步任务，其他存储引擎的配置流程类似。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-45 同步任务信息

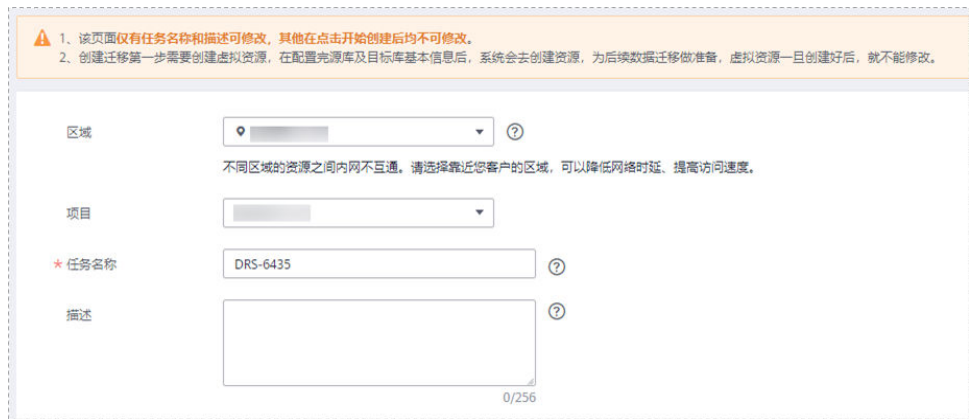


表 4-45 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含!= < > & ' " \ 特殊字符。

- 同步实例信息

图 4-46 同步实例信息



表 4-46 同步实例信息

参数	描述
数据流动方向	选择“出云”，即源端数据库为本云数据库。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“MariaDB”。
网络类型	<p>此处以“公网网络”为示例。可根据业务场景选择“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
源数据库实例	用户所创建的RDS for MySQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 4-47 可用区



表 4-47 任务可用区

参数	描述
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 4-48 可用区</p>

- 企业项目和标签

图 4-49 企业项目和标签



表 4-48 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 同步实例创建成功后，在“源库及目标库”页面，填选源库信息和目标库信息后，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 4-50 源库信息

源库信息

数据库实例名称

数据库用户名: root

数据库密码:

SSL安全连接:

测试连接 ✔ 测试成功 ?

表 4-49 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS for MySQL实例，不可进行修改。
数据库用户名	源数据库对应的用户名。
数据库密码	源数据库用户名对应的密码。
SSL安全连接	如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none">- 最大支持上传500KB的证书文件。- 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

- 目标库信息

图 4-51 目标库信息

目标库信息

IP地址或域名 ?
请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

SSL安全连接
如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 4-50 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL根证书。 说明 <ul style="list-style-type: none">- 最大支持上传500KB的证书文件。- 仅支持后缀为cer和pem的证书。- 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择数据冲突策略和同步对象，单击“下一步”。

图 4-52 同步模式

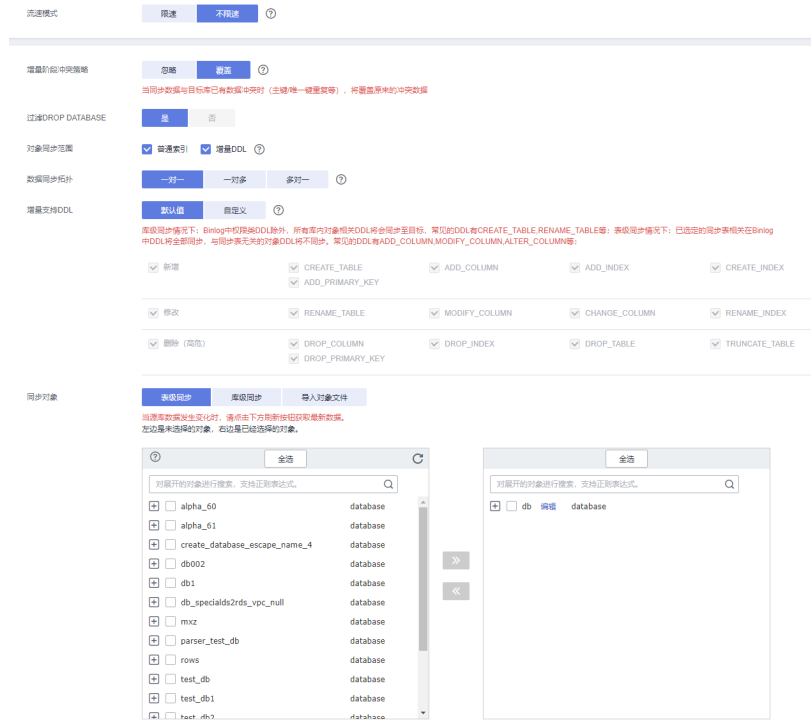




表 4-51 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-53 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>增量阶段冲突策略</p>	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 如果同步的对象在源数据库和目标数据库中的命名不同，可在右侧已选对象框中修改目标库对象映射名称，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 4-54 数据加工



源对象/库	新名称	列名	操作类型	字段类型	操作
		a1	concat_name@database@table为列	varchar(256)	操作
		a2	取以值 123456789	long	
db_has_pk_test1_table1@db_has_pk_test1_table12 db_has_pk_test1_table14 db_has_pk_test1_table16		a3	concat_time为列	datetime	
		a4	update_time为列	timestamp	
		a5	concat@_current_database,@_current_table	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-55 任务启动设置

The screenshot shows the 'Task Start Settings' configuration page. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned on.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned on.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): An input field containing the value '14'. A red asterisk and help icon are present. A red note below reads: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days.)

表 4-52 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.6 将 DDM 同步到 MySQL

支持的源和目标数据库

表 4-53 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • DDM实例 	<ul style="list-style-type: none"> • 本地自建MySQL数据库 • ECS自建MySQL数据库 • 其他云上MySQL数据库 • RDS for MySQL

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-54 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库DDM账户至少需要具备一个权限，比如SELECT；DDM物理分片数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 ● 提供的目标数据库账号必须拥有如下权限：SELECT、CREATE、ALTER、DROP、DELETE、INSERT、UPDATE。RDS for MySQL实例的root账户默认已具备上述权限。
同步对象约束	<ul style="list-style-type: none"> ● 全量同步支持数据、表结构和索引的同步。 ● 源库不允许存在拆分键为timestamp类型的表。 ● 源表的分库分表键要加到目标表的主键和唯一键中（也就是目标表的主键和唯一键中的列应该包含源表的分片列），避免数据冲突出现数据不一致问题。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 增量同步时，源物理分片数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 源分库分表中间件中的库名、表名不能包含：'<>/\以及非ASCII字符。 ● MySQL源数据库建议开启skip-name-resolve，减少连接超时的可能性。 ● 源物理分片数据库GTID状态建议为开启状态。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库为自建MySQL。 ● 目标数据库实例必须有足够的磁盘空间。 ● 除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● DRS同步时会有大量数据写入目标库，目标库max_allowed_packet 参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 数据类型不兼容时，可能引起同步失败。 ● 源数据库中存在主键或唯一键重复的数据时，直接同步将导致目标库数据比源库少，请务必检查并订正数据后启动同步。 ● 目标库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果目标库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 支持断点续传功能，在主机系统崩溃的情况下，对于无主键的表可能会出现重复插入数据的情况。 ● 任务创建后，目标库不能设置为只读。 ● 任务创建后，源数据库不支持增加逻辑库或修改旧逻辑库关联新的RDS，否则会导致数据无法正常同步或任务失败。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源端DDM正在同步的表做改变拆分键的操作；也不允许将单表/广播表改为拆分表，拆分表改为单表/广播表。 ● 增量同步阶段，不支持源数据库进行恢复操作。 ● 增量同步阶段，支持部分DDL操作。 <ul style="list-style-type: none"> - 不支持 DROP_DATABASE、DROP_TABLE、TRUNCATE_TABLE、CREATE_VIEW、DROP_VIEW。 - 不支持使用Online DDL。 - 支持创建表，例如： <pre>create table `ddl_test` (id int, c1 varchar(25), primary key(id)); create table `ddl_test_gho` like `ddl_test`;</pre> - 支持表的重命名，源表和目标表必须都在对象选择里面，例如： <pre>rename table `ddl_test` to `ddl_test_new`;</pre> - 支持表字段的增和改，不支持删列，例如： <pre>alter table `ddl_test` add column `c2` varchar(25); alter table `ddl_test` modify column `c1` varchar(50); alter table `ddl_test` alter c1 set default '***';</pre> - 支持修改表索引，例如：

类型名称	使用和操作限制
	<pre>alter table `ddl_test` drop primary key; alter table `ddl_test` add primary key(id); alter table `ddl_test` add index `ddl_test_uk`(id); alter table `ddl_test` drop index `ddl_test_uk`;</pre> <ul style="list-style-type: none"> - 表级同步支持增加列、修改列、增加主键和普通索引。 - 库级同步支持新建表、rename表、增加列、修改列、增加主键和普通索引。 - 新增和修改表名、列名、索引名时不能超出63字符，否则任务会失败。 - 源库无主键表增加主键的时候，必须含有第一列，否则任务会失败。 <ul style="list-style-type: none"> ● 增量同步阶段，对同一张表或列做DDL操作，需要在业务低峰期，并且时间间隔1分钟以上。 ● 任务再编辑时，对新添加的表，在再编辑变更过程中禁止大批量导出数据，建议在业务低峰期进行再编辑操作。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-56 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 4-55 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。

参数	描述
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

• 同步实例信息

图 4-57 同步实例信息



表 4-56 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“DDM”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的DDM实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>
源端数据库实例个数	源端DDM绑定的实例个数。

- 任务类型

图 4-58 任务类型



表 4-57 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-59 企业项目和标签



表 4-58 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-60 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

✔ 测试成功 ?

表 4-59 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的DDM实例，不可进行修改。
数据库用户名	源数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-61 目标库信息

目标库信息

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功 ?

表 4-60 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。

参数	描述
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> ● 最大支持上传500KB的证书文件。 ● 如果不启用SSL安全连接，请自行承担数据安全风险。


步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 4-62 同步模式



表 4-61 同步模式和对象

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-63 预检查



说明

所有检查项结果均通过时，若存在待确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-64 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-62 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.7 将 DDM 同步到 Oracle

支持的源和目标数据库

表 4-63 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • DDM实例 	<ul style="list-style-type: none"> • 本地自建数据库 • ECS自建数据库

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-64 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库DDM账户至少需要具备一个权限，比如SELECT；DDM物理分片数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 ● 提供的目标数据库账号必须具有如下权限：ALTER ANY INDEX、ALTER ANY TABLE、ALTER SESSION、ANALYZE ANY、COMMENT ANY TABLE、CREATE ANY INDEX、CREATE ANY TABLE、CREATE SESSION、DELETE ANY TABLE、DROP ANY TABLE、INSERT ANY TABLE、SELECT ANY TABLE、SELECT ANY DICTIONARY、SELECT ANY TRANSACTION、UPDATE ANY TABLE、RESOURCE角色。
同步对象约束	<ul style="list-style-type: none"> ● 支持源库数据同步。 ● 源库表结构仅支持全量同步。 ● 不支持同步表结构、索引、约束之外的数据库对象。 ● 目标库Oracle不支持空字符串，同步对象含有非空约束的字段中不能包含空字符串。 ● 源数据库DDM支持的最大列数是1017，目标数据库Oracle支持的最大列数为1000，因为目标数据库限制，同步对象的列数不能大于1000。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 增量同步时，源物理分片数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，binlog保存时间越长越好，建议为3天。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 源分库分表中间件中的库名、表名不能包含：'<>/\以及非ASCII字符。 ● 源物理分片数据库建议开启skip-name-resolve，减少连接超时的可能性。 ● 源物理分片数据库的GTID状态建议为开启状态。 ● 源物理分片数据库不支持枚举类型和set集合类型的。 ● 源数据库表名、字段名不能超过30个字符。 ● 不支持同步无主键表。 ● 源库中需要同步的数据库具有RESOURCE权限。 ● 源库的timestamp列的默认值，需要在目标库的合理取值内，否则会导致同步失败。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。 ● 目标数据库的时区设置必须与源数据库一致。 ● 目标库中需要同步的数据库（用户）具有RESOURCE权限。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 数据类型不兼容时，可能引起同步失败。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 由于数据库引擎支持不同，MySQL中的非空约束支持空字符串，Oracle中的非空约束不支持空字符串。数据同步过程中，如果非空约束的字段出现了空字符串，建议用户在目标端Oracle中将这个非空约束删除。 ● 源数据库中存在主键重复的数据时，直接同步将导致目标库数据比源库少，请务必检查并订正数据后启动同步。 ● 源数据库表中类型为varbinary、binary、timestamp的字段不能含有主键或唯一键约束。 ● 源库varchar类型的长度大于等于667时，该类型到Oracle中将会被转成clob类型。 ● 如果DDM的拆分库或拆分表的拆分键是函数，且不是主键，那么在Oracle里创建该表时，会自动把主键和拆分键作为复合主键。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 任务创建后，源数据库不支持增加逻辑库或修改旧逻辑库关联新的RDS，否则会导致数据无法正常同步或任务失败。 ● 同步过程中，不支持DDL操作。 ● 同步过程中，不允许源端DDM正在同步的表做改变拆分键的操作；也不允许将单表/广播表改为拆分表，拆分表改为单表/广播表。 ● 增量同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步过程中，若需要对源库需要同步的表结构进行修改，则用户必须在目标库同步修改表结构。 ● 增量同步场景下，不支持源数据库进行恢复操作。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-65 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 4-65 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-66 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建迁移任务。

* 数据流动方向: 入云 | 出云 | 自建 | 自建 ?

DRS要求源数据库或目标数据库中至少有一方为本地数据库实例，出云指源数据库为本地数据库实例的场景。

* 源数据库引擎: MySQL | DDM | GaussDB 分布式版 | GaussDB 主备版 | MariaDB | DDS | PostgreSQL | GaussDB(for MySQL)

* 目标数据库引擎: MySQL | Oracle | Kafka

* 网络类型: 公网网络 ?

DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除弹性公网IP。

* 源数据库实例: [下拉菜单] | 查看数据库实例 | 查看不可用实例

同步实例所在子网: [下拉菜单] | 查看子网

* 同步模式: 全量 + 增量 | 全量 | 增量 ?

该模式为数据库持续性实时同步，通过全量过程完成目标数据库的初始化后，增量同步和通过解析日志等技术，将源端和目标端数据保持数据一致。

* 源数据库实例个数: [- 2 +]

需要和源端DDM绑定的实例个数保持一致。

表 4-66 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“DDM”。
目标数据库引擎	选择“Oracle”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的DDM实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>
源端数据库实例个数	源端DDM绑定的实例个数。

● 任务类型

图 4-67 任务类型



表 4-67 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-68 企业项目和标签



表 4-68 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-69 源库信息

表 4-69 源库信息

参数	描述
中间件实例名称	默认为创建同步任务时选择的DDM实例，不可进行修改。
中间件用户名	源数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。
数据库实例	根据源库实际的分片数据库，填写对应的数据库信息。

说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-70 目标库信息

目标库信息

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功 ?

表 4-70 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群，建议使用scanip接入，提高访问性能。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

步骤4 在“设置同步”页面，选择同步对象，单击“下一步”。

图 4-71 同步模式

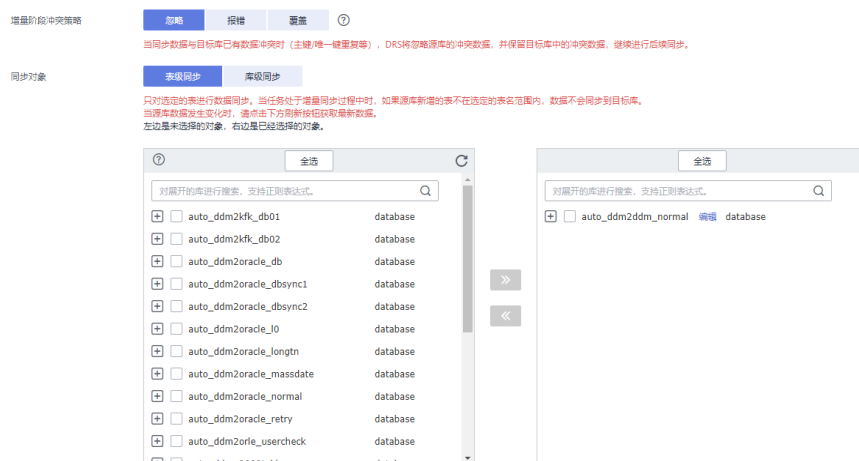



表 4-71 同步模式和对象

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。</p> <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的表对象，填写需要添加的列名、类型、操作类型信息，检查无误后，单击“下一步”。可参考[数据加工](#)中的添加同步时间戳，设置相关规则。

图 4-72 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-73 预检查



📖 说明

所有检查项结果均通过时，若存在待确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-74 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: An input field containing the number '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-72 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.8 将 DDM 同步到 Kafka

支持的源和目标数据库

表 4-73 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">• DDM实例	<ul style="list-style-type: none">• Kafka

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-74 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库DDM账户至少需要具备一个权限，比如SELECT；DDM物理分片数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。
同步对象约束	<ul style="list-style-type: none"> • 支持表数据的同步。
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源物理分片数据库的binlog日志必须打开，且binlog日志格式必须为Row格式，数据库GTID状态建议为开启，binlog_row_image必须为FULL。 • 在磁盘空间允许的情况下，binlog保存时间越长越好，建议为3天。 • 必须设置源数据库的server_id。server_id的取值范围在1 - 4294967296之间。 • 源分库分表中间件中的库名、表名不能包含：'<>/\以及非ASCII字符。 • 源物理分片数据库建议开启skip-name-resolve，减少连接超时的可能性。 • 源物理分片数据库的GTID状态建议为开启状态。 • 源物理分片数据库不支持枚举类型和set集合类型的实时同步。 • 源数据库表名、字段名不能超过30个字符。
目标数据库要求	<ul style="list-style-type: none"> • 目标库为社区Kafka。 • Kafka的auto.create.topics.enable参数建议为false。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> 数据类型不兼容时，可能引起同步失败。 源数据库中存在主键重复的数据时，直接同步将导致目标库数据比源库少，请务必检查并订正数据后启动同步。 任务创建后，源数据库不支持增加逻辑库或修改旧逻辑库关联新的RDS，否则会导致数据无法正常同步或任务失败。 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 同步过程中，不允许源端DDM正在同步的表做改变拆分键的操作；也不允许将单表/广播表改为拆分表，拆分表改为单表/广播表。 增量同步阶段，支持DDL操作。 增量同步过程中，请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-75 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 4-75 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。

参数	描述
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 4-76 同步实例信息

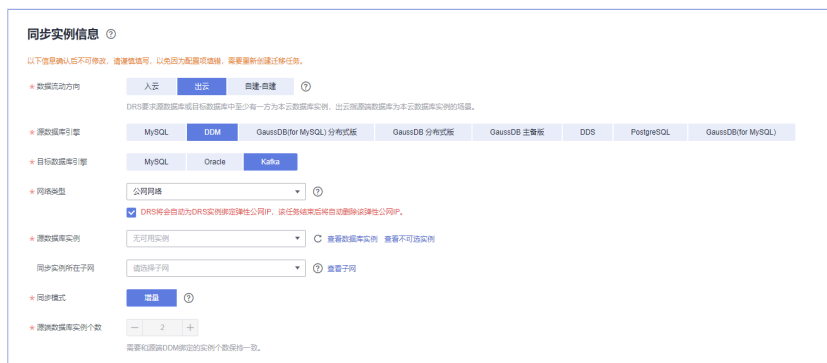


表 4-76 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“DDM”。
目标数据库引擎	选择“Kafka”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	可用的DDM实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	DRS任务支持的同步模式。DRS增量底层实现原理可参考 产品架构和功能原理 。 - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。
源端数据库实例个数	源端DDM绑定的实例个数，和源端实际情况保持一致。

- 任务类型

图 4-77 任务类型



表 4-77 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-78 企业项目和标签



表 4-78 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 4-79 源库信息


表 4-79 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的DDM实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。


 说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-80 目标库信息



目标库信息

IP地址或域名 

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

表 4-80 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 4-81 同步模式

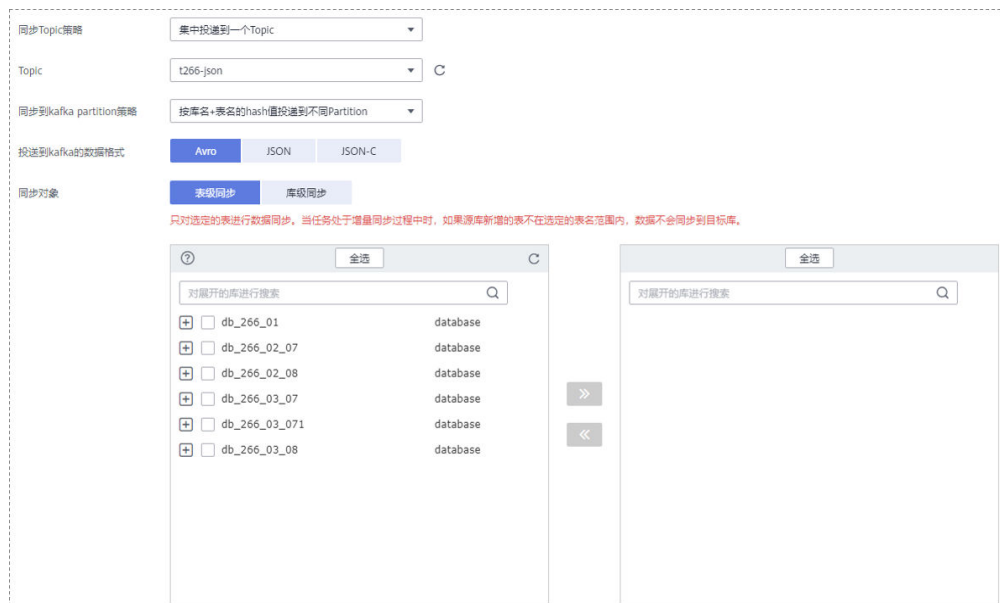



表 4-81 同步对象

参数	描述
同步Topic策略	同步Topic策略, 可选择“集中投递到一个Topic”或者“自动生成Topic名字”。
Topic	选择目标端需要同步到的Topic, 同步Topic策略选择“集中投递到一个Topic”时可见。
Topic名字格式	Topic名字格式, 同步Topic策略选择“自动生成Topic名字”时可见。 由于Kafka的机制限制, Topic名字只能包含ASCII字符、“.”、“_”及“-”, 若超过该范围会导致创建Topic失败, 任务异常。 若Topic名字格式中包含数据库对象名, 请确保对象名的字符在Kafka topic命名机制内。 Topic名字格式支持database和tablename两个变量, 其他字符都当做常量。分别用\$database\$代替数据库名, \$tablename\$代替表名。 例如: 配置成\$database\$-\$tablename\$时, 如果数据库名称为db1, 表名为tab1, 则Topic名字为db1-tab1。如果是DDL语句, \$tablename\$为空, 则Topic名字为db1。
Partition个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的分区个数。每个topic都可以创建多个partition, 越多的partition可以提供更高的吞吐量, 越多的partition会消耗更多的资源, 建议根据broker节点的实际情况来设置partition的数量。
副本个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的副本数。每个topic可以有多个副本, 副本位于集群中不同的broker上, 副本的数量不能超过broker的数量, 否则创建topic时会失败。

参数	描述
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名+表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 全部投递到Partition 0：适用于有事务要求的场景，事务保序，可以保证完全按照事务顺序消费，写入性能比较差，如果没有强事务要求，不推荐使用此选项。 按表的主键值hash值投递到不同的Partition：适用于一个表一个Topic的场景。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。 JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。 JSON-C：一种能够兼容多个批量，流式计算框架的数据格式。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-82 预检查



说明

所有检查项结果均通过时，若存在待确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 4-83 任务启动设置



表 4-82 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.9 将 DDS 同步到 MongoDB

支持的源和目标数据库

表 4-83 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> DDS实例（3.2、3.4、4.0、4.2、4.4版本） 	<ul style="list-style-type: none"> 本地自建Mongo数据库（3.2、3.6、3.4、4.0、4.2、4.4版本） ECS自建Mongo数据库（3.2、3.6、3.4、4.0、4.2、4.4版本） 其他云上Mongo数据库（3.2、3.6、3.4、4.0、4.2、4.4版本） <p>说明 仅支持目标库版本等于或高于源库版本。</p>

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。如果同步不可避免业务高峰期，推荐使用同步限速功能，即“流速模式”选择“限速”。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，会存在3s以内的单表级锁定。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-84 使用须知

类型名称	使用和操作限制
数据库权限设置	<p>源数据库最小权限要求：</p> <ul style="list-style-type: none"> • 增量同步权限要求： <ul style="list-style-type: none"> - 副本集：连接源数据库的用户需要有admin库的readAnyDatabase权限，有local数据库的read权限。 - 单节点：连接源数据库的用户需要有admin库的readAnyDatabase权限，有local数据库的read权限。 - 集群：连接源数据库mongos节点的用户需要有admin库的readAnyDatabase权限，有config数据库的read权限。 <p>目标数据库最小权限要求：连接目标数据库的用户需要有admin数据库的dbAdminAnyDatabase权限，有目标数据库的readWrite权限。对于目标数据库是集群的实例，迁移账号还要有config库的read权限。</p>
同步对象约束	<ul style="list-style-type: none"> • 副本集：目前只支持集合（包括验证器，是否是固定集合），索引和视图的同步。 • 集群：目前只支持集合（包括验证器，是否是固定集合），分片键，索引和视图的同步。 • 单节点：目前只支持集合（包括验证器，是否是固定集合），索引和视图的同步。 • 不支持_id字段没有索引的集合。 • 不支持BinData()的第一个参数为2。 • 不支持范围分片的情况下maxKey当主键。 • 源库为集群时，不支持DBPointer和DBRef数据类型。 • 请勿在源库集合的String类型字段中存储非UTF-8类型字符串，否则会导致同步前后数据不一致。
源数据库要求	<ul style="list-style-type: none"> • 增量同步时，源数据库的Oplog日志必须打开。 • 在磁盘空间允许的情况下，建议源数据库Oplog保存时间越长越好，建议为3天。 • 源数据库的库名不能包含\."\$和空格，集合名和视图名中不能包含\$字符或以system.开头。 • 如果源数据库是集群，则必须关闭Balancer并清理孤儿文档。详细操作可参考如何关闭集合均衡器Balancer和分片集群MongoDB迁移前清除孤儿文档。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none">• 不支持从高版本同步到低版本。• 目标数据库实例的运行状态必须正常。• 目标数据库实例必须有足够的磁盘空间。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 源数据库为非集群实例时，增量同步阶段支持如下操作： <ul style="list-style-type: none"> - 支持数据库（database）新建、删除。 - 支持文档（document）新增、删除、更新。 - 支持集合（collection）新建、删除。 - 支持索引（index）新建、删除。 - 支持视图（view）新建，删除。 - 支持convertToCapped、collMod、renameCollection命令。 ● 源库是集群实例时： <ul style="list-style-type: none"> - 不支持集群版本小于4.0。 - 增量阶段不允许对同步对象做删除操作，否则会导致任务失败。 - 会消耗源数据库一定量的CPU，内存资源，请提前做好源数据库资源评估。 - 受MongoDB Change Streams自身性能影响，如果源库的负载比较大，Change Streams会出现处理速度无法跟上Oplog产生速度，进而导致DRS同步出现时延。 - 目前仅支持drop database，drop collection，rename的DDL，其他DDL均不支持。 - 增量阶段单表最多支持10000条/秒左右的同步速度。 ● 启动时，指定的启动位点必须在Oplog的范围内。 ● 目标数据库的同步账号必须是有写权限。对于目标数据库是集群的实例，同步账号还要有对config数据的读权限。 ● 对于在源数据库已经存在TTL索引的集合，或者在增量同步期间在源库数据创建了TTL索引的集合，由于源数据库和目标库数据库时区，时钟的不一致，不能保证同步完成之后数据的一致性。 ● 压缩参数block_compressor的取值，根据源库对应集合的stats().wiredTiger.creationString.block_compressor来决定，如果目标库已经存在对应的空集合，则不迁移压缩参数；特殊的，对于不支持的压缩参数，按照目标库实例的net.compression.compressors参数决定。当目标库的存储引擎不是wiredTiger时，DRS不同步压缩参数。 ● 出云同步的目标数据库如果是副本集，则建议目标数据库填写所有的主节点和备节点信息，以防主备切换影响同步任务。如果是填写主备节点，注意填写的节点信息必须属于同一个副本集实例。 ● 出云同步的目标数据库如果是集群，则建议填写多个mongos信息，以防单个mongos节点故障影响同步任务，并且多个mongos信息还可以实现负载均衡。如果填写多个mongos信息，注意所有所填的信息必须属于同一个集群实例。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 在整个同步过程中，为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> 在整个同步过程中，不支持源数据库主备切换导致数据回滚的情况。 不支持系统库的同步，用户名和角色需要在目标库手动创建。 不支持增量同步阶段insert、update源库大于16MB的文档。 任务增量阶段，为保证同步的性能，会集合级的进行并发回放。特殊地，以下情况只支持单线程写入，不支持并发回放： <ul style="list-style-type: none"> 集合的索引中有unique key时； 集合属性的capped为true时； 这种情况可能会导致任务时延增高。 为了防止回环，不支持同时启动入云迁移和出云同步任务。 行对比时，集群实例如果存在孤儿文档或者正在进行chunk同步，会导致返回的行数不准确，对比结果不一致。详细情况可参考MongoDB官方文档。

操作步骤

本小节以DDS到MongoDB的出云单增量实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-84 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?
0/256

表 4-85 任务和描述

参数	描述
区域	当前所在区域，可进行切换。

参数	描述
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 4-85 同步实例信息



表 4-86 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“DDS”。
目标数据库引擎	选择“MongoDB”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	可用的DDS实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式。DRS增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> 增量。 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 <p>无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 4-86 任务类型



表 4-87 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-87 企业项目和标签



表 4-88 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 4-88 源库信息

源库信息

数据库实例名称

账号认证数据库

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 4-89 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的DDS实例，不可进行修改。
账号认证数据库	填写的数据库账号所属的数据库名称。例如：华为云DDS实例默认的账号认证数据库为admin。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none">最大支持上传500KB的证书文件。如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-89 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

账号认证数据库

数据库用户名

数据库密码 

SSL安全连接

表 4-90 目标库信息

参数	描述
IP地址或域名	<p>目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。</p> <p>该输入框最多支持填写3组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。同时需要确保所填写的多个IP地址或域名属于同一个实例。</p> <p>说明 此处若填写的是多组IP地址或者域名信息，在进行测试连接的过程中，只要存在一组IP地址或者域名可以连通，那么测试连接就提示成功。所以需要您保证填写的IP地址或域名的正确性。</p>
账号认证数据库	填写的目标数据库账号所属的数据库名称。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择启动位点和同步对象，单击“下一步”。

图 4-90 同步模式

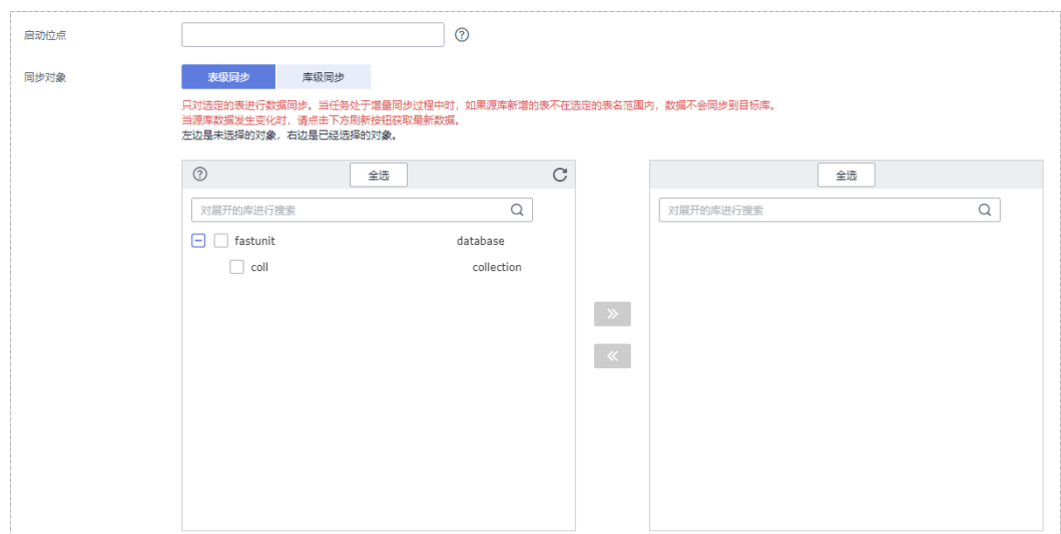



表 4-91 同步对象

参数	描述
启动位点	<p>增量同步的启动位点，格式为timestamp:incre（timestamp代表unix时间戳，单位为秒；incre代表当前命令在同一秒内执行的次序）。同步任务从启动位点开始获取源库增量日志（含当前启动位点）。</p> <ul style="list-style-type: none"> 副本集通过db.getSiblingDB("local").oplog.rs.find();命令查看源库oplog，ts字段的格式即为timestamp:incre。 集群无法从mongos上查询oplog，timestamp可以根据需要启动的时间，转换成unix timestamp格式，incre为1。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

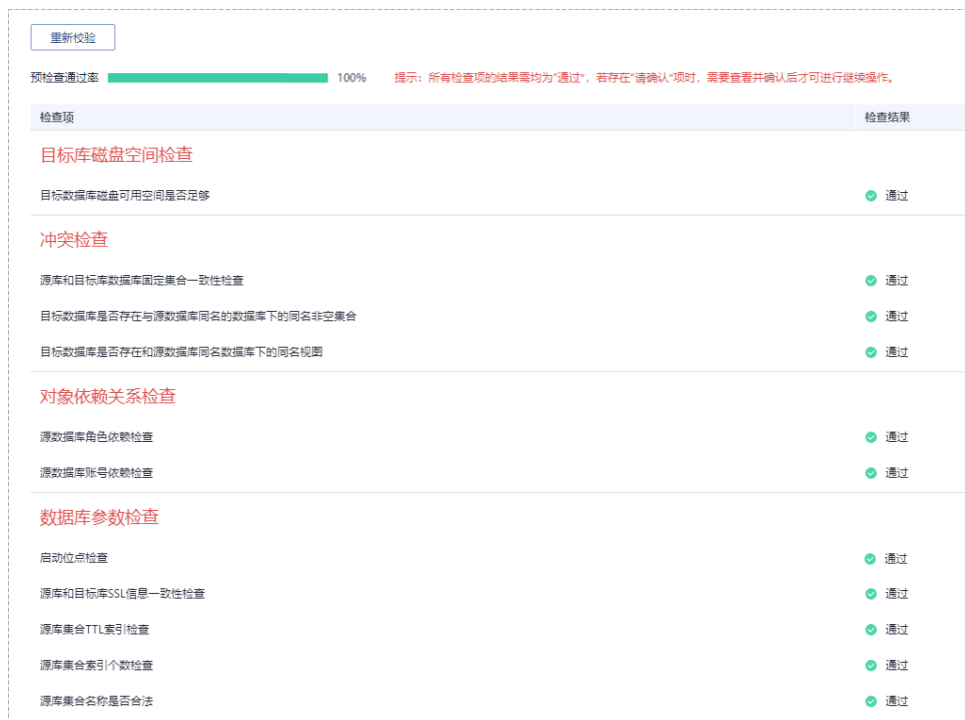
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-91 预检查



说明

所有检查项结果均通过时, 若存在待确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 4-92 任务启动设置



表 4-92 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.10 将 DDS 同步到 Kafka

支持的源和目标数据库

表 4-93 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> DDS实例（4.0、4.2、4.4版本） 	<ul style="list-style-type: none"> Kafka

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-94 使用须知

类型名称	使用和操作限制
数据库权限设置	源数据库最小权限要求： <ul style="list-style-type: none"> 增量同步权限要求：连接源数据库的用户需要有admin库的readAnyDatabase权限。
源数据库要求	<ul style="list-style-type: none"> 增量同步时，源数据库的Oplog日志必须打开。 在磁盘空间允许的情况下，建议源数据库Oplog保存时间越长越好，建议为3天。 源数据库的库名不能包含\."\$和空格，集合名中不能包含\$字符或以system.开头。 如果源数据库是集群，则必须关闭Balancer并清理孤儿文档。详细操作可参考如何关闭集合均衡器Balancer和分片集群MongoDB迁移前清除孤儿文档。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> 目标库为社区Kafka。 Kafka的auto.create.topics.enable参数建议为false。
操作须知	<ul style="list-style-type: none"> 同步会消耗源数据库一定量的CPU，内存资源，请提前做好源数据库资源评估。 受MongoDB Change Streams自身性能影响，如果源库的负载比较大，Change Streams会出现处理速度无法跟上Oplog产生速度，进而导致DRS同步出现时延。 支持的DML：insert、update、replace、delete。 支持的DDL：目前仅支持drop database，drop collection，rename的DDL，其他DDL均不支持。 增量阶段单表最多支持10000条/秒左右的同步速度。 启动时，指定的启动位点必须在Oplog的范围内。 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 在整个同步过程中，为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL、DML操作）。 不支持增量同步阶段insert、update源库大于16MB的文档。 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。 增量同步过程中，请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。

操作步骤

本小节以DDS到Kafka的出云单增量实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-93 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 4-95 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-94 同步实例信息

同步实例信息 ?

以下配置项一旦不可修改，请谨慎填写，以免因配置项错误，影响数据迁移进程。

数据流动方向: [Radio: 入云] [Radio: 出云] ?

源数据库引擎: MySQL DDM GaussDB 分布式 GaussDB 主备版 MariaDB DDS PostgreSQL GaussDB for MySQL

目标数据库引擎: Kafka MongoDB

实例名称: [Text Field] ?

源实例所在子网: [Dropdown] ?

目标实例所在子网: [Dropdown] ?

同步模式: [Radio: 增量] ?

表 4-96 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“DDS”。
目标数据库引擎	选择“Kafka”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	可用的DDS实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>DRS任务支持的同步模式。DRS增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - 增量。 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 <p>无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 4-95 任务类型



表 4-97 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-96 企业项目和标签



表 4-98 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 4-97 源库信息

源库信息

数据库实例名称

账号认证数据库

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 4-99 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的DDS实例，不可进行修改。
账号认证数据库	填写的数据库账号所属的数据库名称。例如：华为云DDS实例默认的账号认证数据库为admin。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-98 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

安全协议 PLAINTEXT ▼

测试连接

表 4-100 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择启动位点和同步对象，单击“下一步”。

图 4-99 同步模式

启动位点 ?

同步Topic策略 集中投递到一个Topic ▼

Topic ... C

同步到kafka partition策略 全部投递到Partition 0 ▼

投递到kafka的数据格式 JSON

同步对象 表级同步 库级同步

只对选定的表进行数据同步。当任务处于增量同步过程中时，如果源库新增的表不在选定的表名范围内，数据不会同步到目标库。当源库数据发生变化时，请点击下方刷新按钮获取最新数据。
左边是未选择的对象，右边是已经选择的对象。

全选

对展开的库进行搜索，支持正则表达式。


- 5万张表 database
- A database
- ZC001 database
- dbz1a_c_42replica2replica42_test_0... database
- dbz2a_c_42replica2replica42_test_0... database
- dbz3a_c_42replica2replica42_test_0... database
- dbz4a_c_42replica2replica42_test_0... database
- quanxuan database
- weidian database
- weidian02 database
- weidian03 database

全选

对展开的库进行搜索，支持正则表达式。

- test database
- test01 database

表 4-101 同步对象

参数	描述
启动位点	<p>增量同步的启动位点，格式为timestamp:incre（timestamp代表unix时间戳，单位为秒；incre代表当前命令在同一秒内执行的次序）。同步任务从启动位点开始获取源库增量日志（含当前启动位点）。</p> <ul style="list-style-type: none"> 副本集通过db.getSiblingDB("local").oplog.rs.find();命令查看源库oplog，ts字段的格式即为timestamp:incre。 集群无法从mongos上查询oplog，timestamp可以根据需要启动的时间，转换成unix timestamp格式，incre为1。
同步Topic策略	同步Topic策略，可选择“集中投递到一个Topic”。
Topic	选择目标端需要同步到的Topic。
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 全部投递到Partition 0：适用于有事务要求的场景，事务保序，可以保证完全按照事务顺序消费，写入性能比较差，如果没有强事务要求，不推荐使用此选项。 按表的主键值hash值投递到不同的Partition：适用于一个表一个Topic的场景。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。详细格式可参考Kafka消息格式。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-100 预检查



说明

所有检查项结果均通过时，若存在待确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-101 任务启动设置




表 4-102 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.11 将 PostgreSQL 同步到 PostgreSQL

支持的源和目标数据库

表 4-103 支持的数据库

源数据库	目标数据库
RDS for PostgreSQL (9.5、9.6、10、11、12、13、14、15版本)	<ul style="list-style-type: none">本地自建数据库 (PostgreSQL 9.5、9.6、10、11、12、13、14、15版本)ECS自建数据库 (PostgreSQL 9.5、9.6、10、11、12、13、14、15版本) <p>说明 仅支持目标库大版本等于或高于源库大版本。</p>

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-104](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-104 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 不支持实例级同步：每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持的字段类型： 数字类型、货币类型、字符类型、二进制数据类型、日期/时间类型、布尔类型、枚举类型、几何类型、网络地址类型、位串类型、文本搜索类型、UUID类型、XML类型、JSON类型、数组、复合类型、范围类型。 ● 全量同步的范围： <ul style="list-style-type: none"> - 库级同步支持：模式、表、索引、约束、视图、物化视图、序列、存储过程、规则、触发器、外键、排序规则、插件、编码转换信息、聚合函数、操作符、统计扩展、转换信息、文本搜索配置、函数、数据类型、类型转换、事件触发器、文本搜索解析器、文本搜索模板的同步。 表级同步时，仅支持表、视图、物化视图、序列、普通索引的同步。 文件导入时，支持表的同步。 - 不支持系统模式、系统表、系统用户、系统函数等系统对象；不支持表空间、外部数据包装器、外部服务器、用户映射、发布、订阅、用户等其他对象；不支持对象权限的同步。 <p>说明 支持同步的对象有如下限制：</p> <ul style="list-style-type: none"> ● 对象名称：库名不可以包含+"%\<>，模式名和表名不可以包含"!<>，列名不可以包含"和'。 ● 表：不同步临时表。表级同步时，表的约束、索引及规则会一起同步，表的触发器不会同步。 ● 函数：不同步C语言函数，不同步带有leakproof属性或带有support属性的函数。 ● 插件：不同步插件的元数据。 ● 数据类型：不同步基本数据类型。 ● 类型转换：不同步二进制强制型的类型转换。 <ul style="list-style-type: none"> ● 增量同步的范围： <ul style="list-style-type: none"> - 支持部分DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 - 支持部分DDL（Data Definition Language）：包括TRUNCATE（仅PostgreSQL 11及以上版本支持）、CREATE SCHEMA、CREATE TABLE、DROP TABLE、ALTER TABLE（包含ADD COLUMN、DROP COLUMN、ALTER COLUMN、RENAME COLUMN、ADD CONSTRAINT、DROP CONSTRAINT、RENAME）、CREATE SEQUENCE、DROP SEQUENCE、ALTER SEQUENCE、CREATE INDEX、ALTER INDEX、DROP INDEX、CREATE VIEW、ALTER VIEW、COMMENT ON COLUMN、COMMENT ON TABLE、COMMENT ON SCHEMA、COMMENT ON SEQUENCE、COMMENT ON INDEX、COMMENT ON VIEW。

类型名称	使用须知
	<p>表级同步时，DDL仅支持TRUNCATE（仅PostgreSQL 11及以上版本支持）、DROP TABLE、ALTER TABLE、COMMENT ON COLUMN、COMMENT ON TABLE，其中ALTER TABLE支持：ADD COLUMN、DROP COLUMN、ALTER COLUMN、RENAME COLUMN、ADD CONSTRAINT、DROP CONSTRAINT、RENAME。</p> <ul style="list-style-type: none"> - 不支持：无日志表（UNLOGGED TABLE）的DML，临时表的DML。 <p>说明 同步DDL的实现原理：在源库通过事件触发器捕获DDL语句，并且记录在特定的表中，因此需要提前在源库创建事件触发器、函数等。详细操作可参考通过创建事件触发器和函数实现PostgreSQL增量DDL同步。</p>

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 4-105](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-105 数据库账号权限

类型名称	全量同步	全量+增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限。	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，序列的SELECT权限，无主键表的UPDATE、DELETE和TRUNCATE权限，REPLICATION连接权限。 说明 无主键表的UPDATE、DELETE和TRUNCATE权限，仅用于对无主键表短暂加锁，以确保迁移后，目标库中无主键表数据与源库保持一致。

类型名称	全量同步	全量+增量同步
目标数据库连接账号	<ul style="list-style-type: none"> ● 库级同步： <ul style="list-style-type: none"> - 如果目标库不是postgres，需要具有CREATEDB权限。 - 如果目标库是postgres，需要具有postgres库的CONNECT和CREATE权限、对模式public的USAGE和CREATE权限。 ● 表级同步： <ul style="list-style-type: none"> - 如果需要同步库，需要具有CREATEDB权限。 - 如果需要同步模式，需要具有模式所在库的CONNECT权限、模式所在库上的CREATE权限。 - 如果需要同步模式下的对象，需要具有模式所在库的CONNECT权限、对象所在模式的USAGE权限、对象所在模式上的CREATE权限。 <p>说明 目标库版本为PostgreSQL 15时，库级同步及表级同步目标库没有建库的场景下，全量或增量同步的过程中，可能会由于没有模式public的USAGE及CREATE权限而导致同步失败。</p>	

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单增量和单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-106 使用须知

类型名称	使用和操作约定
任务启动前使用须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库的分区表触发器不可以设置为disable。 - 全量同步支持源库备机状态，但需要设置 hot_standby_feedback为on；增量同步不支持源库备机状态。 - 若要做增量同步： 源数据库参数wal_level必须配置为logical； 源数据库中无主键表的replica identity属性必须为full； 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量； 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica identity属性必须为full。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不可以存在同名的触发器。 - 同步对象依赖和关联的对象也须一起同步，否则可能导致同步失败。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标数据库的block_size参数值必须大于或等于源库中的对应参数值。 - 目标数据库和源数据库的lc_monetary参数值一致。 - 若要做增量同步，且同步对象包含外键、触发器或事件触发器，则目标数据库的session_replication_role参数必须设置为replica，同步结束后，此参数需改为原来的值。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标库不可以包含，与待同步对象类型相同且名称相同的对象，包括库、模式、表等。系统库、系统模式、系统表等除外。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为此格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 对于全量+增量和增量任务，启动前请确保源库中未启动长事务，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 对于全量+增量和增量任务，如果预校验存在“内部错误”且在任务未启动时结束，请参考PostgreSQL为源强制结束任务章节检查并删除流复制槽，防止源库中流复制槽残留。 - 对于全量+增量和增量任务，任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - 若选择同步DDL，须注意源库执行DDL时，确保在目标库上是兼容的。 <p>说明 同步DDL的实现原理是：在源库通过事件触发器捕获DDL语句，并且记录在特定的表中，然后将DDL语句同步至目标库。因此启动任务前，需要提前在源库创建事件触发器、函数等。详细操作可参考通过创建事件触发器和函数实现PostgreSQL增量DDL同步。</p>
全量同步过程使用须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程使用须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源数据库表的主键或者唯一键（主键不存在时），否则可能导致增量数据不一致或任务失败。 ● 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 ● 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。
同步对比使用须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务使用须知	<ul style="list-style-type: none"> ● 正常结束任务： <ul style="list-style-type: none"> - 将自动重置目标库序列值（自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000），若选择了同步用户，任务结束时将自动同步用户的成员关系。 - 全量实时同步任务结束时会对序列值进行比对，序列值采用逻辑一致性的判断原则，即自增序列当且仅当目标库对应序列值大于等于源库对应序列值，自减序列当且仅当目标库对应序列值小于等于源库对应序列值时，序列值判定为一致。 - 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 - 全量+增量实时同步任务正常结束时，如果目标库 session_replication_role 值为 replica，需手动将目标库 session_replication_role 参数值修改为原来的值。 ● 强制结束任务： <ul style="list-style-type: none"> - 需要手动刷新目标库中的序列值，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 - 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 - 强制结束全量+增量实时同步任务，如果目标库 session_replication_role 值为 replica，需手动将目标库 session_replication_role 参数值修改为原来的值。 - 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。
常见故障排查	<ul style="list-style-type: none"> ● 在任务创建、启动、全量同步、增量同步、结束等过程中，如有遇到问题，可先参考“故障排查”章节进行排查。

操作步骤

本小节以PostgreSQL->PostgreSQL的出云同步为示例，介绍如何使用数据复制服务配置VPC网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-102 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 4-107 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-103 同步实例信息

同步实例信息 ?

以下配置项只读不可修改，请谨慎操作，以免因配置错误，造成数据迁移失败。

数据流动方向: 出云

源数据库引擎: MySQL, ODM, GaussDB 分布式版, GaussDB 主备版, DDS, PostgreSQL, GaussDB(for MySQL)

目标数据库引擎: Kafka, PostgreSQL

网络类型: VPC网络

网络策略: 无网络策略

同步实例名字段: 默认名称

同步模式: 全量 + 增量

表 4-108 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“PostgreSQL”。
目标数据库引擎	选择“PostgreSQL”。

参数	描述
网络类型	<p>此处以“VPC网络”为示例，目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	源数据库为RDS for PostgreSQL数据库实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 4-104 任务类型



表 4-109 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-105 企业项目和标签



表 4-110 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- 源库信息配置

图 4-106 源库信息

源库信息

数据库实例名称: pg-12-for-autotest ()

数据库用户名:

数据库密码:

✔ 测试成功

表 4-111 源库信息

参数	描述
数据库实例名称	默认为创建任务时选择的RDS for PostgreSQL数据库实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

- 目标库信息配置

图 4-107 目标库信息

目标库信息

VPC: [查看虚拟私有云](#)

子网: [查看子网](#)

IP地址或域名:

端口:

数据库用户名:

数据库密码:

SSL安全连接:

✔ 测试成功

表 4-112 目标库信息

参数	描述
VPC	目标数据库所在的虚拟专用网络，可以对不同业务进行网络隔离。您需要创建或选择所需的虚拟私有云。

参数	描述
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用分区内才会有效，创建目标数据库实例的子网需要开启DHCP功能，在创建过程中也不能关闭已选子网的DHCP功能。
IP地址或域名	目标数据库的IP地址或域名。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 4-108 同步模式

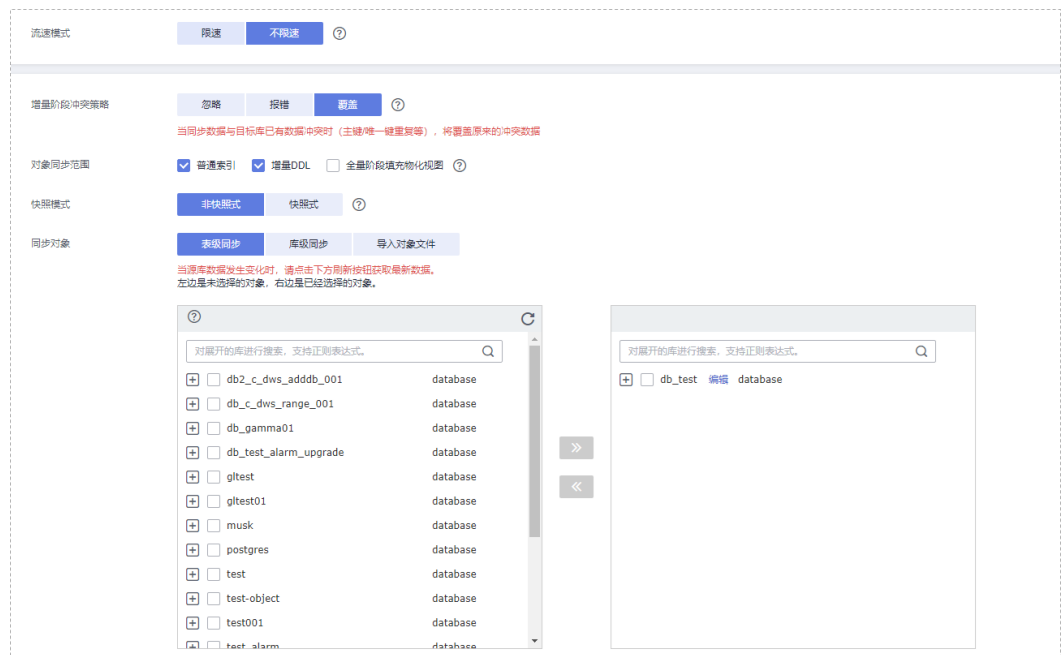



表 4-113 同步对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> <p>限速</p> <p>自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。</p> <p>当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。</p> <p>流速的大小需要根据业务场景来设置，不能超过9999MB/s。</p> <p>图 4-109 设置流速模式</p>  <p>不限速</p> <p>对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。</p> <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
对象同步范围	<p>对象同步范围支持选择普通索引、增量DDL同步和全量阶段填充物化视图，您可以根据业务需求选择是否进行同步。</p> <p>全量阶段填充物化视图：仅对源库已经填充的物化视图生效，该填充操作会影响全量同步性能，建议在全量同步完成后，手动在目标库填充。</p>
快照模式	<p>如果您选择的是全量+增量同步模式的任务，数据复制服务支持设置快照模式。</p> <p>全量阶段使用快照模式导出能够有效提升全量+增量场景下的数据同步效率，但PostgreSQL的快照机制会使导出期间数据库的历史数据不能被回收，可能有空间膨胀的现象。建议在全量或增量数据量大且源库磁盘空间充足的情况下使用该方式。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 选择库级同步：全量同步时，同步对象为所选的库、库中对象的存量数据；增量同步时，同步除了无日志表、临时表外的所有表的DML及部分DDL。 ● 选择表级同步：全量同步时，同步对象为所选的表、序列、视图或物化视图的存量数据；增量同步时，同步所选表的DML及部分DDL。 ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为此格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，同步后表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。
预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。
- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-110 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 4-111 任务启动设置



表 4-114 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.12 将 PostgreSQL 同步到 Kafka

支持的源和目标数据库

表 4-115 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> RDS for PostgreSQL 9.5、9.6、10、11、12、13、14版本 	<ul style="list-style-type: none"> Kafka 0.11及以上版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考表4-116。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-116 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> 不支持实例级同步：每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 支持的字段类型：数字类型、货币类型、字符类型、二进制数据类型、日期/时间类型、布尔类型、枚举类型、几何类型、网络地址类型、位串类型、文本搜索类型、UUID类型、XML类型、JSON类型、数组、复合类型、范围类型。 增量同步的范围： <ul style="list-style-type: none"> 支持部分DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 不支持：DDL（Data Definition Language），无日志表（UNLOGGED TABLE）的DML，临时表的DML。 <p>注意 同步对象的名称有如下限制： 库名不可以包含+"%\<>，模式名和表名不可以包含"!<>，列名不可以包含"和'。</p>

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考表4-117进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-117 数据库账号权限

类型名称	增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，REPLICATION连接权限。

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-118 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库要求： 源数据库参数wal_level必须配置为logical； 源数据库需提前安装test_decoding插件； 源数据库中无主键表的replica identity属性必须为full； 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量； 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica identity属性必须为full。 ● 目标数据库要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 源库必须为华为云RDS for PostgreSQL实例。 - 库级同步时，增量同步中的任务，不支持“编辑”操作。 - 表级同步时，增量同步中的任务，支持“编辑”操作，但是不支持更换数据库。 - 在启动任务前，请确保源库中未启动长事务，源库启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 在任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿删除源数据库表的主键，否则可能导致增量数据缺失或任务失败。 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据缺失或任务失败。 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致增量数据缺失或任务失败。 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。
结束任务须知	<ul style="list-style-type: none"> 正常结束任务： <ul style="list-style-type: none"> 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 强制结束任务： <ul style="list-style-type: none"> 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-112 同步任务信息

表 4-119 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-113 同步实例信息

表 4-120 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“PostgreSQL”。
目标数据库引擎	选择“Kafka”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的RDS for PostgreSQL实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<ul style="list-style-type: none"> - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 <p>无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 4-114 任务类型



表 4-121 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-115 企业项目和标签



表 4-122 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 4-116 源库信息



表 4-123 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS for PostgreSQL实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-117 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

连接方式

✔ 测试成功

表 4-124 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
连接方式	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 4-118 同步模式

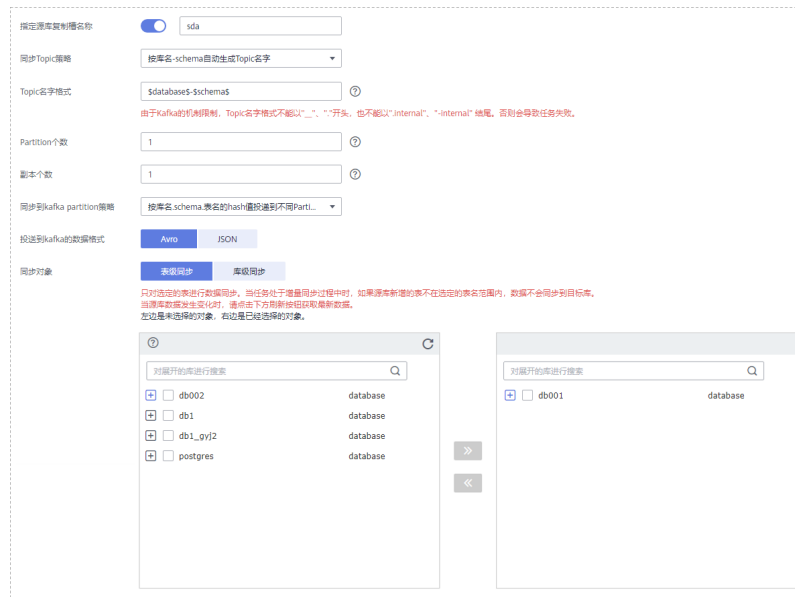



表 4-125 同步对象

参数	描述
指定源库复制槽名称	可选择是否指定源库复制槽，打开后需手动输入复制槽的名称（只能包含小写字母、数字、下划线，长度最大为63，且不能以数字开头）。
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。
Topic名字格式	同步Topic策略选择自动生成Topic名字时可见。 由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"- "，若超过该范围会导致创建Topic失败，任务异常。 若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。
Partition个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。
副本个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。

参数	描述
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名.schema.表名的hash值投递到不同Partition: 适用于单表的查询场景, 表内保序, 表与表之间不保序, 可以提高单表读写性能, 推荐使用此选项。 按主键的hash值投递到不同Partition: 适用于一个表一个Topic的场景, 避免该表都写到同一个分区, 消费者可以并行从各分区获取数据。 对于无主键表, 如果选择“按主键的hash值投递都不同的partition”策略时, 同步任务则会使用默认“按库名.schema.表名的hash值投递到不同的partition”的策略进行同步。 按库名.schema的hash值投递到不同Partition: 适用于一个database一个topic的场景, 避免多个schema下的数据写到一个分区, 消费者可以并行从各分区获取数据。 全部投递到Partition 0: 适用于有事务要求的场景, 事务保序, 可以保证完全按照事务顺序消费, 写入性能比较差, 如果没有强事务要求, 不推荐使用此选项。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro: 可以显示Avro二进制编码, 高效获取数据。 JSON: 为Json消息格式, 方便解释格式, 但需要占用更多的空间。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 同步对象支持表级同步和库级同步, 您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面, 进行同步任务预校验, 校验是否可进行实时同步。

- 查看检查结果, 如有不通过的检查项, 需要修复不通过项后, 单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后, 且所有检查项结果均通过时, 单击“下一步”。

图 4-119 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。

图 4-120 任务启动设置




表 4-126 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.13 将 GaussDB 分布式版同步到 MySQL

支持的源和目标数据库

表 4-127 支持的数据库

源数据库	目标数据库
GaussDB分布式版	<ul style="list-style-type: none">• RDS for MySQL 5.6、5.7版本• 本地自建MySQL数据库 5.5、5.6、5.7版本• ECS自建MySQL数据库 5.5、5.6、5.7版本• 其他云上MySQL数据库 5.5、5.6、5.7版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考表4-128。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-128 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量、增量、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-129](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 4-129 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户需要具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。
目标数据库连接用户	<p>需要具备以下基本权限： INSERT、DELETE、UPDATE、SELECT、SHOW DATABASES。</p>		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-130 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目前支持源库为UTF8字符集，目标库为UTF8MB4字符集。当源库为UTF8字符集，目标库为UTF8MB3字符集，或者当源库为GBK或GB18030字符集，目标库为GBK、GB2312、GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 目标库必须有足够的磁盘空间。 - 同步前保证目标库已创建好对应的database。 - 同步前保证目标库表结构已经建好，且与源端一致或包含源端所有列。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 源数据库和目标数据库表、schema等对象名称大小写要保持一致。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库同步用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持数据对比过程中做限速。 ● 由于MySQL与GaussDB针对零时间 ("0000-00-00 00:00:00") 内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。

类型名称	使用和操作限制
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-121 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。	
项目	<input type="text" value=""/>
* 任务名称	<input type="text" value="DRS-6435"/> ?
描述	<div style="border: 1px solid gray; height: 40px; width: 100%;"></div> <div style="text-align: right; font-size: x-small;">0/256</div> ?

表 4-131 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-122 同步实例信息



表 4-132 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“MySQL”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。 <ul style="list-style-type: none"> - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB分布式版实例。

参数	描述
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-123 任务类型



表 4-133 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更佳。</p>

- 企业项目和标签

图 4-124 企业项目和标签



表 4-134 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 4-125 源库信息

The screenshot shows a form titled '源库信息' (Source Database Information). It contains three input fields: '数据库实例名称' (Database Instance Name), '数据库用户名' (Database Username), and '数据库密码' (Database Password). Below these fields is a button labeled '测试连接' (Test Connection).

表 4-135 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-126 目标库信息

The screenshot shows a form titled '目标库信息' (Target Database Information). It contains four input fields: 'IP地址或域名' (IP Address or Domain Name), '端口' (Port), '数据库用户名' (Database Username), and '数据库密码' (Database Password). There is also a toggle switch for 'SSL安全连接' (SSL Secure Connection) and a '测试连接' (Test Connection) button.

表 4-136 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。

参数	描述
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 4-127 同步模式

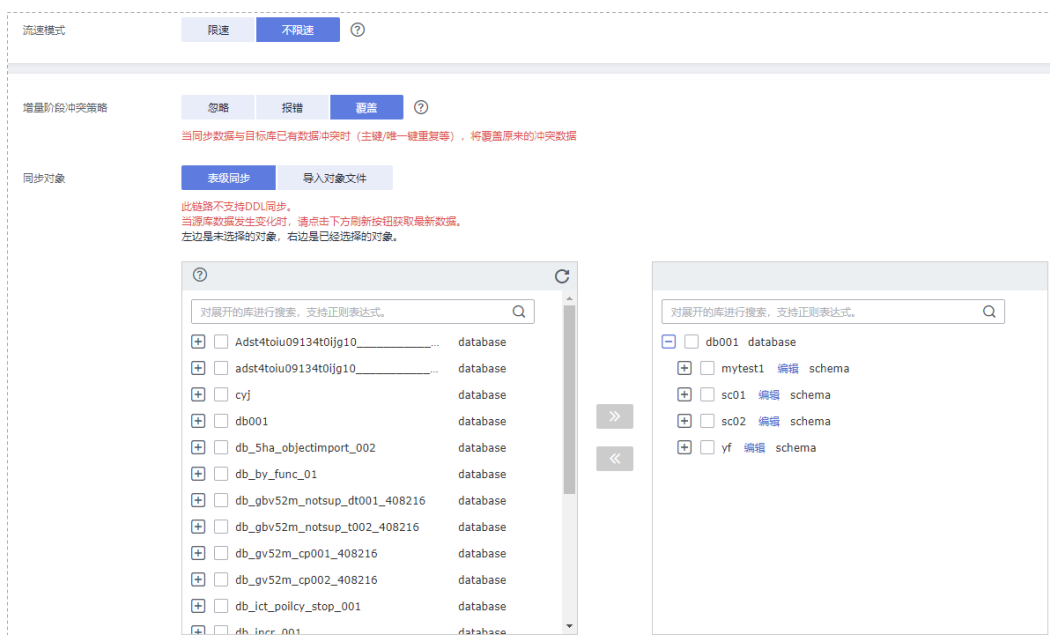



表 4-137 同步对象

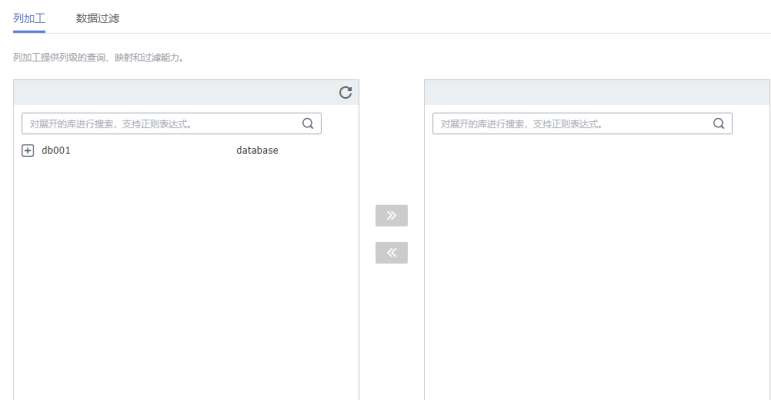
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-128 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 4-129 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-130 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 4-131 任务启动设置



表 4-138 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.14 将 GaussDB 分布式版同步到 Oracle

支持的源和目标数据库

表 4-139 支持的数据库

源数据库	目标数据库
GaussDB分布式	<ul style="list-style-type: none">本地自建Oracle数据库11g、19cECS自建Oracle数据库11g、19c

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-140](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-140 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsith、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\' \"?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-141](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 4-141 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、INSERT 权限，或者具备 RESOURCE角色。</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色。</p>

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量、单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-142 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<>\\'\ ?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目前支持源库为UTF8字符集，目标库为AL32UTF8、UTF8字符集；当源库为GBK或GB18030字符集，目标库为ZHS16GBK、ZHS16CGB231280、ZHS32GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的表结构，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 目标库必须有足够的磁盘空间。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 一般情况下，此链路为Oracle到GaussDB分布式版同步链路的反向逃生链路，并与Oracle到GaussDB分布式版同步链路配合使用，不建议单独使用此链路做其他场景的数据同步。 - 选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。如果此链路作为Oracle到GaussDB分布式的逃生链路使用，建议仅同步Oracle端为大写的schema名和表名，且GaussDB分布式端为小写的schema名和表名。 - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 进行同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 目标数据库Oracle不支持时间类型为1582-10-05至1582-10-14的日期，这种情况会导致同步前后数据不一致。 - 时间类型在源数据库GaussDB公元前闰年为1、5、9等年，在目标数据库Oracle公元前闰年为0、4、8等年，源数据库公元前闰年2月29日在目标数据库是不支持的日期格式，会因为日期解析失败导致同步失败。 - 如果同步表存在可空的联合唯一约束，由于在GaussDB和Oracle对含null值的约束范围差异，会导致最终数据不一致或任务失败。例如，同步表包含可空的联合唯一约束UNIQUE(C1,C2)，在GaussDB可以同时存在多条数据的c1='1'、c2=null，而在Oracle只能存在一条，再插入导致唯一键冲突。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步中过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 含大字段（blob、clob、nclob、long、long raw）的表，大字段不会作为增量数据where条件，建议保证除大字段外的其他字段具有唯一约束的能力。 ● 目标库的表中同时包含lob和long或者同时包含lob和long raw，插入或更新long或long raw的长度大于4000时，会出现插入或更新失败，建议使用clob或blob类型。

类型名称	使用和操作限制
数据加工须知	<ul style="list-style-type: none"> 列加工时，主键、唯一键和分布列不可以被过滤。
同步对比使用须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-132 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 4-143 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-133 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置不正确，需要重新创建迁移任务。

* 数据流动方向: 入库 | 出云 | 双向 | 新建 | ?

DRS要求源数据库或目标数据库中至少有一方为本云数据库实例，出云实例数据库为本云数据库实例的克隆。

* 源数据库引擎: MySQL | DDM | GaussDB for MySQL 分布式版 | GaussDB 分布式版 | GaussDB 主备版 | DDS | PostgreSQL | GaussDB for MySQL

* 目标数据库引擎: MySQL | Oracle | GaussDB(DWS) | GaussDB 分布式版 | Kafka

* 网络类型: 公网网络 | ?

DRS将会自动为DRS实例绑定弹性公网IP，绑定到该实例所属的弹性公网IP。

* 源数据库实例: [下拉菜单] | 新建数据库实例 | 查看不可用实例

同步实例所在子网: [下拉菜单] | ? | 查看子网

* 同步模式: 全量 + 增量 | 全量 | 增量 | ?

该模式为数据库增量实时同步，通过全量过程在目标数据库库的初始化后，增量同步和通过解析日志等技术，将源和目标数据库数据保持一致。

表 4-144 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“Oracle”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 4-134 任务类型



表 4-145 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-135 企业项目和标签



表 4-146 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-136 源库信息

The screenshot shows a form titled "源库信息" (Source Database Information). It contains three input fields: "数据库实例名称" (Database Instance Name), "数据库用户名" (Database Username), and "数据库密码" (Database Password). Below these fields is a button labeled "测试连接" (Test Connection).

表 4-147 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

📖 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-137 目标库信息

目标库信息

IP地址或域名

对于RAC集群, 建议使用SCANIP+SERVICENAME方式创建任务, 提高访问性能。

端口

数据库服务名 Service Name ?

数据库用户名

数据库密码

SSL安全连接

测试成功

表 4-148 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群, 建议使用SCANIP+SERVICENAME方式创建任务, 提高访问性能。
端口	目标数据库服务端口, 可输入范围为1~65535间的整数。
数据库服务名	数据库服务名 (Service Name/SID), 客户端可以通过其连接到Oracle, 具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能, 用户可以选择是否开启对迁移链路的加密。如果开启该功能, 需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接, 请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上, 待该任务删除后会永久清除。

步骤4 在“设置同步”页面, 选择同步对象, 完成后单击“下一步”。

图 4-138 同步模式

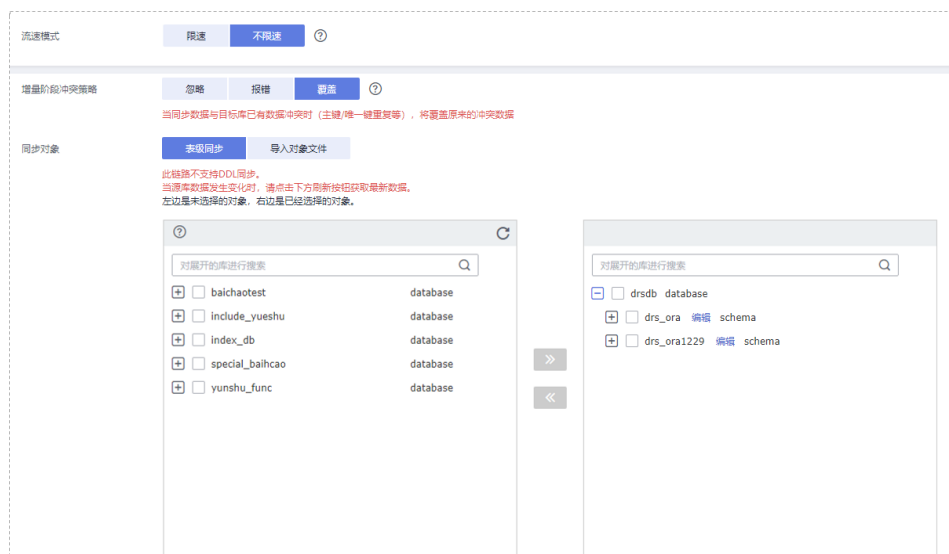




表 4-149 同步对象

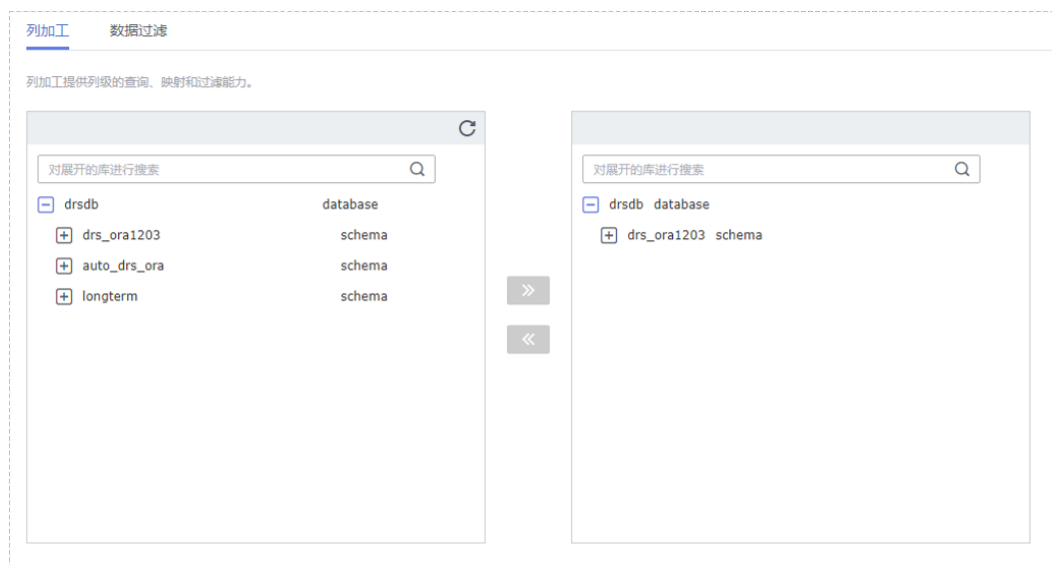
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-139 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 4-140 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-141 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-142 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Start Immediately) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on, with a help icon.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off, with a help icon.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: A text input field contains '14', with a help icon and a red note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (Task will automatically end after a period of abnormal state. Unit is days).

表 4-150 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时会有一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.15 将 GaussDB 分布式版同步到 GaussDB(DWS)

支持的源和目标数据库

表 4-151 支持的数据库

源数据库	目标数据库
GaussDB分布式	GaussDB(DWS)集群

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-152](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-152 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持表数据、表结构和索引约束的同步。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表、二级分区表，增量不建议同步非日志表。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmlDOM、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-153](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因密码填写错误在连接DN过程中导致用户被锁定。

表 4-153 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或被授予以下权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE（仅无主键表需要）权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。
目标数据库连接用户	<p>常需要具备sysadmin角色或以下基本权限：</p> <ul style="list-style-type: none"> • DATABASE上的权限：CONNECT、CREATE。 • SCHEMA上的权限：USAGE。 • 表上的权限：INSERT、DELETE、UPDATE、SELECT。 		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量、单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-154 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<>\\`\"?和!。 - 源数据库分布键中不能含有“，”（逗号+空格）字符。 - 源库表的主键及唯一约束必须包含分布列，否则无法成功同步表结构，需要用户手动在目标库创建表结构后重试任务。 - 不建议在数据库中使用非精确数值类型做主键，该特性影响DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 单独增量同步的任务，必须先为目标库中手动创建SCHEMA和表。 - 单独增量同步的任务，目标库中的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。
<p>结束任务须知</p>	<p>任务正常结束</p> <ul style="list-style-type: none"> ● 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> ● 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 ● 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 ● 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-143 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 4-155 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-144 同步实例信息



表 4-156 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“GaussDB(DWS)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-145 任务类型



表 4-157 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-146 企业项目和标签



表 4-158 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-147 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-159 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-148 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接
✔ 测试成功

表 4-160 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 4-149 同步模式

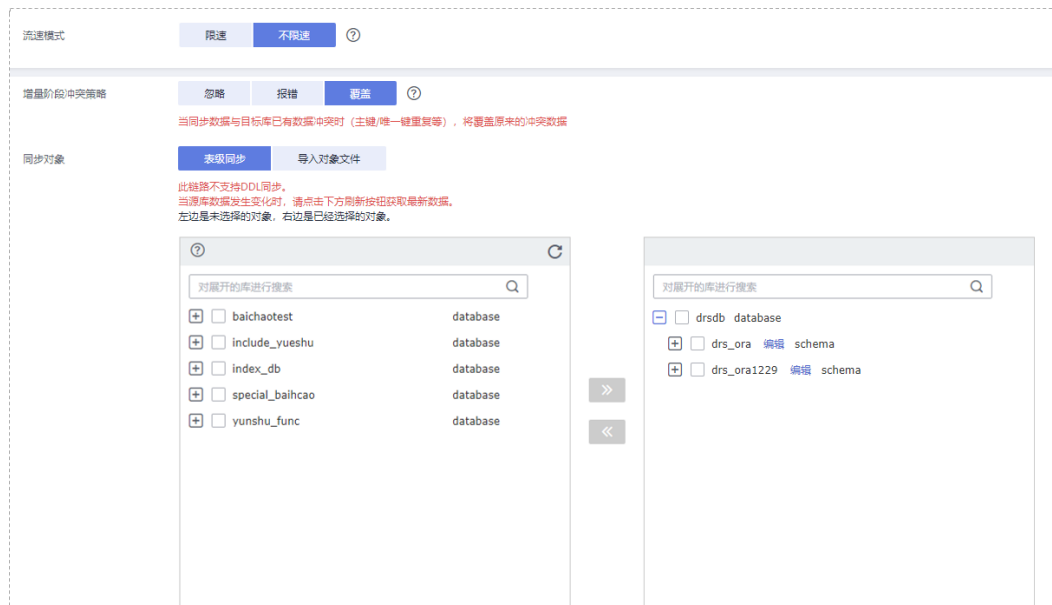



表 4-161 同步对象

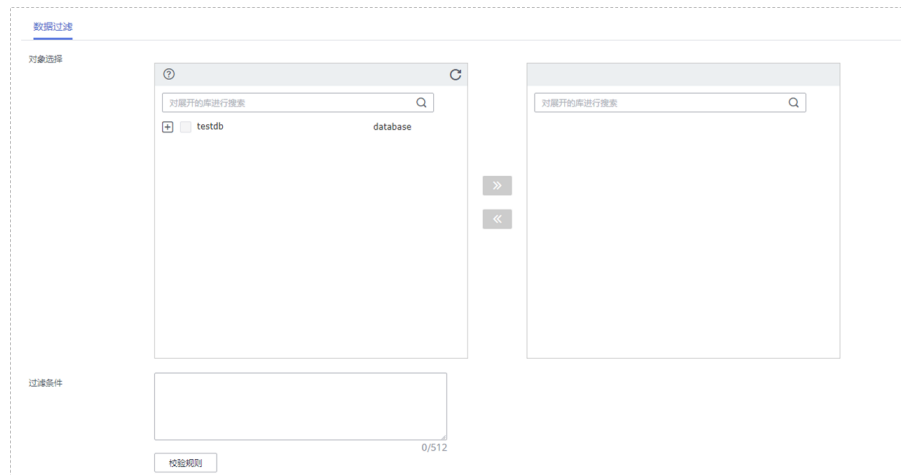
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-150 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 4-151 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-152 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-153 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Start Immediately) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on, with a help icon.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off, with a help icon.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: A text input field contains '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 4-162 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.16 将 GaussDB 分布式版同步到 Kafka

支持的源和目标数据库

表 4-163 支持的数据库

源数据库	目标数据库
GaussDB分布式	Kafka 0.11及以上版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-164](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-164 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持增量同步。 ● 支持的字段：BIGINT, BIT, BIT VARYING, BLOB, BOOLEAN, BYTEA, CHARACTER, CHARACTER VARYING, CLOB, DATE, DOUBLE PRECISION, INTEGER, MONEY, NUMBER, NUMERIC, NVARCHAR2, RAW, REAL, SMALLDATETIME, SMALLINT, TEXT, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, TINYINT。 ● 支持表级同步、schema级同步、库级同步： <ul style="list-style-type: none"> - 仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\\"'\?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考表 4-165进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 4-165 数据库用户权限

类型名称	增量同步
源数据库连接用户	用户具备sysadmin角色或以下最小权限： <ul style="list-style-type: none">• 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。• 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。• 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-166 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 选择增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 将无主键表的REPLICA IDENTITY属性设置为FULL，或给无主键表追加主键。 - 建议将主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 选择表级对象迁移时，同步过程中不建议对表进行重命名操作。 ● 不支持interval partition表复制。 ● 不支持修改主键列名称。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-154 同步任务信息

⚠ 1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text"/>	?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。		
项目	<input type="text"/>	
* 任务名称	<input type="text" value="DRS-6435"/>	?
描述	<div style="border: 1px solid gray; height: 40px; width: 100%;"></div>	?
		0/256

表 4-167 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-155 同步实例信息



表 4-168 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“Kafka”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。 <ul style="list-style-type: none"> 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB分布式版实例。

参数	描述
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>- 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。</p> <p>无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 4-156 任务类型



表 4-169 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-157 企业项目和标签



表 4-170 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 4-158 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-171 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-159 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

连接方式 PLAINTEXT ▼

测试连接
✔ 测试成功

表 4-172 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
连接方式	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 4-160 同步模式

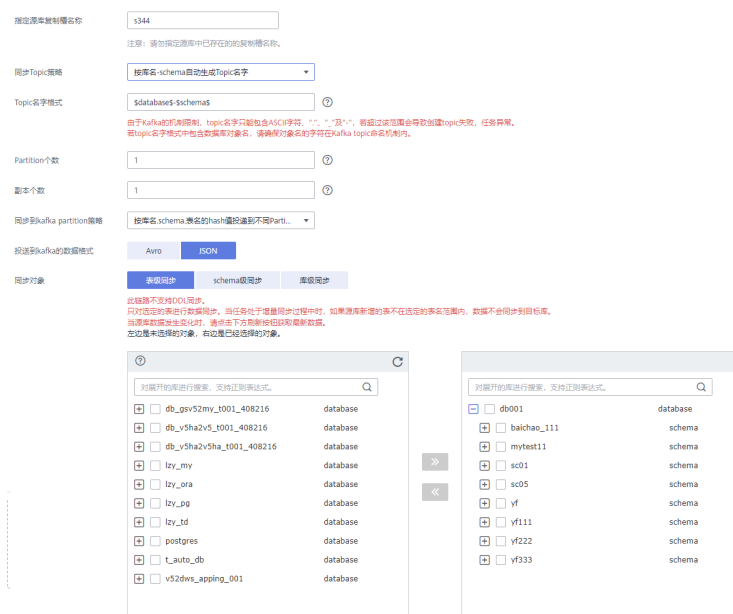



表 4-173 同步对象

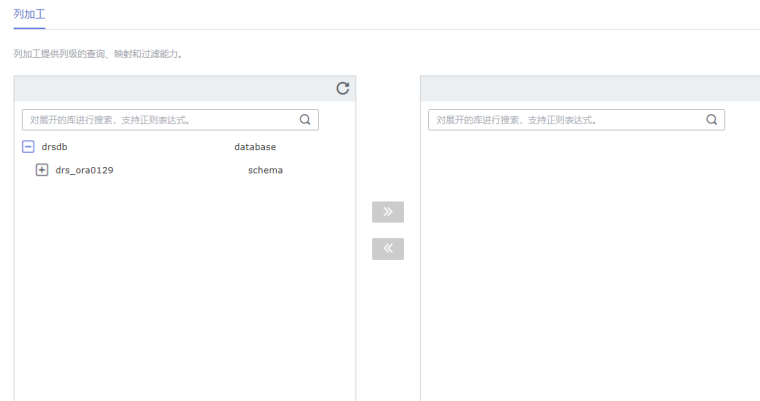
参数	描述
指定源库复制槽名称	可选择是否指定源库复制槽，打开后需手动输入复制槽的名称（只能包含小写字母、数字、下划线，长度63，且不能以数字开头）。
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。
Topic名字格式	同步Topic策略选择自动生成Topic名字时可见。 由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"- "，若超过该范围会导致创建Topic失败，任务异常。 若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。
Partition个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。
副本个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。

参数	描述
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名.schema.表名的hash值投递到不同Partition: 适用于单表的查询场景, 表内保序, 表与表之间不保序, 可以提高单表读写性能, 推荐使用此选项。 按主键的hash值投递到不同Partition: 适用于一个表一个Topic的场景, 避免该表都写到同一个分区, 消费者可以并行从各分区获取数据。 对于无主键表, 如果选择“按主键的hash值投递都不同的partition”策略时, 同步任务则会使用默认“按库名.schema.表名的hash值投递到不同的partition”的策略进行同步。 按库名.schema的hash值投递到不同Partition: 适用于一个database一个topic的场景, 避免多个schema下的数据写到一个分区, 消费者可以并行从各分区获取数据。 全部投递到Partition 0: 适用于有事务要求的场景, 事务保序, 可以保证完全按照事务顺序消费, 写入性能比较差, 如果没有强事务要求, 不推荐使用此选项。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro: 可以显示Avro二进制编码, 高效获取数据。 JSON: 为Json消息格式, 方便解释格式, 但需要占用更多的空间。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 同步对象支持表级同步、schema级同步和库级同步, 您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面, 选择需要加工的数据, 进行列加工。

- 如果不需要数据加工, 单击“下一步”。
- 如果需要进行列加工, 参考[数据加工](#)中的说明, 设置相关规则。

图 4-161 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-162 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-163 任务启动设置

启动时间 立即启动 稍后启动 ?

任务异常通知设置 ?

* SMN主题 ?

时延阈值(s) ?

数据异常通知


* 任务异常自动结束时间 ? 任务处于异常状态一段时间后，将会自动结束。单位为天。

表 4-174 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.17 将 GaussDB 分布式版同步到 GaussDB 分布式版

支持的源和目标数据库

表 4-175 支持的数据库

源数据库	目标数据库
GaussDB分布式	<p>GaussDB分布式</p> <p>说明 仅支持目标库版本等于或高于源库版本。</p>

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-176](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-176 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-177](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 4-177 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限，如果需要使用内容对比，需要拥有用户的MONADMIN权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限，如果需要使用内容对比，需要拥有用户的MONADMIN权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none">● 需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。● 如果需要使用内容对比，需要拥有用户的MONADMIN权限。● 目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-178 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键, 或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 源库要同步的库名、schema名和表名不可以包含/<.>\\` ?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步, 同步前保证目标库表结构已经建好, 且与源端一致; 如果进行了列加工, 请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步, 如果目标库已经存在表结构, 需要保证与源端一致; 如果进行了列加工, 请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景, 但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值, 源端insert数据, 同步到目标端后多的列为null, 不符合目标端要求。 目标端多的列设置固定默认值, 且有唯一约束。源端insert多条数据后, 同步到目标端后多的列为固定默认值, 不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步, 例如从实例A同步到实例B, 再从实例B同步到实例C。 - 实时同步过程中, 不保证分布式事务的一致性。 - 表名映射时, 同步后表上的原索引名称变化为如下格式, 其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到: i_+哈希值+原索引名(可能被截断)+_key。 - 表名映射时, 同步后表上的原约束名称变化为如下格式, 其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索

类型名称	使用和操作约定
	<p>引名” 计算得到: c_+哈希值+原约束名 (可能被截断) +_key。</p> <ul style="list-style-type: none"> - 由于GaussDB分布式版仅支持范围分区表, 所以, 除范围分区外的其他类型分区表作为普通表同步到目标库。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题, 需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束, 同一实例上, 最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 勿修改源库和目标库的端口号, 请勿修改、删除源库和目标库连接用户的密码、权限, 否则可能导致任务失败。 ● 请勿在源库执行任何DDL, 受限于GaussDB逻辑复制功能, 不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作, 否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号, 请勿修改、删除源库和目标库连接用户的密码、权限, 否则可能导致任务失败。 ● 在启动任务后, 进入增量同步阶段前, 请确保在源库中长事务未启动, 启动长事务会阻塞逻辑复制槽的创建, 进而引发任务失败。 ● 请勿在源库执行任何DDL, 受限于GaussDB逻辑复制功能, 不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值, 否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作, 否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后, 该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力, 存在一些使用约束, 具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>数据加工须知</p>	<ul style="list-style-type: none"> ● 列加工时, 主键、唯一键和分布列不可以被过滤。 ● 表上存在列名映射或列过滤时, 索引定义中的部分索引条件表达式将被忽略, 普通索引中的表达式列将被排除, 包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时, 普通索引和外键中对应的列也将被过滤。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对, 防止误报不一致数据, 以及减少对源库和DRS任务的冲击。 ● 源库和目标库的拓扑结构需保持一致, 否则内容对比不可用。 ● 在增量同步过程中做对比时, 源库若存在写入, 则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-164 同步任务信息

表 4-179 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-165 同步实例信息

表 4-180 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPC网络”或“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 - “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。
源端DN数量	需要和分布式源库DN（Data Node）节点数量保持一致。

● 任务类型

图 4-166 任务类型



表 4-181 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-167 企业项目和标签



表 4-182 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-168 源库信息

The screenshot shows a form titled "源库信息" (Source Database Information). It contains three input fields: "数据库实例名称" (Database Instance Name), "数据库用户名" (Database Username), and "数据库密码" (Database Password). Below these fields is a button labeled "测试连接" (Test Connection).

表 4-183 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

📖 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-169 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接

表 4-184 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 4-170 同步模式

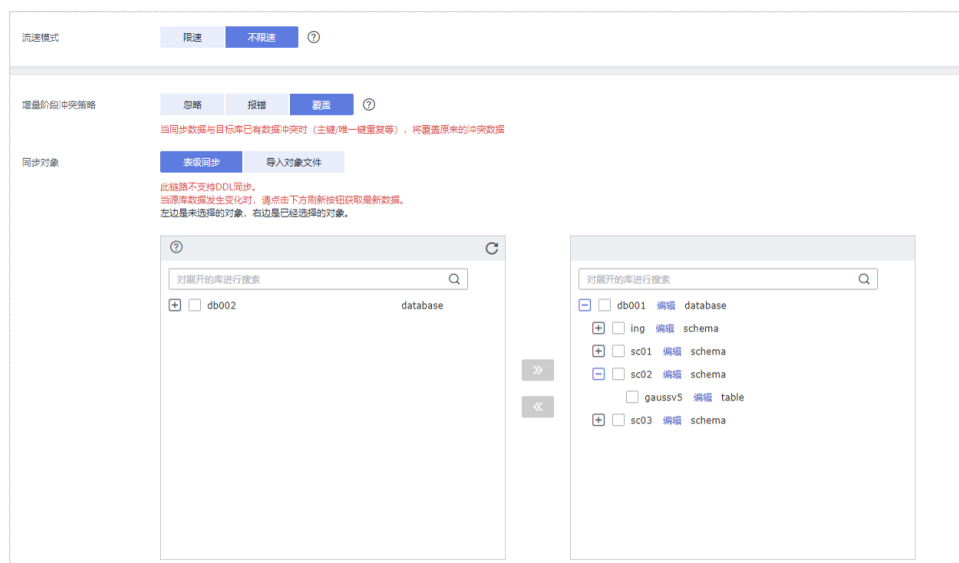



表 4-185 同步对象

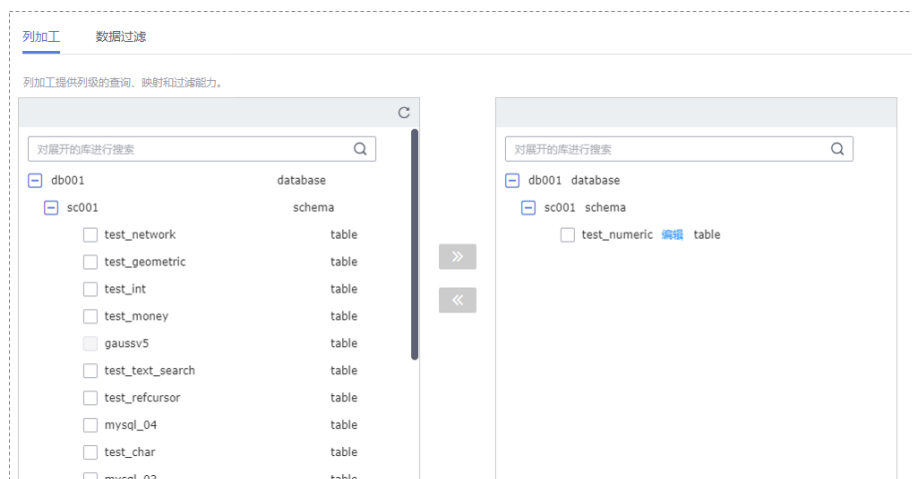
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-171 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和文件导入对象，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要列加工或者数据过滤，可参考[数据加工](#)，设置相关规则。

图 4-172 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-173 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 4-174 任务启动设置

表 4-186 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.18 将 GaussDB 分布式版同步到 GaussDB 主备版

支持的源和目标数据库

表 4-187 支持的数据库

源数据库	目标数据库
GaussDB分布式	GaussDB主备版 说明 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-188](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-188 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\' \"?和!。列名不可以包含\"和.。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-189](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 4-189 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none"> 需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。 目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。 		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-190 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 源库要同步的库名、schema名和表名不可以包含/<.>\\` ?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索

类型名称	使用和操作约定
	<p>引名” 计算得到: c_+哈希值+原约束名 (可能被截断) +_key。</p> <ul style="list-style-type: none"> - 由于GaussDB分布式版仅支持范围分区表, 所以, 除范围分区外的其他类型分区表作为普通表同步到目标库。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题, 需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束, 同一实例上, 最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 勿修改源库和目标库的端口号, 请勿修改、删除源库和目标库连接用户的密码、权限, 否则可能导致任务失败。 ● 请勿在源库执行任何DDL, 受限于GaussDB逻辑复制功能, 不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作, 否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 勿修改源库和目标库的端口号, 请勿修改、删除源库和目标库连接用户的密码、权限, 否则可能导致任务失败。 ● 在启动任务后, 进入增量同步阶段前, 请确保在源库中长事务未启动, 启动长事务会阻塞逻辑复制槽的创建, 进而引发任务失败。 ● 请勿在源库执行任何DDL, 受限于GaussDB逻辑复制功能, 不支持DDL语句的同步, 否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值, 否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作, 否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后, 该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力, 存在一些使用约束, 具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>数据加工须知</p>	<ul style="list-style-type: none"> ● 列加工时, 主键、唯一键和分布列不可以被过滤。 ● 表上存在列名映射或列过滤时, 索引定义中的部分索引条件表达式将被忽略, 普通索引中的表达式列将被排除, 包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时, 普通索引和外键中对应的列也将被过滤。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对, 防止误报不一致数据, 以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时, 源库若存在写入, 则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-175 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 4-191 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-176 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎操作，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向: 入云 | 出云 | 自建 | 自建 ?

DRS要求源数据库和目标数据库中至少有一方为云数据库实例，出云实例数据库为非云数据库实例的场景。

* 源数据库引擎: MySQL | DDM | GaussDB 分布式版 | GaussDB 主备版 | DDS | PostgreSQL | GaussDB(for MySQL)

* 目标数据库引擎: MySQL | Oracle | GaussDB(DW) | GaussDB 分布式版 | GaussDB 主备版 | Kafka

* 网络类型: 公网网络 ?

使用弹性公网IP为同步实例提供弹性公网IP，直到该任务结束会自动删除弹性公网IP。

* 源数据库实例: [下拉菜单] C 查看数据库实例 查看不可达实例

同步实例所在子网: [下拉菜单] ? 查看子网

* 同步模式: 全量 + 增量 | 全量 | 增量 ?

该模式为数据库跨地域实时同步，通过全量过程生成目标数据库的初始化数据，增量同步阶段通过解析日志等技术，将源端和目标数据库的数据保持一致。

表 4-192 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“GaussDB主备版”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB分布式版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 4-177 任务类型



表 4-193 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-178 企业项目和标签



表 4-194 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。 如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-179 源库信息

The screenshot shows a form titled "源库信息" (Source Database Information). It contains three input fields: "数据库实例名称" (Database Instance Name), "数据库用户名" (Database Username), and "数据库密码" (Database Password). Below these fields is a button labeled "测试连接" (Test Connection).

表 4-195 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB分布式版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

📖 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-180 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

✔ 测试成功

表 4-196 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 4-181 同步模式

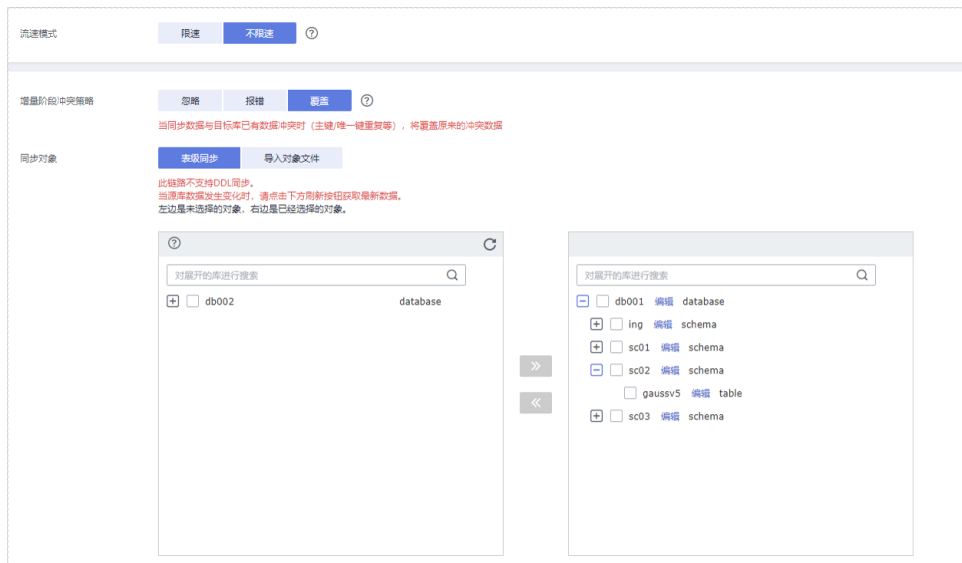




表 4-197 同步对象

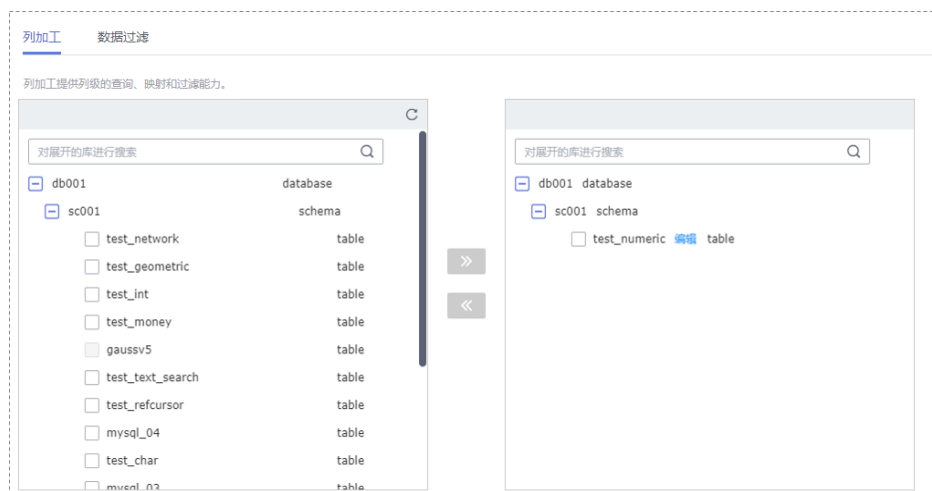
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-182 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，可参考[数据加工](#)，设置相关规则。

图 4-183 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-184 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-185 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-198 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时会有一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.19 将 GaussDB 主备版同步到 MySQL

支持的源和目标数据库

表 4-199 支持的数据库

源数据库	目标数据库
GaussDB主备版	<ul style="list-style-type: none"> • RDS for MySQL 5.6、5.7版本 • 本地自建MySQL数据库 5.5、5.6、5.7版本 • ECS自建MySQL数据库 5.5、5.6、5.7版本 • 其他云上MySQL数据库 5.5、5.6、5.7版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-200](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-200 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-201](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-201 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。
目标数据库连接用户	<p>需要具备以下基本权限： INSERT、DELETE、UPDATE、SELECT、SHOW DATABASES。</p>		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-202 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库必须有足够的磁盘空间。 - 目前支持源库为UTF8字符集，目标库为UTF8MB4字符集。当源库为UTF8字符集，目标库为UTF8MB3字符集，或者当源库为GBK或GB18030字符集，目标库为GBK、GB2312、GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 目标库必须有足够的磁盘空间。 - 同步前保证目标库已创建好对应的database。 - 同步前保证目标库表结构已经建好，且与源端一致或包含源端所有列。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 源数据库和目标数据库表、schema等对象名称大小写要保持一致。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。 - GaussDB中字符类型、二进制类型数据中遇到0x00会截断，与MySQL行为不一致，可能导致最终数据不一致。 - 暂不支持binary类型作为DML匹配条件，即不支持binary类型作为主键。binary为定长类型，会自动填充，GaussDB填充符是0x20，MySQL填充符是0x00。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持数据对比过程中做限速。 ● 由于MySQL与GaussDB针对零时间（"0000-00-00 00:00:00"）内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。

类型名称	使用和操作限制
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-186 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 4-203 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 4-187 同步实例信息



表 4-204 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB主备版实例。

参数	描述
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-188 任务类型



表 4-205 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p>

- 企业项目和标签

图 4-189 企业项目和标签



表 4-206 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 4-190 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-207 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-191 目标库信息

目标库信息

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

表 4-208 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址或域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 4-192 同步模式

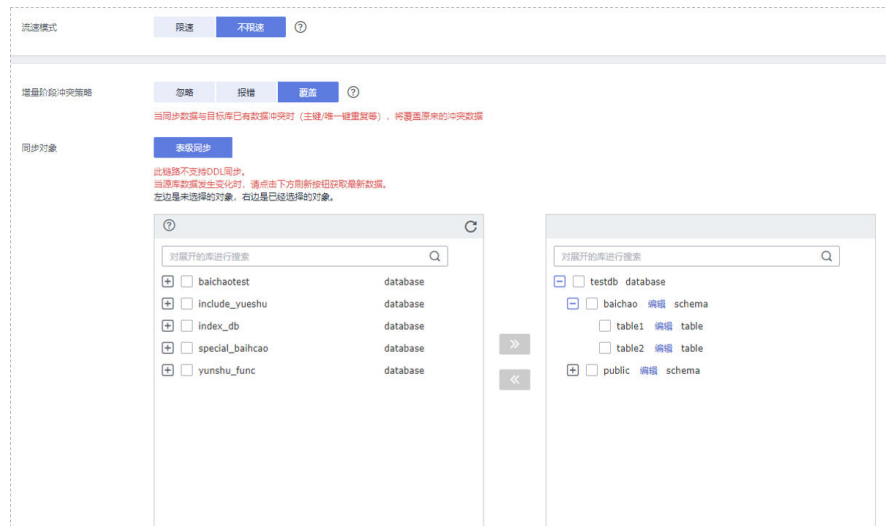



表 4-209 同步对象

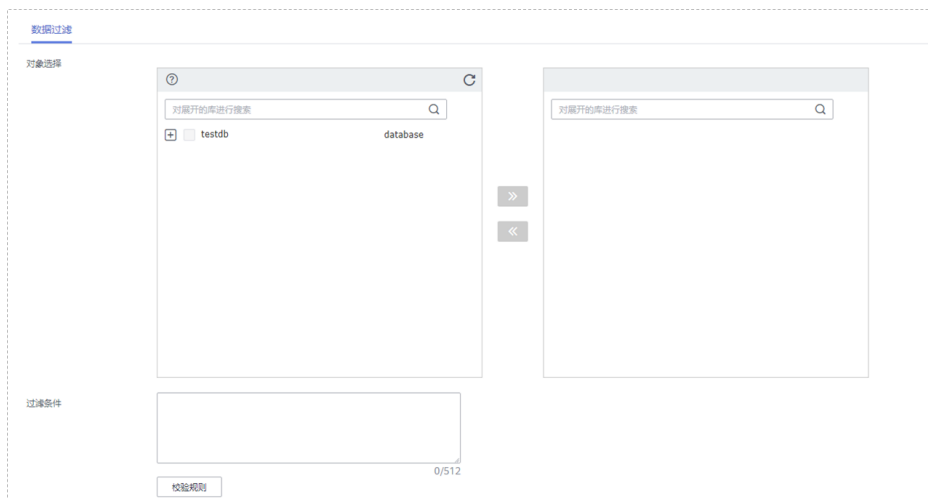
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-193 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。</p> <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 4-194 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-195 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-196 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-210 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.20 将 GaussDB 主备版同步到 Oracle

支持的源和目标数据库

表 4-211 支持的数据库

源数据库	目标数据库
GaussDB主备版	<ul style="list-style-type: none"> • 本地自建Oracle数据库11g、19c • ECS自建Oracle数据库11g、19c

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-212](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-212 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML及部分DDL。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-213](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-213 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、INSERT 权限，或者具备 RESOURCE角色。</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色。</p>

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-214 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<>\\'\ ?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目前支持源库为UTF8字符集，目标库为AL32UTF8、UTF8字符集；当源库为GBK或GB18030字符集，目标库为ZHS16GBK、ZHS16CGB231280、ZHS32GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的表结构，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 目标库必须有足够的磁盘空间。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 一般情况下，此链路为Oracle到GaussDB主备版同步链路的反向逃生链路，并与Oracle到GaussDB主备版同步链路配合使用，不建议单独使用此链路做其他场景的数据同步。 - 选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。如果此链路作为Oracle到GaussDB主备版的逃生链路使用，建议仅同步Oracle端为大写的schema名和表名，且GaussDB主备版端为小写的schema名和表名。 - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 进行同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 目标数据库Oracle不支持时间类型为1582-10-05至1582-10-14的日期，这种情况会导致同步前后数据不一致。 - 时间类型在源数据库GaussDB公元前闰年为1、5、9等年，在目标数据库Oracle公元前闰年为0、4、8等年，源数据库公元前闰年2月29日在目标数据库是不支持的日期格式，会因为日期解析失败导致同步失败。 - 如果同步表存在可空的联合唯一约束，由于在GaussDB和Oracle对含null值的约束范围差异，会导致最终数据不一致或任务失败。例如，同步表包含可空的联合唯一约束UNIQUE(C1,C2)，在GaussDB可以同时存在多条数据的c1='1'、c2=null，而在Oracle只能存在一条，再插入导致唯一键冲突。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。

类型名称	使用和操作限制
增量同步中过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 含大字段（blob、clob、nclob、long、long raw）的表，大字段不会作为增量数据where条件，建议保证除大字段外的其他字段具有唯一约束的能力。 ● 目标库的表中同时包含lob和long或者同时包含lob和long raw，插入或更新long或long raw的长度大于4000时，会出现插入或更新失败，建议使用clob或blob类型。 ● 不支持以下类型的DDL同步：ALTER TABLE MERGE PARTITIONS和CREATE INDEX CONCURRENTLY。 ● 同步DDL语句，需要源和目标库兼容模式一致。 ● 涉及操作表空间的DDL语句，需要目标库迁移用户具备操作表空间的权限。
数据加工须知	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键不可以被过滤。
同步对比使用须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> ● 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> ● 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 ● 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-197 同步任务信息

表 4-215 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-198 同步实例信息

同步实例信息

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建任务。

DRS要求源数据库或目标数据库中至少有一方为多云数据库实例，出云指源数据库为多云数据库实例的场景。

数据流动方向： 入云 出云 自建云库

源数据库引擎： MySQL DDM GaussDB 分布式版 GaussDB 主备版 DDS PostgreSQL GaussDB(for MySQL)

目标数据库引擎： Oracle GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版 Kafka

网络类型：

DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除弹性公网IP。

源数据库实例：

同步实例所在子网：

同步模式： 全量 增量 混合

该模式为数据库持续性实时同步，通过全量过程完成目标数据库的初始化，增量同步和位通过解析日志等技术，将源端和目标数据库数据持续一致。

表 4-216 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“Oracle”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB主备版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-199 任务类型



表 4-217 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-200 企业项目和标签



表 4-218 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 4-201 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-219 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-202 目标库信息



目标库信息

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICENAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功

表 4-220 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCANIP+SERVICENAME方式创建任务，提高访问性能。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。

参数	描述
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 4-203 同步模式

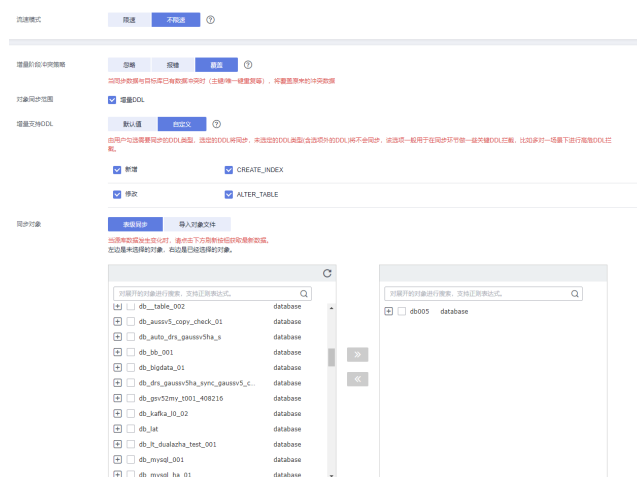



表 4-221 同步对象

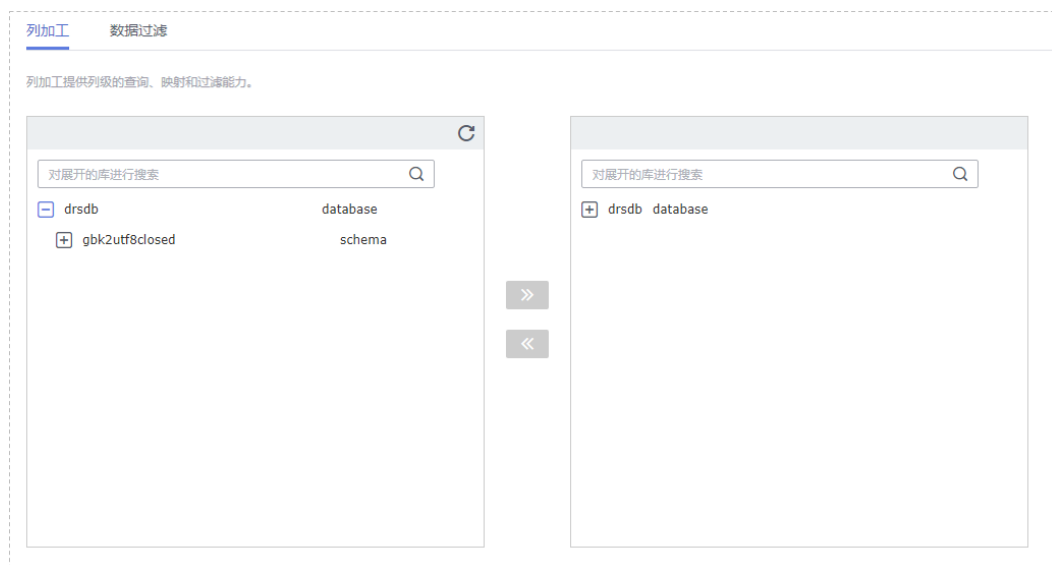
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-204 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
对象同步范围	选择增量阶段是否同步DDL。
增量支持DDL	<p>对象同步范围选择“增量DDL”时可见，用户可根据需求选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> ● “默认值”：已选定同步对象的DDL将会同步至目标，不支持的DDL将不同步。 <ul style="list-style-type: none"> - 表级支持的DDL：ALTER TABLE、CREATE INDEX。 ● “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准，未选定的DDL类型（含选项外的DDL）将不会同步。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 4-205 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-206 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-207 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-222 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.21 将 GaussDB 主备版同步到 GaussDB(DWS)

支持的源和目标数据库

表 4-223 支持的数据库

源数据库	目标数据库
GaussDB主备版	GaussDB(DWS)集群

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-224](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-224 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持表数据、表结构和索引约束的同步。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表，增量不建议同步非日志表。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\''\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表](#)

4-225进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-225 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或被授予以下权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE（仅无主键表需要）权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	<p>通常需要具备 sysadmin 角色或以下基本权限：</p> <ul style="list-style-type: none"> • DATABASE 上的权限：CONNECT、CREATE。 • SCHEMA 上的权限：USAGE、CREATE。 • 表上的权限：INSERT、DELETE、UPDATE、SELECT。 	<p>通常需要具备 sysadmin 角色或以下基本权限：</p> <ul style="list-style-type: none"> • DATABASE 上的权限：CONNECT。 • SCHEMA 上的权限：USAGE。 • 表上的权限：INSERT、DELETE、UPDATE、SELECT。 	<p>通常需要具备 sysadmin 角色或以下基本权限：</p> <ul style="list-style-type: none"> • DATABASE 上的权限：CONNECT、CREATE。 • SCHEMA 上的权限：USAGE、CREATE。 • 表上的权限：INSERT、DELETE、UPDATE、SELECT。

使用建议

注意

- DRS 任务启动和全量数据同步阶段，请不要在源数据库执行 DDL 操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约 50MB/s 的查询压力，以及占用 2~4 个 CPU。
 - 同步无主键表时，为了确保数据一致性，会存在 3s 以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS 并发读取数据库，会占用大约 6-10 个 session 连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多 DRS 对数据库的影响，可参考 [DRS 对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量、单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\\?和!。 - 源数据库分布键中不能含有“，”（逗号+空格）字符。 - 不建议在数据库中使用非精确数值类型做主键，该特性影响DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 单独增量同步的任务，必须先在目标库中手动创建SCHEMA和表。 - 单独增量同步的任务，目标库中的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 模式名或表名映射时，为防止索引和约束名冲突，同步后表上的原索引名称将变为此格式：i_+哈希值+原索引名（可能被截断）+_key。其中哈希值由“原模式名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：c_+哈希值+原约束名（可能被截断）+_key。 - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> • 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 • 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 • 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> • 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 • 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 • 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 • 请勿在目标库做写入操作，否则可能导致数据不一致。 • 不支持interval partition表复制。 • 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 • 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>同步对比须知</p>	<ul style="list-style-type: none"> • 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 • 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 • 不支持在全量同步过程中做对比。 • 不支持数据对比过程中做限速。
<p>结束任务须知</p>	<p>任务正常结束</p> <ul style="list-style-type: none"> • 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> • 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 • 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽名命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽名命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-208 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 4-226 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-209 同步实例信息



表 4-227 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“GaussDB(DWS)”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB主备版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-210 任务类型



表 4-228 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-211 企业项目和标签



表 4-229 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-212 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-230 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 4-213 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

 测试成功

表 4-231 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址或域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 4-214 同步模式

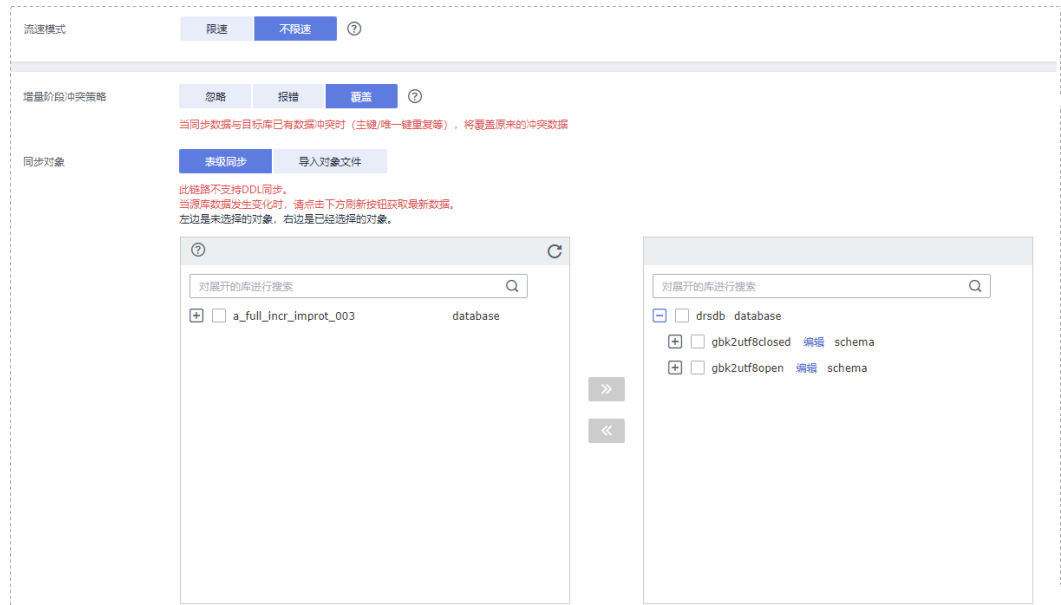



表 4-232 同步对象

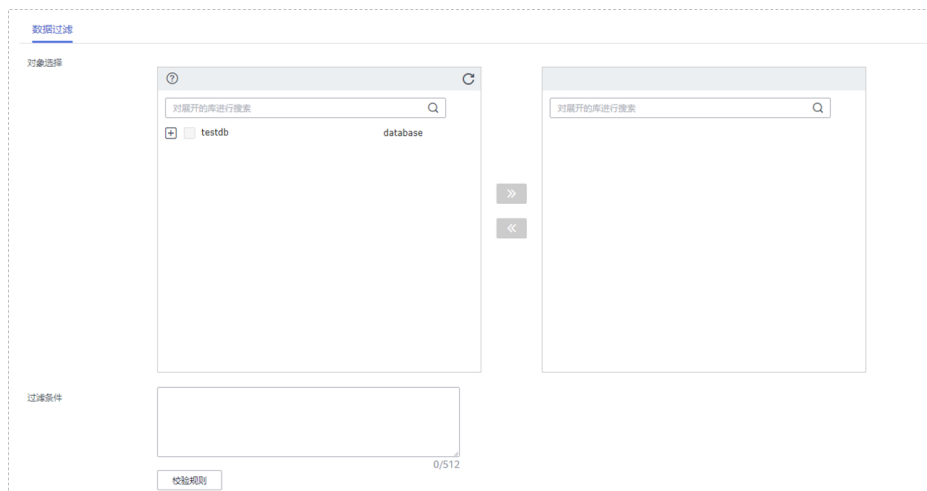
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> <p>限速</p> <p>自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。</p> <p>当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。</p> <p>流速的大小需要根据业务场景来设置，不能超过9999MB/s。</p> <p>图 4-215 设置流速模式</p>  <p>不限速</p> <p>对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。</p> <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>增量阶段冲突策略</p>	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> <p>忽略</p> <p>当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。</p> <p>覆盖</p> <p>当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。</p> <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 4-216 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-217 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。

图 4-218 任务启动设置




表 4-233 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.22 将 GaussDB 主备版同步到 Kafka

支持的源和目标数据库

表 4-234 支持的数据库

源数据库	目标数据库
GaussDB主备版	Kafka 0.11及以上版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考表4-235。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-235 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持增量同步。 ● 支持的字段：BIGINT, BIT, BIT VARYING, BLOB, BOOLEAN, BYTEA, CHARACTER, CHARACTER VARYING, CLOB, DATE, DOUBLE PRECISION, INTEGER, MONEY, NUMBER, NUMERIC, NVARCHAR2, RAW, REAL, SMALLDATETIME, SMALLINT, TEXT, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, TINYINT。 ● 支持表级同步、schema级同步、库级同步： <ul style="list-style-type: none"> - 仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\\"'\?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考表

4-236进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-236 数据库用户权限

类型名称	增量同步
源数据库连接用户	用户具备sysadmin角色或以下最小权限： <ul style="list-style-type: none">• 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。• 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。• 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-237 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 选择增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 将无主键表的REPLICA IDENTITY属性设置为FULL，或给无主键表追加主键。 - 建议将主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\' ?和!。 ● 目标数据库要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 如果DRS实例类型选择主备类型，当DRS任务发生主备切换时，可能会产生无主键表数据重复或不一致的情况。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 选择表级对象迁移时，同步过程中不建议对表进行重命名操作。 不支持interval partition表复制。 不支持修改主键列名称。 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-219 同步任务信息

1. 该页面**仅有任务名称和描述可修改**，其他在点击**开始创建**后均**不可修改**。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 4-238 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-220 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向: 入云 | 出云 | 自建 | 自建 ?

DRS要求源数据库和目标数据库至少有一方为本地数据库实例，出云指源数据库为本地数据库实例的场景。

* 源数据库引擎: MySQL | DDM | GaussDB 分布式版 | GaussDB 主备版 | DDS | PostgreSQL | GaussDB(for MySQL)

* 目标数据库引擎: Oracle | GaussDB(DRS) | GaussDB 分布式版 | GaussDB 主备版 | Kefu

* 网络类型: 公网网络 ?
DRS将会自动为DRS实例创建弹性公网IP，请任意填写公网IP地址。

* DRS实例类型: 单AZ | 双AZ
采用单节点部署架构，与多AZ部署方式相比，每个任务只创建一个节点，也具有高性价比，适用于业务数据量较小、短期同步，并对中断时间无严格要求的场景。

* 源数据库实例: 无可用实例 | 查看数据库实例 | 查看不可用实例

同步实例所在子网: 选择子网 ? | 查看子网 | 查看已占用的IP地址

* 同步模式: 标准 ?
数据同步通过解析日志等技术，将源端产生的增量数据同步至目标端，无需中断业务，实现同步过程中源端和目标端数据的无缝衔接。

表 4-239 同步实例信息

参数	描述
数据流动方向	选择“出云”。

参数	描述
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“Kafka”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
源数据库实例	用户所创建的GaussDB主备版实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<ul style="list-style-type: none"> - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。

- 任务类型

图 4-221 任务类型



表 4-240 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 4-222 可用区</p>

- 企业项目和标签

图 4-223 企业项目和标签



表 4-241 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-224 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-242 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-225 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

连接方式

✔ 测试成功

表 4-243 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
连接方式	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 4-226 同步模式

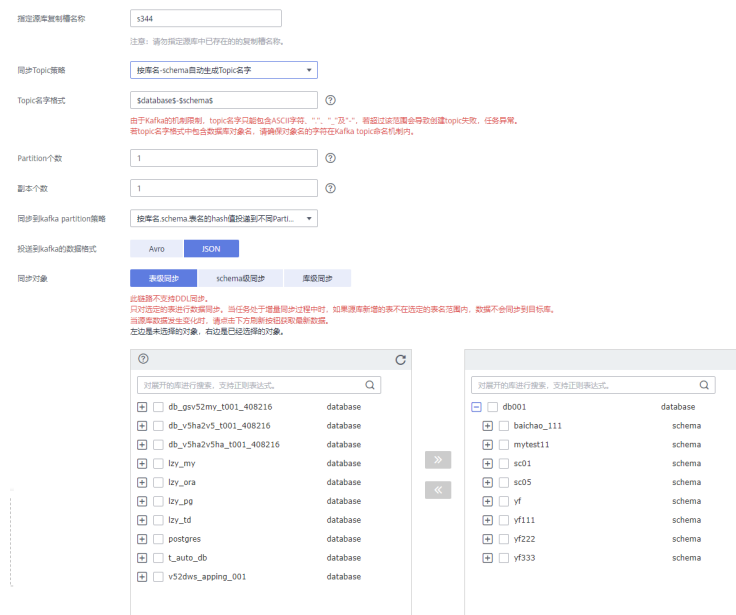



表 4-244 同步对象

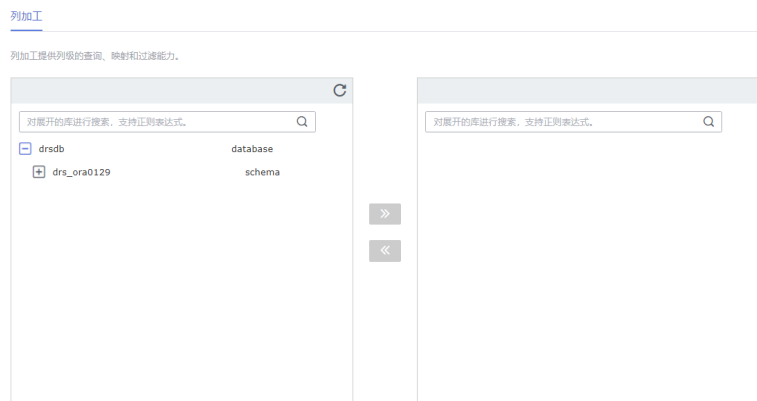
参数	描述
指定源库复制槽名称	可选择是否指定源库复制槽，打开后需手动输入复制槽的名称（只能包含小写字母、数字、下划线，长度63，且不能以数字开头）。
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。
Topic名字格式	同步Topic策略选择自动生成Topic名字时可见。 由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"- "，若超过该范围会导致创建Topic失败，任务异常。 若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。
Partition个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。
副本个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。

参数	描述
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名.schema.表名的hash值投递到不同Partition: 适用于单表的查询场景, 表内保序, 表与表之间不保序, 可以提高单表读写性能, 推荐使用此选项。 按主键的hash值投递到不同Partition: 适用于一个表一个Topic的场景, 避免该表都写到同一个分区, 消费者可以并行从各分区获取数据。 对于无主键表, 如果选择“按主键的hash值投递都不同的partition”策略时, 同步任务则会使用默认“按库名.schema.表名的hash值投递到不同的partition”的策略进行同步。 按库名.schema的hash值投递到不同Partition: 适用于一个database一个topic的场景, 避免多个schema下的数据写到一个分区, 消费者可以并行从各分区获取数据。 全部投递到Partition 0: 适用于有事务要求的场景, 事务保序, 可以保证完全按照事务顺序消费, 写入性能比较差, 如果没有强事务要求, 不推荐使用此选项。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro: 可以显示Avro二进制编码, 高效获取数据。 JSON: 为Json消息格式, 方便解释格式, 但需要占用更多的空间。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 同步对象支持表级同步、schema级同步和库级同步, 您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面, 选择需要加工的数据, 进行列加工。

- 如果不需要数据加工, 单击“下一步”。
- 如果需要进行列加工, 参考[数据加工](#)中的说明, 设置相关规则。

图 4-227 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-228 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-229 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-245 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.23 将 GaussDB 主备版同步到 GaussDB 分布式版

支持的源和目标数据库

表 4-246 支持的数据库

源数据库	目标数据库
GaussDB主备版	GaussDB分布式 说明 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-247](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-247 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 不支持源端表含有表达式唯一约束，否则在写入目标端GaussDB分布式版表结构时会报错，导致任务失败。 - 库名、schema名和表名不可以包含/<.>\\` ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-248](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-248 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none">需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-249 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\\?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：c_+哈希值+原约束名（可能被截断）+_key。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 源数据库的唯一索引是表达式索引时，全量同步阶段目标数据库GaussDB分布式可能会因为无法创建该索引而导致任务失败。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
数据加工须知	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键不可以被过滤。 ● 表上存在列名映射或列过滤时，索引定义中的部分索引条件表达式将被忽略，普通索引中的表达式列将被排除，包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时，普通索引和外键中对应的列也将被过滤。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比，全量完成会自动触发对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-230 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。	
项目	<input type="text" value=""/>
* 任务名称	<input type="text" value="DRS-6435"/> ?
描述	<div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> <div style="text-align: right; font-size: x-small;">0/256</div>

表 4-250 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 4-231 同步实例信息



表 4-251 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“GaussDB分布式版”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB主备版实例。

参数	描述
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPC网络”或“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 4-232 任务类型



表 4-252 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p>

● 企业项目和标签

图 4-233 企业项目和标签



表 4-253 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 4-234 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-254 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-235 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

✔ 测试成功

表 4-255 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 4-236 同步模式

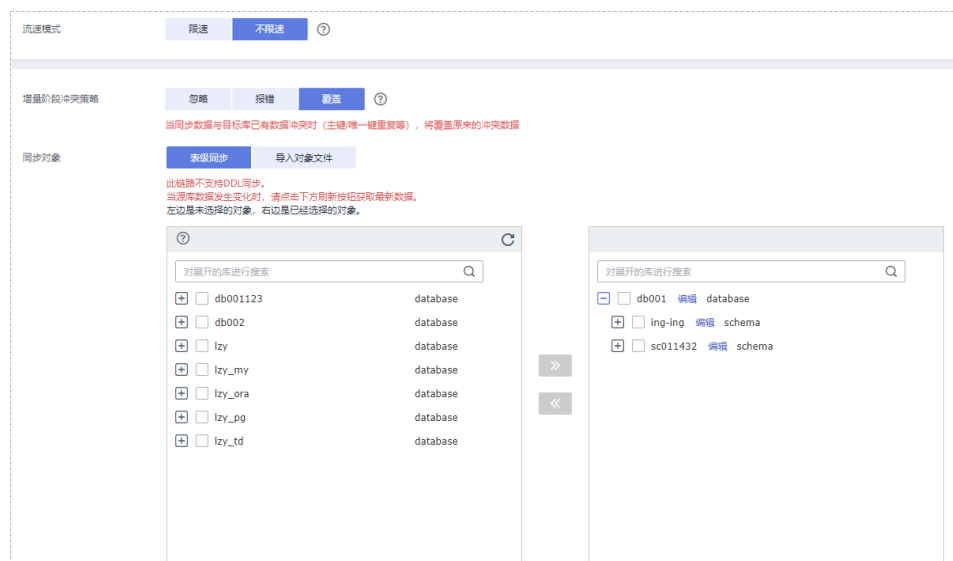




表 4-256 同步对象

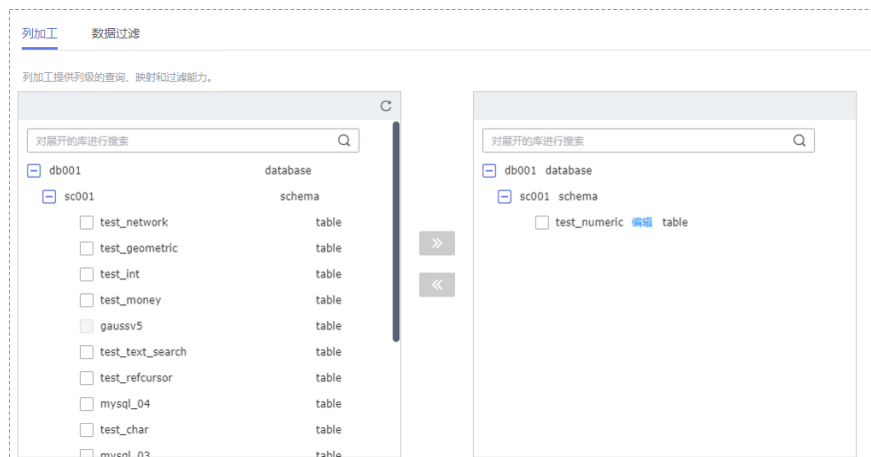
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-237 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和文件导入对象，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或数据过滤，可参考[数据加工](#)，设置相关规则。

图 4-238 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-239 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-240 任务启动设置

启动时间 立即启动 稍后启动 ?

任务异常通知设置 ?

* SMN主题 ?

时延阈值(s) ?

数据异常通知


* 任务异常自动结束时间 ? 任务处于异常状态一段时间后，将会自动结束。单位为天。

表 4-257 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.24 将 GaussDB 主备版同步到 GaussDB 主备版

支持的源和目标数据库

表 4-258 支持的数据库

源数据库	目标数据库
GaussDB主备版	GaussDB主备版 说明 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-259](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-259 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 4-260](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-260 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none">需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-261 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 若选择增量或全量+增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\` ?和!。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：c_+哈希值+原约束名（可能被截断）+_key。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
数据加工须知	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键不可以被过滤。 ● 表上存在列名映射或列过滤时，索引定义中的部分索引条件表达式将被忽略，普通索引中的表达式列将被排除，包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时，普通索引和外键中对应的列也将被过滤。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比，全量完成会自动触发对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-241 同步任务信息

⚠️ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 4-262 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 4-242 同步实例信息



表 4-263 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的GaussDB主备版实例。

参数	描述
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPC网络”或“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 4-243 任务类型



表 4-264 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更佳。</p>

● 企业项目和标签

图 4-244 企业项目和标签



表 4-265 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 4-245 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

表 4-266 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB主备版实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-246 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

✔ 测试成功

表 4-267 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 4-247 同步模式

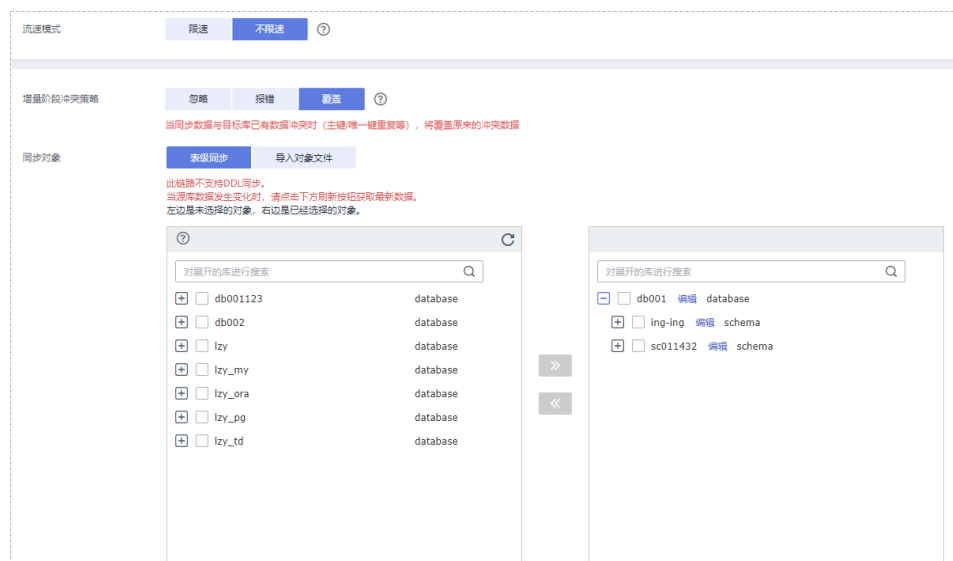




表 4-268 同步对象

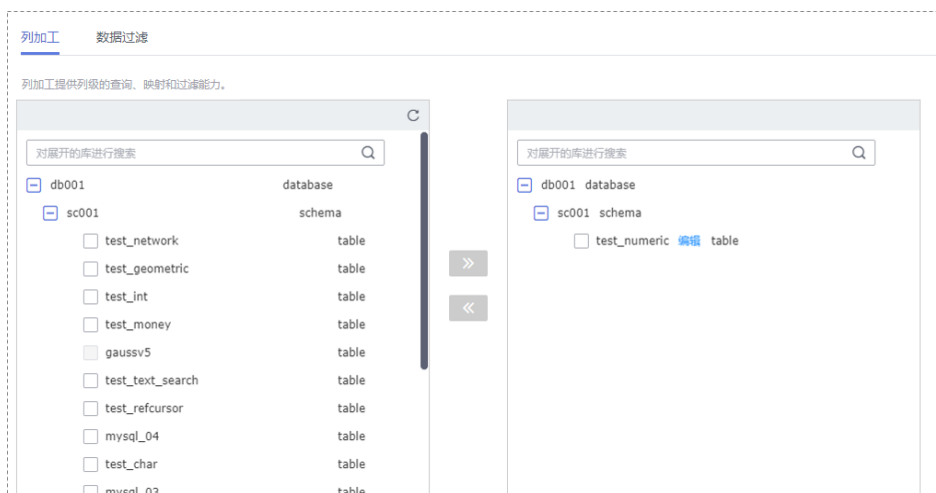
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-248 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和文件导入对象，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或数据过滤，可参考[数据加工](#)中的列加工，设置相关规则。

图 4-249 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-250 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 4-251 任务启动设置

表 4-269 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.25 将 GaussDB(for MySQL)同步到 MySQL

支持的源和目标数据库

表 4-270 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • GaussDB(for MySQL)主备实例 	<ul style="list-style-type: none"> • 本地自建MySQL数据库 • ECS自建MySQL数据库 • 其他云上MySQL数据库 • RDS for MySQL

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-271 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 ● 目标数据库账号必须拥有如下权限：SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES，RDS for MySQL实例的root账户默认已具备上述权限。 当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步。 ● 视图、存储过程、函数依赖的表不支持做表名映射。 ● 任务做表名映射时，不支持该表的外键约束同步。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 ● 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 ● 仅支持MyISAM和InnoDB表的同步。 ● 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> - 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； - 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 - 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 - 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库必须为GaussDB(for MySQL)主备实例的主节点。 ● 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 ● 源数据库expire_logs_days参数值为0，可能会导致同步失败。 ● 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置MySQL源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库中的库名、表名、视图名不能包含：'<'>/\以及非ASCII字符。 ● 不支持非MyISAM和非InnoDB的表同步到RDS。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 数据库映射时，源库中存在视图、存储过程等对象，可能会导致实时同步失败。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库的运行状态必须正常。 ● 目标数据库必须有足够的磁盘空间。 ● 除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 ● 目标数据库的字符集必须与源数据库一致。 ● 目标数据库的时区设置必须与源数据库一致。 ● 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 对于表中的物理生成列，如果列是根据一个时间类型生成的，会存在此列数据不一致的可能性。 ● 在创建DRS任务之前，如果源或者目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 当目标数据库为5.7版本时，因为版本限制，json类型中的浮点数会丢失小数点后末位的0，存在精度损失导致内容对比不一致。 ● 源库不允许存在与目标库同名的无主键表。 ● 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 ● 不支持源数据库恢复到之前时间点的操作(PITR)。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 当在全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 全量同步过程中不支持DDL操作。 ● 增量同步过程中支持部分DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, CHANGE_COLUMN, DROP_COLUMN, DROP_INDEX, ADD_INDEX, CREATE_INDEX, RENAME_INDEX, DROP_TABLE, TRUNCATE_TABLE, DROP_PARTITION, RENAME_COLUMN, DROP_PRIMARY_KEY, ADD_PRIMARY_KEY，客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 增量同步支持表的重命名，源表和目標表必須都在对象选择里。 ● 使用了附加列功能，单表的列数超过500时，对该表添加附加列可能会超过列数上限，会导致任务失败。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-252 同步任务信息

表 4-272 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-253 同步实例信息

表 4-273 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB(for MySQL)”。
目标数据库引擎	选择“MySQL”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户创建的GaussDB(for MySQL)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<ul style="list-style-type: none"> - 全量+增量。 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 增量 增量同步通过解析日志等技术，将源端产生的增量数据至目标端。

● 任务类型

图 4-254 任务类型



表 4-274 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-255 企业项目和标签



表 4-275 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-256 源库信息

表 4-276 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，待该任务删除后会永久清除。

图 4-257 目标库信息

表 4-277 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。

 说明

目标库的IP、用户名和密码，会被系统加密暂存，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象。单击“下一步”。

图 4-258 同步模式

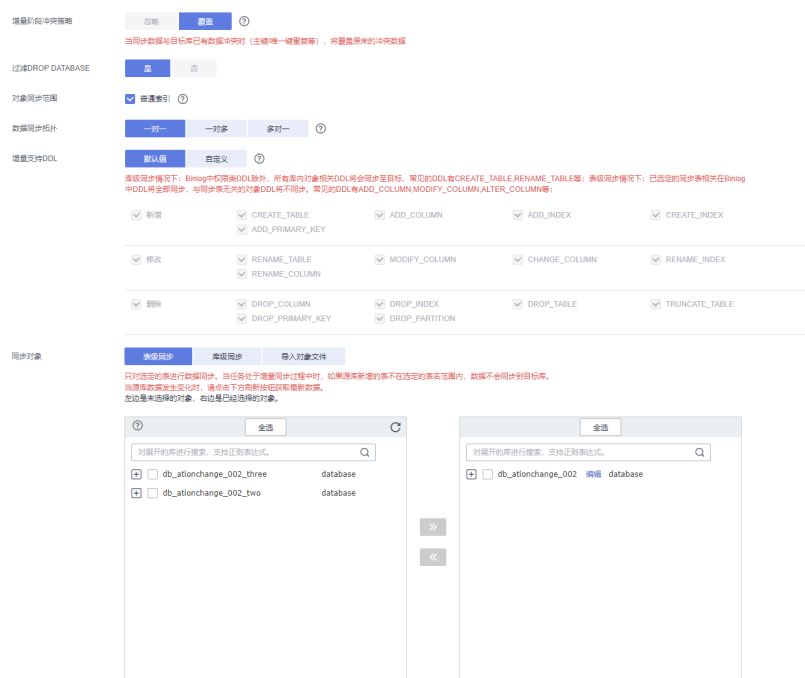



表 4-278 同步模式和对象

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> ● 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 ● 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>根据实际需求进行选择是否同步索引。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明 “数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>

参数	描述
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择“导入对象文件”，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果需要设置数据过滤，选择“数据过滤”，设置相关过滤规则。

- 如果需要设置添加附加列，选择“附加列”，单击“操作”列的“添加”，选填需要添加的列名和操作类型信息。

相关操作参考[数据加工](#)。

图 4-259 数据加工

源表	目标表	列名	操作类型	字符集	操作
db1.pak_011	db1.pak_011	--	--	--	添加
db1.pak02	db1.pak02	--	--	--	添加
db1.pak03	db1.pak03	--	--	--	添加

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-260 预检查

检查项	检查结果
目标库磁盘空间检查	
目标数据库磁盘可用空间是否足够	通过
数据库参数检查	
结构迁移参数一致性校验	请确认 已确认
源迁移无主键表检查	请确认 已确认
源库和目标库的字符集是否一致	通过
时区一致性校验	通过
COLLATION_SERVER的一致性检查	通过
检查目标库的max_allowed_packet参数	通过
迁移对象未选择外键依赖的表	通过
数据库参数INNODB_STRICT_MODE一致性检查	通过
SSL安全连接检查	通过
映射到目标库中的库名是否合法	通过

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-261 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 4-279 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.26 将 GaussDB(for MySQL)同步到 GaussDB(DWS)

支持的源和目标数据库

表 4-280 支持的数据库

源数据库	目标数据库
GaussDB(for MySQL)主备实例	GaussDB(DWS)集群

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-281 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库账号需要具备如下权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 ● 目标数据库账号必须具有每张表的如下权限：INSERT、SELECT、UPDATE、DELETE、CONNECT、CREATE。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> 支持表、普通索引（B-Tree索引）、约束（主键、空、非空）的同步，不支持视图、外键、存储过程、触发器、函数、事件、虚拟列、唯一约束、唯一索引的同步。 全量同步支持comment，增量不支持。 不支持的数据类型有：xml、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon。 仅支持MyISAM和InnoDB表的同步。 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 同步对象列名不能为CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID等GaussDB(DWS)禁止的字段，否则会导致任务失败。 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> 源数据库必须为GaussDB(for MySQL)主备实例的主节点。 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 源数据库expire_logs_days参数值为0，可能会导致同步失败。 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 必须设置源数据库的server-id，server-id的取值范围在2-4294967296之间。 源数据库中的库名、表名不能包含：'<>\ 以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> 目标数据库的运行状态必须正常。 目标数据库必须有足够的磁盘空间。 目标数据库的时区设置必须与源数据库一致。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：索引引用表等。 ● 网络中断在30秒内恢复的，不影响实时同步实时同步，如果超过30秒，则会导致同步任务失败。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB(DWS)中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由“原库名_原表名_原索引名”计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key，或表名_原索引名。 ● 全量阶段基于索引实现上差异，只保留普通B-Tree索引，其他索引默认不同步。GaussDB(DWS)索引过多时会影响存储空间和入库性能，建议用户可根据业务需求按需建索引。 ● 同步无主键表时，如果存在唯一键，可能因数据冲突导致写入失败。 ● 目标数据库与源数据库字符集不一致可能会导致同步后数据不一致或者同步失败。 ● 数据类型不兼容时，可能引起同步失败。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库进行恢复操作。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 实时同步过程中，支持修改端口，修改之后同步任务失败，需要通过重试后继续进行同步。 ● 实时同步过程中，不支持IP、账号、密码修改。 ● 增量同步过程中支持DDL操作。 <ul style="list-style-type: none"> - 一对一的场景下，默认同步的DDL类型有CREATE_TABLE, RENAME_TABLE, ADD_COLUMN, MODIFY_COLUMN, ALTER_COLUMN, DROP_CONSTRAINT、ADD_CONSTRAINT、CREATE_INDEX, DROP_INDEX, RENAME_INDEX, RENAME_COLUMN，客户可以根据自身需求，在对象选择页面选择需要同步的DDL类型。 - 多对一场景下，执行column重命名操作，必须停业务操作，不然会有数据不一致的风险。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 多对一场景下，推荐只同步加列DDL，其他的DDL同步可能会因为目标表被修改而导致任务失败或数据不一致。 - 多对一场景下，执行ADD_COLUMN时，需保证每张表加列的类型一致，否则有可能导致任务失败。 - 新增和修改表名、列名、索引名时不能超出63字符，否则任务会失败。 - 执行DDL语句创建索引时，语句中表名+索引名超出63字符可能出现重名情况，造成索引创建失败。 - 源库无主键表增加主键的时候，必须含有第一列，否则任务会失败。 - 增量阶段执行DDL操作时，如果目标表不存在，则该DDL会被忽略掉。 - 增量阶段，源数据库执行CHANGE COLUMN修改列信息，如果该列在目标数据库GaussDB(DWS)中为分布列，则该语句会被忽略，因为GaussDB(DWS)不支持修改分布列。 - 增量阶段，由于源和目标数据库索引规则不一样，所以不支持RENAME INDEX的DDL操作。 - 增量阶段，不建议CHAR(0)与其他字符拼接进行同步，例如插入CONCAT('a',CHAR(0),'b')，可能造成数据不一致。 - 增量阶段，JSON数据类型里不建议CHAR(34)（双引号）、CHAR(92)（反斜杠）与其他字符拼接进行同步，例如插入JSON_OBJECT('\$0', CONCAT('a', CHAR(34), 'b'))，同步到目标端后会多出一个反斜杠，造成数据不一致。 • 增量同步过程中，库级同步不支持Online DDL，表级同步目前只支持阿里云DMS产生的Online DDL。 • 增量同步过程中，不支持delete和update目标端主键类型为binary, text, blob, clob的表。 • 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-262 同步任务信息

⚠️ 1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述
0/256

表 4-282 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-263 同步实例信息

同步实例信息 ⓘ

以下信息确认后不可修改，请谨慎操作，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向 入云 出云 自建-自建

DRS要求源数据库和目标数据库中至少有一方为本云数据库实例，出云按照源数据库为本云数据库实例的场景。

* 源数据库引擎 MySQL DDM GaussDB 分布式版 GaussDB 主备版 DDS PostgreSQL GaussDB(for MySQL)

* 目标数据库引擎 MySQL Oracle GaussDB(DWS) CSSIES Kafka

* 网络类型 ⓘ

按网段自动为同步实例绑定弹性公网IP，直到任务结束后自动删除弹性公网IP。

* 源数据库实例 ⓘ 查看数据库实例 查看不可选实例

* 同步实例所在子网 ⓘ 查看子网 查看已占用的IP地址

* 同步模式

该模式为数据库的实时同步，通过全量过程生成目标数据库的初始化后，增量同步的增量数据解析日志等技术，将源和目标数据库的数据保持一致。

表 4-283 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB(for MySQL)”。
目标数据库引擎	选择“GaussDB(DWS)”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	可用的GaussDB(for MySQL)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 4-264 任务类型



表 4-284 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-265 企业项目和标签



表 4-285 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-266 源库信息

表 4-286 源库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，待该任务删除后会永久清除。

图 4-267 目标库信息

表 4-287 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。

参数	描述
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

📖 说明

目标库的IP、用户名和密码，会被系统加密暂存，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象。单击“下一步”。

📖 说明

目前，“数据同步拓扑”和“增量支持DDL”功能仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。

图 4-268 同步模式

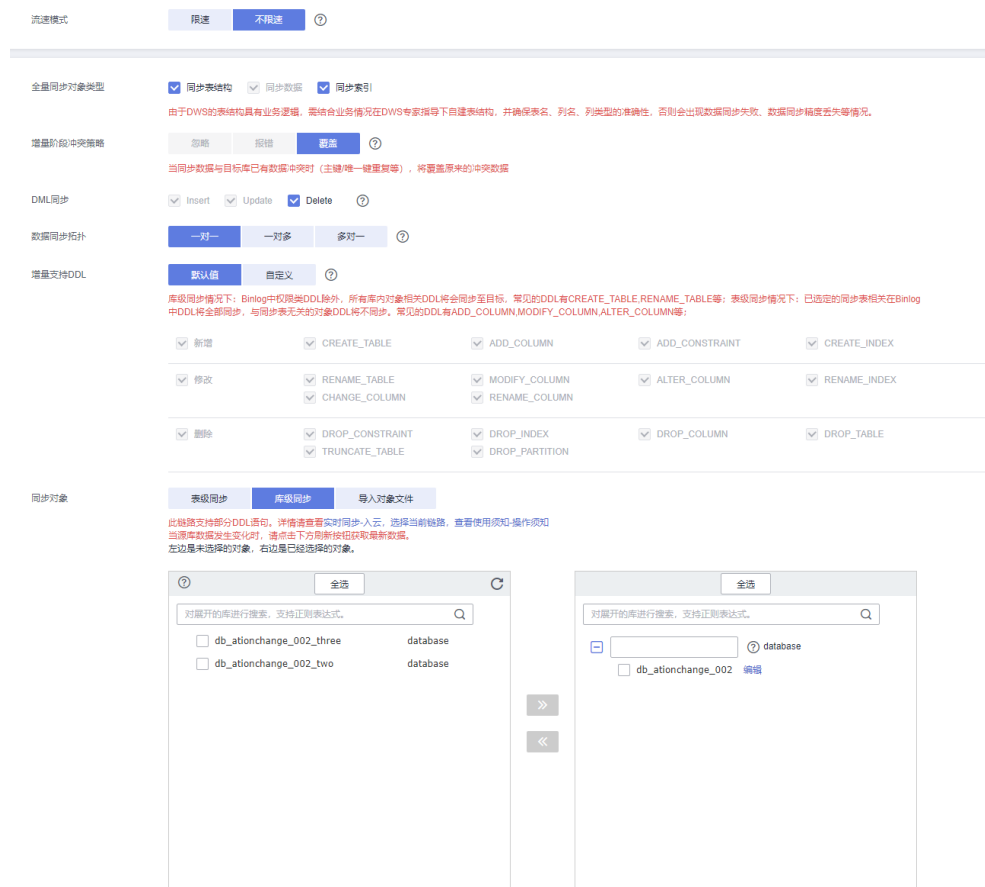



表 4-288 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-269 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> • 同步数据为必选项。 • 选择同步表结构的时候目标库不能有同名的表。 • 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p>
DML同步	<p>选择需要同步的DML操作，默认都勾选。</p>

参数	描述
数据同步拓扑	<p>数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择“导入对象文件”，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的表对象，单击“操作”列添加，填选需要添加的列名、类型、操作类型信息，检查无误后，单击“下一步”。可参考[数据加工](#)中的添加附加列，设置相关规则。

图 4-270 数据加工

源库表/视图	表名称	列名	操作类型	字段类型	操作
db1.jack_011	db1.jack_011	--	--	--	添加
db1.test2	db1.test2	--	--	--	添加
db1.test3	db1.test3	--	--	--	添加

注意

- “以serverName@database@table为列填充列”该操作会使用@符号拼接serverName、源库的库名和表名填充新加的列，并且该列会与源端表的主键形成复合主键。

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-271 预检查

检查项	检查结果
目标库磁盘空间检查	
目标数据库磁盘可用空间是否足够	通过
同步服务器磁盘空间检查	
同步服务器磁盘可用空间是否足够	通过
数据库参数检查	
源库和目标库的字符集是否一致	通过
时钟一致性校验	通过
COLLATION_SERVER的一致性检查	通过
结构迁移参数一致性校验	通过
迁移对象未选择外键依赖的表	通过
数据库参数INNODB_STRICT_MODE一致性检查	通过
SSL安全连接检查	通过
映射到目标库中的库名是否合法	通过
数据库参数SQL_MODE的一致性检查	通过
目标库sql_mode中NO_ENGINE_SUBSTITUTION值检查	通过

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-272 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: "启动时间" (Start Time) and "稍后启动" (Start Later), with "启动时间" selected. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): An input field containing the number "14". A red asterisk indicates it is a required field. A help icon and a note are present: "任务处于异常状态一段时间后，将会自动结束。单位为天。" (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 4-289 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.27 将 GaussDB(for MySQL)同步到 Kafka

支持的源和目标数据库

表 4-290 支持的数据库

源数据库	目标数据库
GaussDB(for MySQL)主备实例	Kafka

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。如果同步不可避免业务高峰期，推荐使用同步限速功能，即“流速模式”选择“限速”。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，会存在3s以内的单表级锁定。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-291 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none">• 源数据库账号需要具备如下权限：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步，增量仅支持表数据和DDL同步。 ● 仅支持MyISAM和InnoDB表的同步。 ● 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> ● 源数据库必须为GaussDB(for MySQL)主备实例的主节点。 ● 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 ● 源数据库expire_logs_days参数值为0，可能会导致同步失败。 ● 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 必须设置MySQL源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库中的库、表名不能包含： '<>\ '以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库为社区Kafka。 ● Kafka的auto.create.topics.enable参数建议为false。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 数据类型不兼容时，可能引起同步失败。 全量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的identifier字段进行去重（同一个shardId）。 增量同步过程中，请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 不支持源数据库恢复到之前时间点的操作(PITR)。 不支持强制清理binlog，否则会导致同步任务失败。 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 选择表级对象同步时，同步过程中不建议对表进行重命名操作。 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-273 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?
	不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。
项目	<input type="text" value=""/>
* 任务名称	<input type="text" value="DRS-6435"/> ?
描述	<div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> <div style="text-align: right; font-size: x-small;">0/256</div> ?

表 4-292 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

● 同步实例信息

图 4-274 同步实例信息



表 4-293 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB(for MySQL)”。
目标数据库引擎	选择“Kafka”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。

参数	描述
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
源数据库实例	用户创建的GaussDB(for MySQL)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<p>此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 4-275 任务类型



表 4-294 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 4-276 可用区</p>

- 企业项目和标签

图 4-277 企业项目和标签



表 4-295 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-278 源库信息

The screenshot shows a form titled '源库信息' (Source Database Information). It contains three input fields: '数据库实例名称' (Database Instance Name), '数据库用户名' (Database Username), and '数据库密码' (Database Password). Below these fields are two buttons: '测试连接' (Test Connection) and '测试成功' (Test Successful). The '测试成功' button has a green checkmark icon.

表 4-296 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

📖 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-279 目标库信息

表 4-297 目标库信息


参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 4-280 同步模式

表 4-298 同步对象

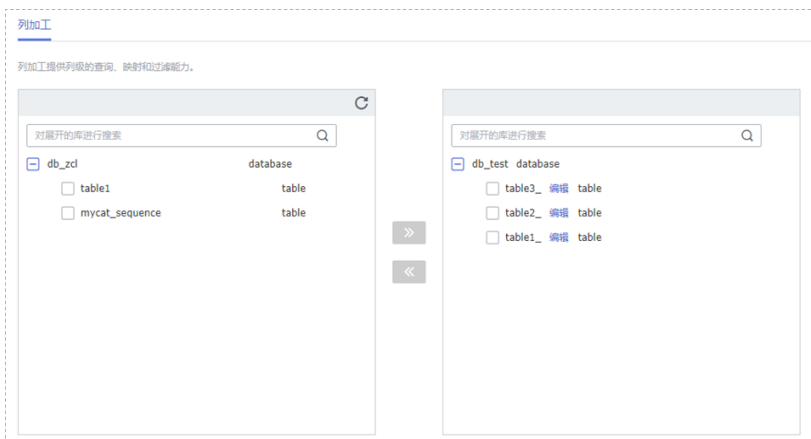
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-281 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
DML同步	选择需要同步的DML操作，默认都勾选。
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。

参数	描述
Topic名字格式	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"-", 若超过该范围会导致创建Topic失败，任务异常。</p> <p>若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。</p>
Partition个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。</p>
副本个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。</p>
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名+表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 按主键的hash值投递到不同的Partition：适用于一个表一个Topic的场景。
投送到kafka的数据格式	<p>选择GaussDB(for MySQL)投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。步骤步骤2的同步模式选择“增量”时可见，Avro格式目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。 JSON-C：一种能够兼容多个批量，流式计算框架的数据格式。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的列，进行列加工。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列，参考[数据加工](#)章节，设置相关规则。

图 4-282 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-283 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-284 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on, with a help icon.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off, with a help icon.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: A text input field contains '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 4-299 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.28 将 GaussDB(for MySQL)同步到 CSS/ES

支持的源和目标数据库

表 4-300 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • GaussDB(for MySQL)主备实例 	<ul style="list-style-type: none"> • ElasticSearch 5.5、6.2、6.5、7.1、7.6、7.9、7.10版本

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-301 使用须知

类型名称	使用和操作限制
数据库权限设置	全量+增量最小同步权限要求： <ul style="list-style-type: none">● 源数据库账号需要具备如下权限： SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。● 目标数据库账号必须拥有如下权限： READ、WRITE。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 支持表数据的同步。 ● 不支持数据库、视图、索引、约束、函数、存储过程、触发器（TRIGGER）和事件（EVENT）的同步。 ● 不支持系统库的同步以及事件状态的同步。 ● 不支持主键为FLOAT类型的表的同步。 ● 不支持无主键表的同步。 ● 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> ● 源数据库必须为GaussDB(for MySQL)主备实例的主节点。 ● 源数据库中的库名不能包含：'<>\'以及非ASCII字符。 ● 源数据库中的表名不能包含：'<>\'以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibttmp1。 ● 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 ● 源数据库expire_logs_days参数值为0，可能会导致同步失败。 ● 增量同步时，必须设置源数据库的server_id。server_id的取值范围在1 - 4294967296之间。 ● 源数据库建议开启skip-name-resolve，减少连接超时的可能性。 ● 源数据库GTID状态建议为开启状态。 ● 源库不支持mysql binlog dump命令。 ● 源数据库和目标数据库字符集需保持一致，否则同步失败。 ● 源数据库log_slave_updates参数需设置为开启状态，否则会导致同步失败。 ● 源数据库的binlog_row_image参数需设置为FULL，否则会导致同步失败。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库恢复到全量同步时间段范围内的PITR操作。 ● 源库时间字段目标库不支持的范围统一转成null。 ● 源库类型是binary类型会截断后面为0的字节，原因是源库是定长的自动补齐长度，目标库是变长类型。 ● binary的值会base64 加密后再写入目标库 ● datetime类型源库无时区到目标库则是用户指定的 ● 表字段名称全部转成小写。 ● 目标库数据_id如果指定源库多列生成需要用“:”分隔。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做DDL操作。 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL操作）。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源库写入binlog格式为statement的数据。 ● 同步过程中，不允许源库执行清除binlog的操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 同步过程中，不允许在源库创建库名为ib_logfile的数据库。 ● 增量同步过程中，若源库存在分布式事务，可能会导致同步失败。 ● 增量同步会过滤所有的DDL操作。 ● 增量同步阶段，支持断点续传功能，在主机系统崩溃的情况下，对于非事务性的无主键的表可能会出现重复插入数据的情况。 ● 选择表级对象同步时，增量同步过程中不支持对表进行重命名操作。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 选择同步对象时，单次选择的库表名称和列名称的大小不能超过4M。如果超过该限制，可以通过再编辑同步对象功能，分批增加同步对象。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-285 同步任务信息

表 4-302 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-286 同步实例信息

表 4-303 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB(for MySQL)”。
目标数据库引擎	选择“CSS/ES”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
源数据库实例	源数据库的GaussDB(for MySQL)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 任务类型

图 4-287 任务类型



表 4-304 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 4-288 可用区</p>

- 企业项目和标签

图 4-289 企业项目和标签



表 4-305 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-290 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

✔ 测试成功

表 4-306 源库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-291 目标库信息

目标库信息

IP地址或域名 ?
请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码 🔒

SSL安全连接
如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

✔ 测试成功

表 4-307 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 仅支持后缀为cer和pem的证书。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 4-292 同步模式

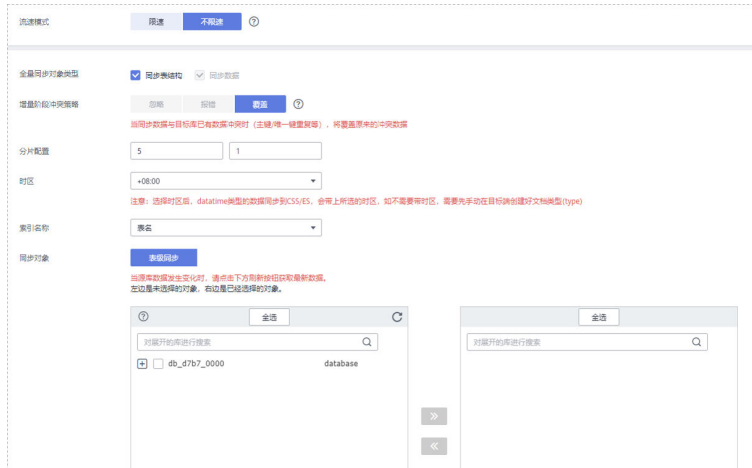



表 4-308 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-293 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> • 同步数据为必选项。 • 选择同步表结构的时候目标库不能有同名的表。 • 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，默认选择覆盖；全量阶段的冲突默认忽略。</p>
分片配置	<p>配置主分片数和分片副本数，主分片数默认为5，分片副本数默认为1。</p>

参数	描述
时区	选择时区后，datetime类型的数据同步到CSS/ES，会带上所选的时区。
索引名称	<ul style="list-style-type: none"> • 表名 选择为表名后，在目标Elasticsearch实例中创建的索引名称和表名一致。 • 库名+表名 选择为库名+表名后，在目标Elasticsearch实例中创建的索引名称为库名_表名。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> • 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 • 选择同步对象时，单次选择的库表名称和列名称的大小不能超过4M。如果超过该限制，可以通过再编辑同步对象功能，分批增加同步对象，参考编辑同步任务。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

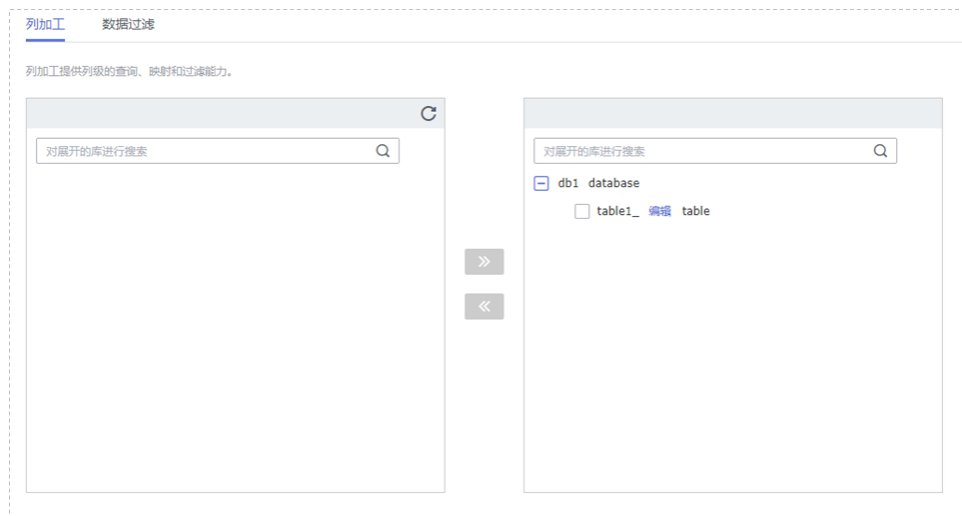
步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”或“列加工”，参考[数据加工](#)章节，设置相关规则。

说明

目前GaussDB(for MySQL)->CSS/ES的列加工仅支持列过滤，暂不支持列映射功能。

图 4-294 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-295 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-296 任务启动设置


The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '启动时间' (Start Time) and '稍后启动' (Start Later), with '立即启动' (Start Immediately) selected. Below this, there are several settings: '任务异常通知设置' (Task Abnormal Notification Settings) is a toggle switch that is turned on; 'SMN主题' (SMN Topic) is a dropdown menu; '时延阈值(s)' (Latency Threshold) is a toggle switch that is turned off; '数据异常通知' (Data Abnormal Notification) is a toggle switch that is turned off; and '任务异常自动结束时间' (Task Abnormal Auto-End Time) is a text input field containing the number '14', with a note indicating the unit is days.

表 4-309 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.29 将 GaussDB(for MySQL)同步到 Oracle

支持的源和目标数据库

表 4-310 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • GaussDB(for MySQL)主备实例 	<ul style="list-style-type: none"> • 本地自建Oracle数据库 • ECS自建Oracle数据库

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-311 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库账号需要具备如下权限： SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 ● 提供的目标数据库账号必须具有如下权限： ALTER ANY INDEX、ALTER ANY TABLE、ALTER SESSION、ANALYZE ANY、COMMENT ANY TABLE、CREATE ANY INDEX、CREATE ANY TABLE、CREATE SESSION、DELETE ANY TABLE、DROP ANY TABLE、INSERT ANY TABLE、SELECT ANY TABLE、SELECT ANY DICTIONARY、SELECT ANY TRANSACTION、UPDATE ANY TABLE、RESOURCE角色。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步支持数据、表结构和索引的同步。 ● 增量同步仅支持数据同步。 ● 不支持geometry, geometrycollection, linestring, multilinestring, multipoint, point, polygon 等地理坐标类型 ● 不支持视图、约束、函数、存储过程、触发器 (TRIGGER) 和事件 (EVENT) 的同步。 ● 不支持系统库的同步以及事件状态的同步。 ● 目标库Oracle不支持空字符串, 同步对象含有非空约束的字段中不能包含空字符串。 ● 源数据库GaussDB(for MySQL)支持的最大列数是1017, 目标数据库Oracle支持的最大列数为1000, 因为目标数据库限制, 同步对象的列数不能大于1000。 ● 全量和增量同步不支持隐藏列 (INVISIBLE), 隐藏列是MySQL 8.0.23开始支持的特性, 例如: <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> ● 源数据库必须为GaussDB(for MySQL)主备实例的主节点。 ● 源数据库中的库名不能包含: '<>/'以及非ASCII字符。 ● 源数据库中的表名不能包含: '<>/'以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头, 也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 增量同步时, 源数据库的binlog日志必须打开, 且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下, 建议源数据库binlog保存时间越长越好, 建议为3天。 ● 源数据库expire_logs_days参数值为0, 可能会导致同步失败。 ● 增量同步时, 必须设置源数据库的server_id, server_id的取值范围在1 - 4294967296之间。 ● 源数据库建议开启skip-name-resolve, 减少连接超时的可能性。 ● 源数据库GTID状态建议为开启状态。 ● 源库不支持mysql binlog dump命令。 ● 源数据库和目标数据库字符集需保持一致, 否则同步失败。 ● 源数据库log_slave_updates参数需设置为开启状态, 否则会导致同步失败。 ● 源数据库的binlog_row_image参数需设置为FULL, 否则会导致同步失败。 ● 源数据库不支持参数lower_case_table_names等于0的同步。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 源库时间字段默认值全0的统一转成1970-01-01 00:00:00。 ● 源库类型是decimal类型精度超过38会有有效数字被截断补0，原因是目标库对应的number类型的最大精度是38。 ● 源库varchar类型的长度大于等于667时，该类型到Oracle中将会被转成clob类型。 ● 表字段名称全部转成大写。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做DDL操作。 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL操作）。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源库写入binlog格式为statement的数据。 ● 同步过程中，不允许源库执行清除binlog的操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 同步过程中，不允许在源库创建库名为ib_logfile的数据库。 ● 增量同步场景下，不支持源数据库进行恢复到某个备份点的操作（PITR）。 ● 增量同步过程中，若源库存在分布式事务，可能会导致同步失败。 ● 增量同步会过滤所有的DDL操作。 ● 增量同步阶段，支持断点续传功能，在主机系统崩溃的情况下，对于非事务性的无主键的表可能会出现重复插入数据的情况。 ● 选择表级对象同步时，增量同步过程中不支持对表进行重命名操作。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-297 同步任务信息

表 4-312 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-298 同步实例信息

表 4-313 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“GaussDB(for MySQL)”。

参数	描述
目标数据库引擎	选择“Oracle”。
网络类型	<p>此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”、和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	源数据库的GaussDB(for MySQL)实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	<ul style="list-style-type: none"> - 全量+增量 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

- 任务类型

图 4-299 任务类型



表 4-314 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-300 企业项目和标签



表 4-315 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 4-301 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码

✔ 测试成功

表 4-316 源库信息

参数	描述
数据库实例名称	默认为创建迁移任务时选择的GaussDB(for MySQL)实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 4-302 目标库信息

目标库信息

IP地址或域名
对于RAC集群，建议使用SCANIP+SERVICENAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ▼ ?

数据库用户名

数据库密码

SSL安全连接

表 4-317 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群，建议使用scanip接入，提高访问性能。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 4-303 同步模式

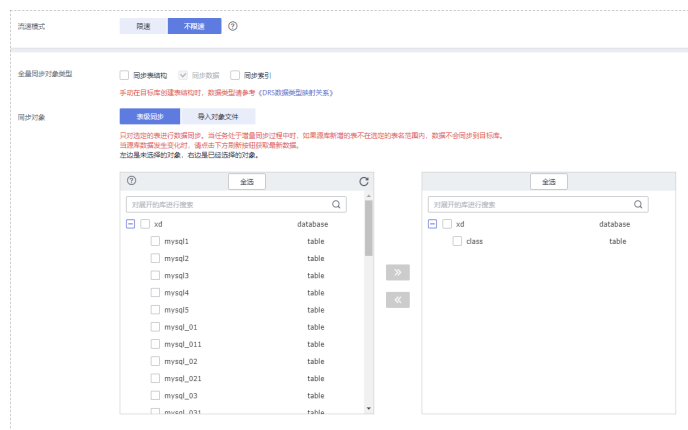



表 4-318 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-304 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。选择导入对象文件，具体步骤和说明可参考导入同步对象。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以保持待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-305 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 4-306 任务启动设置

表 4-319 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

4.30 将 MariaDB 同步到 MariaDB

支持的源和目标数据库

表 4-320 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> RDS for MariaDB 10.5版本 	<ul style="list-style-type: none"> 本地自建MariaDB数据库 10.5版本 ECS自建MariaDB数据库 10.5版本 其他云上MariaDB数据库 10.5版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表4-321](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 4-321 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 仅支持MyISAM和InnoDB表的同步。 不支持事件、触发器的同步。 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 视图、存储过程、函数依赖的表不支持做表名映射。 任务做表名映射时，不支持该表的外键约束同步。 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 已选择的表与未选择的表之间互相rename的DDL操作，在任务同步中会被过滤掉，可能会导致任务失败或数据不一致。 <ul style="list-style-type: none"> 对于一个DDL中存在一个rename的情况（例如：rename A TO B），只支持rename前后库表都在已选择库表中的rename操作（A和B都在已选择库表中），其他情况的rename DDL会被过滤； 对于一个DDL中存在多个rename的情况（例如：rename A TO B, B TO C），只支持rename前后库表都在已选择库表中的部分rename操作，其他情况的rename DDL会被过滤（A和B在已选择库表中，C不在，仅执行rename A TO B）。 不建议在多对一同步场景下的进行rename操作，可能会导致任务失败或数据不一致。 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 4-322](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 4-322 数据库账号权限

类型名称	全量+增量同步
源数据库连接账号	SELECT、SHOW VIEW、EVENT、LOCK TABLES、RELOAD、REPLICATION SLAVE、REPLICATION CLIENT（或BINLOG MONIOTOR，MariaDB 10.5及以上版本更改为此权限）。
目标数据库连接账号	SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES、INDEX。

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-323 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 增量同步时，MariaDB源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 - 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。 - 源数据库expire_logs_days参数值为0，可能会导致同步失败。 - 源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 - 增量同步时，必须设置MariaDB源数据库的server_id，server_id的取值范围在1 - 4294967296之间。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源数据库中的库名不能包含：'<>\'以及非ASCII字符。 - 源数据库中的表名、视图名不能包含：'<>\'以及非ASCII字符。 - 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibttmp1。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 不支持从高版本同步到低版本。 - 目标数据库的字符集必须与源数据库一致。 - 目标数据库的时区设置必须与源数据库一致。 - DRS同步时会有大量数据写入目标库，目标库max_allowed_packet 参数过小会导致无法写入，建议将目标库max_allowed_packet参数值设置为大于100MB。 - 同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 目标数据库实例的运行状态必须正常，若数据库实例是主备实例，复制状态也必须正常。 - 目标数据库实例必须有足够的磁盘空间。 - 除了系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。 - 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“'”。 ● 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 源库和目标库是相同的RDS实例时，不支持没有库映射的实时同步。 - 源库不允许存在与目标库同名的无主键表。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 - 源库和目标库为RDS for MariaDB实例时，不支持带有TDE特性并建立具有加密功能表。 - 如果源库MariaDB不支持TLS1.2协议，需提交运维申请才能使用SSL测试连接。 - 支持断点续传功能，但是对于无主键的表可能会出现重复插入数据的情况。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 - 不支持源数据库恢复到之前时间点的操作(PITR)。 - 不支持目标数据库恢复到全量同步时间段范围内的PITR操作。 - 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 - 不支持强制清理binlog，否则会导致同步任务失败。 - 不支持分区表的分区字段进行列映射。 - 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 - 任务创建后，目标库不能设置为只读。 - 需要DRS任务同步索引的情况下，目标库对应表上不能存在同名但不同列信息的索引。全量阶段DRS会忽略对应表上已存在的同名索引，增量阶段DDL根据索引名操作索引时会触发错位问题。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 全量同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 增量同步支持DDL语句。 ● 增量同步支持表的重命名，源表和目标表必须都在对象选择里。 ● 增量同步支持任务再编辑追加同步对象。

类型名称	使用和操作约定
同步对比须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。

操作步骤

本章节介绍在公网网络场景下，通过数据复制服务配置MariaDB同步出云任务的流程。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-307 同步任务信息

表 4-324 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 4-308 同步实例信息



表 4-325 同步实例信息

参数	描述
数据流动方向	选择出云。 出云指源端数据库为本云数据库。
源数据库引擎	选择“MariaDB”。
目标数据库引擎	选择“MariaDB”。
网络类型	此处以“公网网络”为示例，默认为公网网络类型，可按照需求选择“VPC网络”、“公网网络”和“VPN、专线网络”。 <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	用户所创建的RDS for MariaDB数据库实例。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
同步模式	<p>全量+增量：该模式为数据库持续性同步，适用于对业务中断敏感的场景，通过全量同步过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据库保持数据持续一致。</p> <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 4-309 任务类型



表 4-326 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-310 企业项目和标签



表 4-327 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

- 源库信息

图 4-311 源库信息页面

表 4-328 源库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的RDS for MariaDB数据库实例，不可进行修改。
数据库用户名	源数据库对应的数据库用户名。
数据库密码	源数据库的登录密码。

📖 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该同步任务后自动清除。

- 目标库信息配置

图 4-312 目标库信息

目标库信息

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

● 测试成功 ?

表 4-329 源库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库用户名所对应的密码。支持在任务创建后修改密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。


步骤4 在“同步设置”页面，设置同步对象，单击“下一步”。

图 4-313 设置同步对象



表 4-330 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 4-314 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>增量阶段冲突策略</p>	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”，参考[数据加工](#)章节，设置相关规则。

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且预检查通过率为100%时，单击“下一步”。

说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

图 4-315 预检查



步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-316 任务启动设置




表 4-331 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

4.31 将 Microsoft SQL Server 同步到 Kafka

支持的源和目标数据库

表 4-332 支持的数据库

源数据库	目标数据库
RDS for SQL Server (企业版2012、2014、2016、2017、2019版本, 标准版2016 SP2及以上版本、2017、2019版本)	Kafka 0.11及以上版本

说明

目前仅支持白名单用户使用。

支持的同步对象范围

在使用DRS进行同步时,不同类型的同步任务,支持的同步对象范围不同,详细情况可参考表4-333。DRS会根据用户选择,在“预检查”阶段对同步对象进行自动检查。

表 4-333 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> • 支持的字段类型: TINYINT、SMALLINT、INT、BIGINT、DECIMAL、NUMERIC、FLOAT、REAL、SMALLMONEY、MONEY、BIT、DATE、DATETIME、DATETIME2、DATETIMEOFFSET、TIME、TIMESTAMP、XML、CHAR、VARCHAR、NCHAR、NVARCHAR、BINARY、VARBINARY、IMAGE、HIERARCHYID、NTEXT、TEXT、UNIQUEIDENTIFIER。 • 不支持的字段类型: SQL_VARIANT、GEOMETRY、GEOGRAPHY。 • 增量同步的范围: <ul style="list-style-type: none"> - 支持DML (Data Manipulation Language) : 包括INSERT、UPDATE、DELETE。 - 不支持DDL (Data Definition Language) 。

数据库账号权限要求

在使用DRS进行同步时,连接源库和目标库的数据库账号需要满足以下权限要求,才能启动实时同步任务。不同类型的同步任务,需要的账号权限也不同,详细可参考表4-334进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查,并给出处理建议。

表 4-334 数据库账号权限

类型名称	增量同步
源数据库连接账号	需要具备sysadmin权限，或者view server state权限以及待同步数据库的db_datareader或db_owner权限。

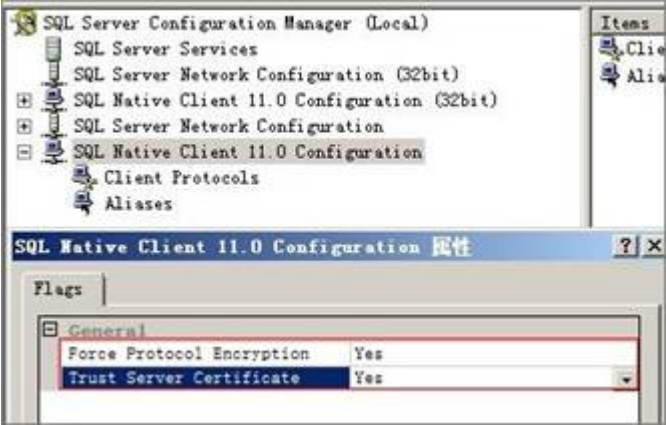
使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步过程一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 4-335 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库模式必须设置为FULL模式。 - 源数据库必须开启SQL Server Agent代理服务。 - 源数据库若开启客户端配置中的“强制协议加密（ Force Protocol Encrypton ）”，必须同时开启“信任服务器证书（ trust server certificate ）”，如下图4-317所示： <p style="text-align: center;">图 4-317 查看客户端属性</p>  <ul style="list-style-type: none"> ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不允许存在被禁用的表的聚簇索引，否则会导致同步失败。 - 源库中不允许存在名称为cdc的用户名或schema。 - 源库待同步对象的库名、schema名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 源库待同步对象的表中列名不能包含如下字符：[]?。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 不支持源数据库主备切换，源数据库主备切换会导致同步任务失败。 - 不支持源数据库Microsoft SQL Server为TLS 1.0、TLS 1.1协议的同步，如果需要同步，建议源库升级到TLS 1.2及以上版本。 - SQL Server为源的增量同步基于SQL Server提供的CDC能力，如果SQL Server源库的CDC功能出现异常（如被关闭、事务日志满等情况），则增量同步就会受到影响。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 不支持DDL操作，源数据库进行的DDL操作不会同步到目标数据库。 ● 不支持大数据类型IMAGE、TEXT、NTEXT的删除操作。 ● 支持任务再编辑追加同步对象。

操作步骤

本小节以Microsoft SQL Server->Kafka的出云同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 4-318 同步任务信息

表 4-336 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' ' " \ 特殊字符。

- 同步实例信息

图 4-319 同步实例信息



表 4-337 同步实例信息

参数	描述
数据流动方向	选择“出云”。
源数据库引擎	选择“Microsoft SQL Server”。
目标数据库引擎	选择“Kafka”。
网络类型	<p>此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
源数据库实例	源数据库为RDS for SQL Server实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
同步模式	增量：增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 4-320 可用区



表 4-338 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 4-321 企业项目和标签



表 4-339 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

- 步骤3** 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 4-322 源库信息

源库信息

数据库实例名称

数据库用户名

数据库密码 

待实例创建成功后再进行测试连接

表 4-340 源库信息

参数	描述
数据库实例名	默认为创建同步任务时选择的RDS for SQL Server数据库实例，不可进行修改。
数据库用户名	源数据库的用户名。
数据库密码	源数据库用户名所对应的密码。

图 4-323 目标库信息

目标库信息

IP地址或域名 

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

表 4-341 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。

参数	描述
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 4-324 同步对象

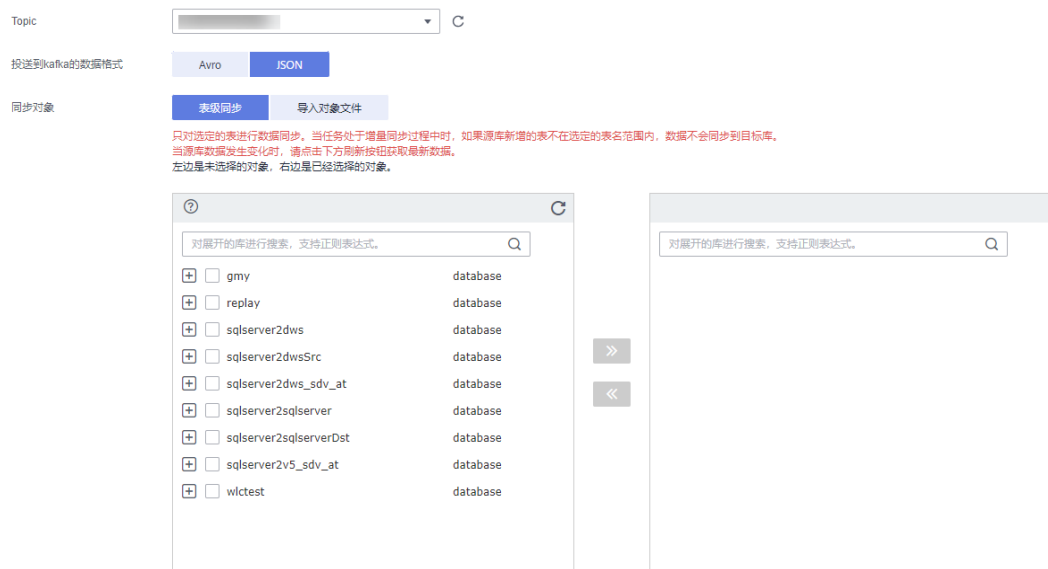



表 4-342 同步模式和对象

参数	描述
Topic	选择目标端需要同步到的Topic。
投送到kafka的数据格式	选择投送到kafka的数据格式。 <ul style="list-style-type: none"> ● Avro：可以显示Avro二进制编码，高效获取数据。 ● Json：为Json消息格式。 详细格式可参考 Kafka消息格式 。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 选择导入对象文件时，支持不同表同步到目标端不同的Topic，具体导入步骤和说明可参考导入同步对象。 使用导入对象功能时，同步Topic策略选择“集中投递到一个Topic”，才能使用对象名映射功能，否则会按Topic名字格式生成。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

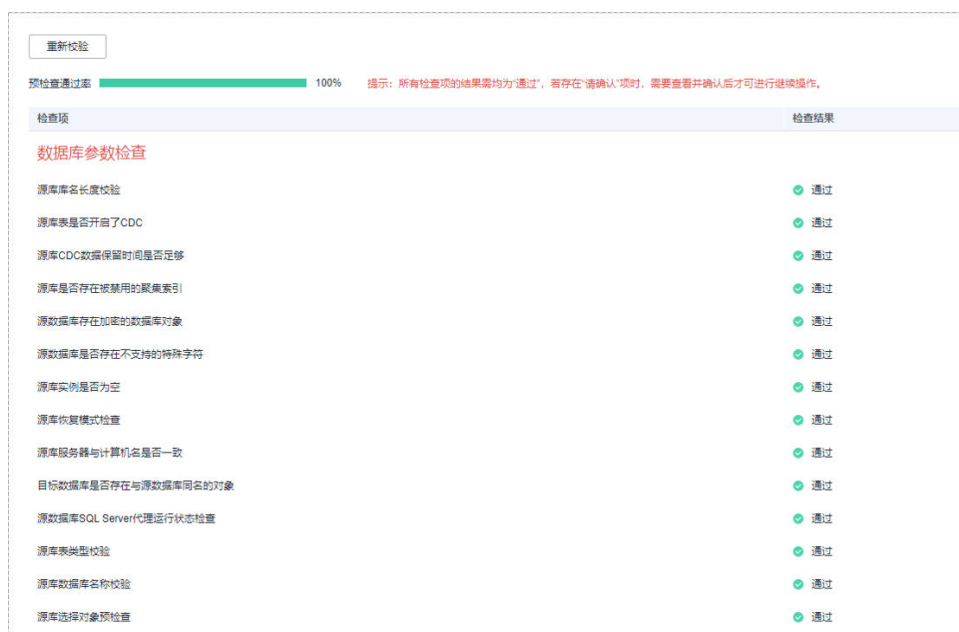
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 4-325 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 4-326 任务启动设置

The screenshot shows a configuration interface for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on, with a help icon.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off, with a help icon.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: An input field contains the value '14', with a help icon and a red note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 4-343 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5 自建到自建

5.1 将 MySQL 同步到 Kafka

支持的源和目标数据库

表 5-1 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MySQL数据库ECS自建MySQL数据库	<ul style="list-style-type: none">Kafka

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。如果同步不可避免业务高峰期，推荐使用同步限速功能，即“流速模式”选择“限速”。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，会存在3s以内的单表级锁定。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。

- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-2 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账号需要具备如下权限：SELECT、LOCK TABLES、SHOW VIEW、EVENT、REPLICATION SLAVE、REPLICATION CLIENT。
同步对象约束	<ul style="list-style-type: none"> • 全量支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步，不支持事件、触发器的同步，增量仅支持表数据和DDL同步。 • 仅支持MyISAM和InnoDB表的同步。 • 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● 源数据库中的库、表名不能包含：'<'>\以及非ASCII字符。 ● 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库为社区Kafka。 ● Kafka的auto.create.topics.enable参数建议为false。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。 ● 全量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的identifier字段进行去重（同一个shardId）。 ● 增量同步过程中，请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持强制清理binlog，否则会导致同步任务失败。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 同步过程中，不允许删除和修改源库的用户名、密码、权限，或修改目标数据库的端口号。 ● 当在同步过程中，对MyISAM表执行修改操作时，可能造成数据不一致。 ● 选择表级对象同步时，同步过程中不建议对表进行重命名操作。 ● 任务做库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-1 同步任务信息

⚠️ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述
0/256

表 5-3 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-2 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向 入云 出云 自建-自建 ?

* 源数据库引擎 MySQL Oracle PostgreSQL

* 目标数据库引擎 CSSDB Kafka

* 网络类型 ?
 DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除弹性公网IP。

* DRS任务类型 单AZ 双AZ
采用单节点部署架构，与多AZ部署方式相比，每个任务只创建一个节点，也具有高性价比，适用于业务数据量较小、短期同步，并对平滑的运维无严格要求的场景。

* VPC ? [查看虚拟私有云](#)
同步实例所在子网 ? [查看子网](#) [查看已占用的IP地址](#)

* 内网安全组 ?

* 同步模式 全量 + 增量 增量 ?
该模式为数据一致性实时同步，通过全量过程或目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源和目标数据库数据保持一致。

表 5-4 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“Kafka”。

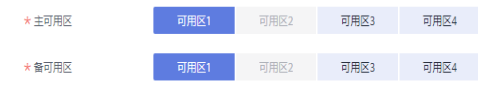
参数	描述
网络类型	此处以“公网网络”为示例。目前支持可选公网网络、VPC网络和VPN、专线网络。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-3 任务类型



表 5-5 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 5-4 可用区</p> 

- 企业项目和标签

图 5-5 企业项目和标签



表 5-6 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 5-6 源库信息

源库信息

不支持数据库参数和系统数据库同步，源数据库参数设定和用户将不会同步至目标数据库中，请在目标数据库中使用参数组修改参数，手工创建用户。

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

表 5-7 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-7 目标库信息

表 5-8 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。


步骤**步骤2**中同步模式选择不同时，“设置同步”页面参数也不同，[图5-8](#)选择“全量+增量”为示例。

图 5-8 同步模式

表 5-9 同步对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-9 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	全量同步对象类型，可选同步表结构、同步数据，根据实际需求进行选择要同步内容。
DML同步	选择需要同步的DML操作，默认都勾选。
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
同步Topic策略	同步Topic策略，可选择“集中投递到一个Topic”或者“自动生成Topic名字”。

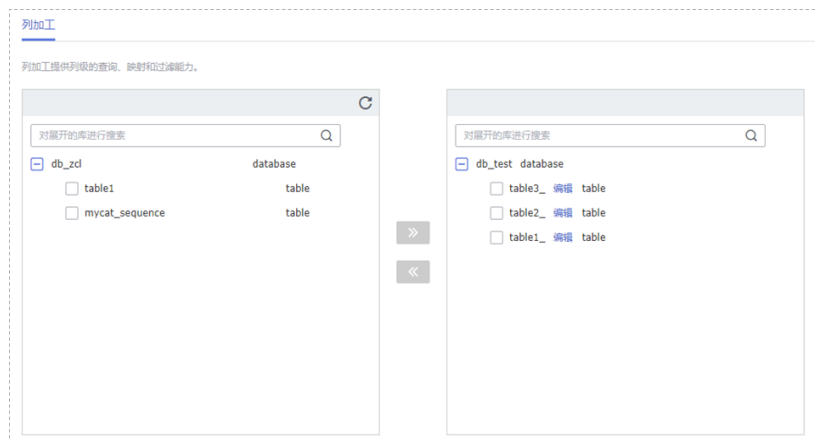
参数	描述
Topic	选择目标端需要同步到的Topic，同步Topic策略选择“集中投递到一个Topic”时可见。
Topic名字格式	<p>Topic名字格式，同步Topic策略选择“自动生成Topic名字”时可见。</p> <p>由于Kafka的机制限制，Topic名字只能包含ASCII字符、“.”、“_”及“-”，若超过该范围会导致创建Topic失败，任务异常。</p> <p>若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。</p> <p>Topic名字格式支持database和tablename两个变量，其他字符都当做常量。分别用\$database\$代替数据库名，\$tablename\$代替表名。</p> <p>例如：配置成\$database\$-\$tablename\$时，如果数据库名称为db1，表名为tab1，则Topic名字为db1-tab1。如果是DDL语句，\$tablename\$为空，则Topic名字为db1。</p>
Partition个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。</p>
副本个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。</p>
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名+表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 全部投递到Partition 0：适用于有事务要求的场景，事务保序，可以保证完全按照事务顺序消费，写入性能比较差，如果没有强事务要求，不推荐使用此选项。 按主键的hash值投递到不同的Partition：适用于一个表一个Topic的场景。
投送到kafka的数据格式	<p>选择MySQL投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。步骤步骤2的同步模式选择“增量”时可见，Avro格式目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。 JSON-C：一种能够兼容多个批量，流式计算框架的数据格式。 <p>详细格式可参考Kafka消息格式。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的列，进行列加工。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列，参考[数据加工](#)章节，设置相关规则。

图 5-10 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-11 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面, 设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间, 并确认同步任务信息无误后, 勾选协议, 单击“启动任务”, 提交同步任务。


图 5-12 任务启动设置



表 5-10 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.2 将 MySQL 同步到 CSS/ES

支持的源和目标数据库

表 5-11 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MySQL数据库 5.5、5.6、5.7、8.0版本ECS自建MySQL数据库 5.5、5.6、5.7、8.0版本	<ul style="list-style-type: none">ElasticSearch 5.5、6.2、6.5、7.1、7.6、7.9、7.10版本

使用建议

⚠ 注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-12 使用须知

类型名称	使用和操作限制
数据库权限设置	<p>全量+增量最小同步权限要求：</p> <ul style="list-style-type: none"> • 源数据库账号需要具备如下权限： SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 • 目标数据库账号必须拥有如下权限： READ、WRITE。
同步对象约束	<ul style="list-style-type: none"> • 支持表数据的同步。 • 不支持数据库、视图、索引、约束、函数、存储过程、触发器（TRIGGER）和事件（EVENT）的同步。 • 不支持系统库的同步以及事件状态的同步。 • 不支持主键为FLOAT类型的表的同步。 • 不支持无主键表的同步。 • 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 源数据库中的库名不能包含：'<>/'以及非ASCII字符。 ● 源数据库中的表名不能包含：'<>/'以及非ASCII字符。 ● 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 ● 增量同步时，MySQL源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 ● 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> - 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 - 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 ● 增量同步时，必须设置MySQL源数据库的server_id。如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间；如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 ● 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 ● MySQL源数据库建议开启skip-name-resolve，减少连接超时的可能性。 ● 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 ● 源库不支持mysql binlog dump命令。 ● 源数据库和目标数据库字符集需保持一致，否则同步失败。 ● 源数据库log_slave_updates参数需设置为开启状态，否则会导致同步失败。 ● 源数据库的binlog_row_image参数需设置为FULL，否则会导致同步失败。
目标数据库要求	<ul style="list-style-type: none"> ● 目标数据库实例的运行状态必须正常。 ● 目标数据库实例必须有足够的磁盘空间。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库恢复到全量同步时间段范围内的PITR操作。 ● 源库时间字段目标库不支持的范围统一转成null。 ● 源库类型是binary类型会截断后面为0的字节，原因是源库是定长的自动补齐长度，目标库是变长类型。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● binary的值会base64 加密后再写入目标库。 ● datetime类型源库无时区到目标库则是用户指定的。 ● 表字段名称全部转成小写。 ● 目标库数据_id如果指定源库多列生成需要用“:”分隔。 ● 在任务启动、任务全量同步阶段，不建议对源数据库做DDL操作。 ● 为了保持数据一致性，不允许对正在同步中的目标数据库进行修改操作（包括但不限于DDL操作）。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，不允许源库写入binlog格式为statement的数据。 ● 同步过程中，不允许源库执行清除binlog的操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 同步过程中，不允许在源库创建库名为ib_logfile的数据库。 ● 增量同步过程中，若源库存在分布式事务，可能会导致同步失败。 ● 增量同步会过滤所有的DDL操作。 ● 增量同步阶段，支持断点续传功能，在主机系统崩溃的情况下，对于非事务性的无主键的表可能会出现重复插入数据的情况。 ● 选择表级对象同步时，增量同步过程中不支持对表进行重命名操作。 ● 建议将expire_log_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证服务中断后的顺利恢复。 ● 选择同步对象时，单次选择的库表名称和列名称的大小不能超过4M。如果超过该限制，可以通过再编辑同步对象功能，分批增加同步对象。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-13 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown menu] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown menu]

* 任务名称: DRS-6435 ?

描述: [Text area] ?
0/256

表 5-13 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-14 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置项填错，需要重新创建迁移任务。

* 数据流动方向: 入云 | 出云 | 自建-自建 ?

* 源数据库引擎: MySQL | Oracle | PostgreSQL

* 目标数据库引擎: CSS/ES | Kafka

* 网络类型: 公网网络 ?

DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除该弹性公网IP。

* VPC: [Dropdown menu] C 查看虚拟私有云

同步实例所在子网: [Dropdown menu] ? 查看子网 查看已占用的IP地址

* 内网安全组: [Dropdown menu] C

* 同步模式: 全量 + 增量 ?

该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

* 可用区: 可用区1 | 可用区2 | 可用区3

DRS实例将创建在此可用区，选择源库或目标库所在的可用区性能更优。

表 5-14 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“CSS/ES”。
网络类型	此处以“公网网络”为示例。目前支持可选公网网络、VPC网络和VPN、专线网络。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 - “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。

- 任务类型

图 5-15 任务类型



表 5-15 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-16 企业项目和标签



表 5-16 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 5-17 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

● 测试成功

表 5-17 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-18 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

● 测试成功

表 5-18 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL根证书。</p> <p>说明</p> <ul style="list-style-type: none"> ● 最大支持上传500KB的证书文件。 ● 仅支持后缀为cer和pem的证书。 ● 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 5-19 同步模式

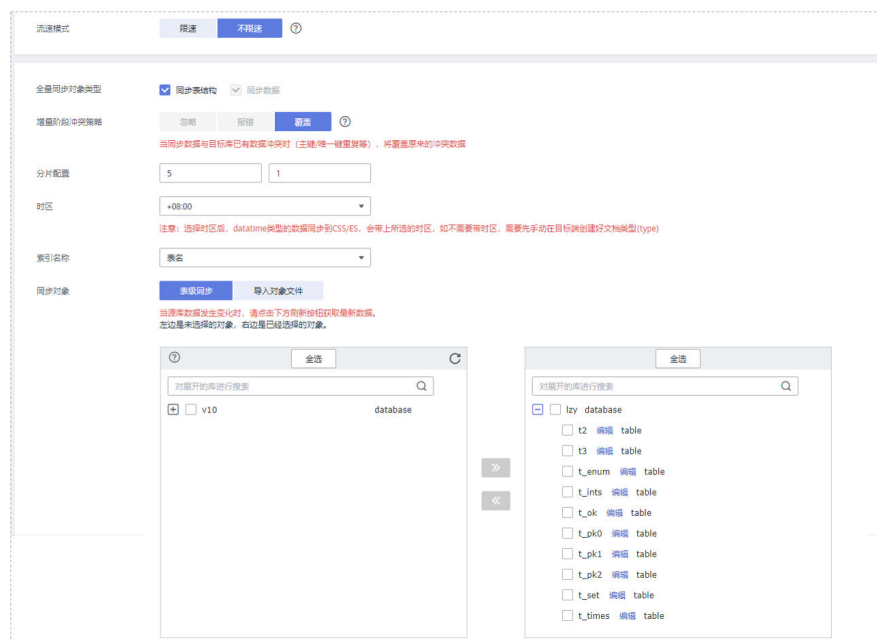



表 5-19 同步模式和对象

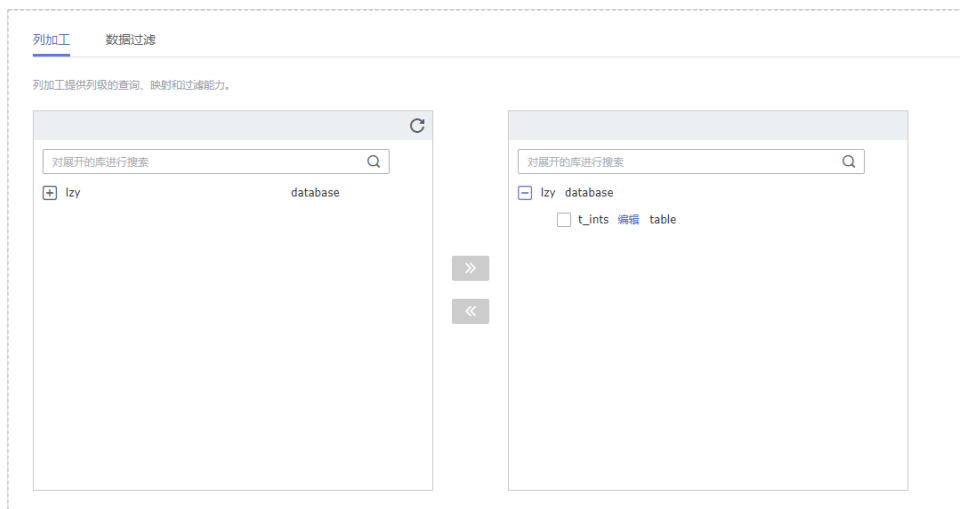
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-20 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
全量同步对象类型	<p>全量同步对象类型，可选同步表结构、同步数据，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> • 同步数据为必选项。 • 选择同步表结构的时候目标库不能有同名的表。 • 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，默认选择覆盖；全量阶段的冲突默认忽略。</p>
分片配置	<p>配置主分片数和分片副本数，主分片数默认为5，分片副本数默认为1。</p>

参数	描述
时区	选择时区后，datetime类型的数据同步到CSS/ES，会带上所选的时区。
索引名称	<ul style="list-style-type: none"> 表名 选择为表名后，在目标Elasticsearch实例中创建的索引名称和表名一致。 库名+表名 选择为库名+表名后，在目标Elasticsearch实例中创建的索引名称为库名_表名。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 可以选择文件导入对象，具体步骤和说明可参考导入同步对象。 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择表级同步时，单次选择的库表名称和列名称的大小不能超过4M。如果超过该限制，可以通过再编辑同步对象功能，分批增加同步对象，参考编辑同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 5-21 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-22 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-23 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-20 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.3 将 MySQL 同步到 GaussDB 主备版

支持的源和目标数据库

表 5-21 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建MySQL数据库 • ECS自建MySQL数据库 	GaussDB主备版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-22 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库账户需要具备如下权限： <ul style="list-style-type: none"> - 全量同步：SELECT。 授权语句：GRANT SELECT ON <database>.<table> to <user>; - 全量+增量和增量同步：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 授权语句：GRANT SELECT, LOCK TABLES ON <database>.<table> to drsUser; GRANT REPLICATION SLAVE, REPLICATION CLIENT ON *.* to <user>; ● 目标库GaussDB账户需要具备如下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>; - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> 仅支持同步表结构、表数据、索引信息，不支持同步存储过程等其他数据库对象。 增量同步不支持同步DDL。 虚拟列会被同步为普通列，且不写入任何数据，也不做增量同步。 不支持的数据类型有：xml、包含bit类型的json、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon、interval。 仅支持MyISAM和InnoDB表的同步。 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 对象名同步到目标库后会转换成小写，因此选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 增量同步时，必须设置MySQL源数据库的server_id。MySQL5.7，server_id的取值范围在1 - 4294967296之间。 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 源数据库中的库名、表名不能包含：.<'>/\以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> 目标数据库实例的运行状态必须正常。 目标数据库实例必须有足够的磁盘空间。 目标数据库的时区设置必须与源数据库一致。 任务配置的映射数据库必须在目标库已经存在。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，在全量同步中，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由"原库名_原表名_原索引名"计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key。 ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 网络中断在30秒内恢复的，不影响实时同步，如果超过30秒，则会导致同步任务失败。 ● 目标数据库与源数据库字符集不一致可能会导致同步后数据不一致或者同步失败。 ● 数据类型不兼容时，可能引起同步失败。 ● 源库表同步至目标库后分布方式为哈希分布，暂不支持复制分布。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 源库和目标库均相同的任务不允许出现重复同步的情况，如：A任务和B任务同时将源库的同一张表的实时同步到目标库的同一张表中，可能导致数据不一致和同步失败。 ● DRS全量同步表结构时，源库中char、varchar、nvarchar、enum、set字符类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），扩大倍数和源库中列的字符集有关。如源库字符集为UTF8的按照字节长扩大3倍，为UTF8MB4的按照字节长扩大4倍。 ● DRS全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 增量同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步过程中，不支持源库DDL的复制。 ● 增量同步场景下，不支持源数据库进行恢复操作。 ● 增量同步过程中，不支持delete和update目标端主键类型为tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary的表。 ● 任务启动、任务全量同步阶段，不建议做删除类型的DDL操作，可能会引起任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库进行恢复操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。 ● 如果无主键表中包含大字段（tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary），增量同步阶段存在大字段数据不一致的可能性。 ● 由于浮点型数据是采用近似值存储，而不是作为精确值存储，此外操作系统平台和底层实现也对浮点型数据有影响，因此FLOAT，DOUBLE类型的数据是不精确的，所以在MySQL和GaussDB同步浮点型数据，存在一定差异的可能性，可参考MySQL的官方文档。 ● 不支持两阶段事务。 ● 由于MySQL与GaussDB针对零时间（"0000-00-00 00:00:00"）内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。 ● 源端同步的表如果有AUTO_INCREMENT属性，结束任务时，DRS会自动刷新同步表的序列整型列所对应的GaussDB自增列（sequence）起始值，刷新值为该列的最大值+10000。

操作步骤

本小节以MySQL到GaussDB主备版的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-24 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-23 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-25 同步实例信息

同步实例信息

数据流动方向: 自建-自建

源数据库引擎: MySQL

目标数据库引擎: GaussDB 主备版

网络类型: 公网网络

VPC: [下拉菜单]

表 5-24 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。

参数	描述
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-26 任务类型



表 5-25 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-27 企业项目和标签



表 5-26 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 5-28 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中[使用参数模板设定](#)

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 5-27 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-29 目标库信息

表 5-28 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 5-30 同步模式

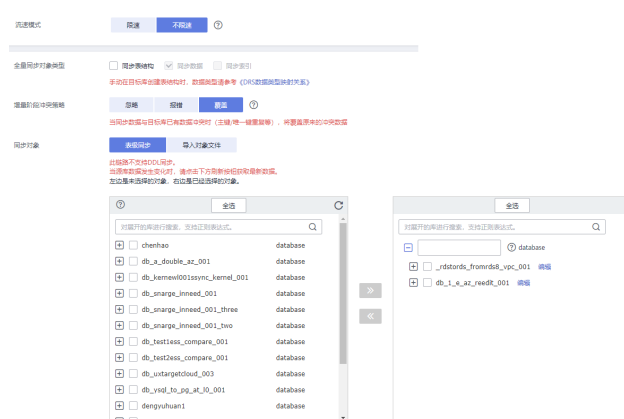



表 5-29 同步模式和对象

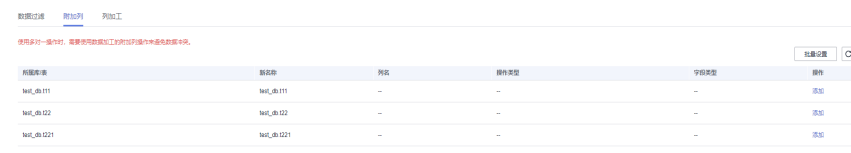
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-31 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 5-32 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-33 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-34 任务启动设置




表 5-30 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.4 将 MySQL 同步到 GaussDB 分布式版

支持的源和目标数据库

表 5-31 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建MySQL数据库ECS自建MySQL数据库	GaussDB分布式

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-32 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> • 源数据库账户需要具备如下权限： <ul style="list-style-type: none"> - 全量同步：SELECT。 授权语句：GRANT SELECT ON <database>.<table> to <user>; - 全量+增量和增量同步：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 授权语句：GRANT SELECT, LOCK TABLES ON <database>.<table> to drsUser; GRANT REPLICATION SLAVE, REPLICATION CLIENT ON *.* to <user>; • 目标库GaussDB账户需要具备如下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>; - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> 仅支持同步表结构、表数据、索引信息，不支持同步存储过程等其他数据库对象。 仅支持同步有主键表，不支持同步无主键表。 增量同步不支持同步DDL。 虚拟列会被同步为普通列，且不写入任何数据，也不做增量同步。 不支持的数据类型有：xml、包含bit类型的json、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon、interval。 仅支持MyISAM和InnoDB表的同步。 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 对象名同步到目标库后会转换成小写，因此选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。 全量和增量同步不支持隐藏列（INVISIBLE），隐藏列是MySQL 8.0.23开始支持的特性，例如： <pre>CREATE TABLE `test11` (`id` int NOT NULL, `c1` int DEFAULT NULL /*!80023 INVISIBLE */, PRIMARY KEY (`id`));</pre>
源数据库要求	<ul style="list-style-type: none"> 增量同步时，源数据库的binlog日志必须打开，且binlog日志格式必须为Row格式。 在磁盘空间允许的情况下，建议源数据库binlog保存时间越长越好，建议为3天。设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> 源数据库为自建MySQL时，通过设置expire_logs_days参数设置binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的binlog尚未过期，以保证任务中断后的顺利恢复。 源数据库为RDS for MySQL时，设置binlog保留时间可参考《RDS用户指南》。 源数据库GTID状态必须为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。 增量同步时，必须设置MySQL源数据库的server_id。MySQL5.7，server_id的取值范围在1 - 4294967296之间。 增量同步时，如果设置session变量character_set_client为binary，可能导致乱码。 源数据库中的库名、表名不能包含：.<'> \以及非ASCII字符。
目标数据库要求	<ul style="list-style-type: none"> 目标数据库实例的运行状态必须正常。 目标数据库实例必须有足够的磁盘空间。 目标数据库的时区设置必须与源数据库一致。 任务配置的映射数据库必须在目标库已经存在。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 ● 在MySQL中，同一个database (schema) 下面的不同表可以存在相同的索引名或约束名。而目标端GaussDB中，同一个schema中索引和约束名是唯一的，且长度存在上限。为防止索引和约束名冲突，在全量同步中，DRS同步到目标表上的索引名称将变为此格式：哈希值+原索引名（可能被截断）+_key。其中哈希值由"原库名_原表名_原索引名"计算得到。同理，表上的原约束名将变为：哈希值+原约束名（可能被截断）+_key。 ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 网络中断在30秒内恢复的，不影响实时同步，如果超过30秒，则会导致同步任务失败。 ● 目标数据库与源数据库字符集不一致可能会导致同步后数据不一致或者同步失败。 ● 数据类型不兼容时，可能引起同步失败。 ● 源库表同步至目标库后分布方式为哈希分布，暂不支持复制分布。 ● 源库为RDS for MySQL实例时，不支持带有TDE特性并建立具有加密功能表。 ● 如果源库MySQL不支持TLS1.2协议，或者为版本较低的自建库（低于5.6.46或在5.7.0-5.7.28之间），需提交运维申请才能使用SSL测试连接。 ● 在创建DRS任务之前，如果源数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 源库和目标库均相同的任务不允许出现重复同步的情况，如：A任务和B任务同时将源库的同一张表的实时同步到目标库的同一张表中，可能导致数据不一致和同步失败。 ● DRS全量同步表结构时，源库中char、varchar、nvarchar、enum、set字符类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），扩大倍数和源库中列的字符集有关。如源库字符集为UTF8的按照字节长扩大3倍，为UTF8MB4的按照字节长扩大4倍。 ● DRS全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 增量同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 增量同步过程中，不支持源库DDL的复制。 ● 增量同步场景下，不支持源数据库进行恢复操作。 ● 增量同步过程中，不支持delete和update目标端主键类型为tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary的表。 ● 任务启动、任务全量同步阶段，不建议做删除类型的DDL操作，可能会引起任务失败。 ● 不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。 ● 不支持源数据库进行恢复操作。 ● 源数据库不支持reset master或reset master to命令，可能会导致DRS任务失败或者数据不一致。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。 ● 如果无主键表中包含大字段（tinyblob、blob、mediumblob、longblob、tinytext、text、mediumtext、longtext、clob、nclob、bytea、binary），增量同步阶段存在大字段数据不一致的可能性。 ● 由于浮点型数据是采用近似值存储，而不是作为精确值存储，此外操作系统平台和底层实现也对浮点型数据有影响，因此FLOAT，DOUBLE类型的数据是不精确的，所以在MySQL和GaussDB同步浮点型数据，存在一定差异的可能性，可参考MySQL的官方文档。 ● 由于MySQL与GaussDB针对零时间（"0000-00-00 00:00:00"）内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。 ● 不支持两阶段事务。 ● 源端同步的表如果有AUTO_INCREMENT属性，结束任务时，DRS会自动刷新同步表的序列整型列所对应的GaussDB自增列（sequence）起始值，刷新值为该列的最大值+10000。

操作步骤

本小节以自建MySQL到GaussDB分布式版的实时同步为示例，介绍如何使用数据复制服务配置实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-35 同步任务信息

表 5-33 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-36 同步实例信息

表 5-34 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“GaussDB分布式版”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。

参数	描述
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考 产品架构和功能原理 。 <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-37 任务类型



表 5-35 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-38 企业项目和标签



表 5-36 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 5-39 源库信息

源库信息

不支持数据库所有参数迁移，DRS将源数据库的部分关键参数迁移至目标数据库，其他参数迁移请在目标数据库中[使用参数模板设定](#)

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 5-37 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none">最大支持上传500KB的证书文件。如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-40 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试成功

表 5-38 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	数据库用户名和密码将被系统加密暂存，直至该任务删除后清除。

步骤4 在“设置同步”页面，选择同步对象类型和同步对象，单击“下一步”。

图 5-41 同步模式

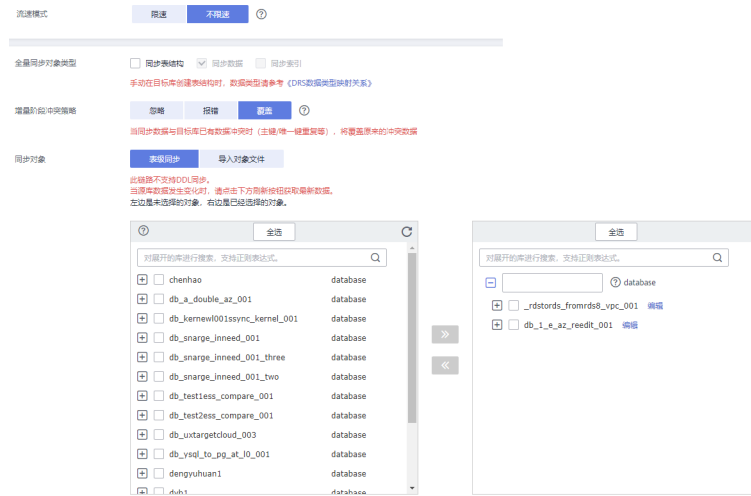



表 5-39 同步模式和对象

参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-42 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
<p>全量同步对象类型</p>	<p>全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。</p> <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 5-43 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-44 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-45 任务启动设置




表 5-40 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。

参数	描述
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.5 将 Oracle 同步到 Kafka

支持的源和目标数据库

表 5-41 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none">本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本）ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本）	<ul style="list-style-type: none">Kafka

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-42 环境要求

类型名称	使用限制 (DRS自动检查)
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 12c及以上版本租户模式： 12c及以上版本CDB数据库同步时，需要具有CREATE SESSION， SELECT ANY DICTIONARY， EXECUTE_CATALOG_ROLE， SELECT ANY TRANSACTION， LOGMINING权限， 以及针对单表的SELECT权限（ GRANT SELECT ON <userName.tbName> to drsUser; ）。 - 12c及以上版本PDB数据库同步时， 除了需要具有PDB内的 CREATE SESSION， SELECT ANY DICTIONARY， EXECUTE_CATALOG_ROLE， SELECT ANY TRANSACTION权限， 以及针对单表的SELECT权限（ GRANT SELECT ON <userName.tbName> to drsUser; ）， 还需要具有CDB的 CREATE SESSION， SELECT ANY DICTIONARY， EXECUTE_CATALOG_ROLE， SELECT ANY TRANSACTION， LOGMINING， SET CONTAINER（ GRANT SET CONTAINER TO <userName> CONTAINER=ALL; ） 权限。 - 12c及以上版本非租户模式： 需要具有CREATE SESSION， SELECT ANY DICTIONARY， EXECUTE_CATALOG_ROLE， SELECT ANY TRANSACTION， LOGMINING权限， 以及针对单表的SELECT权限（ GRANT SELECT ON <userName.tbName> to drsUser; ）。 - 11g及以下版本数据库同步时， 需要具有CREATE SESSION， SELECT ANY DICTIONARY， EXECUTE_CATALOG_ROLE， SELECT ANY TRANSACTION权限， 以及针对单表的SELECT权限（ GRANT SELECT ON <userName.tbName> to drsUser; ）。 - 增量同步时， 源库Oracle需要开启日志归档模式和最小补充日志， 所需同步表必须开启PK/UI或以ALL级别的补充日志， 不限制库级或表级补充日志方式， 若只开启表级补充日志， 重建或者RENAME表后需要重新设置； 请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步（ system/sys除外 ）， 因为该属性的账号无日志解析权限。

类型名称	使用限制 (DRS自动检查)
同步对象约束	<ul style="list-style-type: none"> • 支持表的同步，其他数据库对象暂不支持。 • 支持VARCHAR、VARCHAR2、NVARCHAR2、NUMBER、FLOAT、LONG、DATE、BINARY_FLOAT、BINARY_DOUBLE、CHAR、NCHAR、ROWID、TIMESTAMP、TIMESTAMP WITH TIME ZONE、TIMESTAMP WITH LOCAL TIME ZONE类型。 • 不支持接入的列类型：GEOMETRY以及自定义类型。 • 不支持同步但以过滤方式接入的列类型：INTERVAL_YEAR_TO_MONTH、INTERVAL_DAY_TO_SECOND、UROWID、BFILE、XML。 • 支持同步但默认过滤的列类型：RAW、BLOB、CLOB、NCLOB、LONG、LONG RAW。 • 增量同步LOB类型仅支持BasicFiles属性，不支持SecureFiles属性，支持的LOB类型大小限10M以内。 • 增量阶段源库为Oracle物理备库 (PHYSICAL STANDBY) 时不支持解析lob类型数据 (无法生成数据字典) ，如果增量同步的表中出现lob类型会导致增量同步异常。 • 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 • 不支持同步源库中的临时表。 • 不支持默认值含有表达式的函数的表的同步。 • 不支持同步源库中有虚拟列的表。 • 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。
源数据库要求	<ul style="list-style-type: none"> • 库名、表名不支持的字符有：非ASCII字符、“.”、“>”、“<”、“\”、“.”、“ ”、“,”、“?”、“!”、“”和“'”。 • 不支持选择源数据库的空库进行同步。 • 源库为RAC时，不支持增加、减少节点数量。 • 源库为RAC时，如果需要使用SCAN IP，需要drs node能够连接全部节点的VIP，否则无法通过连接检查。 • 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> • 目标库为社区Kafka。 • Kafka的auto.create.topics.enable参数建议为false。

类型名称	使用限制 (DRS自动检查)
操作须知	<ul style="list-style-type: none"> ● 如有中文、日文等特殊字符，业务连接Oracle数据库使用的编码需和Oracle数据库服务端编码一致，否则目标库会出现乱码。 ● Oracle中实时同步到kafka后的字符集为UTF8。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 附加日志级别为all或者pk+ui。 ● 日志中未出现的列在传递的消息中不会出现，表示该列未更新。 ● 同步过程中，不允许删除连接源和目标数据库的用户的用户名、密码、权限，或修改目标数据库的端口号。 ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 增量同步过程中，请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 单增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。

类型名称	使用限制 (DRS自动检查)
	<ul style="list-style-type: none"> 选择表级对象同步时，增量同步过程中不建议对表进行重命名操作。 选择表级对象同步时，仅支持所有表同步到目标端的同一个 Topic；选择导入对象文件时，支持不同表同步到目标端不同的 Topic。 支持表级DDL操作。 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-46 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-43 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-47 同步实例信息



表 5-44 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“Kafka”。
网络类型	此处以“公网网络”为示例。目前支持可选公网网络、VPC网络和VPN、专线网络。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<ul style="list-style-type: none"> 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。

- 任务类型

图 5-48 任务类型



表 5-45 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 5-49 可用区</p>

- 企业项目和标签

图 5-50 企业项目和标签



表 5-46 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

图 5-51 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 5-47 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用scanip接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

 说明

源数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-52 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

表 5-48 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择Topic和同步对象，单击“下一步”。

图 5-53 同步模式

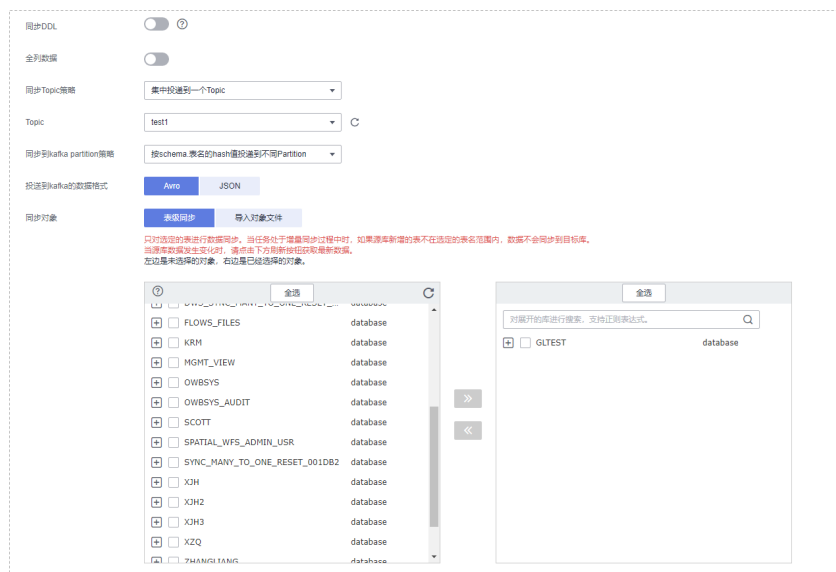



表 5-49 同步模式和对象

参数	描述
同步DDL	选择是否将DDL语句投递到Kafka中。若选择同步DDL，在选择partition策略为按照主键hash时，因DDL中无主键值，DDL会按照表名hash后投递到topic中，其余情况下投递方式与partition策略一致。
全列数据	选择是否需要单行的数据完整，DRS增量通过解析源库日志实现，单行数据是否完整取决于日志中是否记录了所有列的值。 若需要全列数据，则需要源库打开ALL级别的补全日志让日志中记录单行数据所有列的信息。该选项联动任务预检查阶段对源库补全日志级别的校验，DRS增量同步对于补全日志最低的要求为表级PK/UI级别的补全日志，补全日志检查方法可参考 源库Oracle补全日志检查方法 。

参数	描述
同步Topic策略	<p>同步Topic策略，可选择</p> <ul style="list-style-type: none"> 集中投递到一个Topic：适合源库业务量不大的场景。 按schema自动生成Topic名字：如果每个schema数据量很大，建议选择按schema自动生成Topic名字。 按schema-表名自动生成Topic名字：如果每张表数据量很大，建议选择schema-表名自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。
Topic名字格式	<p>Topic名字格式，同步Topic策略选择自动生成Topic名字时可见。</p> <p>由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"-"，若超过该范围会导致创建Topic失败，任务异常。</p> <p>若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。</p> <p>Topic名字格式支持schema和tablename两个变量，其他字符都当做常量。分别用\$schema\$代替模式名，\$tablename\$代替表名。</p> <p>例如：配置成 \$schema\$-\$tablename\$时，如果以oracle为源，模式名为schema1，表名为tab1，则Topic名字为schema1-tab1。</p>
Partition个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。</p>
副本个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。</p>
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按schema.表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 按主键的hash值投递到不同Partition：适用于一个表一个Topic的场景，避免该表都写到同一个分区，消费者可以并行从各分区获取数据。 对于无主键表，如果选择“按主键的hash值投递到不同的partition”策略时，同步任务则会使用默认“schema.表名的hash值投递到不同的partition”的策略进行同步。 按schema的hash值投递到不同Partition：适用于一个schema一个topic的场景，避免多个schema下的数据写到一个分区，消费者可以并行从各分区获取数据。 全部投递到Partition 0：数据固定发送到Topic的0号分区，任务默认多线程发送数据，能够保证表级保序，表与表之间不保序，并发写入性能比较差，若有强事务一致性要求，推荐选择此选项并找运维人员修改为单线程写Kafka或配置Topic策略按表名自动生成。

参数	描述
投送到kafka的数据格式	<p>选择Oracle投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro: 可以显示Avro二进制编码, 高效获取数据。 Json: 为Json消息格式。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 同步对象支持表级同步、导入对象文件, 您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 选择导入对象文件时, 支持不同表同步到目标端不同的Topic, 具体导入步骤和说明可参考导入同步对象。 使用导入对象功能时, 同步Topic策略选择“集中投递到一个Topic”, 才能使用对象名映射功能, 否则会按Topic名字格式生成。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面, 进行同步任务预校验, 校验是否可进行实时同步。

- 查看检查结果, 如有不通过的检查项, 需要修复不通过项后, 单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后, 且所有检查项结果均通过时, 单击“下一步”。

图 5-54 预检查



说明

所有检查项结果均通过时, 若存在请确认项, 需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-55 任务启动设置

The screenshot shows a configuration interface for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned on.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned on.
- * 任务异常自动结束时间** (Task Abnormal Auto End Time): An input field containing the number '14', with a help icon and a red note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-50 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.6 将 Oracle 同步到 GaussDB 主备版

支持的源和目标数据库

表 5-51 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建数据库 (Oracle 10g、11g、12c、18c、19c、21c版本) ECS自建数据库 (Oracle 10g、11g、12c、18c、19c、21c版本) 	GaussDB主备版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-52 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量和增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。</p> <p>12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> <p>11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> </p> - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>;

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 全量阶段不支持bfile、xml、sdo_geometry、urowid和自定义类型。增量阶段不支持bfile、xml、sdo_geometry、urowid、interval和自定义类型。 ● timestamp和interval day to second类型支持的最大精度是6。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 全量同步结构不支持位图索引、倒排索引、函数索引。 ● 增量同步LOB类型仅支持BasicFiles属性，不支持SecureFiles属性，支持的LOB类型大小限10M以内。 ● 对于TIMESTAMP WITH TIME ZONE类型，根据目标库时区做转换后不得大于“9999-12-31 23:59:59.999999”。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会因为无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。 ● 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 ● 选择手动创建表结构时，目标库中的时间类型是否带有时区需要与源库中保持一致，否则可能会因为时区转换导致时间数据不一致。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“\`”、“ ”、“/”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是GaussDB主备版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。主键支持的类型可参考Oracle数据库->GaussDB数据库映射关系。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 当Oracle字符集是WE8MSWIN1252、WE8ISO8859P1时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 当每行的行内存储类型数据的总长度超过GaussDB的限制时（默认是8192字节），可能会导致任务失败。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 由于无主键表缺乏行的唯一性标志，网络不稳定时涉及少量重试，表数据存在少量不一致的可能性。 ● 如果源库和目标库字符集不一致，如源库是ZHS16GBK，目标库是UTF8，由于ZHS16GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 使用DRS同步表结构的场景下，同一个schema中，同步的表中，表、约束、索引等不能有忽略大小写后的同名对象，比如表"A"中有索引名"inx1"，表B中有索引名"a"，表A和索引"a"忽略大小写重名了，会导致结构同步失败。多个schema映射到一个schema的场景，源库的多个schema中也不能包含同名但字母大小写不同的表、约束和索引。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步表结构时，如果目标库存在与源库同名约束会导致建表失败。 ● 全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量或单增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果转换失败或执行失败会忽略错误，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify

类型名称	使用和操作限制
	<p>column以及truncate table的基本DDL，不支持默认值等的修改。</p> <ul style="list-style-type: none"> - 增量DDL新增的字段不会自动扩展字符长度。 ● 增量同步阶段，修改抓取任务的启动位点主要用于重新同步数据。 <ul style="list-style-type: none"> - 修改抓取位点之后，上一次对象级对比结果不会展示出来。 - 单独修改抓取任务的启动位点，会把该位点同步到回放任务的启动位点上，即回放任务的启动位点和抓取任务的启动位点一致，不影响用户单独修改回放任务的启动位点。 ● 全量和增量同步不支持隐藏列（UNUSED, INVISIBLE）。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。 ● 当双AZ任务发生切换（例如主任务AZ级故障）时，如果主任务延迟较大，或源数据库有长事务未提交，任务倒换后可能因源数据库日志被清理导致任务无法启动，或着因为延迟、长事务期间的DDL变更导致同步异常。 ● 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-56 同步任务信息

⚠ 1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。	
项目	<input type="text" value=""/>
* 任务名称	<input type="text" value="DRS-6435"/> ?
描述	<div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> <div style="text-align: right; font-size: x-small;">0/256</div> ?

表 5-53 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-57 同步实例信息



表 5-54 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	此处以公网网络为示例。目前支持可选公网网络和VPN、专线网络。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	全量，增量，全量+增量，三种选择类型。

- 任务类型

图 5-58 任务类型

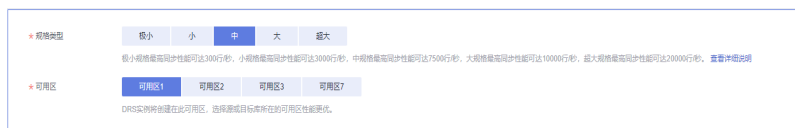


表 5-55 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更佳。

- 企业项目和标签

图 5-59 企业项目和标签



表 5-56 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过，单击“下一步”。

图 5-60 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICENAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

PDB名称 ?

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 5-57 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。

参数	描述
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

图 5-61 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接
✔ 测试成功

表 5-58 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 5-62 同步模式

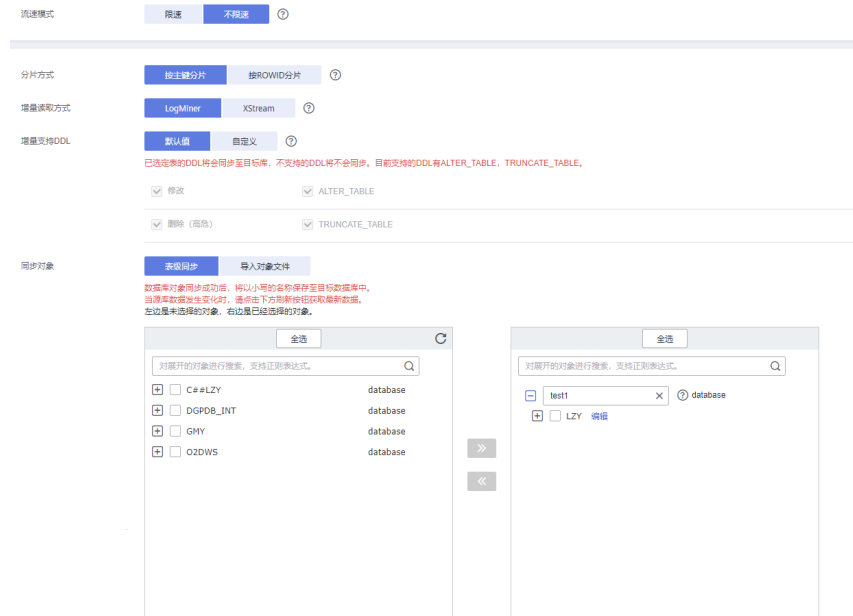



表 5-59 同步对象

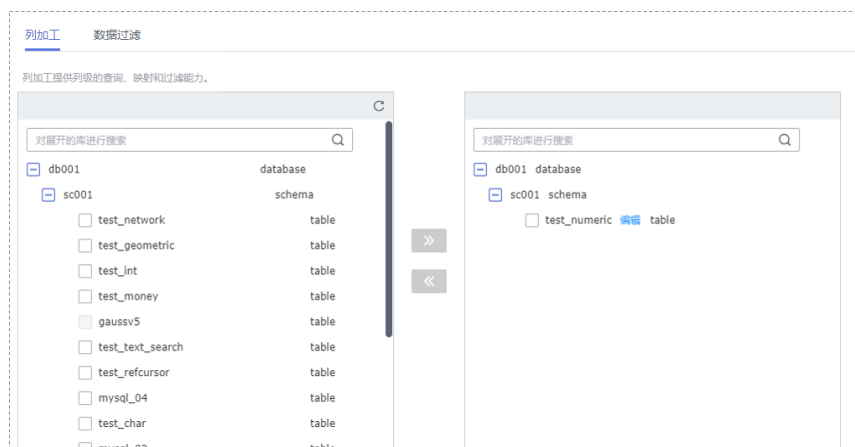
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-63 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
增量读取方式	<p>选择增量阶段读取日志方式。</p> <ul style="list-style-type: none"> LogMiner：采用Oracle官方接口读取重做日志文件，稳定性好。 Xstream：需要开启Xstream接口。 <p>目前建议使用LogMiner。</p>
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： 表级同步情况下，已选定的同步表DDL将会同步至目标，不支持的DDL将不同步。目前支持的DDL有CREATE_TABLE，ALTER_TABLE，TRUNCATE_TABLE。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的列加工，设置相关规则。

图 5-64 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-65 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-66 任务启动设置




表 5-60 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.7 将 Oracle 同步到 GaussDB 分布式版

支持的源和目标数据库

表 5-61 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> • 本地自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） • ECS自建数据库（Oracle 10g、11g、12c、18c、19c、21c版本） 	GaussDB分布式版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。

- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-62 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步：需要具有CREATE SESSION, SELECT ANY DICTIONARY, 针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 全量+增量和增量同步： <ul style="list-style-type: none"> 12c及以上版本租户模式： <p>12c及以上版本CDB数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> 12c及以上版本PDB数据库同步时，除了需要具有PDB内的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) ，还需要具有CDB的CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING, SET CONTAINER (GRANT SET CONTAINER TO <userName> CONTAINER=ALL;) 权限。 12c及以上版本非租户模式： <p>需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION, LOGMINING权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。</p> 11g及以下版本数据库同步时，需要具有CREATE SESSION, SELECT ANY DICTIONARY, EXECUTE_CATALOG_ROLE, SELECT ANY TRANSACTION权限，以及针对单表的SELECT权限 (GRANT SELECT ON <userName.tbName> to drsUser;) 。 - 增量同步时，源库Oracle需要开启日志归档模式和最小补充日志，所需同步表必须开启PK/UI或以ALL级别的补充日志，不限制库级或表级补充日志方式，若只开启表级补充日志，重建或者RENAME表后需要重新设置；请确保以上配置在同步过程中始终保持开启状态。 - 12c及以上版本不支持使用ORACLE_MAINTAINED=Y的用户账号进行增量同步 (system/sys除外) ，因为该属性的账号无日志解析权限。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有Sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 授权语句：GRANT CREATE, CONNECT ON DATABASE <database> TO <user>;

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - SCHEMA级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 授权语句：GRANT CREATE, USAGE ON SCHEMA <schema> TO <user>; - 表级权限：需要使用 root、或其他有Sysadmin角色的 DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下表的DML相关权限（SELECT权限在处理无主键表时需要）。 授予SCHEMA下所有表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON ALL TABLES IN SCHEMA <schema> TO <user>; 授予SCHEMA下指定表的DML权限：GRANT SELECT, UPDATE, INSERT, DELETE, INDEX, ALTER ON TABLE <schema.table> TO <user>; - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。

类型名称	使用和操作限制
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 全量阶段不支持bfile, xmltype、sdo_geometry、urowid和自定义类型。增量阶段不支持XMLTYPE、bfile, xmltype、sdo_geometry、urowid、interval year to month、interval day to second和自定义类型。 ● 增量阶段源库为Oracle物理备库（PHYSICAL STANDBY）时不支持解析lob类型数据（无法生成数据字典），如果增量同步的表中出现lob类型会导致增量同步异常。 ● 增量阶段不支持Oracle字符集扩展的字符，标准字符集无法解析Oracle自定义扩展字符。 ● 增量同步LOB类型仅支持BasicFiles属性，不支持SecureFiles属性，支持的LOB类型大小限10M以内。 ● 全量同步结构不支持位图索引、倒排索引、函数索引。 ● timestamp和interval day to second类型支持的最大精度是6。 ● 增量同步时，源端或者目标端数据库异常会触发任务失败，数据库恢复后重试启动任务时触发全局启动，此时会忽略原有抓取或回放组件的状态，同时回放也会按照抓取的中断点位再启动。 ● 对于TIMESTAMP WITH TIME ZONE类型，根据目标库时区做转换后不得大于“9999-12-31 23:59:59.999999”。 ● 源库支持to_date和sys_guid函数做默认值。将其他函数作为default值时，需要目标库也有相同功能的函数。对于目标库不存在对应函数的情况，可能会出现以下结果： <ul style="list-style-type: none"> - 默认值函数可能会被置空。 - 创建表失败，导致对象对比不一致或者任务失败。 ● 如果表中只有LOB字段，可能出现数据不一致性情况。 ● 如果Oracle中使用LOB类型各自的empty函数写入数据时，通过JDBC查询出来的值是空字符串，写入到目标库后是空字符串还是NULL取决于目标库对空字符串值的处理。 ● 针对无主键且无索引的表，非大字段的列必须大于3列，否则会无法全列匹配导致增量异常。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中有虚拟列的表。 ● 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 ● 选择手动创建表结构时，目标库中的时间类型是否带有时区需要与源库中保持一致，否则可能会因为时区转换导致时间数据不一致。

类型名称	使用和操作限制
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“`”、“ ”、“/”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：ZHS16GBK、AL32UTF8、UTF8、US7ASCII、WE8MSWIN1252、WE8ISO8859P1、WE8ISO8859P2、WE8ISO8859P4、WE8ISO8859P5、WE8ISO8859P7、WE8ISO8859P9、WE8ISO8859P13、WE8ISO8859P15。
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是GaussDB分布式版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 <ul style="list-style-type: none"> - 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 - 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。

类型名称	使用和操作限制
操作须知	<ul style="list-style-type: none"> ● 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，同时也会导致内容比对不可用。主键支持的类型可参考Oracle数据库->GaussDB数据库映射关系。 ● Oracle归档日志文件大小必须大于单条数据最大值，避免单条数据日志跨文件（超过2个日志文件）导致的增量数据解析异常。 ● 当Oracle字符集是WE8MSWIN1252、WE8ISO8859P1时，CLOB列同步到目标库可能出现乱码，建议先修改源库字符集为AL32UTF8再同步数据。 ● 当使用PDB数据库同步时，由于Oracle LogMiner组件的限制，增量同步时必须打开全部PDB。 ● Oracle 12.2及以上版本，由于Oracle LogMiner组件的限制，增量同步不支持Schema名、表名或列名超过30个字符。 ● 当每行的行内存储类型数据的总长度超过GaussDB的限制时（默认是8192字节），可能会导致任务失败。 ● 对于Oracle RAC集群，建议使用SCAN IP+ SERVICE_NAMES方式创建任务，SCAN IP具有更强的容错性，更好的负载能力，更快的同步体验。 ● 源库为Oracle RAC时，如果需要使用SCAN IP配置DRS任务，需要保证SCAN IP、DRS节点的IP同时能与源库的所有VIP互通（Oracle内部机制），否则无法通过连通性检查。若不使用SCAN IP，可以使用某一节点的VIP，这种情况下DRS日志解析只会在VIP指定的RAC节点上进行。 ● 若源库为RAC，增量同步首次启动时所有RAC节点必须正常在线，否则增量启动会出现异常。 ● 若源库为RAC，增量同步时，不支持增加、减少节点数量，避免导致增量同步异常（为保证数据的强一致性）。 ● 无主键表缺乏行的唯一性标志，网络不稳定时涉及重试，表数据存在少量不一致的可能性。 ● 如果源库和目标库字符集不一致，如源库是ZHS16GBK，目标库是UTF8，由于ZHS16GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 使用DRS同步表结构的场景下，同一个schema中，同步的表中，表、约束、索引等不能有忽略大小写后的同名对象，比如表"A"中有索引名"inx1"，表B中有索引名"a"，表A和索引"a"忽略大小写重名了，会导致结构同步失败。多个schema映射到一个schema的场景，源库的多个schema中也不能包含同名但字母大小写不同的表、约束和索引。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> ● 同步过程中，禁止对Oracle源库做resetlogs操作，否则会导致数据无法同步且任务无法恢复。 ● 同步过程中，不支持LOB类型的rollback操作，否则会导致同步任务失败。 ● 同步过程中，不支持修改源数据库Oracle用户名（SCHEMA名），包括11.2.0.2之前版本通过修改USER\$字典表方式及11.2.0.2之后通过ALTER USER username RENAME TO new_username修改SCHEMA名称的场景。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致 ● 全量同步分为表结构同步（含索引）、实时同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入实时同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● DRS全量同步表结构时，源库中的char、varchar2类型长度在目标库会按照字节长自动扩大（因为目标库为字节长），至少扩大1.5倍。扩大倍数和源库目标库的字符集有关，例如同为UTF8的情况下，默认扩大3倍，同为GBK的情况下，默认扩大2倍。 ● 全量同步表结构时，如果目标库存在与源库同名约束会导致建表失败。 ● 全量同步表结构时，列名不能是CTID、XMIN、CMIN、XMAX、CMAX、TABLEOID、XC_NODE_ID、TID、OID、GS_TUPLE_UID、TABLEBUCKETID这些GaussDB禁止的字段，否则会导致建表失败。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表，增量同步时，源库跟分区表相关的操作，可能无法同步到目标库或者在目标库执行失败。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步时，不建议选择混合分区表，因为混合分区表中的外部分区数据变更不产生DML日志，DRS增量数据同步无法获取变更信息，会存在数据不一致的风险。 ● 增量同步时，增量解析时间类型仅支持在Oracle时间类型合法范围内的数据，超范围的数据可能导致任务异常，例如年份大于9999时会出现数据截断。 ● 全量+增量或单增量任务场景，不支持直接连PDB数据库，用户需要提供CDB的Service Name/SID、用户名和密码。 ● 增量同步过程中，支持部分DDL操作。异构数据库DDL转换需要分析语义，考虑语法兼容性，无法做到完美转换，所以目前仅支持部分DDL操作，仅在转换成功并在以下范围内时会进行同步。如果转换失败或执行失败会忽略错误，需要手动在目标库执行相应的DDL来规避。 <ul style="list-style-type: none"> - 表级同步支持alter table add column、alter table drop column、alter table rename column、alter table modify

类型名称	使用和操作限制
	<p>column以及truncate table的基本DDL，不支持默认值等的修改。</p> <ul style="list-style-type: none"> - 增量DDL新增的字段不会自动扩展字符长度。 ● 增量同步阶段，修改抓取任务的启动位点主要用于重新同步数据。 <ul style="list-style-type: none"> - 修改抓取位点之后，上一次对象级对比结果不会展示出来。 - 单独修改抓取任务的启动位点，会把该位点同步到回放任务的启动位点上，即回放任务的启动位点和抓取任务的启动位点一致，不影响用户单独修改回放任务的启动位点。 ● 全量和增量同步不支持隐藏列（UNUSED, INVISIBLE）。 ● 当目标库为GaussDB分布式版时，增量同步阶段源库Oracle对分布列的update操作会导致在GaussDB数据库执行失败，造成数据不一致，所以业务上要尽量规避对分布列的更新。 ● 建议目标库禁用全局二级索引（GSI），否则可能会导致增量同步失败。 ● 任务再编辑增加新表时，请确保新增的表的事务都已提交，否则未提交的事务可能无法同步到目标库。建议在业务低峰期做增加表的操作。 ● lob类型及扩展字符类型（字节长度超过4000）不建议作为增量数据过滤条件，Oracle日志中可能出现不记录update旧值的情况。 ● 当双AZ任务发生切换（例如主任务AZ级故障）时，如果主任务延迟较大，或源数据库有长事务未提交，任务倒换后可能因源数据库日志被清理导致任务无法启动，或着因为延迟、长事务期间的DDL变更导致同步异常。 ● 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-67 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置先源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-63 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-68 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置或选择，需要重新创建迁移任务。

* 数据流动方向: 入云 出云 自建-自建 ?

* 源数据库引擎: MySQL Oracle GaussDB 主备版 PostgreSQL

* 目标数据库引擎: GaussDB 分布式版 GaussDB 主备版 Kafka

* 网络类型: 公网网络 ?
 我同意自动为实例绑定弹性公网IP，直到该任务结束后自动删除该弹性公网IP。

* VPC: [下拉菜单] C 查看虚拟私有云

同步实例所在子网: [下拉菜单] ? 查看子网 查看已占用的IP地址

* 内网安全组: [下拉菜单] C

* 同步模式: 全量 + 增量 全量 增量 ?

该模式为数据持续性实时同步，通过全量过程完成目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

表 5-64 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“Oracle”。
目标数据库引擎	选择“GaussDB分布式版”。

参数	描述
网络类型	此处以公网网络为示例。目前支持可选公网网络和VPN、专线网络。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	全量，增量，全量+增量，三种选择类型。

- 任务类型

图 5-69 任务类型



表 5-65 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更佳。

- 企业项目和标签

图 5-70 企业项目和标签



表 5-66 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

图 5-71 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name

PDB名称 ⓘ

数据库用户名

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

表 5-67 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCAN IP接入，提高访问性能。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
PDB名称	PDB同步仅在Oracle12c及以后的版本支持，该功能为选填项，当需要迁移PDB中的表时开启。 PDB功能开启后，只能迁移该PDB中的表，并且需要提供CDB的service name/sid及用户名和密码，不需要PDB的用户名和密码。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

图 5-72 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接
✔ 测试成功

表 5-68 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 5-73 同步模式

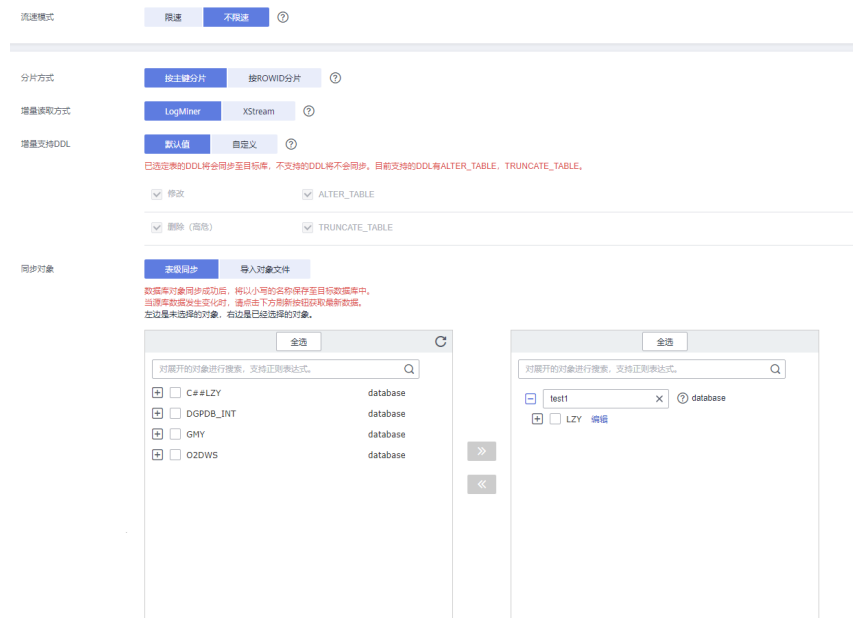



表 5-69 同步对象

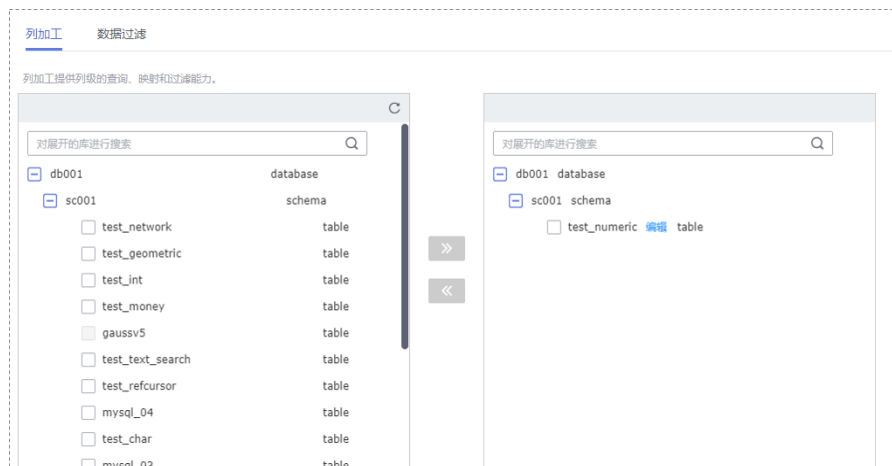
参数	描述
<p>流速模式</p>	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-74 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在建任务后修改流速模式。具体方法请参见修改流速模式。
<p>分片方式</p>	<p>可选“按主键分片”或“按ROWID分片”，根据实际需求进行选择。</p> <ul style="list-style-type: none"> “按主键分片”： 有主键表利用主键值分片，分片效率低于按ROWID分片方式。 “按ROWID分片”： 针对无主键大表分片更友好，分片效率更高。 <p>选择“按ROWID分片”的使用限制：全量期间限制源端导致ROWID变化的操作，例如Export/import of the table、ALTER TABLE XXXX MOVE、ALTER TABLE XXXX SHRINK SPACE、FLASHBACK TABLE XXXX、Splitting a partition、Updating a value so that it moves to a new partition、Combining two partitions等，否则可能导致数据不一致，限制方法：ALTER TABLE XXXX DISABLE ROW MOVEMENT。</p>

参数	描述
增量读取方式	<p>选择增量阶段读取日志方式。</p> <ul style="list-style-type: none"> LogMiner：采用Oracle官方接口读取重做日志文件，稳定性好。 Xstream：需要开启Xstream接口。 <p>目前建议使用LogMiner。</p>
增量支持DDL	<p>选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”：表级同步情况下，已选定的同步表DDL将会同步至目标，不支持的DDL将不同步。目前支持的DDL有CREATE_TABLE，ALTER_TABLE，TRUNCATE_TABLE。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的列加工，设置相关规则。

图 5-75 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-76 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-77 任务启动设置

启动时间 立即启动 稍后启动 ?

任务异常通知设置 ?

* SMN主题 ?

时延阈值(s) ?

数据异常通知


* 任务异常自动结束时间 ? 任务处于异常状态一段时间后，将会自动结束。单位为天。

表 5-70 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.8 将 PostgreSQL 同步到 Kafka

支持的源和目标数据库

表 5-71 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本） ECS自建数据库（PostgreSQL 9.4、9.5、9.6、10、11、12、13、14版本） 	<ul style="list-style-type: none"> Kafka 0.11及以上版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-72](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-72 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> ● 不支持实例级同步：每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持的字段类型： 数字类型、货币类型、字符类型、二进制数据类型、日期/时间类型、布尔类型、枚举类型、几何类型、网络地址类型、位串类型、文本搜索类型、UUID类型、XML类型、JSON类型、数组、复合类型、范围类型。 ● 增量同步的范围： <ul style="list-style-type: none"> - 支持部分DML（Data Manipulation Language）：包括INSERT、UPDATE、DELETE。 - 不支持：DDL（Data Definition Language），无日志表（UNLOGGED TABLE）的DML，临时表的DML。 <p>注意 同步对象的名称有如下限制： 库名不可以包含+"%\<>，模式名和表名不可以包含'!\<>，列名不可以包含"和'。</p>

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考表 5-73进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-73 数据库账号权限

类型名称	增量同步
源数据库连接账号	数据库的CONNECT权限，模式的USAGE权限，表的SELECT权限，REPLICATION连接权限。

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。

- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-74 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> • 源数据库要求： <ul style="list-style-type: none"> 源数据库参数wal_level必须配置为logical； 源数据库需提前安装test_decoding插件； 源数据库中无主键表的replica identity属性必须为full； 源数据库的max_replication_slots参数值必须大于当前已使用的复制槽数量； 源数据库的max_wal_senders参数值必须等于或大于max_replication_slots参数值。 源数据库中表的主键列toast属性为main、external、extended时，其replica identity属性必须为full。 • 目标数据库要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 • 其他使用须知： <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 库级同步时，增量同步中的任务，不支持“编辑”操作。 - 表级同步时，增量同步中的任务，支持“编辑”操作，但是不支持更换数据库。 - 在启动任务前，请确保源库中未启动长事务，源库启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 在任务启动后，仅支持以下版本的源库发生主备倒换：源库为RDS for PostgreSQL 12.6及以上的小版本、RDS for PostgreSQL 13及以上版本的所有小版本。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿删除源数据库表的主键，否则可能导致增量数据缺失或任务失败。 请勿修改源数据库中表的replica identity属性，否则可能导致增量数据缺失或任务失败。 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 库级同步时，源库新增无主键表时，请务必同时将该无主键表的replica identity属性设置为full，然后再写入数据，否则可能导致增量数据缺失或任务失败。 库级同步时，源库新增主键表时，如果主键列toast属性为main、external、extended时，请务必同时将该表的replica identity属性设置为full，然后再写入数据，否则可能导致数据不一致或任务失败。 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。
结束任务须知	<ul style="list-style-type: none"> 正常结束任务： <ul style="list-style-type: none"> 全量+增量实时同步任务正常结束时，将自动删除任务在源库创建的流复制槽。 强制结束任务： <ul style="list-style-type: none"> 强制结束全量+增量实时同步任务，需要手动删除源库可能残留的复制槽，可参考PostgreSQL为源强制结束任务章节的指导进行操作。 逻辑复制槽命名规则为“drs_唯一性标识”，其中唯一性标识需要将任务node id中的“-”替换为“_”获取。node id可在DRS任务同步日志页面查看task node id is ***日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-78 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 5-75 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-79 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎操作，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向: 入云 出云 自建-自建 ?

* 源数据库引擎: MySQL Oracle GaussDB 主副本 PostgreSQL

* 目标数据库引擎: GaussDB 分布式版 GaussDB 主副本 Kafka

* 网络类型: 公网网络 ?
DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除该弹性公网IP。

* VPC: [Dropdown] 查看虚拟私有云

同步实例所在子网: [Dropdown] 查看子网

* 内网安全组: [Dropdown] C

* 同步模式: 增量 ?

* 可用区: 可用区1 可用区2 可用区3 可用区4
DRS实例将创建在此可用区，选择源库或目标库所在的可用区进行变更。

表 5-76 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“PostgreSQL”。

参数	描述
目标数据库引擎	选择“Kafka”。
网络类型	此处以公网网络为示例。目前支持可选公网网络和VPN、专线网络。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<ul style="list-style-type: none"> 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。

- 任务类型

图 5-80 任务类型



表 5-77 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-81 企业项目和标签



表 5-78 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 5-82 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

端口

指定数据库名称

数据库名称

数据库用户名

数据库密码 🗨

SSL安全连接

表 5-79 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
指定数据库名称	选择是否指定数据库，开启后需手动输入数据库名称。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
SSL安全连接	通过该功能，用户可以选择是否开启对同步链路的加密。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-83 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

连接方式 ▼

✔ 测试成功

表 5-80 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
连接方式	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 5-84 同步模式

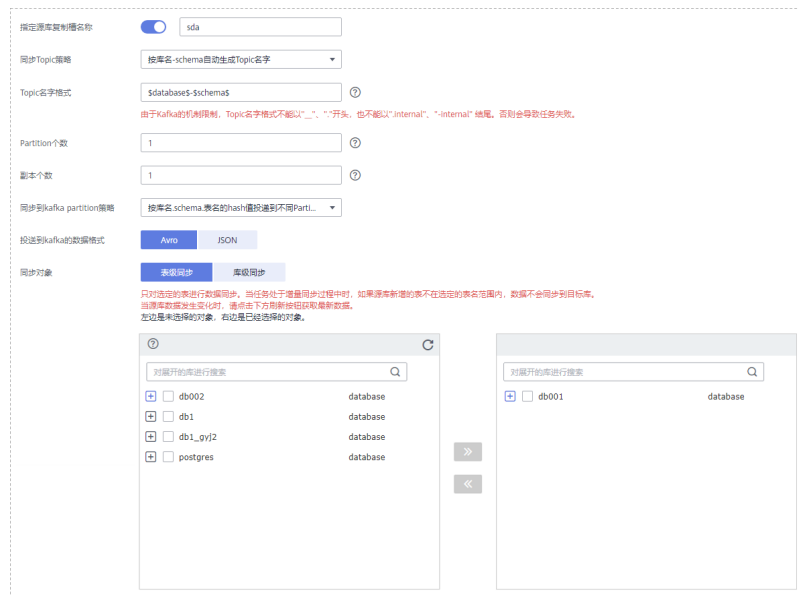



表 5-81 同步对象

参数	描述
指定源库复制槽名称	可选择是否指定源库复制槽，打开后需手动输入复制槽的名称（只能包含小写字母、数字、下划线，长度最大为63，且不能以数字开头）。
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。

参数	描述
Topic名字格式	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"-", 若超过该范围会导致创建Topic失败，任务异常。</p> <p>若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。</p>
Partition个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。</p>
副本个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。</p>
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名.schema.表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 按主键的hash值投递到不同Partition：适用于一个表一个Topic的场景，避免该表都写到同一个分区，消费者可以并行从各分区获取数据。 对于无主键表，如果选择“按主键的hash值投递到不同的partition”策略时，同步任务则会使用默认“按库名.schema.表名的hash值投递到不同的partition”的策略进行同步。 按库名.schema的hash值投递到不同Partition：适用于一个database一个topic的场景，避免多个schema下的数据写到一个分区，消费者可以并行从各分区获取数据。 全部投递到Partition 0：适用于有事务要求的场景，事务保序，可以保证完全按照事务顺序消费，写入性能比较差，如果没有强事务要求，不推荐使用此选项。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。 JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。 <p>详细格式可参考Kafka消息格式。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-85 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 5-86 任务启动设置

表 5-82 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.9 将 GaussDB 主备版同步到 MySQL

支持的源和目标数据库

表 5-83 支持的数据库

源数据库	目标数据库
GaussDB主备版	<ul style="list-style-type: none"> • 本地自建MySQL数据库 5.5、5.6、5.7版本 • ECS自建MySQL数据库 5.5、5.6、5.7版本 • 其他云上MySQL数据库 5.5、5.6、5.7版本

说明

目前仅支持白名单用户使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-84](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-84 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\\"'\ ?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-85](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-85 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。
目标数据库连接用户	<p>需要具备以下基本权限： INSERT、DELETE、UPDATE、SELECT、SHOW DATABASES。</p>		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-86 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库必须有足够的磁盘空间。 - 目前支持源库为UTF8字符集，目标库为UTF8MB4字符集。当源库为UTF8字符集，目标库为UTF8MB3字符集，或者当源库为GBK或GB18030字符集，目标库为GBK、GB2312、GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 目标库必须有足够的磁盘空间。 - 同步前保证目标库已创建好对应的database。 - 同步前保证目标库表结构已经建好，且与源端一致或包含源端所有列。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 源数据库和目标数据库表、schema等对象名称大小写要保持一致。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。 - GaussDB中字符类型、二进制类型数据中遇到0x00会截断，与MySQL行为不一致，可能导致最终数据不一致。 - 暂不支持binary类型作为DML匹配条件，即不支持binary类型作为主键。binary为定长类型，会自动填充，GaussDB填充符是0x20，MySQL填充符是0x00。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持数据对比过程中做限速。 ● 由于MySQL与GaussDB针对零时间（"0000-00-00 00:00:00"）内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。

类型名称	使用和操作限制
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-87 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?

不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?

0/256

表 5-87 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-88 同步实例信息



表 5-88 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“MySQL”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。 网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-89 任务类型



表 5-89 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-90 企业项目和标签



表 5-90 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 5-91 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中用参数组修改参数，手工导出导入用户和作业。

容灾同步 ?

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例
请填写主备实例的所有IP

数据库用户名

数据库密码

表 5-91 源库信息

参数	描述
容灾同步	开启后表示源数据库是容灾集群部署形态，IP地址或域名必须填入容灾集群所有节点的信息。
IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例，且请填写所有主备实例的IP。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 5-92 目标库信息

目标库信息

IP地址或域名

端口

数据库用户名

数据库密码 

SSL安全连接

表 5-92 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址或域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。

参数	描述
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 5-93 同步模式

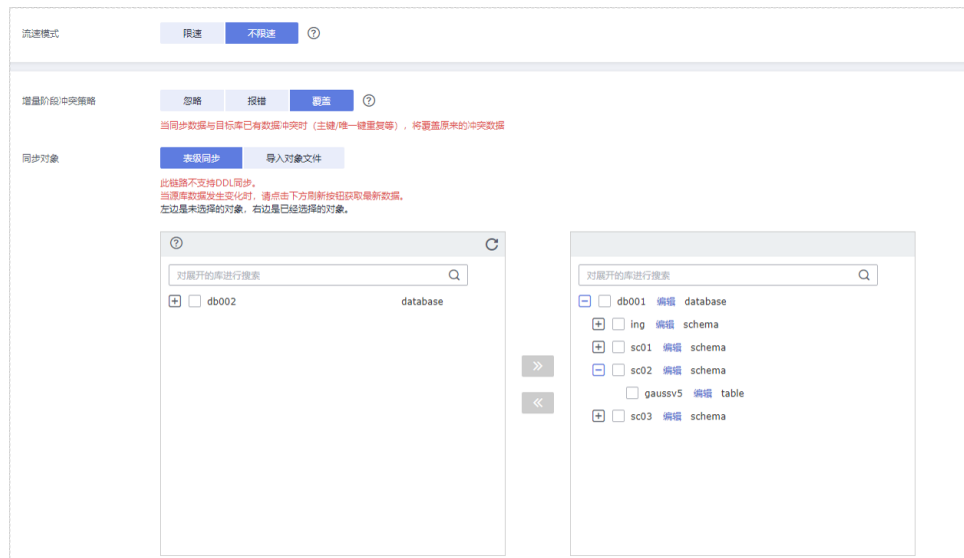



表 5-93 同步对象

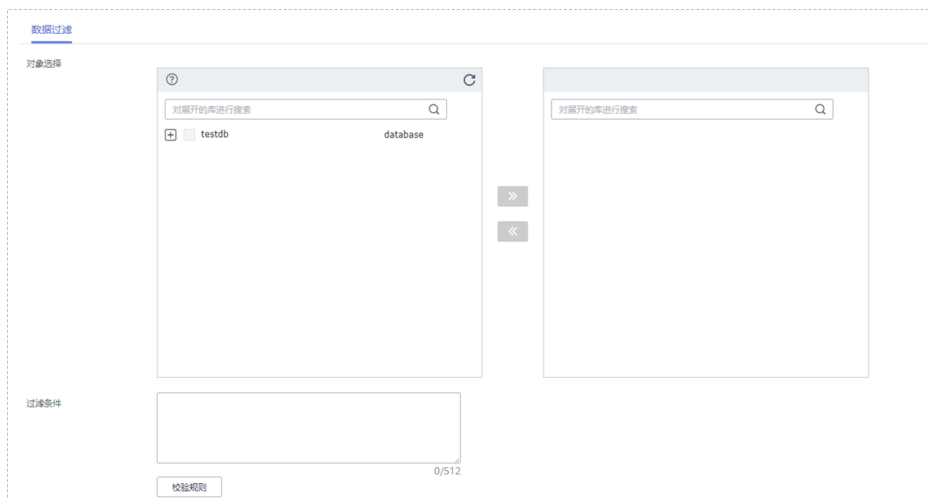
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-94 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 5-95 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-96 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-97 任务启动设置

启动时间 立即启动 稍后启动 ?

任务异常通知设置 ?

* SMN主题 ?

时延阈值(s) ?

数据异常通知


* 任务异常自动结束时间 ? 任务处于异常状态一段时间后，将会自动结束。单位为天。

表 5-94 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.10 将 GaussDB 主备版同步到 Oracle

支持的源和目标数据库

表 5-95 支持的数据库

源数据库	目标数据库
GaussDB主备版	<ul style="list-style-type: none"> • 本地自建Oracle数据库 11g、19c • ECS自建Oracle数据库11g、19c

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-96](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-96 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML及部分DDL。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-97](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-97 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、INSERT 权限，或者具备 RESOURCE角色。</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色。</p>

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-98 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<>\\'\ ?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目前支持源库为UTF8字符集，目标库为AL32UTF8、UTF8字符集；当源库为GBK或GB18030字符集，目标库为ZHS16GBK、ZHS16CGB231280、ZHS32GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的表结构，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 目标库必须有足够的磁盘空间。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 一般情况下，此链路为Oracle到GaussDB主备版同步链路的反向逃生链路，并与Oracle到GaussDB主备版同步链路配合使用，不建议单独使用此链路做其他场景的数据同步。 - 选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。如果此链路作为Oracle到GaussDB主备版的逃生链路使用，建议仅同步Oracle端为大写的schema名和表名，且GaussDB主备版端为小写的schema名和表名。 - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 进行同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 目标数据库Oracle不支持时间类型为1582-10-05至1582-10-14的日期，这种情况会导致同步前后数据不一致。 - 时间类型在源数据库GaussDB公元前闰年为1、5、9等年，在目标数据库Oracle公元前闰年为0、4、8等年，源数据库公元前闰年2月29日在目标数据库是不支持的日期格式，会因为日期解析失败导致同步失败。 - 如果同步表存在可空的联合唯一约束，由于在GaussDB和Oracle对含null值的约束范围差异，会导致最终数据不一致或任务失败。例如，同步表包含可空的联合唯一约束UNIQUE(C1,C2)，在GaussDB可以同时存在多条数据的c1='1'、c2=null，而在Oracle只能存在一条，再插入导致唯一键冲突。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。

类型名称	使用和操作限制
增量同步中过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 含大字段（blob、clob、nclob、long、long raw）的表，大字段不会作为增量数据where条件，建议保证除大字段外的其他字段具有唯一约束的能力。 ● 目标库的表中同时包含lob和long或者同时包含lob和long raw，插入或更新long或long raw的长度大于4000时，会出现插入或更新失败，建议使用clob或blob类型。 ● 不支持以下类型的DDL同步：ALTER TABLE MERGE PARTITIONS和CREATE INDEX CONCURRENTLY。 ● 同步DDL语句，需要源和目标库兼容模式一致。 ● 涉及操作表空间的DDL语句，需要目标库迁移用户具备操作表空间的权限。
数据加工须知	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键不可以被过滤。
同步对比使用须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> ● 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> ● 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 ● 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填写区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-98 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?
0/256

表 5-99 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-99 同步实例信息

同步实例信息

以下信息确认后不可修改，请谨慎填写，以免因为配置错误，需要重新创建迁移任务。

- 数据流动方向: 入云, 出云, **自建-自建**
- 源数据库引擎: MySQL, Oracle, DB2 for LUW, GaussDB 分布式版, **GaussDB 主备版**, PostgreSQL
- 目标数据库引擎: **Oracle**, Kafka
- 网络类型: 公网网络 (选中), 虚拟私有云
- VPC: vpc-xxxxxx (选中), 查看虚拟私有云
- 同步实例所在子网: 子网-xxxxxx (选中), 查看子网
- 内网安全组: 安全组-xxxxxx (选中)
- 同步模式: **全量 + 增量**, 全量, 增量

该模式为数据持续性实时同步，通过全量过程先将目标数据库初始化后，增量同步阶段通过解析日志等技术，将源端和目标数据库数据持续一致。

表 5-100 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“Oracle”。
网络类型	此处以公网网络为示例。目前支持可选“公网网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-100 任务类型

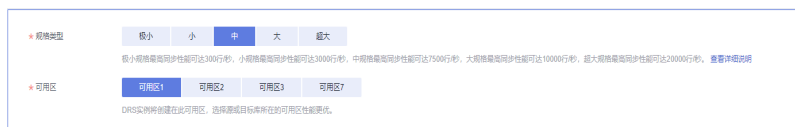


表 5-101 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-101 企业项目和标签



表 5-102 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-102 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中用参数组修改参数，手工导出导入用户和作业。

容灾同步 ?

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例
请填写主备实例的所有IP

数据库用户名

数据库密码

表 5-103 源库信息

参数	描述
容灾同步	开启后表示源数据库是容灾集群部署形态，IP地址或域名必须填入容灾集群所有节点的信息。
IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例，且请填写所有主备实例的IP。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 5-103 目标库信息

目标库信息

IP地址或域名

对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。

端口

数据库服务名 Service Name ?

数据库用户名

数据库密码

SSL安全连接

● 测试成功

表 5-104 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群，建议使用SCANIP+SERVICE_NAME方式创建任务，提高访问性能。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。

参数	描述
数据库服务名	数据库服务名（Service Name/SID），客户端可以通过其连接到Oracle，具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 5-104 同步模式

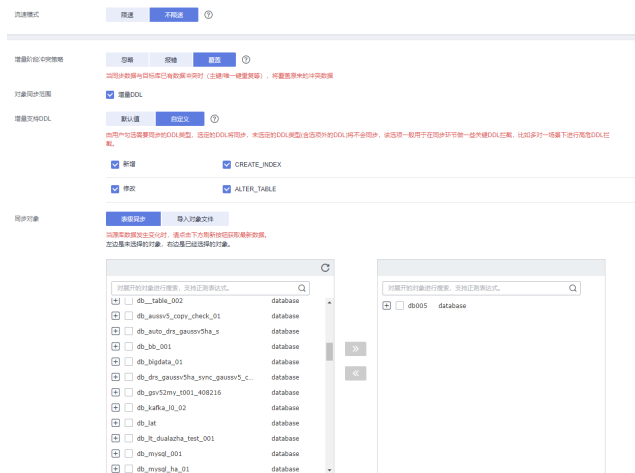



表 5-105 同步对象

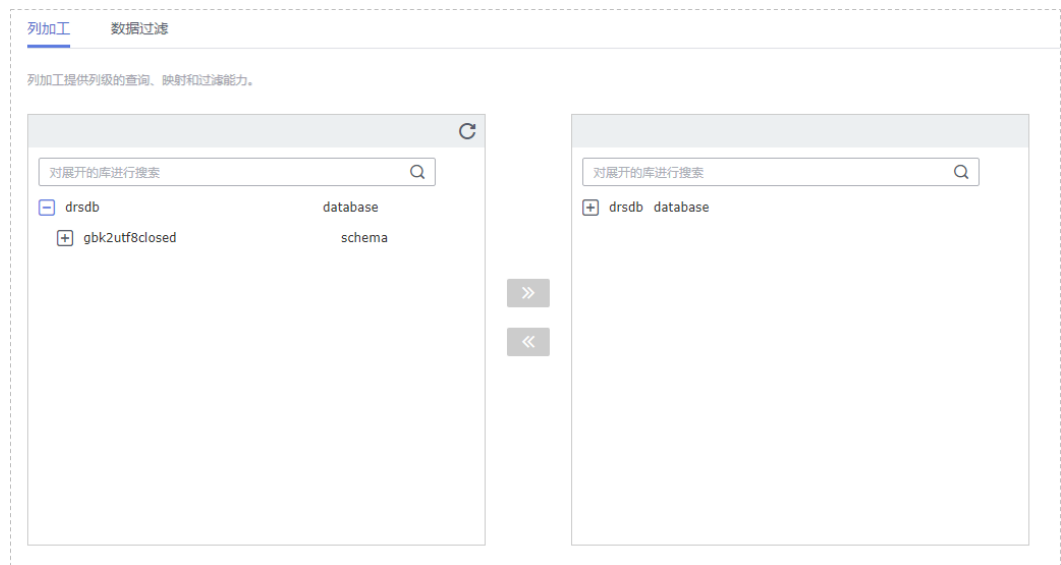
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-105 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
对象同步范围	选择增量阶段是否同步DDL。
增量支持DDL	<p>对象同步范围选择“增量DDL”时可见，用户可根据需求选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> ● “默认值”：已选定同步对象的DDL将会同步至目标，不支持的DDL将不同步。 <ul style="list-style-type: none"> - 表级支持的DDL：ALTER TABLE、CREATE INDEX。 ● “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准，未选定的DDL类型（含选项外的DDL）将不会同步。
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 5-106 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-107 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 5-108 任务启动设置

表 5-106 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.11 将 GaussDB 主备版同步到 Kafka

支持的源和目标数据库

表 5-107 支持的数据库

源数据库	目标数据库
GaussDB主备版	Kafka 0.11及以上版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-108](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-108 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持增量同步。 ● 支持的字段：BIGINT, BIT, BIT VARYING, BLOB, BOOLEAN, BYTEA, CHARACTER, CHARACTER VARYING, CLOB, DATE, DOUBLE PRECISION, INTEGER, MONEY, NUMBER, NUMERIC, NVARCHAR2, RAW, REAL, SMALLDATETIME, SMALLINT, TEXT, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, TINYINT。 ● 支持表级同步、schema级同步、库级同步： <ul style="list-style-type: none"> - 仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\\"'?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-109](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

📖 说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因为数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-109 数据库用户权限

类型名称	增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-110 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 选择增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 将无主键表的REPLICA IDENTITY属性设置为FULL，或给无主键表追加主键。 - 建议将主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\'\ ?和!。 ● 目标数据库要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 如果DRS实例类型选择主备类型，当DRS任务发生主备切换时，可能会产生无主键表数据重复或不一致的情况。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 选择表级对象迁移时，同步过程中不建议对表进行重命名操作。 ● 不支持interval partition表复制。 ● 不支持修改主键列名称。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> ● 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> ● 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 ● 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-109 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-111 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-110 同步实例信息

同步实例信息 ?

以下信息确认后可不修改，请谨慎操作，以免因配置错误，需要重新创建迁移任务。

* 数据流动方向: [单选按钮] 自建-自建 ?

* 源数据库引擎: MySQL Oracle DB2 for LUW GaussDB 主备版 PostgreSQL

* 目标数据库引擎: Oracle Kafka

* 网络类型: 公网网络 ?
 DRS将会创建与DRS实例同地域的公网IP，请谨慎操作并定期检查公网IP。

* DRS实例名称: A-Z

实例每个节点的部署策略，每个节点只创建一个角色，由具有高性价比、适用于业务数据量较小、短期任务，并符合实时同步要求的信息。

* VPC: [下拉菜单] C 查看公网IP地址

同步实例所在子网: [下拉菜单] ? 查看子网 查看已占用的IP地址

* 内网安全组: [下拉菜单] C

* 同步模式: [单选按钮] 同步 ?
增量同步通过解析日志数据，根据产生的增量数据同步至目标库，无需中断业务，实现同步过程中业务和数据库连续对外提供服务。

表 5-112 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“Kafka”。

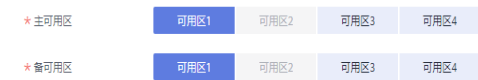
参数	描述
网络类型	此处以公网网络为示例。目前支持可选“公网网络”和“VPN、专线网络”。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<ul style="list-style-type: none"> - 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 <p>无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p>

- 任务类型

图 5-111 任务类型



表 5-113 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 5-112 可用区</p> 

- 企业项目和标签

图 5-113 企业项目和标签



表 5-114 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>

参数	描述
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

📖 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 5-114 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户、作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

容灾同步 ?

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例
请填写主备实例的所有IP

数据库用户名

数据库密码 🔒

表 5-115 源库信息

参数	描述
容灾同步	开启后表示源数据库是容灾集群部署形态，IP地址或域名必须填入容灾集群所有节点的信息。
IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例，且请填写所有主备实例的IP。
数据库用户名	源数据库的用户名。

参数	描述
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-115 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

连接方式

✔ 测试成功

表 5-116 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
连接方式	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 5-116 同步模式

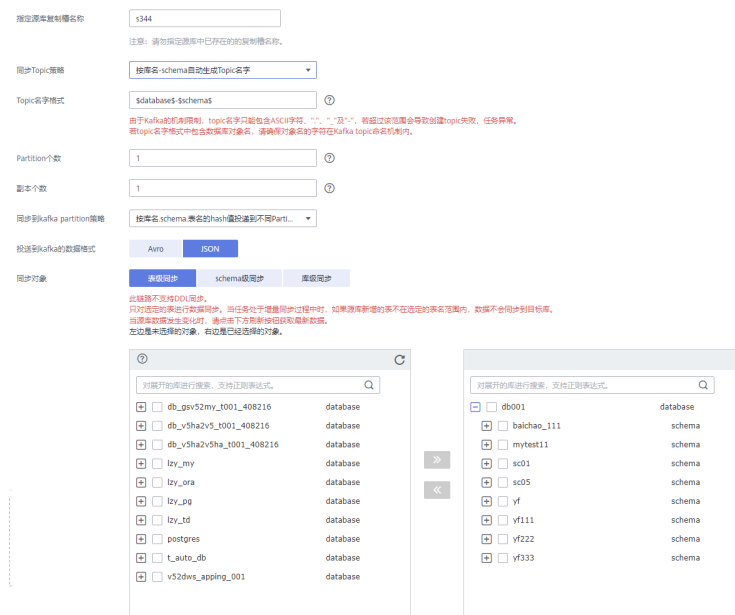



表 5-117 同步对象

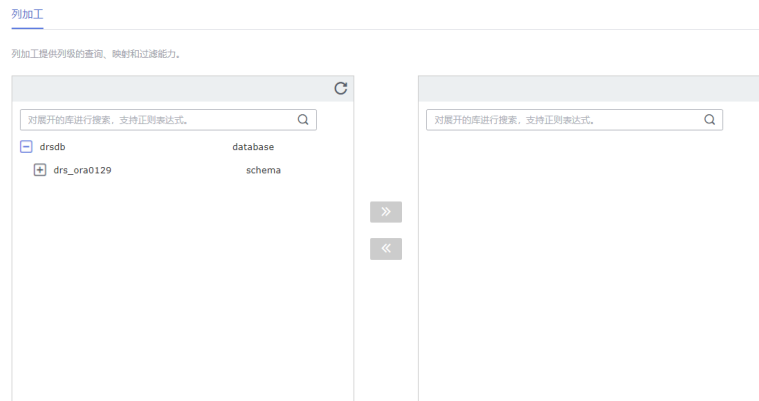
参数	描述
指定源库复制槽名称	可选择是否指定源库复制槽，打开后需手动输入复制槽的名称（只能包含小写字母、数字、下划线，长度63，且不能以数字开头）。
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。
Topic名字格式	同步Topic策略选择自动生成Topic名字时可见。 由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"-", 若超过该范围会导致创建Topic失败，任务异常。 若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。
Partition个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。
副本个数	同步Topic策略选择自动生成Topic名字时可见。 用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。

参数	描述
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名.schema.表名的hash值投递到不同Partition: 适用于单表的查询场景, 表内保序, 表与表之间不保序, 可以提高单表读写性能, 推荐使用此选项。 按主键的hash值投递到不同Partition: 适用于一个表一个Topic的场景, 避免该表都写到同一个分区, 消费者可以并行从各分区获取数据。 对于无主键表, 如果选择“按主键的hash值投递都不同的partition”策略时, 同步任务则会使用默认“按库名.schema.表名的hash值投递到不同的partition”的策略进行同步。 按库名.schema的hash值投递到不同Partition: 适用于一个database一个topic的场景, 避免多个schema下的数据写到一个分区, 消费者可以并行从各分区获取数据。 全部投递到Partition 0: 适用于有事务要求的场景, 事务保序, 可以保证完全按照事务顺序消费, 写入性能比较差, 如果没有强事务要求, 不推荐使用此选项。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro: 可以显示Avro二进制编码, 高效获取数据。 JSON: 为Json消息格式, 方便解释格式, 但需要占用更多的空间。 <p>详细格式可参考Kafka消息格式。</p>
同步对象	<p>左侧框是源数据库对象, 右侧框是已经选择的对象, 同步对象支持表级同步、schema级同步和库级同步, 您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索, 以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况, 请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时, 名称前后的空格不显示, 中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面, 选择需要加工的数据, 进行列加工。

- 如果不需要数据加工, 单击“下一步”。
- 如果需要进行列加工, 参考[数据加工](#)中的说明, 设置相关规则。

图 5-117 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-118 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-119 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-118 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.12 将 GaussDB 主备版同步到 GaussDB 分布式版

支持的源和目标数据库

表 5-119 支持的数据库

源数据库	目标数据库
GaussDB主备版	GaussDB分布式 说明 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-120](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-120 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\\?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-121](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-121 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none"> 需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。 目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。 		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-122 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 若选择增量或全量+增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\\?和!。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：c_+哈希值+原约束名（可能被截断）+_key。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 源数据库的唯一索引是表达式索引时，全量同步阶段目标数据库GaussDB分布式可能会因为无法创建该索引而导致任务失败。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
数据加工须知	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键不可以被过滤。 ● 表上存在列名映射或列过滤时，索引定义中的部分索引条件表达式将被忽略，普通索引中的表达式列将被排除，包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时，普通索引和外键中对应的列也将被过滤。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比，全量完成会自动触发对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-120 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 5-123 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-121 同步实例信息



表 5-124 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“GaussDB分布式版”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-122 任务类型



表 5-125 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-123 企业项目和标签



表 5-126 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-124 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中用参数组修改参数，手工导出导入用户和作业。

容灾同步 ?

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例
请填写主备实例的所有IP

数据库用户名

数据库密码

表 5-127 源库信息

参数	描述
容灾同步	开启后表示源数据库是容灾集群部署形态，IP地址或域名必须填入容灾集群所有节点的信息。
IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例，且请填写所有主备实例的IP。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-125 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接
✔
测试成功

表 5-128 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 5-126 同步模式

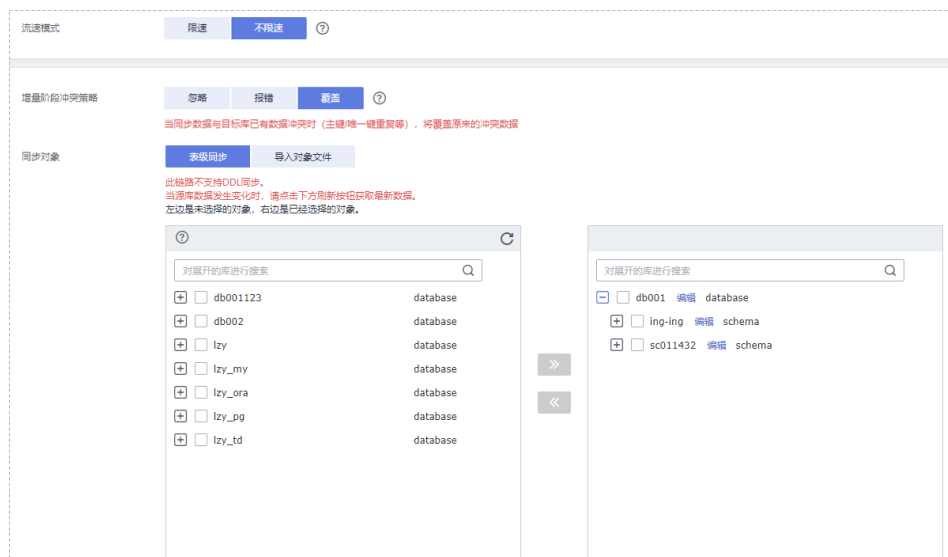




表 5-129 同步对象

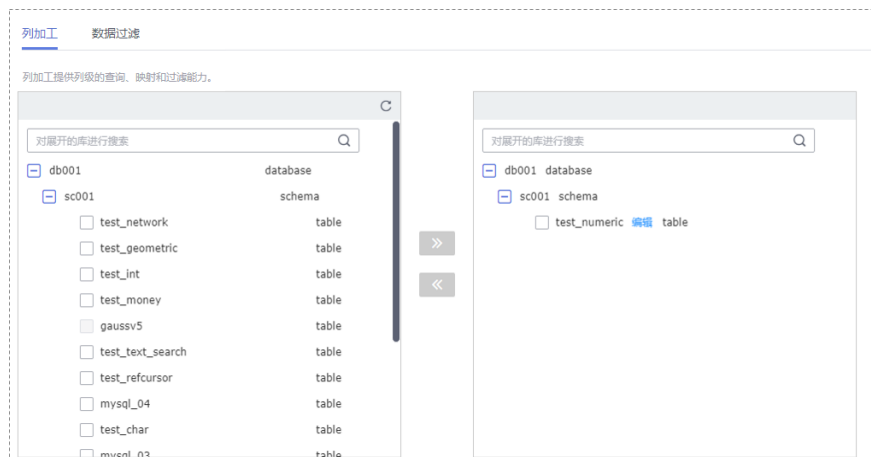
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-127 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和文件导入对象，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或数据过滤，可参考[数据加工](#)，设置相关规则。

图 5-128 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-129 预检查



📖 说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-130 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-130 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.13 将 GaussDB 主备版同步到 GaussDB 主备版

支持的源和目标数据库

表 5-131 支持的数据库

源数据库	目标数据库
GaussDB主备版	GaussDB主备版 说明 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-132](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-132 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\\?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-133](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-133 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none">需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-134 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 若选择增量或全量+增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\\?和!。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：c_+哈希值+原约束名（可能被截断）+_key。

类型名称	使用和操作约定
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
全量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
数据加工须知	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键不可以被过滤。 ● 表上存在列名映射或列过滤时，索引定义中的部分索引条件表达式将被忽略，普通索引中的表达式列将被排除，包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时，普通索引和外键中对应的列也将被过滤。
同步对比须知	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比，全量完成会自动触发对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB主备版为源强制结束任务章节的指导进行操作。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-131 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?
0/256

表 5-135 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-132 同步实例信息



表 5-136 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB主备版”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 任务类型

图 5-133 任务类型



表 5-137 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-134 企业项目和标签



表 5-138 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-135 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中用参数组修改参数，手工导出导入用户和作业。

容灾同步 ?

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例
请填写主备实例的所有IP

数据库用户名

数据库密码

表 5-139 源库信息

参数	描述
容灾同步	开启后表示源数据库是容灾集群部署形态，IP地址或域名必须填入容灾集群所有节点的信息。
IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例，且请填写所有主备实例的IP。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。

 说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-136 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接

✔

测试成功

表 5-140 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

📖 说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 5-137 同步模式

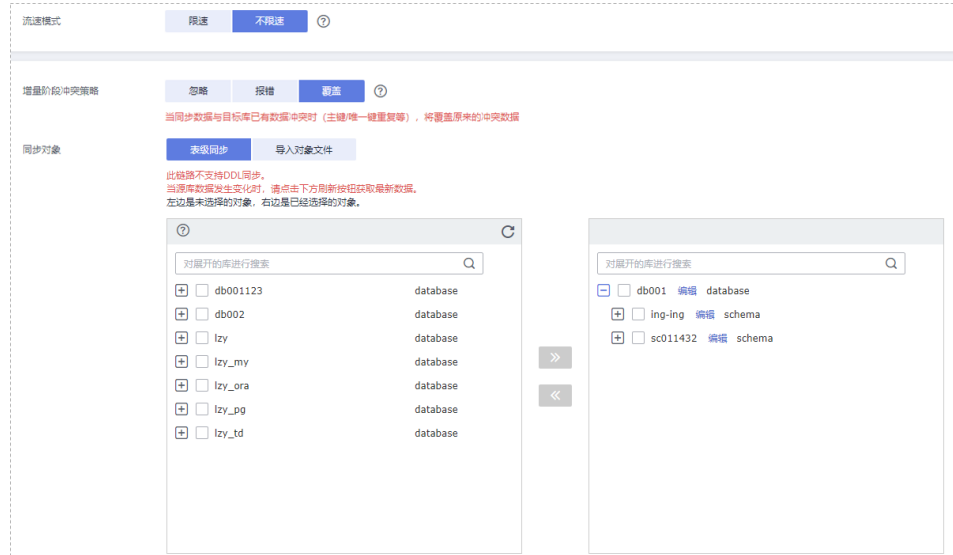




表 5-141 同步对象

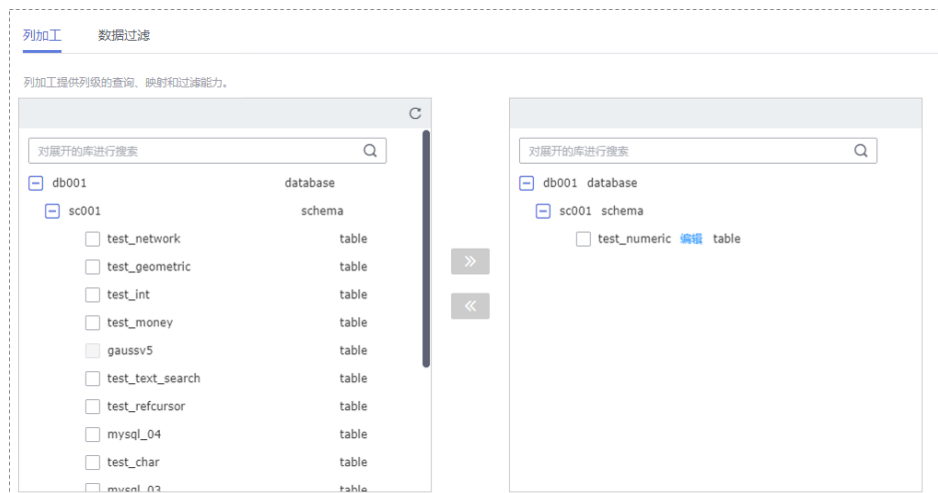
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-138 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步和文件导入对象，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或数据过滤，可参考[数据加工](#)中的列加工，设置相关规则。

图 5-139 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-140 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-141 任务启动设置

启动时间 立即启动 稍后启动 ?

任务异常通知设置 ?

* SMN主题 ?

时延阈值(s) ?

数据异常通知


* 任务异常自动结束时间 ? 任务处于异常状态一段时间后，将会自动结束。单位为天。

表 5-142 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.14 将 GaussDB 分布式版同步到 MySQL

支持的源和目标数据库

表 5-143 支持的数据库

源数据库	目标数据库
GaussDB分布式版	<ul style="list-style-type: none"> • 本地自建MySQL数据库 5.5、5.6、5.7版本 • ECS自建MySQL数据库 5.5、5.6、5.7版本 • 其他云上MySQL数据库 5.5、5.6、5.7版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-144](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-144 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量、增量、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-145](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 5-145 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户需要具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。
目标数据库连接用户	<p>需要具备以下基本权限： INSERT、DELETE、UPDATE、SELECT、SHOW DATABASES。</p>		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-146 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目前支持源库为UTF8字符集，目标库为UTF8MB4字符集。当源库为UTF8字符集，目标库为UTF8MB3字符集，或者当源库为GBK或GB18030字符集，目标库为GBK、GB2312、GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 目标库必须有足够的磁盘空间。 - 同步前保证目标库已创建好对应的database。 - 同步前保证目标库表结构已经建好，且与源端一致或包含源端所有列。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 若专属计算集群不支持4vCPU/8G或以上规格实例，则无法创建同步任务。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 源数据库和目标数据库表、schema等对象名称大小写要保持一致。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 在创建DRS任务之前，如果目标数据库实例开启并设置了SQL限流规则，可能会导致DRS任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库同步用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持数据对比过程中做限速。 ● 由于MySQL与GaussDB针对零时间 ("0000-00-00 00:00:00") 内置的处理方式不同，当进行内容对比时，如果源、目标数据库存在零时间，对比结果会显示数据不一致。

类型名称	使用和操作限制
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-142 同步任务信息

⚠ 1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域 ?

不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目

* 任务名称 ?

描述 ?

0/256

表 5-147 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-143 同步实例信息



表 5-148 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“MySQL”。
网络类型	此处以公网网络为示例。目前支持可选“公网网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<p>此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。
源端DN数量	需要和分布式源库DN（Data Node）节点数量保持一致。

- 任务类型

图 5-144 任务类型



表 5-149 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-145 企业项目和标签



表 5-150 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 5-146 源库信息

源库信息

CN IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

DN IP地址或域名 ?

IP地址或域名

待实例创建成功后再进行测试连接

表 5-151 源库信息

参数	描述
CN IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
DN IP地址或域名	DN IP信息可以登录CN节点通过SQL查询。 由于不同数据库实例的资源类型间存在差异，建议咨询数据库运维人员获取。

说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 5-147 目标库信息

目标库信息

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

表 5-152 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
端口	目标数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对迁移链路的加密。如果开启该功能，需要用户上传SSL CA根证书。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步对象，完成后单击“下一步”。

图 5-148 同步模式

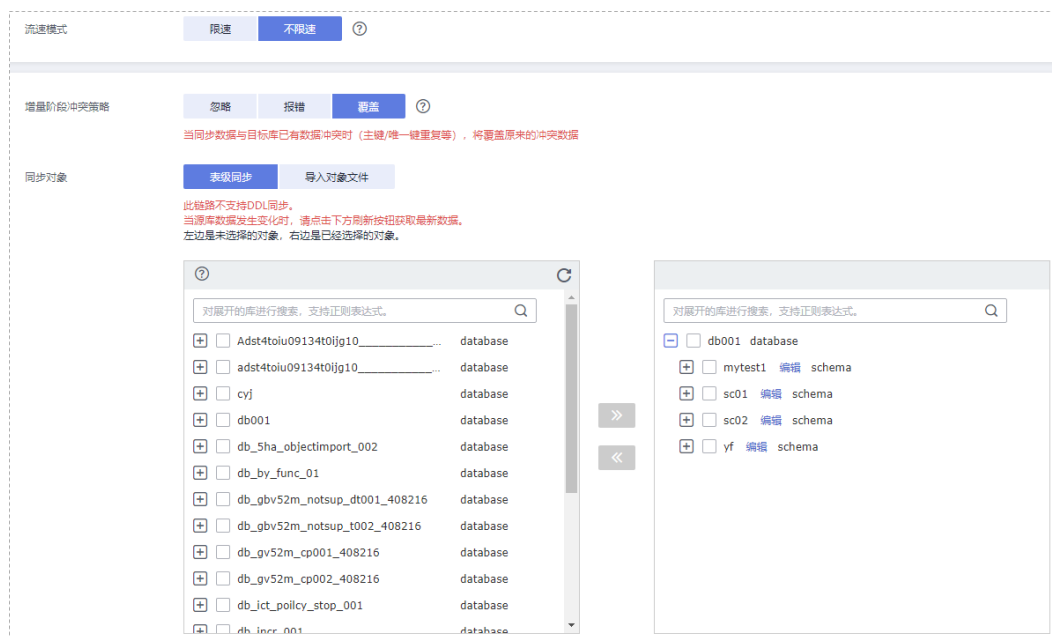




表 5-153 同步对象

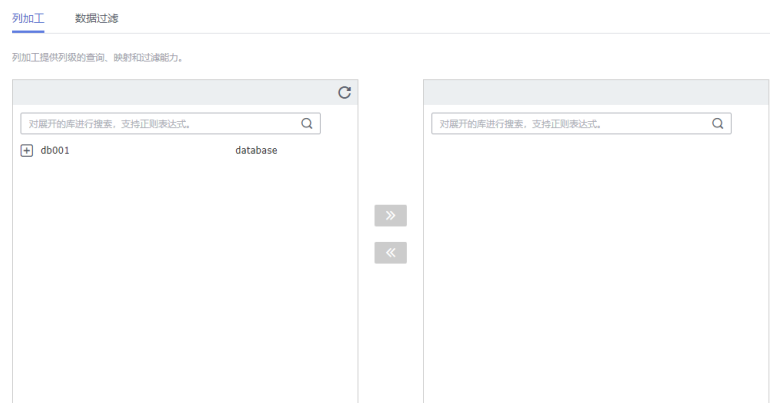
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-149 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 5-150 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-151 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 5-152 任务启动设置



表 5-154 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会出现一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.15 将 GaussDB 分布式版同步到 Oracle

支持的源和目标数据库

表 5-155 支持的数据库

源数据库	目标数据库
GaussDB分布式	<ul style="list-style-type: none">本地自建Oracle数据库11g、19cECS自建Oracle数据库11g、19c

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-156](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-156 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, MONEY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的数据。 - 增量同步过程仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，增量不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-157](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

 说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 5-157 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、INSERT 权限，或者具备 RESOURCE角色。</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色</p>	<p>用户需要具备DBA角色或者以下最小权限： CREATE SESSION 权限，表的 SELECT、 INSERT、 UPDATE、DELETE 权限，或者具备 RESOURCE角色。</p>

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量、单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-158 使用须知

类型名称	使用和操作限制
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步时: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库要同步的库名、schema名和表名不可以包含/<>\\'\ ?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目前支持源库为UTF8字符集，目标库为AL32UTF8、UTF8字符集；当源库为GBK或GB18030字符集，目标库为ZHS16GBK、ZHS16CGB231280、ZHS32GB18030字符集时，由于数据库差异，可能存在部分字符不兼容，导致部分数据乱码、数据同步不一致或目标库写入失败等问题。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的表结构，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 目标库必须有足够的磁盘空间。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 一般情况下，此链路为Oracle到GaussDB分布式版同步链路的反向逃生链路，并与Oracle到GaussDB分布式版同步链路配合使用，不建议单独使用此链路做其他场景的数据同步。 - 选择的源库表中不能存在表名称字母相同但大小写不同的表，否则，会导致同步失败。如果此链路作为Oracle到GaussDB分布式的逃生链路使用，建议仅同步Oracle端为大写的schema名和表名，且GaussDB分布式端为小写的schema名和表名。 - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。

类型名称	使用和操作限制
	<ul style="list-style-type: none"> - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 进行同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 - 目标数据库Oracle不支持时间类型为1582-10-05至1582-10-14的日期，这种情况会导致同步前后数据不一致。 - 时间类型在源数据库GaussDB公元前闰年为1、5、9等年，在目标数据库Oracle公元前闰年为0、4、8等年，源数据库公元前闰年2月29日在目标数据库是不支持的日期格式，会因为日期解析失败导致同步失败。 - 如果同步表存在可空的联合唯一约束，由于在GaussDB和Oracle对含null值的约束范围差异，会导致最终数据不一致或任务失败。例如，同步表包含可空的联合唯一约束UNIQUE(C1,C2)，在GaussDB可以同时存在多条数据的c1='1'、c2=null，而在Oracle只能存在一条，再插入导致唯一键冲突。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程中须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步中过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限于GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 含大字段（blob、clob、nclob、long、long raw）的表，大字段不会作为增量数据where条件，建议保证除大字段外的其他字段具有唯一约束的能力。 ● 目标库的表中同时包含lob和long或者同时包含lob和long raw，插入或更新long或long raw的长度大于4000时，会出现插入或更新失败，建议使用clob或blob类型。

类型名称	使用和操作限制
数据加工须知	<ul style="list-style-type: none"> 列加工时，主键、唯一键和分布列不可以被过滤。
同步对比使用须知	<ul style="list-style-type: none"> 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 不支持在全量同步过程中做对比。 不支持数据对比过程中做限速。
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-153 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-159 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-154 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎操作，以免因为配置错误，需要重新创建任务。

* 数据流动方向: [下拉菜单: 自建-自建] ?

* 源数据库引擎: MySQL, Oracle, DB2 for LUW, GaussDB 分布式版, GaussDB 主备版, PostgreSQL

* 目标数据库引擎: MySQL, Oracle, GaussDB(DWS), GaussDB 分布式版, GaussDB 主备版, Kafka

* 网络类型: [下拉菜单: 公网] ?
DB2将会自动为DB2实例创建公网IP，请在表格末尾将自动创建公网IP。

* VPC: [下拉菜单] ? 查看该VPC

* 同步实例所在子网: [下拉菜单] ? 查看子网

* 内网安全组: [下拉菜单] ?

* 同步模式: [按钮: 全量、增量] [按钮: 全量、增量] ?

* 源DB引擎: [输入框: 2] ?

表 5-160 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“Oracle”。

参数	描述
网络类型	此处以公网网络为示例。目前支持可选“公网网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。 “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。
源端DN数量	需要和分布式源库DN（Data Node）节点数量保持一致。

- 任务类型

图 5-155 任务类型



表 5-161 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-156 企业项目和标签



表 5-162 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-157 源库信息

The screenshot shows a web form titled "源库信息" (Source Database Information). It contains the following fields and elements:

- CN IP地址或域名**: A text input field with a help icon. Below it is a red warning message: "请确保所填写的多个IP地址或域名属于同一个实例" (Please ensure the multiple IP addresses or domains entered belong to the same instance).
- 数据库用户名**: A text input field.
- 数据库密码**: A password input field with a visibility toggle icon.
- DN IP地址或域名**: A section with a help icon, containing a blue header "IP地址或域名" with a copy icon, and two empty text input fields below it.
- 测试连接**: A button at the bottom left.
- 待实例创建成功后再进行测试连接**: A note at the bottom right.

表 5-163 源库信息

参数	描述
CN IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
DN IP地址或域名	DN IP信息可以登录CN节点通过SQL查询。 由于不同数据库实例的资源类型间存在差异，建议咨询数据库运维人员获取。

说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 5-158 目标库信息

目标库信息

IP地址或域名

对于RAC集群, 建议使用SCANIP+SERVICENAME方式创建任务, 提高访问性能。

端口

数据库服务名 Service Name ?

数据库用户名

数据库密码

SSL安全连接

测试连接 测试成功

表 5-164 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。 说明 对于RAC集群, 建议使用SCANIP+SERVICENAME方式创建任务, 提高访问性能。
端口	目标数据库服务端口, 可输入范围为1~65535间的整数。
数据库服务名	数据库服务名 (Service Name/SID), 客户端可以通过其连接到Oracle, 具体查询方法请参照界面提示。
数据库用户名	目标数据库的用户名。
数据库密码	目标数据库的用户名所对应的密码。
SSL安全连接	通过该功能, 用户可以选择是否开启对迁移链路的加密。如果开启该功能, 需要用户上传SSL CA根证书。 说明 <ul style="list-style-type: none"> 最大支持上传500KB的证书文件。 如果不启用SSL安全连接, 请自行承担数据安全风险。

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上, 待该任务删除后会永久清除。

步骤4 在“设置同步”页面, 选择同步对象, 完成后单击“下一步”。

图 5-159 同步模式

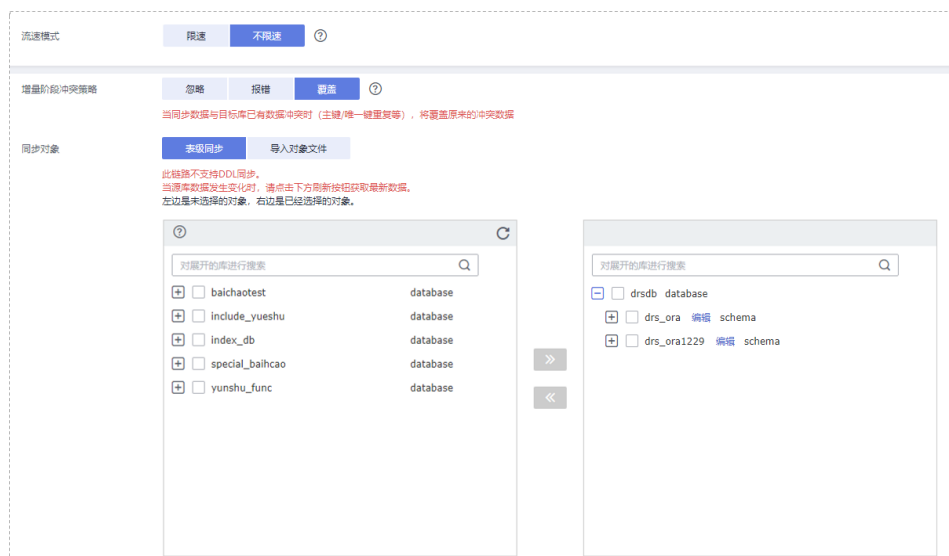




表 5-165 同步对象

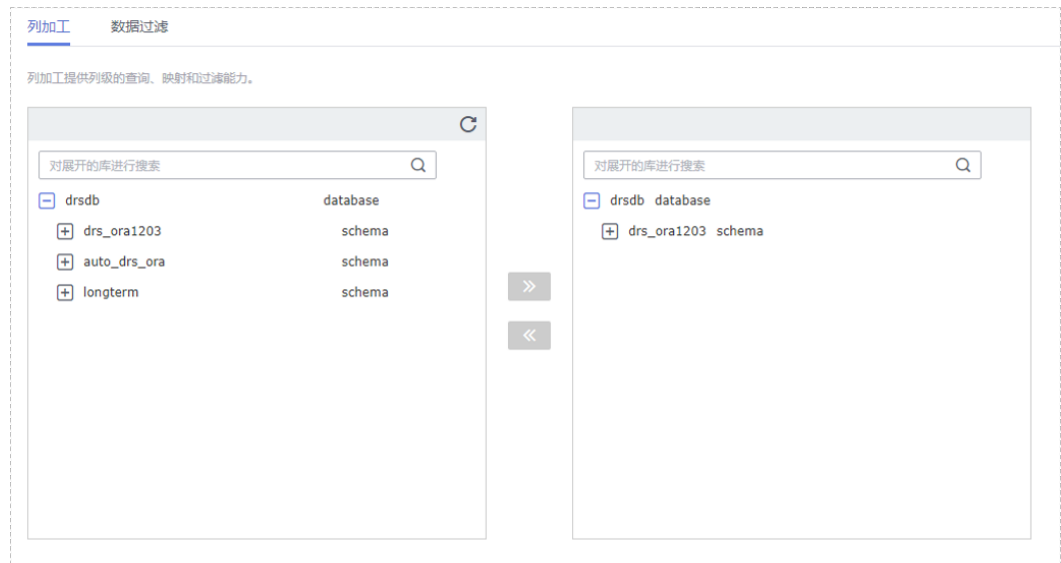
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-160 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下三种形式：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 ● 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> ● 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 ● 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 ● 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 ● 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，参考[数据加工](#)中的说明，设置相关规则。

图 5-161 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-162 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-163 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-166 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.16 将 GaussDB 分布式版同步到 Kafka

支持的源和目标数据库

表 5-167 支持的数据库

源数据库	目标数据库
GaussDB分布式	Kafka 0.11及以上版本

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-168](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-168 支持的同步对象

类型名称	同步范围
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持增量同步。 ● 支持的字段：BIGINT, BIT, BIT VARYING, BLOB, BOOLEAN, BYTEA, CHARACTER, CHARACTER VARYING, CLOB, DATE, DOUBLE PRECISION, INTEGER, MONEY, NUMBER, NUMERIC, NVARCHAR2, RAW, REAL, SMALLDATETIME, SMALLINT, TEXT, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, TINYINT。 ● 支持表级同步、schema级同步、库级同步： <ul style="list-style-type: none"> - 仅支持同步选中表的DML。 - 不支持同步没有schema的database。 - 不支持同步没有表的schema。 - 不支持同步列存表、压缩表、延迟表、临时表，不建议同步非日志表。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\\"'\?和!。列名不可以包含"和。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-169](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因密码填写错误在连接DN过程中导致用户被锁定。

表 5-169 数据库用户权限

类型名称	增量同步
源数据库连接用户	<p>用户具备sysadmin角色或以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步任务一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-170 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 选择增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 将无主键表的REPLICA IDENTITY属性设置为FULL，或给无主键表追加主键。 - 建议将主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
增量同步过程须知	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标连接用户的密码、权限，否则可能导致任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 ● 选择表级对象迁移时，同步过程中不建议对表进行重命名操作。 ● 不支持interval partition表复制。 ● 不支持修改主键列名称。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限于GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。 ● 增量同步过程中暂停、任务异常续传，目标Kafka中可能会出现重复数据，请使用Kafka数据中的id字段进行去重。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-164 同步任务信息

⚠ 1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
 2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域	<input type="text" value=""/> ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。	
项目	<input type="text" value=""/>
* 任务名称	<input type="text" value="DRS-6435"/> ?
描述	<div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> <div style="text-align: right; font-size: x-small;">0/256</div> ?

表 5-171 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-165 同步实例信息

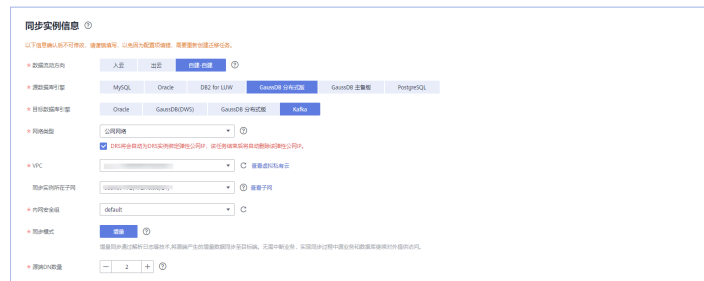


表 5-172 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“Kafka”。
网络类型	此处以公网网络为示例。目前支持可选“公网网络”和“VPN、专线网络”。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。

参数	描述
同步模式	<ul style="list-style-type: none"> 增量 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。 无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。
源端DN数量	需要和分布式源库DN（Data Node）节点数量保持一致。

- 任务类型

图 5-166 任务类型



表 5-173 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-167 企业项目和标签



表 5-174 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

 说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性：**确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性：**确保源库和目标库允许DRS通过用户名和密码访问。

图 5-168 源库信息

源库信息

CN IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

DN IP地址或域名 ?

IP地址或域名

待实例创建成功后再进行测试连接

表 5-175 源库信息

参数	描述
CN IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
DN IP地址或域名	DN IP信息可以登录CN节点通过SQL查询。 由于不同数据库实例的资源类型间存在差异，建议咨询数据库运维人员获取。

说明

源数据库的数据库用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-169 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

连接方式

✔ 测试成功

表 5-176 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
连接方式	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

步骤4 在“设置同步”页面，选择同步策略、数据格式和同步对象，单击“下一步”。

图 5-170 同步模式

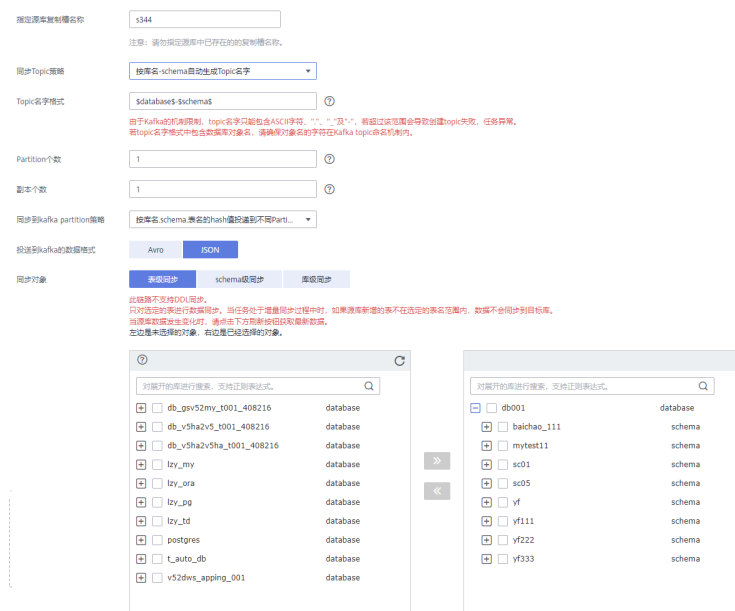



表 5-177 同步对象

参数	描述
指定源库复制槽名称	可选择是否指定源库复制槽，打开后需手动输入复制槽的名称（只能包含小写字母、数字、下划线，长度63，且不能以数字开头）。
同步Topic策略	同步Topic策略，可选择集中投递到一个Topic或者按照格式自动生成Topic名字。
Topic	选择目标端需要同步到的Topic，同步Topic策略选择集中投递到一个Topic时可见。

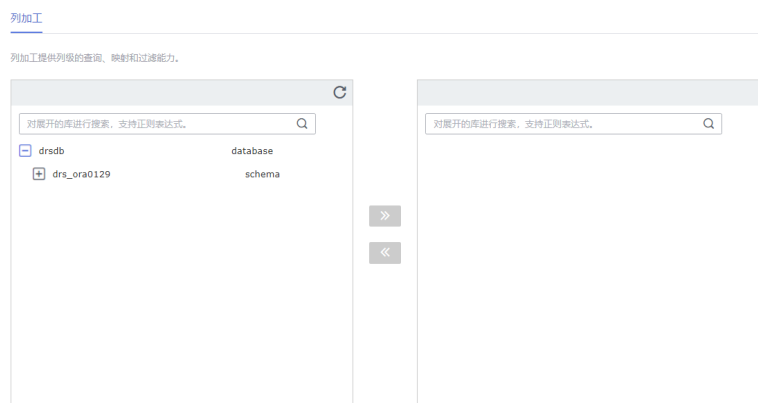
参数	描述
Topic名字格式	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>由于Kafka的机制限制，Topic名字只能包含ASCII字符、"."、"_"及"-", 若超过该范围会导致创建Topic失败，任务异常。</p> <p>若Topic名字格式中包含数据库对象名，请确保对象名的字符在Kafka topic命名机制内。</p>
Partition个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的分区个数。每个topic都可以创建多个partition，越多的partition可以提供更高的吞吐量，越多的partition会消耗更多的资源，建议根据broker节点的实际情况来设置partition的数量。</p>
副本个数	<p>同步Topic策略选择自动生成Topic名字时可见。</p> <p>用来设置topic的副本数。每个topic可以有多个副本，副本位于集群中不同的broker上，副本的数量不能超过broker的数量，否则创建topic时会失败。</p>
同步到kafka partition策略	<p>同步到kafka partition策略。</p> <ul style="list-style-type: none"> 按库名.schema.表名的hash值投递到不同Partition：适用于单表的查询场景，表内保序，表与表之间不保序，可以提高单表读写性能，推荐使用此选项。 按主键的hash值投递到不同Partition：适用于一个表一个Topic的场景，避免该表都写到同一个分区，消费者可以并行从各分区获取数据。 对于无主键表，如果选择“按主键的hash值投递到不同的partition”策略时，同步任务则会使用默认“按库名.schema.表名的hash值投递到不同的partition”的策略进行同步。 按库名.schema的hash值投递到不同Partition：适用于一个database一个topic的场景，避免多个schema下的数据写到一个分区，消费者可以并行从各分区获取数据。 全部投递到Partition 0：适用于有事务要求的场景，事务保序，可以保证完全按照事务顺序消费，写入性能比较差，如果没有强事务要求，不推荐使用此选项。
投送到kafka的数据格式	<p>选择投送到kafka的数据格式。</p> <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。 JSON：为Json消息格式，方便解释格式，但需要占用更多的空间。 <p>详细格式可参考Kafka消息格式。</p>

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、schema级同步和库级同步，您可以根据业务场景选择对应的数据进行同步。</p> <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工。

- 如果不需要数据加工，单击“下一步”。
- 如果需要进行列加工，参考[数据加工](#)中的说明，设置相关规则。

图 5-171 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-172 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-173 任务启动设置

The screenshot shows a configuration interface for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Latency Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto End Time): An input field containing the number '14', with a help icon and a note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 5-178 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.17 将 GaussDB 分布式版同步到 GaussDB 分布式版

支持的源和目标数据库

表 5-179 支持的数据库

源数据库	目标数据库
GaussDB分布式	GaussDB分布式 说明 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-180](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-180 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\\ ?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-181](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 5-181 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限，如果需要使用内容对比，需要拥有用户的MONADMIN权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限，如果需要使用内容对比，需要拥有用户的MONADMIN权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none"> 需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。 如果需要使用内容对比，需要拥有用户的MONADMIN权限。 目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log, public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。 		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
 - 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。
-
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
 - 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
 - 数据对比

建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-182 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求: 若选择增量或全量+增量同步: <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求: <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 源库要同步的库名、schema名和表名不可以包含/<.>\\` \?和!。 ● 目标数据库参数要求: <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求: <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知: <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索

类型名称	使用和操作约定
	<p>引名” 计算得到：c_+哈希值+原约束名（可能被截断）+_key。</p> <ul style="list-style-type: none"> - 由于GaussDB分布式版仅支持范围分区表，所以，除范围分区外的其他类型分区表作为普通表同步到目标库。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限与GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限与GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限与GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>数据加工须知</p>	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键和分布列不可以被过滤。 ● 表上存在列名映射或列过滤时，索引定义中的部分索引条件表达式将被忽略，普通索引中的表达式列将被排除，包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时，普通索引和外键中对应的列也将被过滤。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 源库和目标库的拓扑结构需保持一致，否则内容对比不可用。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-174 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-183 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-175 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎操作，以免影响业务连续性，请谨慎新建任务。

数据流动方向: [入站] [出站] [自建-自建] ?

源数据库引擎: MySQL Oracle DB2 for LUW GaussDB 分布式版 GaussDB 主备版 PostgreSQL

目标数据库引擎: MySQL Oracle GaussDB(DWS) GaussDB 分布式版 GaussDB 主备版 Kafka

网络类型: [公网网络] ?

DRS将自动为您创建网络并配置公网IP，您在创建实例时会看到公网IP。

VPC: [下拉菜单] ? 查看默认VPC

同步实例所在子网: [下拉菜单] ? 查看子网

内网安全组: [下拉菜单] ?

同步模式: [全量+增量] [全量] [增量] ?

源端DB数量: [2] ?

表 5-184 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“GaussDB分布式版”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。

参数	描述
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。
源端DN数量	需要和分布式源库DN（Data Node）节点数量保持一致。

- 任务类型

图 5-176 任务类型



表 5-185 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-177 企业项目和标签



表 5-186 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-178 源库信息

表 5-187 源库信息

参数	描述
CN IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
DN IP地址或域名	DN IP信息可以登录CN节点通过SQL查询。 由于不同数据库实例的资源类型间存在差异，建议咨询数据库运维人员获取。

说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 5-179 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

表 5-188 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 5-180 同步模式

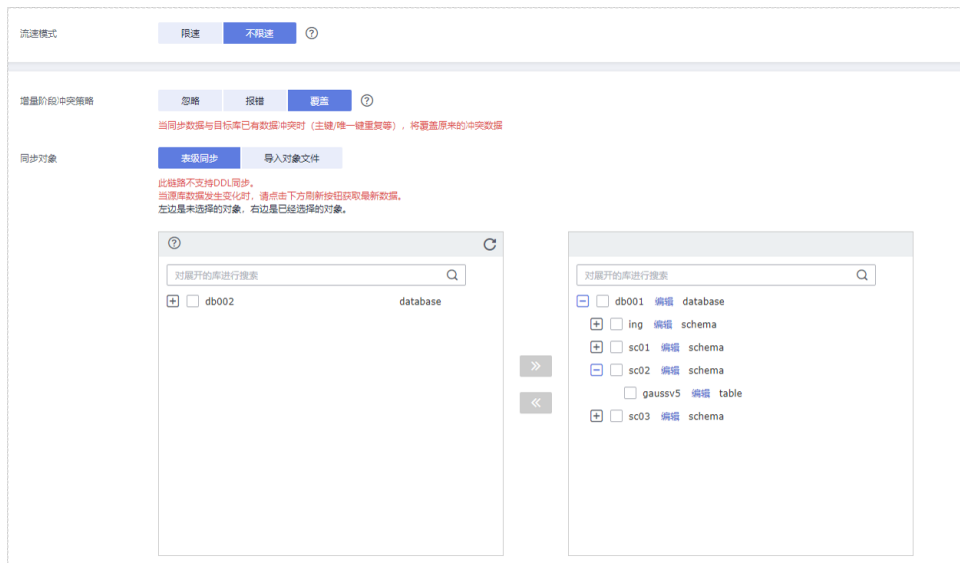




表 5-189 同步对象

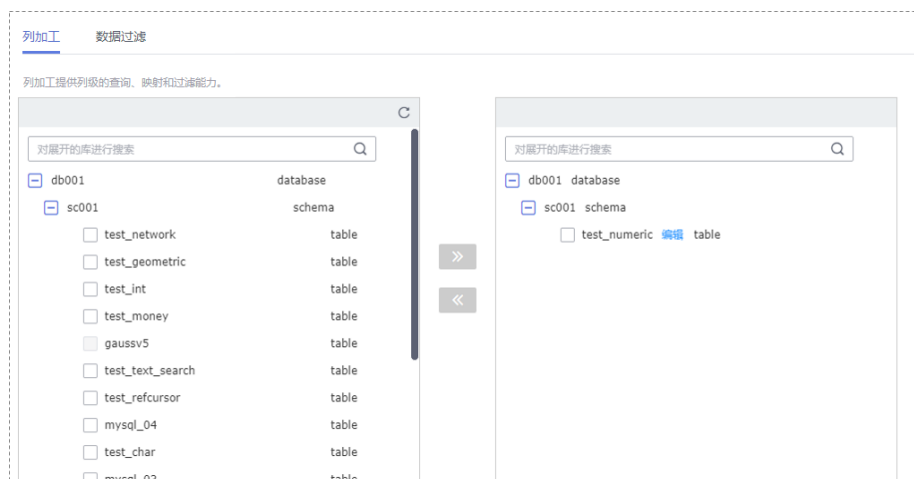
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-181 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和文件导入对象，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要列加工或者数据过滤，可参考[数据加工](#)，设置相关规则。

图 5-182 数据加工



步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-183 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-184 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-190 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.18 将 GaussDB 分布式版同步到 GaussDB 主备版

支持的源和目标数据库

表 5-191 支持的数据库

源数据库	目标数据库
GaussDB分布式	<p>GaussDB主备版</p> <p>说明</p> <ul style="list-style-type: none"> 仅支持目标库版本等于或高于源库版本。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-192](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-192 支持的同步对象

类型名称	使用和操作约定
同步范围须知	<ul style="list-style-type: none"> ● 不支持实例级同步：不支持实例级别同步，每次至多同步一个库（database），同步多个库需要使用多个DRS任务。 ● 支持场景：支持全量同步、增量同步、全量+增量同步。 ● 支持的字段：INTEGER, TINYINT, SMALLINT, BIGINT, NUMBER, NUMERIC, REAL, DOUBLE PRECISION, CHARACTER, CHARACTER VARYING, NVARCHAR2, BIT, BIT VARYING, BLOB, BYTEA, CLOB, RAW, TEXT, JSON, BOOLEAN, DATE, SMALLDATETIME, TIME WITH TIME ZONE, TIME WITHOUT TIME ZONE, TIMESTAMP WITH TIME ZONE, TIMESTAMP WITHOUT TIME ZONE, INTERVAL, BOX, CIDR, CIRCLE, INET, LSEG, MACADDR, MONEY, PATH, POINT, POLYGON, TSQUERY, TSVECTOR, REFCURSOR, UUID, ARRAY。 ● 支持表级同步或导入对象文件： <ul style="list-style-type: none"> - 全量同步过程仅支持同步选中表的表结构、序列、数据、约束和索引。 - 增量同步过程仅支持同步选中表的DML。 - 不支持选择没有schema的database。 - 不支持选择没有表的schema。 - 不支持列存表、压缩表、延迟表、临时表、二级分区表、无主键复制表，增量不建议同步非日志表。 - 不支持表字段为生成列的默认值。 - 不支持同步既是无主键表，又是分区表的自建表，可能会导致任务失败。 - 不支持同步系统schema（pg_toast、cstore、snapshot、sys、dbms_job、dbms_perf、pg_catalog、information_schema、utl_file、dbms_output、dbms_random、utl_raw、dbms_sql、dbms_lob、dbe_perf、pkg_service、pkg_util、dbe_file、dbe_random、dbe_output、dbe_raw、dbe_sql、dbe_lob、dbe_task、blockchain、db4ai、dbe_pldebugger、sqladvisor、dbe_application_info、dbe_match、dbe_pldeveloper、dbe_scheduler、dbe_session、dbe_utility、dbe_sql_util、dbe_xml、dbe_xmldom、dbe_xmlparser、dbe_compression、dbe_heat_map、dbe_ilm、dbe_ilm_admin、prvt_ilm、dbe_profiler、dbe_stats、rdsBackup、rdsMetric、rdsRepl、public）。 - 不支持同步系统表（public schema下的redis_progress_detail、redis_status、pgxc_redsitb、redis_progress）。 - 库名、schema名和表名不可以包含/<.>\\'\\?和!。列名不可以包含"和。 - 使用导入对象文件选表的方式时，要保证导入的表在源库存在或对同步用户可见。

数据库用户权限要求

在使用DRS进行同步时，连接源库和目标库的数据库用户需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的用户权限也不同，详细可参考[表 5-193](#)进行赋权。DRS会在“预检查”阶段对数据库用户权限进行自动检查，并给出处理建议。

说明

- 建议创建单独用于DRS任务连接的数据库用户，避免因为数据库用户密码修改，导致的任务连接失败。
- 连接源或目标数据库的用户密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库用户被锁定影响使用。
- 由于分布式数据库为源链路需要检查每个DN的连通性，建议使用非root数据库用户进行同步，避免因为密码填写错误在连接DN过程中导致用户被锁定。

表 5-193 数据库用户权限

类型名称	全量同步	增量同步	全量+增量同步
源数据库连接用户	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。 	<p>具备sysadmin角色或者以下最小权限：</p> <ul style="list-style-type: none"> • 需要REPLICATION权限或继承了内置角色gs_role_replication的权限、DATABASE的CONNECT权限、SCHEMA的USAGE权限、表的SELECT或UPDATE权限（无主键表锁表需要UPDATE权限）、SEQUENCE的SELECT权限。 • 需要具备远程连接权限，添加方法可参考GaussDB远程连接配置方法。 • 需要保证安全组和防火墙策略放开对应连接端口，详细说明可参考GaussDB为源链路增量同步连接及端口说明。

类型名称	全量同步	增量同步	全量+增量同步
目标数据库连接用户	具备sysadmin角色或者以下最小权限： <ul style="list-style-type: none">需要DATABASE的CONNECT和CREATE权限、SCHEMA的USAGE和CREATE权限、表的INSERT、UPDATE、DELETE、SELECT、ALTER、INDEX、COMMENT权限。目标库使用gsloader等工具创建相关系统表（如public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的all privilege权限，详细信息可参考《GaussDB工具参考》。		

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。
 - 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
 - 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
 - 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
 - DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
 - 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
 - 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

DRS全量+增量同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量和单增量任务则包含三个阶段）。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-194 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： 若选择增量或全量+增量同步时： <ul style="list-style-type: none"> - 源库的wal_level参数设置为logical。 - 源库的enable_slot_log参数设置为on。 - 源库的max_replication_slots参数值必须大于当前已使用的复制槽数量。 - 建议对无主键表追加主键，或将无主键表的REPLICA IDENTITY属性设置为FULL。 ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库必须是GaussDB分布式版实例。 - 源库要同步的库名、schema名和表名不可以包含/<.>\\`\"?和!。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标库database的字符集与源库要保持一致。 - 建议目标库database兼容的数据库的类型与源库保持一致。 ● 目标数据库对象要求： <ul style="list-style-type: none"> - 同步前保证目标库已创建好对应的database。 - 选择增量同步，同步前保证目标库表结构已经建好，且与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 选择全量或全量+增量同步，如果目标库已经存在表结构，需要保证与源端一致；如果进行了列加工，请保证跟列加工后的表结构保持一致。 - 目标库已启用的触发器不可以与同步表关联。 - 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。 ● 其他使用须知： <ul style="list-style-type: none"> - 该链路不支持SSL安全连接。 - 暂不支持级联单向同步，例如从实例A同步到实例B，再从实例B同步到实例C。 - 实时同步过程中，不保证分布式事务的一致性。 - 表名映射时，同步后表上的原索引名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索引名”计算得到：i_+哈希值+原索引名（可能被截断）+_key。 - 表名映射时，同步后表上的原约束名称变化为如下格式，其中哈希值由“原模式名_原表名_映射后模式名_映射后表名_原索

类型名称	使用和操作约定
	<p>引名” 计算得到：c_+哈希值+原约束名（可能被截断）+_key。</p> <ul style="list-style-type: none"> - 由于GaussDB分布式版仅支持范围分区表，所以，除范围分区外的其他类型分区表作为普通表同步到目标库。 - 由于长事务导致的创建逻辑复制槽失败或逻辑复制槽不存在的问题，需要通过重置任务来重新启动。 - 受GaussDB逻辑复制能力约束，同一实例上，最多支持同时开启20个逻辑复制槽进行解码。
<p>全量同步过程须知</p>	<ul style="list-style-type: none"> ● 勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 请勿在源库执行任何DDL，受限与GaussDB逻辑复制功能，不支持DDL语句的同步。否则可能导致数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。
<p>增量同步过程须知</p>	<ul style="list-style-type: none"> ● 勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 ● 在启动任务后，进入增量同步阶段前，请确保在源库中长事务未启动，启动长事务会阻塞逻辑复制槽的创建，进而引发任务失败。 ● 请勿在源库执行任何DDL，受限与GaussDB逻辑复制功能，不支持DDL语句的同步，否则可能导致数据不一致或任务失败。 ● 请勿修改源库中表的REPLICA IDENTITY值，否则可能导致增量数据不一致或任务失败。 ● 请勿在目标库做写入操作，否则可能导致数据不一致。 ● 不支持interval partition表复制。 ● 在事务中执行DDL语句后，该DDL语句与之后的语句不会被同步。 ● 逻辑日志解码受限与GaussDB内核的解码能力，存在一些使用约束，具体约束请参考对应版本GaussDB逻辑解码章节的注意事项说明。
<p>数据加工须知</p>	<ul style="list-style-type: none"> ● 列加工时，主键、唯一键和分布列不可以被过滤。 ● 表上存在列名映射或列过滤时，索引定义中的部分索引条件表达式将被忽略，普通索引中的表达式列将被排除，包含表达式列的唯一索引将被排除。 ● 表上存在列过滤时，普通索引和外键中对应的列也将被过滤。
<p>同步对比须知</p>	<ul style="list-style-type: none"> ● 建议在源库的业务低峰期进行数据比对，防止误报不一致数据，以及减少对源库和DRS任务的冲击。 ● 在增量同步过程中做对比时，源库若存在写入，则对比结果可能不一致。 ● 不支持在全量同步过程中做对比。 ● 不支持数据对比过程中做限速。

类型名称	使用和操作约定
结束任务须知	<p>任务正常结束</p> <ul style="list-style-type: none"> 结束时将自动删除任务在源库创建的流复制槽。 结束任务时，如果目标库同步表中已关联序列，DRS将自动重置目标库序列值。重置的自增序列值为源库序列值+安全余度，自减序列值为源库序列值-安全余度，安全余度默认为10000。 <p>任务强制结束</p> <ul style="list-style-type: none"> 强制结束同步任务，需要手动删除源库可能残留的复制槽，可参考GaussDB分布式版为源强制结束任务章节的指导进行操作。 强制结束同步任务，目标数据库中可能存在以drs_为前缀命名的无主键临时表，需要用户手动删除。 复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务同步日志页面查看“task node id is ***”日志找到。

前提条件

- 已登录数据复制服务控制台。
- 账户余额大于等于0美元。
- 满足实时同步支持的数据库类型和版本，详情请参见[支持的数据库](#)。
- 当用户创建DRS任务使用的是子账号时，请先确认已添加过委托。创建委托方法，可参考[委托管理](#)。
- 已阅读以上[使用建议](#)和[使用须知](#)。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-185 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [下拉菜单] ?
不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [下拉菜单]

* 任务名称: DRS-6435 ?

描述: [文本框] ?
0/256

表 5-195 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-186 同步实例信息

同步实例信息 ?

以下配置项创建后不可修改，请谨慎填写，以免因配置项错误，造成数据迁移失败。

数据流动方向: [下拉菜单: 自建-自建] ?

源数据库引擎: [下拉菜单: GaussDB 分布式版] ?

目标数据库引擎: [下拉菜单: GaussDB 主备版] ?

网络类型: [下拉菜单: 公网网络] ?

VPC: [下拉菜单] ?

同步模式: [下拉菜单: 全量] ?

表 5-196 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“GaussDB分布式版”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	此处以“公网网络”为示例。目前支持可选“公网网络”、“VPC网络”和“VPN、专线网络”。

参数	描述
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p> <p>网络类型选择“VPN、专线网络”时，支持输入指定IP。需要注意，同时创建多个任务时，不能指定同一组未使用IP，否则会导致任务创建失败。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<p>全量+增量、全量、增量三种选择类型，此处以“全量+增量”为示例。</p> <ul style="list-style-type: none"> “全量+增量”： <p>该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。</p> <p>说明</p> <p>选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> “全量”： <p>该模式为数据库一次性同步，全量同步将源库数据一次性同步至目标端数据库。</p> “增量”： <p>增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。</p>
源端DN数量	需要和分布式源库DN（Data Node）节点数量保持一致。

- 任务类型

图 5-187 任务类型

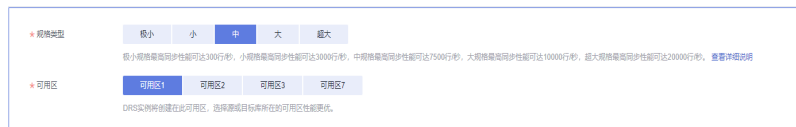


表 5-197 任务类型信息

参数	描述
规格类型	选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考 实时同步链路规格说明 。 说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考 同步规格变更 。
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-188 企业项目和标签



表 5-198 企业项目和标签

参数	描述
企业项目	企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。 更多关于企业项目的信息，请参见《 企业管理用户指南 》。如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《 企业管理用户指南 》中“ 创建企业项目 ”的内容。
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，待同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-189 源库信息

The screenshot shows a form titled "源库信息" (Source Database Information). It contains the following fields and elements:

- CN IP地址或域名**: A text input field with a help icon (question mark) to its right. Below it is a red warning message: "请确保所填写的多个IP地址或域名属于同一个实例" (Please ensure that the multiple IP addresses or domains entered belong to the same instance).
- 数据库用户名**: A text input field.
- 数据库密码**: A text input field with a password visibility icon (eye) to its right.
- DN IP地址或域名**: A text input field with a help icon (question mark) to its right. Below it is a blue button labeled "IP地址或域名" with a plus icon, and two more empty text input fields stacked vertically.
- At the bottom, there is a grey button labeled "测试连接" (Test Connection) and a grey text label "待实例创建成功后再进行测试连接" (Test connection after instance creation is successful).

表 5-199 源库信息

参数	描述
CN IP地址或域名	输入GaussDB的IP，请确保所填写的多个IP地址或域名属于同一个实例。
数据库用户名	源数据库的用户名。
数据库密码	源数据库的用户名所对应的密码。
DN IP地址或域名	DN IP信息可以登录CN节点通过SQL查询。 由于不同数据库实例的资源类型间存在差异，建议咨询数据库运维人员获取。

📖 说明

源数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

图 5-190 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接 ✔ 测试成功

表 5-200 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8080,192.168.0.2:8080。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，完成后单击“下一步”。

图 5-191 同步模式

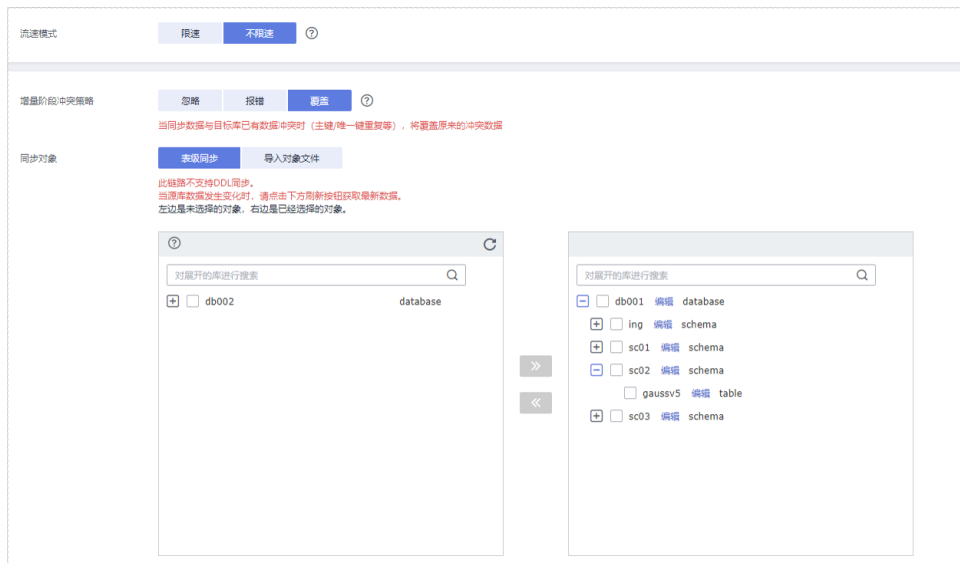



表 5-201 同步对象

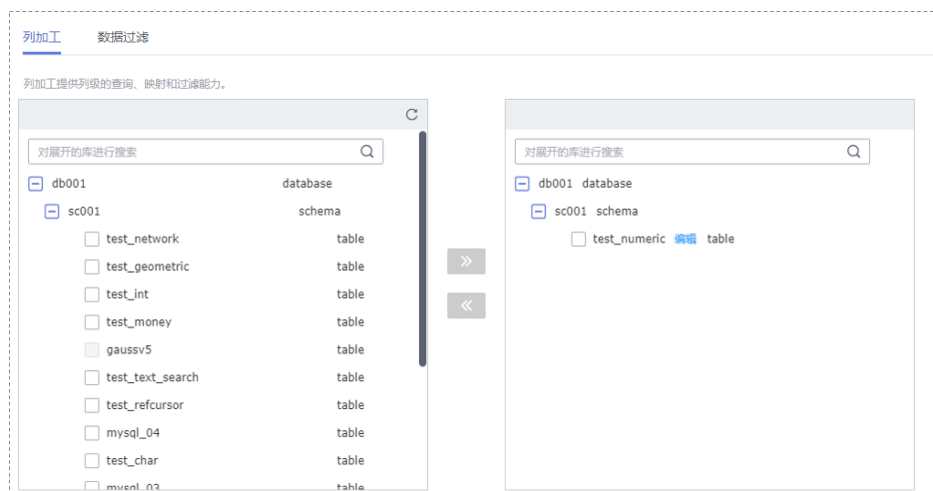
参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 5-192 设置流速模式</p>  <ul style="list-style-type: none"> 不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> 限速模式只对全量阶段生效，增量阶段不生效。 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。

参数	描述
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。</p> <p>目前支持的冲突策略有：</p> <ul style="list-style-type: none"> ● 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以目标数据为准，DRS将跳过冲突数据，继续进行后续同步。 ● 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），以源数据为准，DRS将覆盖原来的冲突数据。 ● 报错 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），同步任务将失败并立即中止。 <p>当同步数据与目标库已有数据冲突时，选择“忽略”无法保证最终一致性，如需保证数据一致性，可选择“覆盖”策略。</p>
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> ● 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 对象名映射时，对schema名或者表名进行映射后，为防止索引/约束名称冲突，被映射表上的索引和约束名称将按照“前缀+基于全名称的哈希值+原索引/约束名+_key”的规则进行映射，其中索引的前缀为“i_”，约束的前缀为“c_”，全名称为“模式名_表名_索引/约束名”，原索引/约束名可能由于长度限制而发生截断。 ● 选择导入对象文件也支持对象名映射，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> - 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 - 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 - 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 - 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行列加工或者数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要加工列或者数据过滤，可参考[数据加工](#)，设置相关规则。

图 5-193 数据加工



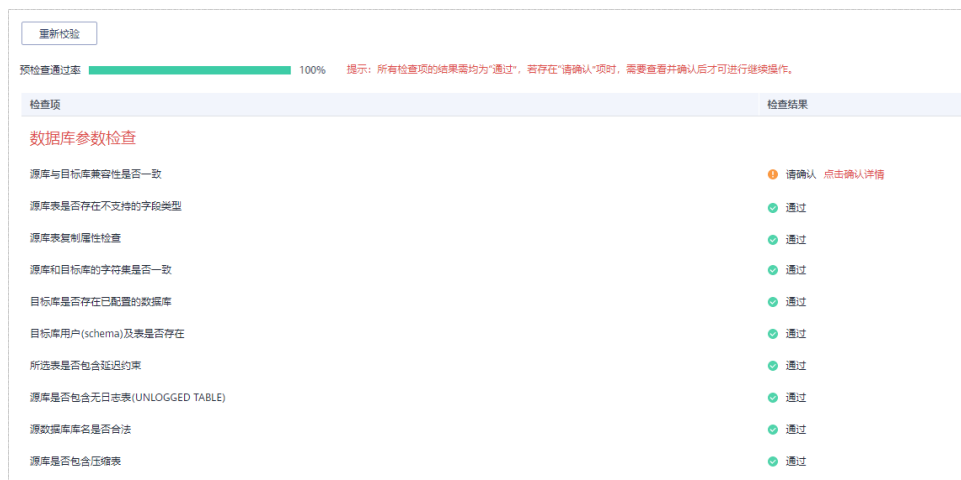
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-194 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-195 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): A text input field containing '14', with a help icon and a note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。' (After a period of time in an abnormal state, it will automatically end. Unit is days).

表 5-202 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求, 设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后, 会对源数据库和目标数据库的性能产生影响, 建议选择业务低峰期, 合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数, 开启之后, 选择对应的SMN主题。当同步任务状态、时延指标或数据异常时, 系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见, 需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段, 源数据库和目标数据库之间的同步有时存在一个时间差, 称为时延, 单位为秒。 时延阈值设置是指时延超过一定的值后 (时延阈值范围为0到3600s), DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送, 避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时, 不会发送通知给收件人。 首次进入增量同步阶段, 会有较多数据等待同步, 存在较大的时延, 属于正常情况, 不在此功能的监控范围之内。 设置时延阈值之前, 需要设置任务异常通知。
数据异常通知	该项为可选参数, 开启之后, 当同步任务数据异常时, 系统将发送通知。

参数	描述
任务异常自动结束时间 (天)	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.19 将 DB2 for LUW 同步到 GaussDB 主备版

支持的源和目标数据库

表 5-203 支持的数据库

源数据库	目标数据库
DB2 for LUW 9.7、10.1、10.5、11.1、11.5版本	GaussDB主备版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-204 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步需要用户具有CONNECT、DATAACCESS权限。 - 全量+增量同步需要用户具有DBADM权限。 - 选择全量+增量同步时，需要确保开启源数据库的归档日志，确保需要同步的表的Datacapture属性为Y。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 - SCHEMA级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 - 表级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下所有表的SELECT，UPDATE，INSERT和DELETE权限。 - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 不支持自定义类型。 ● timestamp类型支持的最大精度是6。 ● 全量同步结构不支持位图索引、倒排索引、函数索引、xml索引。 ● 增量同步支持的LOB类型大小限10M以内。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中执行的DDL语句。 ● 全量+增量任务不支持带XML字段的无主键表。 ● 浮点数类型作为主键时，增量同步可能会导致数据不一致。
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS不支持非ASCII字符、“.”、“>”、“<”、“\”、“.”、“ ”、“”、“?”、“!”、“”和“”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：GBK、UTF8。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是GaussDB主备版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。
操作须知	<ul style="list-style-type: none"> ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步。 ● 不支持ARM虚拟机环境。 ● 如果无主键表中包含LOB、LONG等大字段，增量同步阶段存在大字段数据不一致的可能性。 ● 如果源库和目标库字符集不一致，如源库是GBK，目标库是UTF8，由于GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 任务启动后，对于分区表的detach操作可能导致数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，源库不允许有未提交的DDL事务。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束，如有需要，用户必须手动在目标库创建对应表的默认值约束。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步阶段，支持任务再编辑追加同步对象。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-196 同步任务信息

1、该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2、创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 5-205 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-197 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，谨慎性填写，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向: 入云 | 出云 | 自建-自建 ?

* 源数据库引擎: MySQL | Oracle | DB2 for LUW | GaussDB 分布式版 | GaussDB 主备版 | PostgreSQL

* 目标数据库引擎: GaussDB 分布式版 | GaussDB 主备版 | Kafka

* 网络类型: 公网网络 ?
 DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除弹性公网IP。

* VPC: vpc-[ID] ? 查看虚拟私有云

同步实例所在子网: subnet-[ID] ? 查看子网

* 内网安全组: default ?

* 同步模式: 全量+增量 | 全量 ?

该模式为数据的一致性实时同步，通过全量过程生成目标数据库的初始化后，增量同步的数据通过日志解析技术，将源端和目标数据库中的数据保持一致。

* 可用区: 可用区1 | 可用区2 | 可用区3 | 可用区4
DRS实例创建在此可用区，选择源端或目标库所在的可用区性能更佳。

表 5-206 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“DB2 for LUW”。
目标数据库引擎	选择“GaussDB主备版”。
网络类型	此处以公网网络为示例。目前支持可选公网网络和VPN、专线网络。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量： 该模式为数据库一次性同步，适用于可中断业务的数据数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。

- 企业项目和标签

图 5-198 企业项目和标签



表 5-207 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过，单击“下一步”。

图 5-199 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

数据库名

SSL安全连接

同步驱动 已同步驱动: jjar

表 5-208 源库信息

参数	描述
IP地址或域名	源数据库DB2 for LUW的IP地址或域名。

参数	描述
端口	源数据库DB2 for LUW服务端口，可输入范围为1~65535间的整数。
数据库用户名	源DB2 for LUW的数据库账号。
数据库密码	源数据库账号所对应的密码。
数据库名	源DB2 for LUW中同步对象所属数据库的名称。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。目前仅DB2 for LUW 10.5及以上的版本支持SSL安全连接，DB2 for LUW 9.7和10.1版本暂不支持。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不使用SSL安全连接，请自行承担数据安全风险。 • DB2 for LUW SSL安全连接需要上传三个证书文件：jksi证书、kdbi证书和sth证书。 • kdbi证书和sth证书测试连接时DRS不会进行校验，需要用户保证文件正确性，避免后续任务失败。
同步驱动	用户需要手动添加并上传源数据库DB2 for LUW对应的JDBC驱动包后，再进行测试连接。

 说明

源数据库的数据库IP地址、用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-200 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接
✔ 测试成功

表 5-209 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的数据库的用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。


步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 5-201 同步模式



表 5-210 同步对象

参数	描述
全量同步对象类型	全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。 <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步或者导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的数据过滤，设置相关规则。

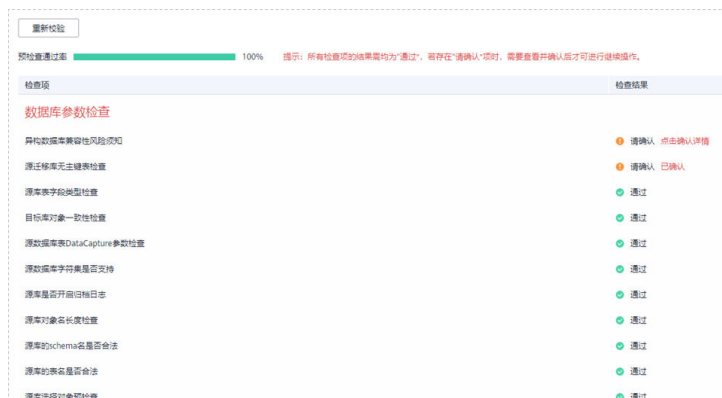
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-202 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-203 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: "立即启动" (Immediately Start) and "稍后启动" (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置** (Task Abnormal Notification Settings): A toggle switch is turned on.
- * SMN主题** (SMN Topic): A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)** (Delay Threshold (s)): A toggle switch is turned off.
- 数据异常通知** (Data Abnormal Notification): A toggle switch is turned off.
- * 任务异常自动结束时间** (Task Abnormal Auto-End Time): An input field containing the value "14". A red asterisk and help icon are present. A red note below reads: "任务处于异常状态一段时间后，将会自动结束。单位为天。" (After a period of time in an abnormal state, it will automatically end. Unit is days.)

表 5-211 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

---结束

5.20 将 DB2 for LUW 同步到 GaussDB 分布式版

支持的源和目标数据库

表 5-212 支持的数据库

源数据库	目标数据库
DB2 for LUW 9.7、10.1、10.5、11.1、11.5版本	GaussDB分布式版

使用建议

注意

- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 为保证同步前后数据一致性，确保同步期间目标数据库无业务写入。

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 基于以下原因，建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。

- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- DRS并发读取数据库，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 数据对比
建议您结合数据对比的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

使用须知

在创建同步任务前，请务必阅读以下使用须知。

说明

- 建议创建单独用于DRS任务连接的数据库账号，避免因数据库账号密码修改，导致的任务连接失败。
- 连接源和目标数据库的账号密码修改后，请尽快[修改DRS任务中的连接信息](#)，避免任务连接失败后自动重试，导致数据库账号被锁定影响使用。

表 5-213 使用须知

类型名称	使用和操作限制
数据库权限设置	<ul style="list-style-type: none"> ● 源数据库端： <ul style="list-style-type: none"> - 全量同步需要用户具有CONNECT、DATAACCESS权限。 - 全量+增量同步需要用户具有DBADM权限。 - 选择全量+增量同步时，需要确保开启源数据库的归档日志，确保需要同步的表的Datacapture属性为Y。 ● 目标数据库端必须同时拥有以下权限： <ul style="list-style-type: none"> - 库级权限：需要使用root或其他有sysadmin角色的DATABASE用户登录postgres基库，赋予用户DATABASE的CREATE、CONNECT权限。 - SCHEMA级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA的CREATE、USAGE权限。 - 表级权限：需要使用 root、或其他有sysadmin角色的DATABASE用户、或使用数据库的OWNER用户登录数据库，赋予用户SCHEMA下所有表的SELECT，UPDATE，INSERT和DELETE权限。 - 目标库使用gsloader等工具创建相关系统表（如 public.pgxc_copy_error_log，public.gs_copy_summary）时，DRS访问相关系统表需要有系统表的所有权限，详细信息可参考《GaussDB工具参考》。
同步对象约束	<ul style="list-style-type: none"> ● 全量同步时支持表、普通索引、主键与唯一约束、数据的同步。增量同步时支持表的实时同步。 ● 不支持自定义类型。 ● timestamp类型支持的最大精度是6。 ● 全量同步结构不支持位图索引、倒排索引、函数索引、xml索引。 ● 增量同步支持的LOB类型大小限10M以内。 ● 不支持默认值含有表达式的函数的表的同步。 ● 不支持同步源库中的临时表。 ● 不支持同步源库中执行的DDL语句。 ● 全量+增量任务不支持带XML字段的无主键表。 ● 浮点数类型作为主键时，增量同步可能会导致数据不一致。
源数据库要求	<ul style="list-style-type: none"> ● 库名、表名等数据库对象名称支持英文字符、“#”、“\$”、“_”等符号，DRS 不支持非ASCII字符、“.”、“>”、“<”、“\”、“.”、“ ”、“?”、“!”、“”和“'”等字符。对象名同步到目标库后会转换成小写，为避免同步失败，选择的源库表中不能存在表名称字母相同但大小写不同的表。 ● 目前仅支持同步如下字符集：GBK、UTF8。

类型名称	使用和操作限制
目标数据库要求	<ul style="list-style-type: none"> ● 目标库必须是GaussDB分布式版实例。 ● 确保目标库已创建以小写字母命名的数据库。 ● 同步的表要禁用外键，因为DRS并行回放会使得不同表之间的写入顺序和源库不一致，可能会触发外键约束限制，造成同步失败。 ● 支持目标数据库中的表比源数据库多列场景，但是需要避免以下场景可能导致的任务失败。 目标端多的列要求非空且没有默认值，源端insert数据，同步到目标端后多的列为null，不符合目标端要求。 目标端多的列设置固定默认值，且有唯一约束。源端insert多条数据后，同步到目标端后多的列为固定默认值，不符合目标端要求。
操作须知	<ul style="list-style-type: none"> ● 不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步。 ● 不支持ARM虚拟机环境。 ● 如果无主键表中包含LOB、LONG等大字段，增量同步阶段存在大字段数据不一致的可能性。 ● 如果源库和目标库字符集不一致，如源库是GBK，目标库是UTF8，由于GBK字符集单个中文字符占用2个字节，而UTF8字符集单个中文字符占用3个字节，可能会导致CHAR或VARCHAR类型数据同步到目标库后超出字段定义长度，所以客户需要根据实际情况对目标库CHAR和VARCHAR类型字段长度进行扩充（如扩大为源库的1.5倍）。 ● 任务启动后，对于分区表的detach操作可能导致数据不一致。 ● 同步过程中，不允许修改、删除连接源和目标数据库的用户的用户名、密码、权限，或修改源和目标数据库的端口号。 ● 同步过程中，源库不允许有未提交的DDL事务。 ● 全量同步分为表结构同步（含索引）、数据同步两个阶段，任务中只要有一个表的结构在目标库中创建成功即进入数据同步阶段。若同步完成产生失败表，再启动时将只同步数据，不同步表结构信息，用户必须手动在目标库中建表。 ● 全量同步分区表的结构时会将该对象转为非分区的普通表。 ● 全量同步表结构时只支持字符串或数字类型的普通默认值约束，不支持函数、序列等类型的默认值约束，如有需要，用户必须手动在目标库创建对应表的默认值约束。 ● 对于全量同步中的目标数据库表对象，不能进行写入操作，否则会导致数据不一致。 ● 增量同步时，BLOB末尾的0x00、CLOB末尾的空格会被截断。 ● 增量同步阶段，支持任务再编辑追加同步对象。 ● 增量同步阶段，源库DB2 for LUW对主键列或无主键表的第一列的update操作会导致在GaussDB数据库更新分布列，有可能会造成数据不一致，所以业务上要尽量规避对上述列的更新。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-204 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置光源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间网不互通，请选择靠近您客户的区域，可以降低网络时延，提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 5-214 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 5-205 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免因为配置项错误，需要重新创建迁移任务。

* 数据流动方向: 入云 出云 **自建-自建** ?

* 源数据库引擎: MySQL Oracle **DB2 for LUW** GaussDB 分布式版 GaussDB 主备版 PostgreSQL

* 目标数据库引擎: **GaussDB 分布式版** GaussDB 主备版 Kafka

* 网络类型: 公网网络 ?
 DRS将会自动为DRS实例绑定弹性公网IP，该任务结束后将自动删除该弹性公网IP。

* VPC: vpc- [Dropdown] ? 查看虚拟私有云

同步实例所在子网: subnet- [Dropdown] ? 查看子网

* 内网安全组: default ?

* 同步模式: **全量 + 增量** 全量 ?
该模式为数据持续性实时同步，通过全量过程先将目标数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。

* 可用区: **可用区1** 可用区2 可用区3 可用区4
DRS实例创建在此可用区，选择源库或目标库所在的可用区进行更改。

表 5-215 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“DB2 for LUW”。
目标数据库引擎	选择“GaussDB分布式版”。
网络类型	此处以公网网络为示例。目前支持可选公网网络和VPN、专线网络。
VPC	选择可用的虚拟私有云。
同步实例所在子网	请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。 默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	<ul style="list-style-type: none"> - 全量+增量： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 - 全量： 该模式为数据库一次性同步，适用于可中断业务的数据数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。
可用区	选择DRS实例创建在哪个可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-206 企业项目和标签



表 5-216 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通过，单击“下一步”。

图 5-207 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

数据库名

SSL安全连接

同步驱动 已同步驱动: jjar

表 5-217 源库信息

参数	描述
IP地址或域名	源数据库DB2 for LUW的IP地址或域名。

参数	描述
端口	源数据库DB2 for LUW服务端口，可输入范围为1~65535间的整数。
数据库用户名	源DB2 for LUW的数据库账号。
数据库密码	源数据库账号所对应的密码。
数据库名	源DB2 for LUW中同步对象所属数据库的名称。
SSL安全连接	<p>通过该功能，用户可以选择是否开启对同步链路的加密。如果开启该功能，需要用户上传SSL CA根证书。目前仅DB2 for LUW 10.5及以上的版本支持SSL安全连接，DB2 for LUW 9.7和10.1版本暂不支持。</p> <p>说明</p> <ul style="list-style-type: none"> • 最大支持上传500KB的证书文件。 • 如果不使用SSL安全连接，请自行承担数据安全风险。 • DB2 for LUW SSL安全连接需要上传三个证书文件：jksi证书、kdbi证书和sth证书。 • kdbi证书和sth证书测试连接时DRS不会进行校验，需要用户保证文件正确性，避免后续任务失败。
同步驱动	用户需要手动添加并上传源数据库DB2 for LUW对应的JDBC驱动包后，再进行测试连接。

📖 说明

源数据库的数据库IP地址、用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。

图 5-208 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

数据库用户名

数据库密码

测试连接
✔ 测试成功

表 5-218 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名，格式为IP地址/域名:端口。其中目标数据库服务端口，可输入范围为1~65535间的整数。 该输入框最多支持填写10组目标数据库的IP地址或者域名信息，多个值需要使用英文逗号隔开。例如： 192.168.0.1:8000,192.168.0.2:8000。
数据库用户名	目标数据库对应的数据库用户名。
数据库密码	目标数据库的用户名所对应的密码

说明

目标数据库的数据库的用户名和密码，会被系统加密暂存，直至删除该迁移任务后自动清除。


步骤4 在“设置同步”页面，选择同步对象，此处必须手动输入目标数据库名称，完成后单击“下一步”。

图 5-209 同步模式



表 5-219 同步对象

参数	描述
全量同步对象类型	全量同步对象类型，可选同步表结构、同步数据、同步索引，根据实际需求进行选择要同步内容。 <ul style="list-style-type: none"> 同步数据为必选项。 选择同步表结构的时候目标库不能有同名的表。 不选同步表结构的时候目标库必须有相应的表，且要保证表结构与所选表结构相同。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，支持表级同步和导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 在同步对象右侧已选对象框中，可以使用对象名映射功能进行源数据库和目标数据库中的同步对象映射，具体操作可参考对象名映射。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的, 以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，选择需要加工的数据，进行数据过滤。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据过滤，参考[数据加工](#)中的数据过滤，设置相关规则。

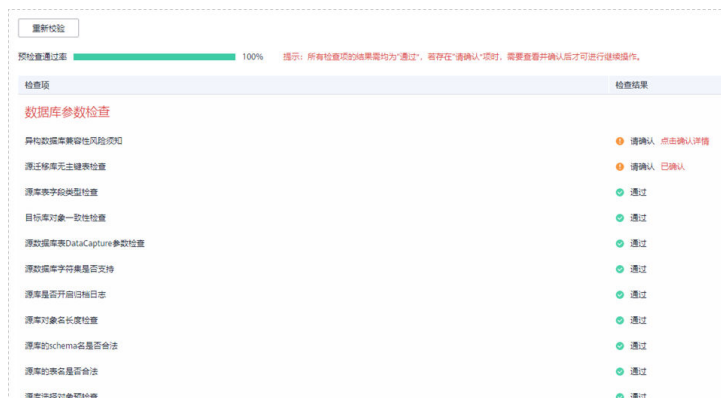
步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-210 预检查



📖 说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-211 任务启动设置

The screenshot shows a configuration panel for task start settings. At the top, there are two tabs: '立即启动' (Start Immediately) and '稍后启动' (Start Later), with a help icon. Below this, there are several settings:


- 任务异常通知设置**: A toggle switch is turned on.
- * SMN主题**: A dropdown menu with a refresh icon and a help icon.
- 时延阈值(s)**: A toggle switch is turned off.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: A text input field containing '14', with a help icon and a red note: '任务处于异常状态一段时间后, 将会自动结束。单位为天。'

表 5-220 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《消息通知服务用户指南》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none"> 当时延阈值设置为0时，不会发送通知给收件人。 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。

参数	描述
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

5.21 将 Microsoft SQL Server 同步到 Kafka

支持的源和目标数据库

表 5-221 支持的数据库

源数据库	目标数据库
<ul style="list-style-type: none"> 本地自建Microsoft SQL Server（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本） ECS自建Microsoft SQL Server数据库（企业版2012、2014、2016、2017、2019版本，标准版2016 SP2及以上版本、2017、2019版本） 	Kafka 0.11及以上版本

说明

目前仅支持白名单用户使用。

支持的同步对象范围

在使用DRS进行同步时，不同类型的同步任务，支持的同步对象范围不同，详细情况可参考[表5-222](#)。DRS会根据用户选择，在“预检查”阶段对同步对象进行自动检查。

表 5-222 支持的同步对象

类型名称	使用须知
同步对象范围	<ul style="list-style-type: none"> • 支持的字段类型： TINYINT、SMALLINT、INT、BIGINT、DECIMAL、NUMERIC、FLOAT、REAL、SMALLMONEY、MONEY、BIT、DATE、DATETIME、DATETIME2、DATETIMEOFFSET、TIME、TIMESTAMP、XML、CHAR、VARCHAR、NCHAR、NVARCHAR、BINARY、VARBINARY、IMAGE、HIERARCHYID、NTEXT、TEXT、UNIQUEIDENTIFIER。 • 不支持的字段类型：SQL_VARIANT、GEOMETRY、GEOGRAPHY。 • 增量同步的范围： <ul style="list-style-type: none"> - 支持DML (Data Manipulation Language) : 包括INSERT、UPDATE、DELETE。 - 不支持DDL (Data Definition Language) 。

数据库账号权限要求

在使用DRS进行同步时，连接源库和目标库的数据库账号需要满足以下权限要求，才能启动实时同步任务。不同类型的同步任务，需要的账号权限也不同，详细可参考[表 5-223](#)进行赋权。DRS会在“预检查”阶段对数据库账号权限进行自动检查，并给出处理建议。

表 5-223 数据库账号权限

类型名称	增量同步
源数据库连接账号	需要具备sysadmin权限，或者view server state权限以及待同步数据库的db_datareader或db_owner权限。

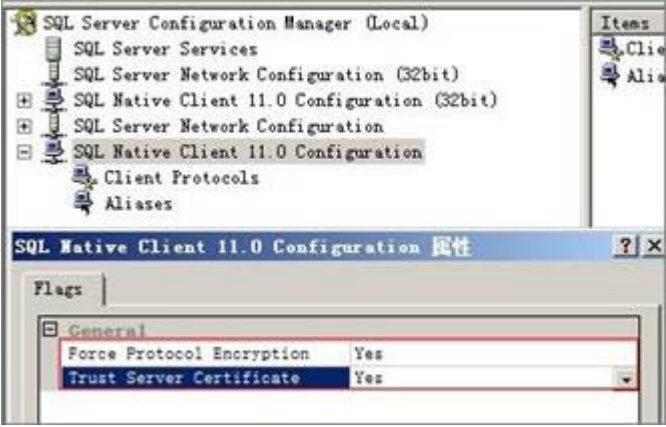
使用建议

- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，避免同步对业务造成性能影响。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。

使用须知

DRS增量同步过程一般包含三个阶段：任务启动阶段、增量阶段、结束任务阶段。为了确保同步各个阶段的平顺，在创建同步任务前，请务必阅读以下使用须知。

表 5-224 使用须知

类型名称	使用和操作约定
任务启动前须知	<ul style="list-style-type: none"> ● 源数据库参数要求： <ul style="list-style-type: none"> - 源数据库模式必须设置为FULL模式。 - 源数据库必须开启SQL Server Agent代理服务。 - 源数据库若开启客户端配置中的“强制协议加密（ Force Protocol Encrypton ）”，必须同时开启“信任服务器证书（ trust server certificate ）”，如下图5-212所示： <p style="text-align: center;">图 5-212 查看客户端属性</p>  <ul style="list-style-type: none"> ● 源数据库对象要求： <ul style="list-style-type: none"> - 源库中不允许存在被禁用的表的聚簇索引，否则会导致同步失败。 - 源库中不允许存在名称为cdc的用户名或schema。 - 源库待同步对象的库名、schema名、表名只能包含如下字符：字母、数字、下划线和中划线，库名长度不能超过64个字符。 - 源库待同步对象的表中列名不能包含如下字符：[]?。 ● 目标数据库参数要求： <ul style="list-style-type: none"> - 目标库为社区Kafka。 - Kafka的auto.create.topics.enable参数建议为false。 ● 其他使用须知： <ul style="list-style-type: none"> - 不支持源数据库主备切换，源数据库主备切换会导致同步任务失败。 - 不支持源数据库Microsoft SQL Server为TLS 1.0、TLS 1.1协议的同步，如果需要同步，建议源库升级到TLS 1.2及以上版本。 - SQL Server为源的增量同步基于SQL Server提供的CDC能力，如果SQL Server源库的CDC功能出现异常（如被关闭、事务日志满等情况），则增量同步就会受到影响。

类型名称	使用和操作约定
增量同步过程须知	<ul style="list-style-type: none"> 请勿修改源库和目标库的端口号，请勿修改、删除源库和目标库连接用户的密码、权限，否则可能导致任务失败。 请勿在目标库做写入操作，否则可能导致数据不一致。 请勿在Kafka上删除接收DRS数据的topic，否则可能导致任务失败。 不支持DDL操作，源数据库进行的DDL操作不会同步到目标数据库。 不支持大数据类型IMAGE、TEXT、NTEXT的删除操作。 支持任务再编辑追加同步对象。

操作步骤

本小节以自建Microsoft SQL Server->Kafka的同步为示例，介绍如何使用数据复制服务配置公网网络场景下的实时同步任务。

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 5-213 同步任务信息

表 5-225 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' ' " \ 特殊字符。

- 同步实例信息

图 5-214 同步实例信息



表 5-226 同步实例信息

参数	描述
数据流动方向	选择“自建-自建”。
源数据库引擎	选择“Microsoft SQL Server”。
目标数据库引擎	选择“Kafka”。
网络类型	<p>此处以“公网网络”为示例，可选“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
VPC	选择可用的虚拟私有云。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
内网安全组	请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。
同步模式	增量：增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

- 可用区

图 5-215 可用区



表 5-227 任务可用区

参数	描述
可用区	DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。

- 企业项目和标签

图 5-216 企业项目和标签



表 5-228 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 在“源库及目标库”页面，同步实例创建成功后，填写源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

测试连接时需要确保DRS任务与源库、目标库的连通性。

- **网络连通性**：确保源库和目标库允许DRS访问，通常至少需要放通DRS的IP。公网网络的DRS任务放通DRS的公网IP，VPC网络、VPN、专线网络的DRS任务放通DRS的内网IP，可参考[网络准备](#)章节进行网络打通。
- **账号连通性**：确保源库和目标库允许DRS通过用户名和密码访问。

图 5-217 源库信息

源库信息

不支持数据库参数和系统数据库迁移，源数据库参数设定和用户，作业将不会迁移至目标数据库中，请在目标数据库中使用参数组修改参数，手工导出导入用户和作业。

IP地址或域名

端口

数据库用户名

数据库密码

SSL安全连接

待实例创建成功后再进行测试连接

表 5-229 源库信息

参数	描述
IP地址或域名	源数据库Microsoft SQL Server的IP地址或域名。
端口	源数据库Microsoft SQL Server服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库Microsoft SQL Server的用户名。
数据库密码	源数据库Microsoft SQL Server的用户名所对应的密码。

图 5-218 目标库信息

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

表 5-230 目标库信息

参数	描述
IP地址或域名	目标数据库的IP地址或域名。
安全协议	支持四种认证方式的选择，PLAINTEXT、SSL、SASL_PLAINTEXT和SASL_SSL，详细说明可参考 Kafka认证方式 。

说明

源和目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择同步策略和同步对象，单击“下一步”。

图 5-219 同步对象

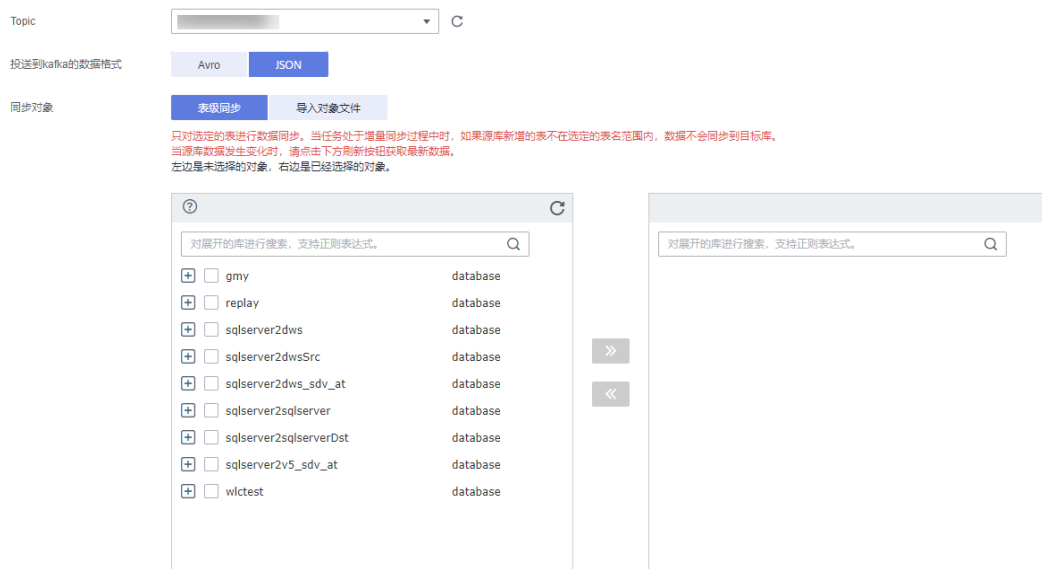



表 5-231 同步模式和对象

参数	描述
Topic	选择目标端需要同步到的Topic。
投送到kafka的数据格式	选择投送到kafka的数据格式。 <ul style="list-style-type: none"> Avro：可以显示Avro二进制编码，高效获取数据。 Json：为Json消息格式。 详细格式可参考 Kafka消息格式 。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，同步对象支持表级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 选择导入对象文件时，支持不同表同步到目标端不同的Topic，具体导入步骤和说明可参考导入同步对象。 使用导入对象功能时，同步Topic策略选择“集中投递到一个Topic”，才能使用对象名映射功能，否则会按Topic名字格式生成。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

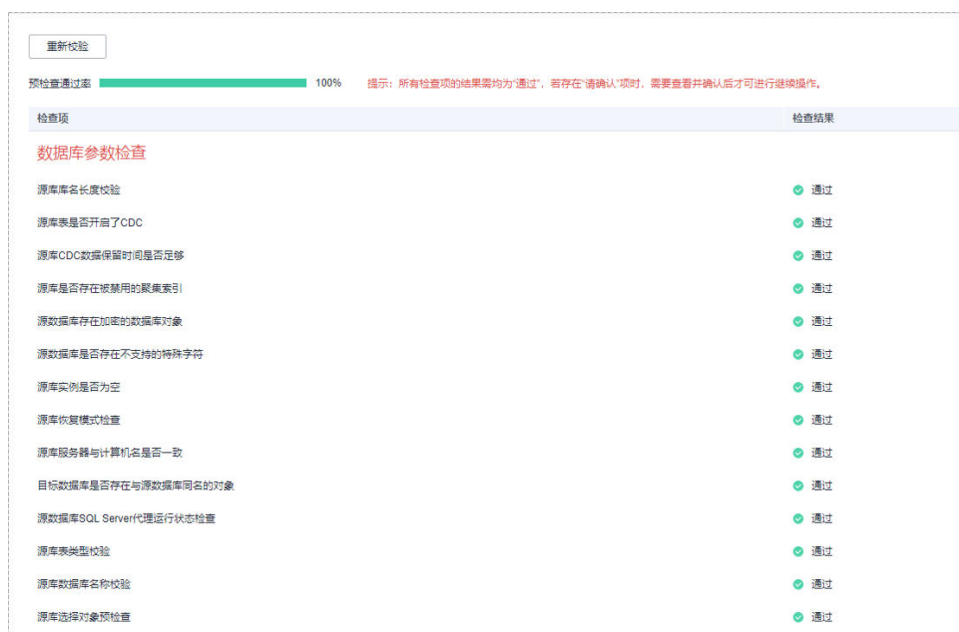
步骤5 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 5-220 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤6 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。

图 5-221 任务启动设置

The screenshot displays the 'Task Start Settings' interface. At the top, there are two tabs: '立即启动' (Immediately Start) and '稍后启动' (Start Later), with a help icon. Below this, several settings are listed:


- 任务异常通知设置**: A toggle switch is turned on.
- * SMN主题**: A dropdown menu is present, with a refresh and help icon.
- 时延阈值(s)**: A toggle switch is turned off.
- 数据异常通知**: A toggle switch is turned off.
- * 任务异常自动结束时间**: An input field contains the value '14', with a help icon and a note: '任务处于异常状态一段时间后，将会自动结束。单位为天。' (After a period of time in an abnormal state, the task will automatically end. Unit is days).

表 5-232 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。

参数	描述
时延阈值	<p>在增量同步阶段，源数据库和目标数据库之间的同步有时存在一个时间差，称为时延，单位为秒。</p> <p>时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。</p> <p>说明</p> <ul style="list-style-type: none"> • 当时延阈值设置为0时，不会发送通知给收件人。 • 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。 • 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	<p>该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。</p>
任务异常自动结束时间（天）	<p>设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。</p> <p>说明</p> <ul style="list-style-type: none"> • 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。 • 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤7 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

6 任务管理

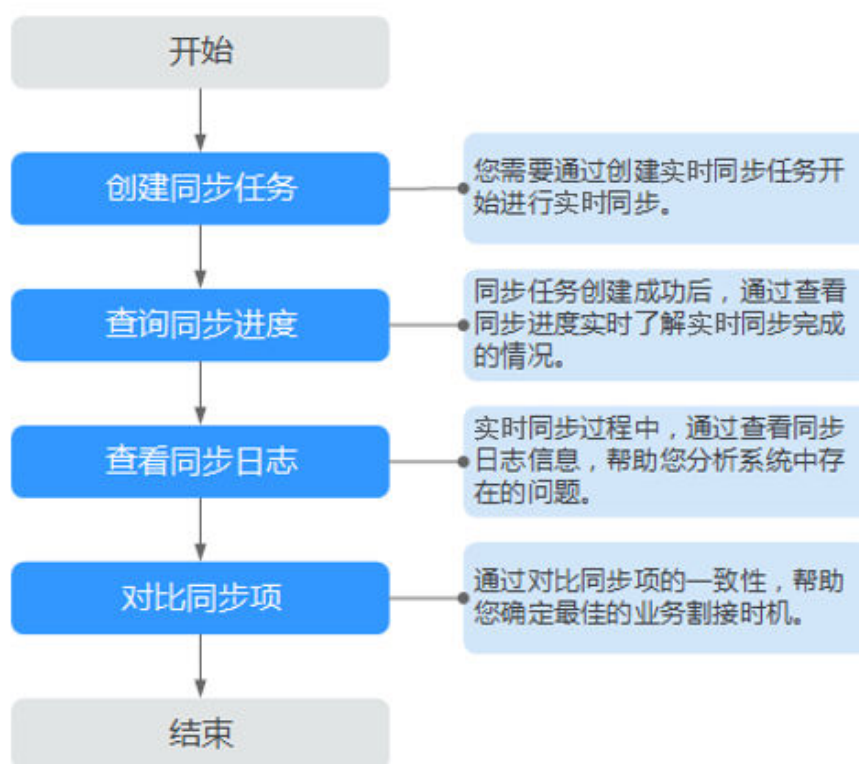
6.1 创建同步任务

使用流程

一次完整的实时同步，是通过创建实时同步任务，以同步任务作为导向，依次进行同步进度观察、同步日志分析、同步数据一致性对比等多项操作。通过多项指标和数据的对比分析，可以帮助您实现不同业务系统间的数据实时同步。

一次完整的实时同步包含如下过程：

图 6-1 实时同步流程



- **步骤一：创建同步任务。** 根据需要，选择源和目标数据库，创建同步任务。
- **步骤二：查询同步进度。** 同步过程中，可以通过查看同步进度了解实时同步的完成情况。
- **步骤三：查看同步日志。** 同步日志包含告警、错误和提示等类型的信息，可根据此类信息分析系统存在的问题。
- **步骤四：对比同步项。** 实时同步提供对比功能，可根据需要查看对象级对比、数据级对比等，来确保源和目标数据库的数据一致性。

本小节以MySQL->RDS for MySQL的实时同步为示例，介绍如何使用数据复制服务创建两个数据库实例之间的实时同步任务，其他存储引擎的配置流程类似。

操作步骤

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选区域、项目、任务名称、描述、同步实例信息，单击“开始创建”。

- 任务信息

图 6-2 同步任务信息

1. 该页面仅有任务名称和描述可修改，其他在点击开始创建后均不可修改。
2. 创建迁移第一步需要创建虚拟资源，在配置完源库及目标库基本信息后，系统会去创建资源，为后续数据迁移做准备，虚拟资源一旦创建好后，就不能修改。

区域: [Dropdown] ?
不同区域的资源之间内网不互通，请选择靠近您客户的区域，可以降低网络时延、提高访问速度。

项目: [Dropdown]

* 任务名称: DRS-6435 ?

描述: [Text Area] ?
0/256

表 6-1 任务和描述

参数	描述
区域	当前所在区域，可进行切换。
项目	当前区域对应的项目，可进行切换。
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含! = < > & ' " \ 特殊字符。

- 同步实例信息

图 6-3 同步实例信息

同步实例信息 ?

以下信息确认后不可修改，请谨慎填写，以免影响数据迁移、恢复等操作。

* 数据流动方向: [Radio: In-Cloud] [Radio: Out-Cloud] [Radio: Bidirectional] ?
DRS默认源数据库和目标数据库中至少有一方为本云数据库实例，入云目标数据库为本云数据库实例的场景，本云数据库实例之间同步，请勾选入云。

* 源数据库引擎: [MySQL] [Oracle] [DB2 for LUW] [DDM] [MariaDB] [MongoDB] [PostgreSQL] [Microsoft SQL Server] [GaussDB(for MySQL)] [TDG]

* 目标数据库引擎: [MySQL] [GaussDB(DWS)] [GaussDB 分布式版] [GaussDB 主备版] [PostgreSQL] [GaussDB(for MySQL)]

* 网络类型: [Public Network] ?
DRS将会自动为DRS实例创建弹性公网IP，请任意结束页将数据源网络地址公网IP。

* DRDS任务类型: [HA2] [HA2] ?
采用单个节点部署架构，与多AZ部署方式对比，每个任务只提供一个节点，但具有高性价比，适用于业务数据量较小，短期同步，并对中等的成本有要求的情况。

* 目标数据库实例: [Dropdown] [Refresh] [Reset] [Cancel] ?
同步实例所在子网: [Dropdown] [Refresh] [Reset] [Cancel] ?

* 内网安全组: [default] ?

* 同步模式: [Full] [Incremental] [Full] [Incremental] ?
该模式为高实时性的实时同步，通过全量过程为后续增量数据的初始化，增量同步的数据通过解译日志等技术，将源端和目标端的数据保持一致。

表 6-2 同步实例信息

参数	描述
数据流动方向	选择“入云”，即目标端数据库为本云数据库。
源数据库引擎	选择“MySQL”。
目标数据库引擎	选择“MySQL”。

参数	描述
网络类型	<p>此处以“公网网络”为示例。可根据业务场景选择“公网网络”、“VPC网络”和“VPN、专线网络”。</p> <ul style="list-style-type: none"> - VPC网络：适合云上同账号同Region数据库之间的同步。 - 公网网络：适合通过公网网络把其他云下或其他平台的数据库同步到目标数据库，该类型要求数据库绑定弹性公网IP（EIP）。 - VPN、专线网络：适合通过专线网络，实现其他云下自建数据库与云上数据库同步、云上同Region跨账号、或云上跨Region的数据库之间的同步。
DRS任务类型	<p>选择DRS任务类型，可选“单AZ”和“双AZ”。</p> <ul style="list-style-type: none"> - 双AZ：一主一备的经典高可用架构，提高了DRS任务的可靠性。双AZ任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。适用于业务数据量较大，需要长期同步，并对中断时间有严格要求的场景。 - 单AZ：采用单个节点部署架构，每个任务只创建一个节点，性价比较高。适用于业务数据量较小，短期同步，并对中断时间无严格要求的场景。 <p>目前仅部分链路支持DRS任务类型的选择，详细说明可参考双AZ任务倒换。</p>
目标数据库实例	<p>用户所创建的关系型数据库实例。</p> <p>说明</p> <ul style="list-style-type: none"> - 目标数据库实例不支持选择只读实例。 - 目标数据库实例可以和源数据库选择同一个实例。
同步实例所在子网	<p>请选择同步实例所在的子网。也可以单击“查看子网”，跳转至“网络控制台”查看实例所在子网帮助选择。</p> <p>默认值为当前所选数据库实例所在子网，请选择有可用IP地址的子网。为确保同步实例创建成功，仅显示已经开启DHCP的子网。</p>
内网安全组	<p>请选择内网安全组。内网安全组限制实例的安全访问规则，加强安全访问。</p>

参数	描述
同步模式	<p>DRS任务支持的同步模式，此处以“全量+增量”为示例。DRS全量、增量底层实现原理可参考产品架构和功能原理。</p> <ul style="list-style-type: none"> “全量+增量”： 该模式为数据持续性实时同步，通过全量过程完成目标端数据库的初始化后，增量同步阶段通过解析日志等技术，将源端和目标端数据保持数据持续一致。 <p>说明 选择“全量+增量”同步模式，增量同步可以在全量同步完成的基础上实现数据的持续同步，无需中断业务，实现同步过程中源业务和数据库继续对外提供访问。</p> <ul style="list-style-type: none"> “全量”： 该模式为数据库一次性同步，适用于可中断业务的数据库同步场景，全量同步将非系统数据库的全部数据库对象和数据一次性同步至目标端数据库。 “增量”： 增量同步通过解析日志等技术，将源端产生的增量数据同步至目标端。

● 任务类型

图 6-4 任务类型



表 6-3 任务类型信息

参数	描述
规格类型	<p>选择DRS实例的规格类型，不同规格类型对应不同性能的性能上限，详细可参考实时同步链路规格说明。</p> <p>说明 目前，DRS同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持降低规格，具体操作及约束限制可参考同步规格变更。</p>
可用区	<p>DRS任务创建的可用区，选择跟源或目标库相同的可用区性能更优。</p> <p>DRS任务类型选择“双AZ”时，可用区分为“主可用区”和“备可用区”。</p> <p>图 6-5 可用区</p>

- 企业项目和标签

图 6-6 企业项目和标签



表 6-4 企业项目和标签

参数	描述
企业项目	<p>企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。请在下拉框中选择所在的企业项目，其中，default为默认项目。</p> <p>更多关于企业项目的信息，请参见《企业管理用户指南》。</p> <p>如果需要自定义企业项目，请在控制台右上角单击“企业”，进入“企业项目管理”页面创建，具体请参见《企业管理用户指南》中“创建企业项目”的内容。</p>
标签	<ul style="list-style-type: none"> - 可选配置，对同步任务的标识。使用标签可方便管理您的任务。每个任务最多支持20个标签配额。 - 如果您的组织已经设定DRS的相关标签策略，则需按照标签策略规则为任务添加标签。标签如果不符合标签策略的规则，则可能会导致任务创建失败，请联系组织管理员了解标签策略详情。 - 任务创建成功后，您可以单击任务名称，在“标签”页签下查看对应标签。关于标签的详细操作，请参见标签管理。

说明

对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。

步骤3 同步实例创建成功后，在“源库及目标库”页面，填选源库信息和目标库信息后，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库和目标库连通后，勾选协议，单击“下一步”。

入云和出云场景中源库和目标库需要填选的信息不同，具体以界面为准。

- 源库信息

图 6-7 源库信息

源库信息

IP地址或域名 请确保本云安全组允许访问源数据库公网网址，[点击这里](#)

端口

数据库用户名

不支持数据库参数和系统数据库同步，源数据库参数设置和用户将不会同步至目标数据库中，请在目标数据库中使用参数组修改参数，手工创建用户。

数据库密码

SSL安全连接

如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确，并上传SSL证书

加密证书

● 测试成功

表 6-5 源库信息

参数	描述
IP地址或域名	源数据库的IP地址或域名。
端口	源数据库服务端口，可输入范围为1~65535间的整数。
数据库用户名	源数据库的用户名。
数据库密码	源数据库用户名所对应的密码。支持在任务创建后修改密码。 任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面的，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。
SSL安全连接	如启用SSL安全连接，请在源库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

源数据库IP地址、端口、用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

- 目标库信息

图 6-8 目标库信息

目标库信息

数据库实例名称

数据库用户名

数据库密码

SSL安全连接

✔ 测试成功 ?

表 6-6 目标库信息

参数	描述
数据库实例名称	默认为创建同步任务时选择的关系型数据库实例，不可进行修改。
数据库用户名	目标数据库对应的用户名。
数据库密码	目标数据库用户名对应的密码。支持在任务创建后修改密码。 任务为启动中、全量同步、增量同步、增量同步失败状态时，可在“基本信息”页面，单击“连接信息”后的“修改连接信息”，在弹出的对话框中修改密码。
SSL安全连接	如启用SSL安全连接，请在目标库开启SSL，并确保相关配置正确后上传SSL证书。 说明 <ul style="list-style-type: none"> - 最大支持上传500KB的证书文件。 - 如果不启用SSL安全连接，请自行承担数据安全风险。

说明

目标数据库用户名和密码将在同步过程中被加密暂存到数据库和同步实例主机上，待该任务删除后会永久清除。

步骤4 在“设置同步”页面，选择数据冲突策略和同步对象，单击“下一步”。

图 6-9 同步模式

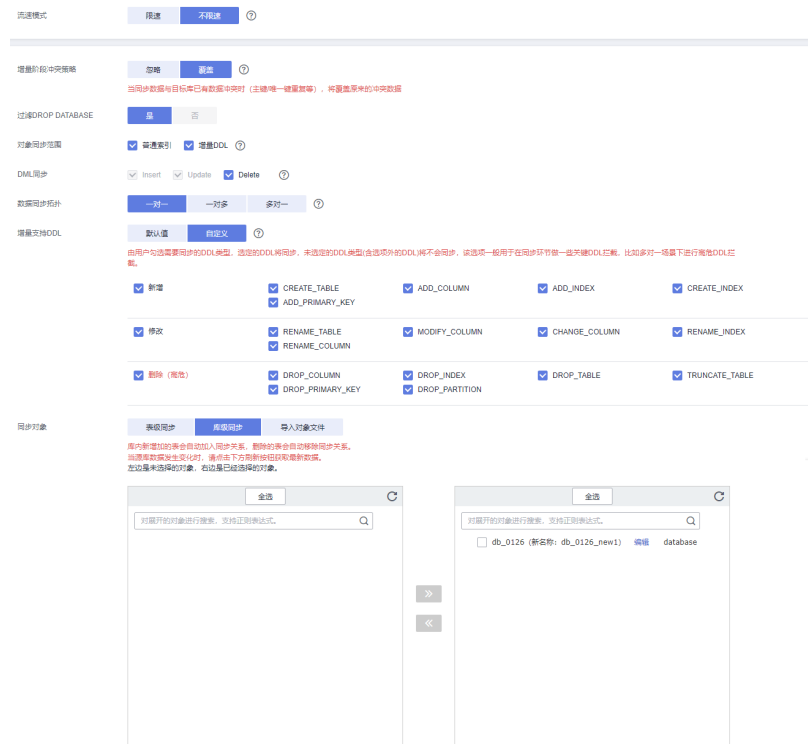




表 6-7 同步模式和对象

参数	描述
流速模式	<p>流速模式支持限速和不限速，默认为不限速。限速模式只对全量阶段生效，增量阶段不生效。</p> <ul style="list-style-type: none"> 限速 自定义的最大同步速度，全量同步过程中的同步速度将不会超过该速度。 当流速模式选择了“限速”时，你需要通过流速设置来定时控制同步速度。流速设置通常包括限速时间段和流速大小的设置。默认的限速时间段为“全天限流”，您也可以根据业务需求选择“时段限流”。自定义的时段限流支持最多设置3个定时任务，每个定时任务之间不能存在交叉的时间段，未设定在限速时间段的时间默认为不限速。 流速的大小需要根据业务场景来设置，不能超过9999MB/s。 <p>图 6-10 设置流速模式</p>  不限速 对同步速度不进行限制，通常会最大化使用源数据库的出口带宽。该流速模式同时会对源数据库造成读消耗，消耗取决于源数据库的出口带宽。比如源数据库的出口带宽为100MB/s，假设高速模式使用了80%带宽，则同步对源数据库将造成80MB/s的读操作IO消耗。 <p>说明</p> <ul style="list-style-type: none"> - 限速模式只对全量阶段生效，增量阶段不生效。 - 您也可以在创建任务后修改流速模式。具体方法请参见修改流速模式。
增量阶段冲突策略	<p>该冲突策略特指增量同步中的冲突处理策略，全量阶段的冲突默认忽略。冲突策略目前支持如下形式：</p> <ul style="list-style-type: none"> 忽略 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将跳过冲突数据，继续进行后续同步。选择忽略可能导致源库与目标库数据不一致。 覆盖 当同步数据与目标数据库已有数据冲突时（主键/唯一键存在重复等），将覆盖原来的冲突数据。

参数	描述
过滤DROP DATABASE	<p>实时同步过程中，源数据库端执行的DDL操作在一定程度上会影响数据的同步能力，为了降低同步数据的风险，数据复制服务提供了过滤DDL操作的功能，目前支持默认过滤删除数据库的操作。</p> <ul style="list-style-type: none"> 是，表示过程中不会同步用户在源数据库端执行的删除数据库的操作。 否，则表示过程中将相关操作同步到目标库。
对象同步范围	<p>对象同步范围支持普通索引和增量DDL同步。您可以根据业务需求选择是否进行同步。</p>
启动位点	<p>步骤步骤2的同步模式选择“增量”时可见，增量同步的启动位点，任务的源库日志从位点后开始获取（不含当前启动位点）。</p> <p>通过show master status命令获取源库位点，根据提示分别填写File、Position、Executed_Gtid_Set。</p>
数据同步拓扑	<p>对象同步范围勾选“增量DDL”时可见，数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。数据同步拓扑说明可参考数据同步拓扑介绍。</p> <p>说明</p> <p>“数据同步拓扑”选择目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。</p>
增量支持DDL	<p>对象同步范围勾选“增量DDL”时可见，选择增量同步的DDL类型，用户根据需求可选择“默认值”或“自定义”。</p> <ul style="list-style-type: none"> “默认值”： <ul style="list-style-type: none"> 库级同步情况下，Binlog中权限类DDL除外，所有库内对象相关DDL将会同步至目标，常见的DDL有CREATE_TABLE，RENAME_TABLE等。 表级同步情况下，已选定的同步表相关在Binlog中DDL将全部同步，与同步表无关的对象DDL将不同步。常见的DDL有ADD_COLUMN，MODIFY_COLUMN，ALTER_COLUMN等。 “自定义”：用户根据需求，选择需要同步的DDL类型，不同链路支持的DDL类型以界面显示为准。当“对象同步范围”勾选“增量DDL”，但是“自定义”未选择任一DDL类型时，默认不同步DDL。 <p>说明</p> <ul style="list-style-type: none"> “增量支持DDL”功能目前仅支持白名单用户，需要提交工单申请才能使用。您可以在管理控制台右上角，选择“工单 > 新建工单”，完成工单提交。 一对一、一对多场景：如果业务上认为源和目标应该使用保持严格一致，那么高危类DDL也应该勾选并同步。如果业务上确定某个高危DDL不应该发生，则可以不勾选同步高危类DDL，这样DRS将拦截过滤这个DDL，从而起到保护目标数据的作用。但需要知晓，过滤DDL的附带问题是可能导致同步失败，例如过滤删除列动作。 多对一数据聚合场景：最佳方式是推荐只选择同步加列DDL，其他大部分DDL同步都可能因目标表修改而导致数据不一致或多对一中其他任务失败的情况发生。

参数	描述
同步对象	<p>左侧框是源数据库对象，右侧框是已经选择的对象，可选表级同步、库级同步、导入对象文件，您可以根据业务场景选择对应的数据进行同步。</p> <ul style="list-style-type: none"> 如果同步的对象在源数据库和目标数据库中的命名不同，可在右侧已选对象框中修改目标库对象映射名称，具体操作可参考对象名映射。 <ul style="list-style-type: none"> 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 多对一场景下，选择表级同步时，如果要对已保存的库映射名称进行修改，需展开库进行操作。 选择导入对象文件，具体步骤和说明可参考导入同步对象。 <p>说明</p> <ul style="list-style-type: none"> 选择对象的时候支持对展开的库进行搜索，以便您快速选择需要的数据库对象。 如果有切换源数据库的操作或源库同步对象变化的情况，请在选择同步对象前单击右上角的 ，以确保待选择的对象为最新源数据库对象。 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。 选择的同步对象名称中不能包含空格。

步骤5 在“数据加工”页面，根据需要选择数据加工的方式。

- 如果不需要数据加工，单击“下一步”。
- 如果需要数据加工，可选择“数据过滤”、“附加列”或“列加工”，参考[数据加工](#)章节，设置相关规则。

图 6-11 数据加工



源库/库	新名称	列名	操作类型	字符类型	操作
		a1	concat	varchar(256)	
		a2	concat	long	
db_has_pk_test1_table16 db_has_pk_test1_table12 db_has_pk_test1_table14 db_has_pk_test1_table10		a3	create_time	datetime	
		a4	update_time	timestamp	
		a5	concat	varchar(256)	

步骤6 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 6-12 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

步骤7 在“任务确认”页面，设置同步任务的启动时间、任务异常通知设置、SMN主题、时延阈值、任务异常自动结束时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。


图 6-13 任务启动设置



表 6-8 任务启动设置

参数	描述
启动时间	同步任务的启动时间可以根据业务需求，设置为“立即启动”或“稍后启动”。 说明 预计同步任务启动后，会对源数据库和目标数据库的性能产生影响，建议选择业务低峰期，合理设置同步任务的启动时间。
任务异常通知设置	该项为可选参数，开启之后，选择对应的SMN主题。当同步任务状态、时延指标或数据异常时，系统将发送通知。
SMN主题	“任务异常通知设置”项开启后可见，需提前在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	在增量同步阶段，源数据库和目标数据库之间的同步有时会有一个时间差，称为时延，单位为秒。 时延阈值设置是指时延超过一定的值后（时延阈值范围为0到3600s），DRS可以发送告警通知。告警通知将在时延稳定超过设定的阈值6min后发送，避免出现由于时延波动反复发送告警通知的情况。 说明 <ul style="list-style-type: none">• 当时延阈值设置为0时，不会发送通知给收件人。• 首次进入增量同步阶段，会有较多数据等待同步，存在较大的时延，属于正常情况，不在此功能的监控范围之内。• 设置时延阈值之前，需要设置任务异常通知。
数据异常通知	该项为可选参数，开启之后，当同步任务数据异常时，系统将发送通知。
任务异常自动结束时间（天）	设置任务异常自动结束天数，输入值必须在14到100之间，默认值14。 说明 <ul style="list-style-type: none">• 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。• 异常状态下的任务仍然会计费，而长时间异常的任务无法续传和恢复。设置任务异常自动结束天数后，异常且超时的任务将会自动结束，以免产生不必要的费用。

步骤8 同步任务提交后，您可在“实时同步管理”页面，查看并[管理自己的任务](#)。

- 您可查看任务提交后的状态，状态请参见[任务状态说明](#)。
- 在任务列表的右上角，单击  刷新列表，可查看到最新的任务状态。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源，此时DRS任务IP会发生改变。

----结束

相关链接

- [实时同步支持的数据库类型和版本](#)
- [准备工作概览](#)
- [同步方案概览](#)
- [数据同步拓扑介绍](#)

6.2 查询同步进度

实时同步进度展示了同步过程中，任务的表同步进度，DRS通过流式进度展示，可以帮助您实时了解同步完成的情况。

- 在全量同步过程中，DRS展示进度总览，您可以查看结构、数据、索引迁移的进度。同步进度是根据同步对象的数量来统计百分比进行展示，当显示为100%时，表示该项同步完成。同步过程中，数据和索引的同步相对较慢。
- 在全量同步过程中，DRS展示同步进度同明细，您可以查看具体同步对象的进度，当“对象数目”和“已迁移数目”相等时，表示该对象已经同步完成，可通过“查看详情”查看每个对象的同步进度。增量同步中，“进度明细”将不再显示，您可以使用“同步对比”页签查看一致性情况。
- 在增量同步过程中，DRS展示增量时延，可通过时延大小判断源和目标同步情况，时延为0代表源和目标瞬时一致，无新的事务需要同步。

前提条件

已登录数据复制服务控制台。

操作步骤

步骤1 在“实时同步管理”界面，选中指定同步任务，单击任务名称，进入“基本信息”页面。

步骤2 单击“同步进度”页签，查看表同步进度。

- 查看同步的百分比进度。
 - 当全量同步完成时，显示全量已完成100%。
 - 当全量同步完成后，开始进行增量同步，可在“同步进度”页签下，查看增量同步时延。
 - 增量时延也可在“实时同步管理”界面查看，当增量时延超过用户设置或系统默认的时延阈值时，任务管理界面增量时延显示为红色。
 - 当时延为0s时，说明源数据库和目标数据库的数据是实时同步的。

📖 说明

时延 = 源库当前系统时间 - 成功同步到目标库的最后一个事务在源库的提交成功时间。

一个事务同步的完整过程如下：

1. 源端数据库的抽取；
2. 经过网络的传输；
3. 由DRS进行日志解析；
4. 最终在目标数据库上的执行完成。

这样完成了事务从源到目标的同步，时延为该事务最后在目标数据库上执行完成时的源库当前系统时间（current_time）与该事务在源库的提交成功时间（committed_time）的时差，时延为0代表源和目标瞬时一致，无新的事务需要同步。

⚠️ 注意

长时间未提交事务和频繁DDL操作均可以造成高时延。

- 查看同步对象的进度。在“进度明细”下，单击同步对象“详细信息”列的“查看详情”，查看对象的同步进度，进入增量同步后，该数据将不再显示，您可以使用“同步对比”页签，进行一致性对比。

图 6-14 同步进度明细

进度明细 ^

提示：当任务进入增量同步中时，请停用同步对比功能查看数据同步详情。

同步对象	同步数据	状态	已同步数据	查看详情
index	5	🔄 同步中	0	查看详情
table_indexes	0	🔄 同步中	0	查看详情
schema	1	✅ 已完成	1	查看详情
sequence	0	✅ 已完成	0	查看详情
table	1	✅ 已完成	1	查看详情
table_structure	1	✅ 已完成	1	查看详情

步骤3 针对MySQL的同步场景，您可以在“同步进度”页签中查看源位点和已消费位点的信息。

📖 说明

展示的位点信息刷新频率为每10s一次。

图 6-15 同步位点信息



步骤4 针对Oracle为源、MySQL->Kafka、GaussDB->Kafka的同步场景，您可以在“同步进度”页签的“链路监控”表中，查看该同步任务执行过程中对源库分别做DML操作（Insert、Delete、Update）的次数。在“链路监控”列表的右上角，单击刷新列表，可查看到最新的监控数据。

图 6-16 链路监控

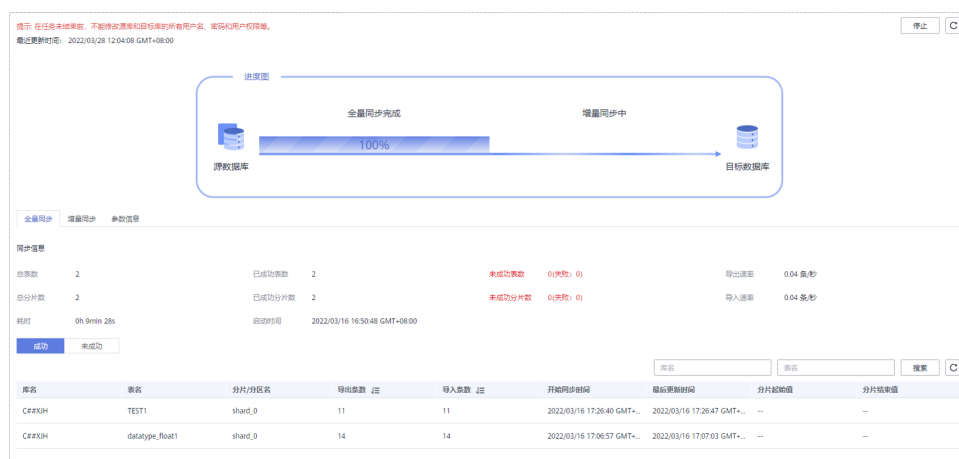


说明

1. 需要在源库进行DML操作并执行Commit后，该次操作数据才能够生效并统计到“链路监控”表中。
2. DRS 根据重做日志的信息统计操作次数。
3. Oracle 对于含 LOB 类型的表进行 INSERT、UPDATE 操作时会记录更多的重做日志信息：先基于非 LOB 的普通字段进行 INSERT、UPDATE操作，再基于 LOB 类型的字段执行。但 DRS 仅根据重做日志产生的操作信息进行统计，因此在链路监控操作审计时可能产生与实际影响行数不符的操作次数。
4. Oracle 的 MERGE 语句作为一种方言化的 DML，在重做日志层面转化为 INSERT，UPDATE，DELETE 三种操作。DRS 会根据重做日志的信息将 MERGE 影响的行数归并到 INSERT，UPDATE，DELETE 三种操作的统计中。

步骤5 针对Oracle->GaussDB主备版和分布式、GaussDB主备版->MySQL、GaussDB分布式版->MySQL、Oracle->MySQL（单增量）、Oracle->GaussDB(DWS)（单增量）的同步场景，您可以在“同步进度”页签中搜索查看详细的同步对象信息，对象搜索区分大小写。

图 6-17 同步信息



----结束

6.3 查看同步日志

同步日志记录了实时同步过程中的信息，包含告警、错误和提示等类型的信息。实时同步过程中，可以通过查看同步日志信息，帮助您分析系统中存在的问题。

前提条件

已登录数据复制服务控制台。

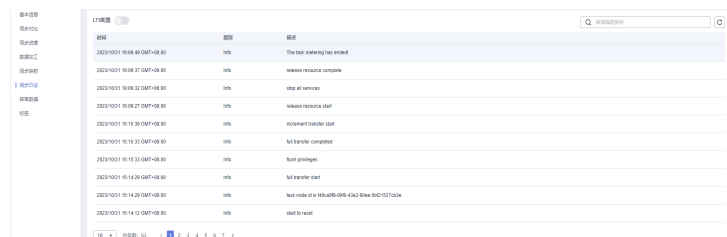
操作步骤

步骤1 在“实时同步管理”界面，选择指定的同步任务，单击任务名称。

步骤2 在“同步日志”页签，查看当前同步任务的日志。

您可查看到日志对应的时间，级别和描述。

图 6-18 同步日志



时间	级别	描述
2023/03/15 10:08:40 GMT+08:00	info	The task monitoring has ended.
2023/03/15 10:08:27 GMT+08:00	info	Release resource completely.
2023/03/15 10:08:20 GMT+08:00	info	Stop all services.
2023/03/15 10:08:20 GMT+08:00	info	Release resource start.
2023/03/15 10:14:30 GMT+08:00	info	Incremental transfer start.
2023/03/15 10:13:20 GMT+08:00	info	Full transfer completed.
2023/03/15 10:13:20 GMT+08:00	info	Full progress.
2023/03/15 10:14:20 GMT+08:00	info	Full transfer start.
2023/03/15 10:14:20 GMT+08:00	info	Task made via HdfsURL=HDFS_43a23D9e5002702743a
2023/03/15 10:14:02 GMT+08:00	info	Start to read.

除此以外，DRS支持对接云日志服务（Log Tank Service，简称LTS），进行LTS日志配置后，DRS实例新生成的所有日志记录会上传到云日志服务进行管理，详情可参考[日志配置管理](#)。

----结束

6.4 对比同步项

对比使用场景

对比实时同步项可以清晰反馈出源数据库和目标数据库的数据是否存在差异。为了尽可能减少业务的影响和业务中断时间，实时同步场景提供了以下几种对比方式：

- 对象级对比：支持对数据库、索引、表、视图、存储过程和函数、表的排序规则等对象进行对比。
- 数据级对比分为行对比和内容对比。
 - 行对比：对比同步的表的行数是否一致，只查询表的行数，对比速度较快，建议优先使用。
 - 内容对比：对比同步的表内数据是否一致，需要查询每条数据进行对比，相较于行对比，内容对比比较慢。
- 用户对比：对源库和目标库的账号名称、权限进行对比。
- 周期性对比：定时对源数据库与目标数据库表的行数进行对比并展示结果，需要用户[开启并设置对比策略](#)。

在使用对比功能查看数据一致性时，建议优先使用行数对比，确定行数是否一致。如果出现行数不一致的情况，可以针对不一致的表，使用内容对比，来确定具体有哪些数据不一致。

表 6-9 支持的对比方式

同步方向	数据流向	对象级对比	行对比	内容对比	动态内容对比	用户对比
入云	MySQL->MySQL	支持	支持	支持	支持	不支持
入云	MySQL->PostgreSQL	支持	支持	不支持	不支持	不支持
入云	MySQL->GaussDB分布式版	支持	支持	不支持	不支持	不支持
入云	MySQL->GaussDB主备版	支持	支持	不支持	不支持	不支持
入云	MySQL->GaussDB(for MySQL)	支持	支持	支持	支持	不支持
入云	MySQL->GaussDB(DWS)	支持	支持	不支持	不支持	不支持
入云	MySQL->MariaDB	支持	支持	不支持	不支持	不支持
入云	PostgreSQL->PostgreSQL	支持	支持	支持	支持	支持
入云	PostgreSQL->GaussDB(DWS)	支持	支持	不支持	不支持	不支持
入云	PostgreSQL->GaussDB 主备版	支持	支持	不支持	不支持	不支持
入云	PostgreSQL->GaussDB分布式版	支持	支持	不支持	不支持	不支持
入云	DDM->MySQL	支持	支持	不支持	不支持	不支持
入云	DDM->GaussDB(DWS)	不支持	支持	不支持	不支持	不支持
入云	DDM->DDM	支持	支持	不支持	不支持	不支持
入云	Oracle->MySQL	支持	支持	支持	支持	不支持
入云	Oracle->GaussDB(for MySQL)	支持	支持	支持	支持	不支持
入云	Oracle->GaussDB主备版	支持	支持	支持	支持	不支持
入云	Oracle->GaussDB分布式版	支持	支持	支持	支持	不支持
入云	Oracle->DDM	支持	支持	不支持	不支持	不支持
入云	Oracle->GaussDB(DWS)	支持	支持	不支持	不支持	不支持
入云	Oracle->PostgreSQL	支持	支持	不支持	不支持	不支持
入云	DB2 for LUW->GaussDB主备版	支持	支持	不支持	不支持	不支持
入云	DB2 for LUW->GaussDB分布式版	支持	支持	不支持	不支持	不支持
入云	DB2 for LUW->GaussDB(DWS)	支持	支持	不支持	不支持	不支持

同步方向	数据流向	对象级对比	行对比	内容对比	动态内容对比	用户对比
入云	TiDB->GaussDB(for MySQL)	支持	支持	不支持	不支持	不支持
入云	Microsoft SQL Server->GaussDB(DWS)	支持	支持	不支持	不支持	不支持
入云	Microsoft SQL Server->GaussDB主备版	支持	支持	支持	不支持	不支持
入云	Microsoft SQL Server->GaussDB分布式版	支持	支持	支持	不支持	不支持
入云	Microsoft SQL Server->Microsoft SQL Server	支持	支持	不支持	不支持	不支持
入云	MongoDB->DDS	支持	支持	不支持	不支持	不支持
入云	MariaDB->MariaDB	支持	支持	支持	支持	不支持
入云	MariaDB->MySQL	支持	支持	不支持	不支持	不支持
入云	MariaDB->GaussDB(for MySQL)	支持	支持	不支持	不支持	不支持
入云	GaussDB(for MySQL)->GaussDB(for MySQL)	支持	支持	支持	支持	不支持
出云	MySQL->MySQL	支持	支持	支持	支持	不支持
出云	MySQL->kafka	不支持	不支持	不支持	不支持	不支持
出云	MySQL->CSS/ES	支持	支持	不支持	不支持	不支持
出云	MySQL->MariaDB	支持	支持	不支持	不支持	不支持
出云	DDM->MySQL	支持	支持	不支持	不支持	不支持
出云	DDM->Oracle	支持	支持	不支持	不支持	不支持
出云	DDM->Kafka	不支持	不支持	不支持	不支持	不支持
出云	DDS->MongoDB	支持	支持	支持	不支持	不支持
出云	DDS->Kafka	不支持	不支持	不支持	不支持	不支持
出云	PostgreSQL->PostgreSQL	支持	支持	支持	支持	不支持
出云	PostgreSQL->Kafka	不支持	不支持	不支持	不支持	不支持
出云	GaussDB主备版->MySQL	支持	支持	不支持	不支持	不支持
出云	GaussDB主备版->Oracle	支持	支持	不支持	不支持	不支持
出云	GaussDB主备版->Kafka	不支持	不支持	不支持	不支持	不支持
出云	GaussDB主备版->GaussDB(DWS)	支持	支持	支持	不支持	不支持

同步方向	数据流向	对象级对比	行对比	内容对比	动态内容对比	用户对比
出云	GaussDB主备版->GaussDB分布式版	支持	支持	不支持	不支持	不支持
出云	GaussDB主备版->GaussDB主备版	支持	支持	不支持	不支持	不支持
出云	GaussDB分布式版->MySQL	支持	支持	不支持	不支持	不支持
出云	GaussDB分布式版->Oracle	支持	支持	不支持	不支持	不支持
出云	GaussDB分布式版->GaussDB(DWS)	支持	支持	支持	不支持	不支持
出云	GaussDB分布式版->Kafka	不支持	不支持	不支持	不支持	不支持
出云	GaussDB分布式版->GaussDB布式版	支持	支持	不支持	不支持	不支持
出云	GaussDB分布式版->GaussDB主备版	支持	支持	不支持	不支持	不支持
出云	GaussDB(for MySQL)->MySQL	支持	支持	支持	支持	不支持
出云	GaussDB(for MySQL)->GaussDB(DWS)	支持	支持	不支持	不支持	不支持
出云	GaussDB(for MySQL)->Kafka	不支持	不支持	不支持	不支持	不支持
出云	GaussDB(for MySQL)->CSS/ES	支持	支持	不支持	不支持	不支持
出云	GaussDB(for MySQL)->Oracle	支持	支持	不支持	不支持	不支持
出云	MariaDB->MariaDB	支持	支持	支持	支持	不支持
出云	Microsoft SQL Server->Kafka	不支持	不支持	不支持	不支持	不支持
自建-自建	Oracle->Kafka	不支持	不支持	不支持	不支持	不支持
自建-自建	Oracle->GaussDB主备版	支持	支持	支持	支持	不支持
自建-自建	Oracle->GaussDB分布式版	支持	支持	支持	支持	不支持

同步方向	数据流向	对象级对比	行对比	内容对比	动态内容对比	用户对比
自建-自建	MySQL->Kafka	不支持	不支持	不支持	不支持	不支持
自建-自建	MySQL->CSS/ES	支持	支持	不支持	不支持	不支持
自建-自建	MySQL->GaussDB主备版	支持	支持	不支持	不支持	不支持
自建-自建	MySQL->GaussDB分布式版	支持	支持	不支持	不支持	不支持
自建-自建	PostgreSQL->Kafka	不支持	不支持	不支持	不支持	不支持
自建-自建	GaussDB主备版->MySQL	支持	支持	不支持	不支持	不支持
自建-自建	GaussDB主备版->Oracle	支持	支持	不支持	不支持	不支持
自建-自建	GaussDB主备版->Kafka	不支持	不支持	不支持	不支持	不支持
自建-自建	GaussDB主备版->GaussDB主备版	支持	支持	不支持	不支持	不支持
自建-自建	GaussDB分布式版->Oracle	支持	支持	不支持	不支持	不支持
自建-自建	GaussDB分布式版->Kafka	不支持	不支持	不支持	不支持	不支持
自建-自建	GaussDB分布式版->GaussDB分布式版	支持	支持	不支持	不支持	不支持
自建-自建	DB2 for LUW->GaussDB主备版	支持	支持	不支持	不支持	不支持

同步方向	数据流向	对象级对比	行对比	内容对比	动态内容对比	用户对比
自建-自建	DB2 for LUW->GaussDB分布式版	支持	支持	不支持	不支持	不支持
自建-自建	Microsoft SQL Server->Kafka	不支持	不支持	不支持	不支持	不支持

约束限制

- 对比需要在任务处于增量阶段才能手动创建。
- 对比是大小写敏感的，如果源数据库或目标数据库一端为非大小写敏感，一端为大小写敏感，对比结果可能出现不一致的情况。
- 对于单全量，在任务结束的时候，DRS会自动创建对象对比和行对比。对比时如果源库有数据操作，可能会导致对比结果不一致。
- 如果在对比时源库进行DDL操作，为保证对比结果的准确性，需重新进行内容对比。
- 如果单独对目标库进行数据修改操作，可能会导致对比结果不一致。
- 如果源数据库字符类型存在编码异常的情况，通过DRS迁移或对比时会因数据库驱动转码转换为异常码点，最终可能导致内容对比一致，实际字节不一致。
- 内容对比功能目前只支持带有主键的表，对于不支持内容对比的表可以使用行数对比功能。所以数据级对比功能需要结合业务场景，选用行数对比或者内容对比。
- DRS进行内容比对期间不能暂停DRS任务，否则可能导致比对任务失败。
- 部分数据类型不支持内容对比，详情参考[内容对比不支持哪些数据类型](#)。
- 为避免占用资源，DRS限制对比任务的时长，超过限制时长，对比任务自动停止。
 - 对于单全量任务结束时自动创建的对比任务，超时限制时长30分钟，超过后行对比和对象对比任务自动停止，DRS单全量任务结束。
 - 对于增量阶段用户手动创建的行对比任务，当源库是关系型数据库时，行对比限制时长为60分钟；源库为非关系数据库，比如MongoDB，行对比限制时长为30分钟。
- 为避免占用资源，DRS任务的对比结果限制保留60天，60天后自动清除。
- 对于支持内容对比的同步链路，如果创建任务时可以选择实例规格，那么仅支持大规格及以上规格进行内容对比。
- 行对比的多对一场景，是将源库中的表与映射到目标库中的聚合表中对应的那部分数据做行数比对。
- 对于多对一同步任务场景，不建议使用内容对比，因为内容对比无法保证一致性。
- 不支持一个库中的表映射到多个库的任务进行内容对比。
- PostgreSQL为源的链路，在表映射场景下，索引和约束名称也会被修改，导致索引和约束的对比不一致。

- Oracle数据库中插入的空字符会被当成NULL来处理，因此对于Oracle为目标数据库的链路，DRS对比时会把空字符串和NULL做等价处理。源数据库是空字符串，目标数据库是NULL，对比结果显示一致。
- Oracle->GaussDB分布式内容比对，因为GaussDB分布式的BLOB用DBE_LOB.LOB_GET_LENGTH查询失败，当LOB比对策略选择比对LOB长度时，对比会忽略BLOB比对。

对数据库影响

- 对象对比：会查询源库及目标库的系统表，占用10个左右的session的连接数，正常情况下不会对数据库产生影响。但是如果对象数量巨大（比如几十万张表），可能会对数据库产生一定的查询压力。
- 行数对比：会查询源库及目标库的表行数，占用10个左右的session的连接数，正常的select count语句不会对数据库产生影响。但是如果表数据量巨大（亿级）会对数据库产生一定的查询压力，返回查询结果会比较慢。
- 内容对比：会查询源库及目标库的全部数据，涉及每个字段都会对比，所以对数据库产生一定的查询压力，主要体现在IO上，查询速度受限于源库和目标库的IO和网络带宽。占用1-2个CPU，占用10个左右的session的连接数。
- 用户对比：会查询源库及目标库的账户和权限，基本不会对数据库产生影响。

对比耗时预估

- 对象对比：根据源库查询性能，一般会在几分钟内返回结果，如果对象数据量特别巨大，对比时长可能达到几十分钟。
- 行数对比：使用select count方式，查询速度跟数据库性能相关。
- 内容对比：在数据库没有压力并且网络正常的情况下，对比速度大概是5M/s。
- 用户对比：与对象对比同时返回结果，如果对象不多的情况下，基本会在几分钟就会返回结果。

前提条件

- 已登录数据复制服务控制台。
- 已启动同步任务。

创建对比任务

步骤1 在“实时同步管理”界面，选中指定同步任务，单击任务名称，进入“基本信息”页签。

步骤2 单击“同步对比”页签，进入“同步对比”信息页面。

步骤3 对比同步项。


- 创建对象级对比：选择“对象级对比”页签，单击“开始对比”后稍等一段时间再单击，观察源数据库和目标数据库的各个对比项结果是否一致。若需要查看结果详情，可单击指定对比项操作列的“详情”按钮。

图 6-19 同步对象对比

对比项	源数据库	目标数据库	对比结果	操作
数据库	1	1	一致	详情
表	2	1	一致	详情
约束	1	1	一致	详情
索引	2	1	一致	详情
视图	4	1	一致	详情

- 创建数据级对比：选择“数据级对比”页签，单击“创建对比任务”，选择“对比类型”、“对比方式”、“对比时间”和“对象选择”，单击“是”提交对比任务。

图 6-20 创建数据对比任务

对比类型：分为行数对比和内容对比。

- 行数对比：用于对比源和目标端的表的行数是否相等。

说明

- 任务进入增量阶段后，用户可以创建行对比任务。
- Oracle->GaussDB同步任务，当全量任务结束的时候会自动触发行对比任务。

- 内容对比：用于对比源和目标端的表的数据是否一致。

说明

- 任务进入增量阶段后，用户可以创建内容对比任务。全量同步完成后，源库数据不能发生变更，否则内容对比结果会不一致。
- Oracle->GaussDB主备版、Oracle->GaussDB分布式进入增量校验状态后，对比服务会开始通过解析的该任务涉及的表的增量日志抽取数据，持续对比。如果用户想要对比其他表的同步状态，可以取消正在对比的任务，创建新的对比任务。

由于内容对比功能目前只支持带有主键的表，不支持内容对比的表可以使用行数对比功能。所以数据级对比功能需要结合业务场景，选用行数对比或者内容对比。

- 对比策略：分为普通对比和多对一对比两种。
 - 普通对比策略：源库中的一张表跟映射到目标库中的那张表做整表的行数比对。
 - 多对一对比策略：源库中的一张表跟映射到目标库中的聚合表中对应的那部分数据做行数比对。

说明

对比类型为行数对比时，支持对比策略选择。

- 对比方式：分为静态对比和动态对比两种。
 - 静态对比：对源数据库和目标数据库进行一次全量内容对比，内容对比完成后对比任务结束，适用于无数据变化的非业务时间。
 - 动态对比：先对源数据库和目标数据库进行一次全量内容对比，对比任务完成后进入增量对比阶段，实时比对源数据库和目标数据库的增量数据，适用于有数据变化的业务时间。

说明

- 对比类型为内容对比时，支持对比方式的选择。
- 选择库级同步时，动态对比过程中不支持在源库创建新表。如需创建，请先取消动态对比，待新表创建并同步完成后，再重新启动动态对比。

- 对比时间：可设置为“立即启动”和“稍后启动”。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现符合实际情况的少量数据不一致对比结果，推荐结合对比定时功能，选择在业务低峰期进行对比，得到更为具有参考性的对比结果。
- LOB对比策略：分为忽略LOB比对和比对长度对比两种。

说明

目前仅支持Oracle->GaussDB设置LOB对比策略。

- 忽略LOB比对：内容比对时忽略LOB字段。
- 比对长度：内容比对时对LOB字段的长度进行比对。
LOB比对长度会对数据库有负载，建议选择忽略LOB比对。
- 进行数据过滤：开启后，可根据配置的过滤条件进行对象比对。

说明

目前仅支持Oracle->GaussDB、GaussDB->Oracle、MySQL-GaussDB设置数据过滤对比。

- i. 选择开启数据过滤后，需要为选择对比的表对象添加过滤条件。
- ii. 在过滤条件区域，填写过滤条件后，单击“校验规则”。

说明

- 过滤表达式不支持使用某种数据库引擎特有的package、函数、变量、常量等写法，须使用通用SQL标准。
 - 请直接输入SQL语句中WHERE之后的部分（不包含WHERE和分号，例如：sid > 3 and sname like "G %"）。
 - 过滤条件不支持使用隐式转换规则，请填写正确的数据类型条件。例如Oracle的c列是字符类型varchar2，过滤条件应该设置为 c > '10'，而不是c > 10。
 - 不支持对LOB字段设置过滤条件，如CLOB、BLOB、BYTEA等大字段类型。
 - 建议不要对非精确类型字段设置过滤条件，如FLOAT、DECIMAL、DOUBLE等。
 - 建议不要对带有特殊字符的字段设置过滤条件。
 - 暂不支持对库名、shema名、表名大小写混用的对象进行过滤比对。
- iii. 校验通过后，单击“生成加工规则”，即可在加工规则表格中看到该规则。
- iv. 检查无误后，单击“确定”，创建对比任务。
- 对象选择：可根据具体的业务场景选择需要进行对比的对象。
 - 用户对比：单击“用户对比”页签，可查看数据库账号及权限的对比结果。

图 6-21 用户对比

源数据库对象名	目标数据库对象名	对比时间	对比结果
guser1	guser1	2021/07/29 16:02:41 GMT+08:00	一致
guser1	guser1	2021/07/29 16:02:41 GMT+08:00	一致
guser2	guser2	2021/07/29 16:02:41 GMT+08:00	一致
guser2	guser2	2021/07/29 16:02:41 GMT+08:00	一致

说明

- 全量同步中的任务无法进行用户对比。
- 实时同步目前仅PostgreSQL->PostgreSQL的同步支持用户对比。

步骤4 对比任务提交成功后，返回“数据级对比”页签，单击 刷新列表，可以查看到所选对比类型的对比结果。

图 6-22 同步数据对比

对比类型	对比开始时间	对比结束时间	对比状态	导出对比报告	操作
内容对比	2021/08/27 14:07:09 GMT+08:00	2021/08/27 14:08:03 GMT+08:00	完成	无	查看对比报表 导出报告
行数对比	2021/08/26 22:31:24 GMT+08:00	2021/08/26 22:32:01 GMT+08:00	完成	无	查看对比报表 导出报告

由于内容对比功能目前只支持带有主键的表，不支持内容对比的表可以使用行数对比功能。所以数据级对比功能需要结合业务场景，选用行数对比或者内容对比。

若需要查看行数对比或者内容对比详情，可单击指定对比类型操作列的“查看对比报表”，页面将跳转至新的窗口，可观察对比结果的详细情况。

若需要下载行数对比或者内容对比结果，可单击指定对比类型操作列的“导出报告”。

图 6-23 行数对比详情

源库表名	目标库表名	源库表行数	目标库表行数	对比结果	操作
t_test	t_test	2	2	一致	

图 6-24 内容对比详情

源库表名	目标库表名	内容对比结果	操作
t_test	t_test	一致	

说明

- 运行中的任务可以随时取消，已取消的对比任务也支持查看对比报表。
- 行对比结果可按照“源库表行数”或“目标库表行数”，对当前页面显示的结果进行升序或者降序排列，方便进行筛选。
- 行对比差异显示为负数时，代表目标库表行数比源库表行数多；显示为正数时，代表源库表行数比目标库表行数多。

----结束

周期性对比

周期性对比是DRS定时对源数据库与目标数据库表的行数进行对比，并展示结果。

- 步骤1** 在“实时同步管理”界面，选中指定同步任务，单击任务名称，进入“基本信息”页签。
- 步骤2** 单击“同步对比”页签，进入“同步对比”信息页面。
- 步骤3** 选择“周期性对比”页签，单击“修改对比策略”

说明

目前仅支持MySQL->MySQL、MySQL->GaussDB(for MySQL)、MySQL->GaussDB(DWS)、GaussDB(for MySQL)->MySQL同步支持周期性对比设置。

- 步骤4** 在“修改对比策略”对话框中，开启周期性对比并设置对比频率、时间等策略，单击“是”完成设置。

说明

- 开启后，DRS会按照设置的定时策略进行行数对比，对比结果需要在数据级对比中查看。
- 关闭周期性对比后，下一次的周期性对比将不会再进行，历史对比结果仍可正常查看。
- 修改检查设置不会影响当前在进行的周期性对比任务，修改后的设置下次生效。
- 周期性对比会在源和目标端执行一定的读取操作，请选择无业务期时间点进行对比。
- 周期性对比自动剔除超大表（行数超过一亿行），该类大表建议采用数据级对比功能进行抽查，不建议作为周期性对比表。

图 6-25 修改对比策略

✕

修改对比策略

开启状态 请在数据级对比中查看对比结果

对比频率 每周对比 对比频率较高可能会影响业务性能，请根据实际情况选择。

对比时间

周一

周二

周三

周四

周五

周六

周日

时区 GMT+08:00

生效时间 00 : 00 - 02 : 00

建议在业务低峰期进行周期性对比，对业务性能影响较小；相对静止的数据对比结果更具有参考性；生效时间内未完成的对比会自动中断，已完成的对比结果仍可查看。

对比类型 行数对比

对比策略 普通对比 多对一对比 ?

*修改检查设置不会影响当前在进行的周期性对比任务，修改后的设置下次生效。
*周期性对比会在源和目标端执行一定的读取操作，请选择无业务期时间点进行对比。
*周期性对比自动剔除超大表（行数超过一亿行），该类大表建议采用数据级对比功能进行抽查，不建议作为周期性对比表。

是
否

----结束

6.5 对象管理

6.5.1 编辑同步对象

数据复制服务提供的编辑同步对象功能，可以对已启动并在增量中的同步任务对象进行修改。

- 对于正常增量中的任务，您可以通过编辑同步对象功能来增加或者移除同步的数据库和表。
- 对于增量失败中的任务，也可通过编辑同步对象功能修改同步对象。在错选同步对象导致增量失败的场景下，可通过该功能移除数据库或者表，重新提交启动后恢复任务。
- 对于增量失败中的任务通过再编辑修改同步对象的时候，在任务未启动前，请不要进行续传或重置操作，否则会导致同步对象缓存被清理，再编辑修改同步对象失败。

目前支持增量阶段再编辑的同步链路可参考[支持编辑的链路](#)。

前提条件

已登录数据复制服务控制台。

方法一

步骤1 在“实时同步管理”页面，选择需要修改同步对象的任务，单击操作列的“编辑”按钮。

图 6-26 编辑实时同步对象



步骤2 进入“设置同步”页面，修改需要同步的对象，单击“下一步”。

- 选择对象时可对展开的库进行搜索，搜索支持正则表达式。
- 当对象名称包含空格时，名称前后的空格不显示，中间如有多个空格只显示一个空格。
- 选择的同步对象名称中不能包含空格。

步骤3 在“数据加工”页面，可参考[数据加工](#)章节，为新增加表设置相关规则。

说明

任务再编辑时，已经同步的表不支持修改加工规则。

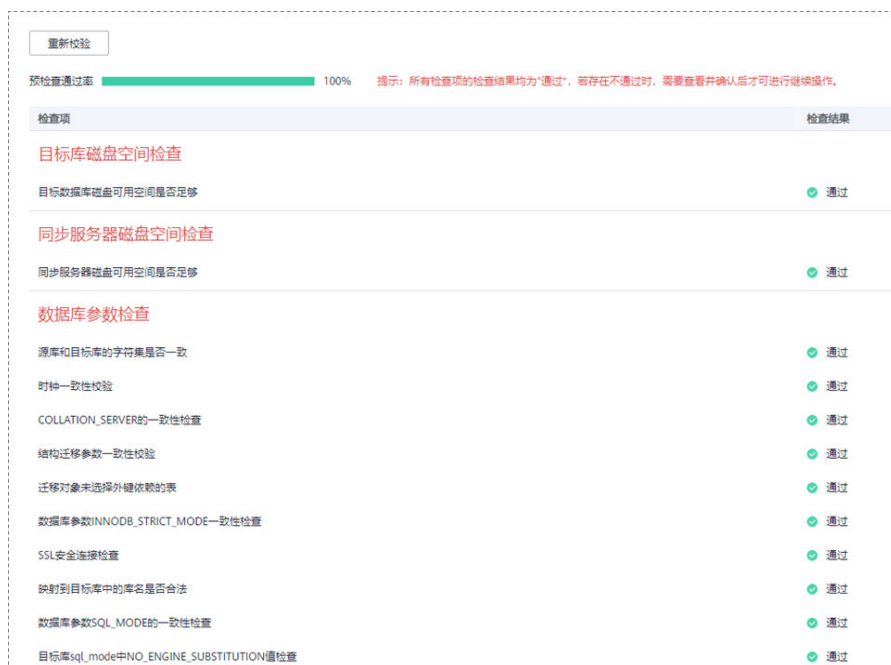
步骤4 在“预检查”页面，进行同步任务预校验，校验是否可进行。

- 查看检查结果，如有不通过的检查项，需要修复不通过项后，单击“重新校验”按钮重新进行任务预校验。

预检查不通过项处理建议请参见《数据复制服务用户指南》中的“[预检查不通过项修复方法](#)”。

- 预检查完成后，且所有检查项结果均通过时，单击“下一步”。

图 6-27 预检查



说明

所有检查项结果均通过时，若存在请确认项，需要阅读并确认详情后才可以继续执行下一步操作。

- 步骤5** 在“任务确认”页面，确认同步任务信息无误后，单击“启动任务”，提交同步任务。
- 步骤6** 返回“实时同步管理”，在同步任务列表中，当前任务状态显示为“增量同步”，并生成任务状态为“任务变更中”的子任务，待子任务变更完成后，开始对编辑后的同步对象进行增量同步。

----结束

方法二

- 步骤1** 在“实时同步管理”页面，选择需要修改同步对象的任务，单击任务名称。
- 步骤2** 页面跳转至“基本信息”，切换到“同步映射”页签，单击指定同步对象后面的“编辑”按钮。
- 步骤3** 执行方法一的步骤[步骤2](#)到[步骤6](#)。

----结束

支持编辑的链路

入云：

- MySQL->MySQL
- MySQL->GaussDB(for MySQL)
- MySQL->GaussDB主备版
- MySQL->GaussDB分布式版
- MySQL->MariaDB
- DDM->DDM
- DDM->GaussDB(DWS)
- Oracle->MySQL
- Oracle->PostgreSQL
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式版
- Oracle->GaussDB(DWS)
- Oracle->GaussDB(for MySQL)
- PostgreSQL->PostgreSQL
- PostgreSQL->GaussDB分布式
- PostgreSQL->GaussDB主备版
- MariaDB->MariaDB
- MariaDB->MySQL
- MariaDB->GaussDB(for MySQL)
- DB2 for LUW->GaussDB分布式
- DB2 for LUW->GaussDB主备版
- DB2 for LUW->GaussDB(DWS)
- Microsoft SQL Server->GaussDB(DWS)
- Microsoft SQL Server->GaussDB分布式
- Microsoft SQL Server->GaussDB主备版
- GaussDB(for MySQL)->GaussDB(for MySQL)

出云:

- MySQL->MySQL
- MySQL->CSS/ES
- MySQL->Kafka
- MySQL->Oracle
- MySQL->MariaDB
- DDM->Kafka
- DDS->Kafka
- PostgreSQL->PostgreSQL
- PostgreSQL->Kafka
- MariaDB->MariaDB
- GaussDB(for MySQL)->MySQL

- GaussDB(for MySQL)->CSS/ES
- GaussDB(for MySQL)->GaussDB(DWS)
- GaussDB(for MySQL)->Oracle
- GaussDB(for MySQL)->Kafka
- GaussDB主备版->MySQL
- GaussDB主备版->Oracle
- GaussDB主备版->Kafka
- GaussDB主备版->GaussDB(DWS)
- GaussDB主备版->GaussDB分布式版
- GaussDB主备版->GaussDB主备版
- GaussDB分布式版->MySQL
- GaussDB分布式版->Oracle
- GaussDB分布式版->Kafka
- GaussDB分布式版->GaussDB(DWS)
- GaussDB分布式版->GaussDB分布式版
- GaussDB分布式版->GaussDB主备版
- Microsoft SQL Server->Kafka

自建-自建

- MySQL->CSS/ES
- MySQL->Kafka
- MySQL->GaussDB主备版
- Oracle->Kafka
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式版
- PostgreSQL->Kafka
- DB2 for LUW->GaussDB分布式
- DB2 for LUW->GaussDB主备版
- GaussDB主备版->Oracle
- GaussDB主备版->Kafka
- GaussDB主备版->GaussDB主备版
- GaussDB分布式版->Oracle
- GaussDB分布式版->Kafka
- GaussDB分布式版->GaussDB分布式版
- Microsoft SQL Server->Kafka

6.5.2 导入同步对象

实时同步支持通过文件导入同步对象，创建实时同步任务后，在“设置同步”页面可选择“导入对象文件”。

使用须知

- 文件导入仅支持导入Windows Microsoft Excel 97-2003版本 (*.xls)，2007及以上版本 (*.xlsx) 的文件，下载的压缩包提供上述两个版本模板。
- 文件名支持的有效字符范围：汉字，空格，大小写字母，数字，“-”，“_”，“（”，“）”。
- 模板中的对象信息需按照格式进行填写，不允许包含“<”，“>”，"."和“"”，不支持以空格开头或结尾的对象，且大小写敏感。
- 配置中的任务，可使用“表级同步”，“库级同步”或“文件导入”三种方式，每次切换新的方式后，当前选择或者导入的数据库对象被清空，需重新选择或导入。
- 可映射的文件导入，必须严格按照模板分别填写第一列和第二列，均不可为空，如果某行的两列均为空时，会忽略该行。
- 选择文件导入功能所创建的任务，任务启动之后再编辑时，不支持切换到“表级同步”和“库级同步”功能。
- 任务再编辑时，导入的文件必须包含全量对象信息，不支持单独导入增删的对象信息。
- 任务再编辑时，已经同步的对象不支持再次映射，请保持其映射后的对象名称不变。
- 任务再编辑时导出的对象信息，是已经同步的对象信息。
- 上传文件后校验失败，可单击“查看失败详情”下载错误信息。
- Excel中填写的对象名称的大小写须与源库中保持一致。

操作步骤

步骤1 在实时同步任务的“设置同步”页面，点击“同步对象”处的“导入对象文件”。

图 6-28 导入对象文件



步骤2 单击“下载模板”。

步骤3 在下载Excel模板中，填写需要导入的对象信息。

步骤4 单击“添加文件”，在对话框中选择编辑完成的模板。

步骤5 单击“上传文件”。

----结束

6.5.3 对象名映射

实时同步功能可以支持源数据库和目标数据库中的同步对象（包括数据库、schema、表）的名称不同。如果同步的数据库对象在源数据库和目标数据库中的命名不同，那

么可以使用数据复制服务提供的对象名映射功能进行映射。其中包括：库映射、schema映射、表映射。

对象名映射功能支持在如下场景使用：

- 首次创建实时同步任务时，在选择同步对象阶段可设置映射，后续不可编辑映射。
- 通过编辑同步对象功能来增加或者减少同步对象时，首次编辑的对象可以设置映射关系，已经在同步关系中的对象不可设置映射。

本小节主要介绍如何在实时同步任务配置过程中，使用对象名映射功能，该功能的映射关系详情可参见[查看同步映射](#)章节。

使用须知

- 不支持库名、表名带有换行符的对象进行映射。
- 若目标库是PG类（如GaussDB(DWS)、GaussDB和PostgreSQL），目的库的schema名称不能设置为以pg_开头，否则会迁移失败。
- 以下链路，不支持包含smallserial, serial, bigserial 这3个类型的表进行schema映射。
 - GaussDB主备版->GaussDB分布式版
 - GaussDB主备版->GaussDB主备版
 - GaussDB分布式版->GaussDB分布式版
 - GaussDB分布式版->GaussDB主备版
- 以下链路，支持多对一的对象映射。
 - MySQL->MySQL
 - MySQL->GaussDB(for MySQL)
 - MySQL->GaussDB(DWS)
 - MySQL->Kafka
 - GaussDB(for MySQL)->MySQL
 - GaussDB(for MySQL)->GaussDB(for MySQL)
 - GaussDB(for MySQL)->GaussDB(DWS)
 - Oracle->GaussDB
 - Oracle->GaussDB(DWS)

库映射

实时同步过程中，如果待同步的库在源数据库和目标数据库中存储名称不同，可以使用实时同步提供的库映射功能进行库名映射。例如：将源数据库中的A库同步到目标数据库，变成目标数据库中的B库，此时就可以使用库映射功能来实现。

步骤1 在“设定同步”页面，同步对象右侧已选对象框中，选择需要进行映射的数据库，单击“编辑”按钮。

图 6-29 库映射



步骤2 修改库名。

在“编辑库名”的弹出框中，填写新的数据库名，修改后的名称即为保存在目标数据库中的库名。

图 6-30 修改库名



步骤3 查看修改结果。

库名修改成功后，您会看到修改前的库名和修改后的新数据库名，至此，表示完成库映射的配置。

图 6-31 查看库名修改结果



----结束

schema 映射

schema指数据库对象的集合，这个集合包含了各种对象如：表、视图、存储过程、索引等。

实时同步过程中，如果待同步的schema在源数据库和目标数据库中的名称不同，可以使用实时同步提供的schema映射功能进行schema名映射。例如：将源数据库中的schemaA同步到目标数据库，变成目标数据库中的schemaB，此时就可以使用schema映射功能来实现。

步骤1 在“设定同步”页面，同步对象右侧已选对象框中，选择需要进行映射的schema，单击“编辑”按钮。

图 6-32 schema 映射



步骤2 修改schema名。

在“编辑schema名”的弹出框中，填写新的schema名，修改后的名称即为保存在目标数据库中的schema名。

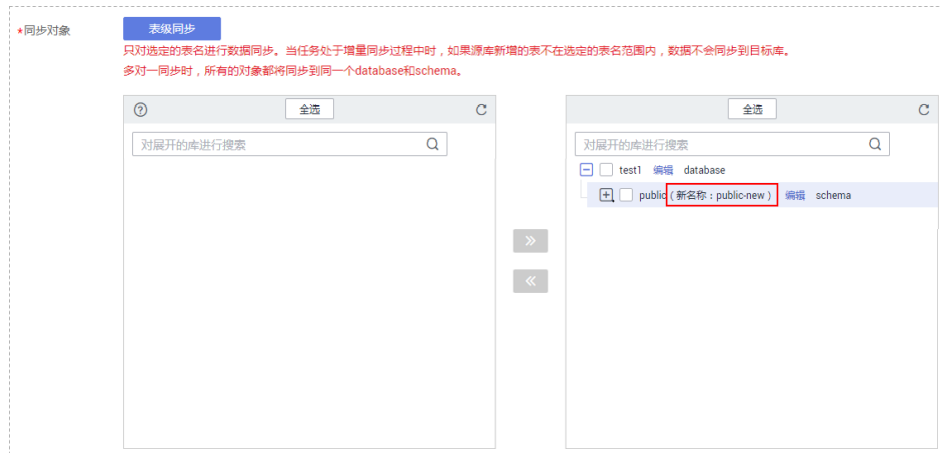
图 6-33 修改 schema 名



步骤3 查看修改结果。

schema名修改成功后，您会看到修改前的schema名和修改后的新schema名，至此，表示完成schema映射的配置。

图 6-34 查看 schema 修改结果



----结束

表映射

实时同步过程中，如果待同步的表在源数据库和目标数据库中的名称不同，可以使用实时同步提供的表映射功能进行表名映射。例如：将源数据库中的A表同步到目标数据库，变成目标数据库中的B表，此时就可以使用表映射功能来实现。

步骤1 在“设定同步”页面，同步对象右侧已选对象框中，选择需要进行映射的表，单击“编辑”按钮。

图 6-35 表映射



步骤2 修改表名。

在“编辑表名”的弹出框中，填写新的表名，修改后的名称即为保存在目标数据库中的表名。

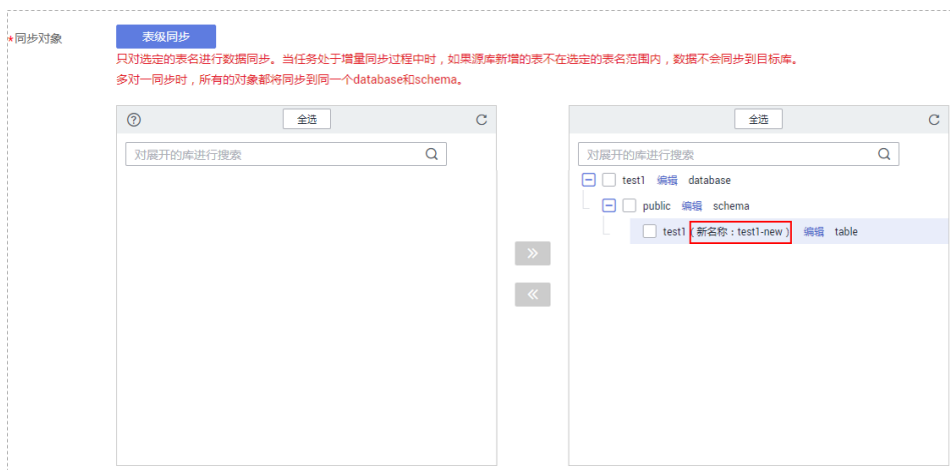
图 6-36 修改表名



步骤3 查看修改结果。

表名修改成功后，您会看到修改前的表名和修改后的新表名，至此，表示完成表映射的配置。

图 6-37 查看表名修改结果



----结束

6.5.4 查看同步映射

实时同步映射关系包括对象选择时的库映射、schema映射、表映射和数据加工时的列映射。当同步映射关系配置成功后，启动实时同步任务即可查看同步映射详情。

前提条件

已登录数据复制服务控制台。

操作步骤

- 步骤1** 在“实时同步管理”界面，选择指定的同步任务，单击任务名称。
- 步骤2** 页面跳转至“基本信息”，切换到“同步映射”页签，查看当前同步任务的映射关系。

图 6-38 查看同步映射信息

源数据库	表名	目标数据库	表名	操作
gaussdb0501	test1	gaussdb0501	test1	编辑
gaussdb0504	test1	gaussdb050803	test1	编辑
mysql_02	test1	mysql_020803	test1	编辑 数据加工

说明

选择对象的时候，对象名称的前后空格不显示，中间如有多个空格只显示一个空格。

- 步骤3** 当前同步任务的映射关系包括同步对象映射和列映射，可通过右上角进行筛选和搜索。

----结束

6.5.5 数据加工

数据复制服务支持对同步的对象进行加工，即可以为选择的对象添加规则。各链路支持的加工规则不一样，目前仅部分链路支持数据加工，支持数据加工链路参考表 6-10，具体操作方法可参考对应的内容。

表 6-10 支持数据加工的链路

同步方向	数据流向	数据过滤	附加列	列加工
入云	MySQL->MySQL	支持	支持	支持
入云	MySQL->GaussDB(DWS)	支持	支持	暂不支持
入云	MySQL->GaussDB(for MySQL)	支持	支持	暂不支持
入云	MySQL->MariaDB	支持	支持	支持
入云	DDM->GaussDB(DWS)	暂不支持	支持	暂不支持
入云	Oracle->GaussDB(DWS)	支持	支持	暂不支持
入云	Oracle->MySQL	支持	暂不支持	暂不支持
入云	Oracle->GaussDB(for MySQL)	支持	暂不支持	暂不支持
入云	Oracle->GaussDB主备版	支持	暂不支持	支持
入云	Oracle->GaussDB分布式版	支持	暂不支持	支持
入云	DB2 for LUW->GaussDB主备版	支持	暂不支持	暂不支持
入云	DB2 for LUW->GaussDB分布式版	支持	暂不支持	暂不支持
入云	MariaDB->MariaDB	支持	暂不支持	暂不支持
入云	MariaDB->MySQL	支持	支持	支持
入云	MariaDB->GaussDB(for MySQL)	支持	支持	支持
出云	MySQL->MySQL	支持	支持	支持
出云	MySQL->kafka	暂不支持	暂不支持	支持
出云	MySQL->CSS/ES	支持	暂不支持	支持
出云	MySQL->MariaDB	支持	支持	支持
出云	GaussDB主备版->MySQL	支持	暂不支持	暂不支持
出云	GaussDB主备版->Oracle	支持	暂不支持	支持
出云	GaussDB主备版->Kafka	暂不支持	暂不支持	支持
出云	GaussDB主备版->GaussDB(DWS)	支持	暂不支持	暂不支持
出云	GaussDB主备版->GaussDB分布式版	支持	暂不支持	支持

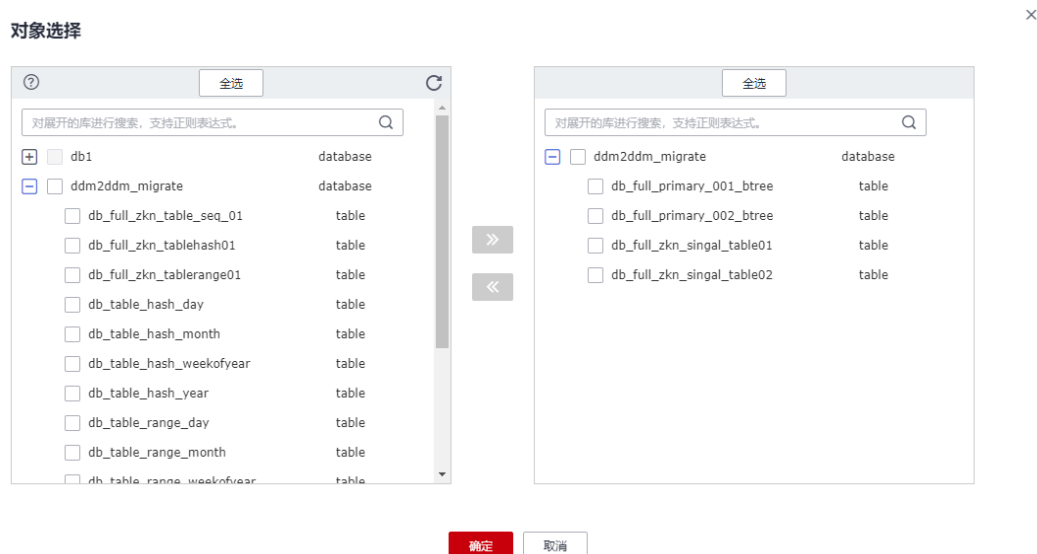
同步方向	数据流向	数据过滤	附加列	列加工
出云	GaussDB主备版->GaussDB主备版	支持	暂不支持	支持
出云	GaussDB分布式版->MySQL	支持	暂不支持	暂不支持
出云	GaussDB分布式版->Oracle	支持	暂不支持	支持
出云	GaussDB分布式版->GaussDB(DWS)	支持	暂不支持	暂不支持
出云	GaussDB分布式版->Kafka	暂不支持	暂不支持	支持
出云	GaussDB分布式版->GaussDB分布式版	支持	暂不支持	支持
出云	GaussDB分布式版->GaussDB主备版	支持	暂不支持	支持
出云	GaussDB(for MySQL)->MySQL	支持	支持	暂不支持
出云	GaussDB(for MySQL)->GaussDB(DWS)	暂不支持	支持	暂不支持
出云	GaussDB(for MySQL)->CSS/ES	支持	暂不支持	支持
出云	MariaDB->MariaDB	支持	暂不支持	暂不支持
自建-自建	MySQL->Kafka	暂不支持	暂不支持	支持
自建-自建	MySQL->CSS/ES	支持	暂不支持	支持
自建-自建	Oracle->GaussDB主备版	支持	暂不支持	支持
自建-自建	Oracle->GaussDB分布式版	支持	暂不支持	支持
自建-自建	GaussDB主备版->Kafka	暂不支持	暂不支持	支持
自建-自建	GaussDB分布式版->Kafka	暂不支持	暂不支持	支持
自建-自建	DB2 for LUW->GaussDB主备版	支持	暂不支持	暂不支持
自建-自建	DB2 for LUW->GaussDB分布式版	支持	暂不支持	暂不支持

添加同步时间戳

步骤1 在实时同步任务的“数据加工”的“附加列”页面，单击“选择对象”。

步骤2 在弹出的对话框中选择需要加工的表对象，单击“确定”。

图 6-39 对象选择



步骤3 在“数据加工区”区域，填写需要添加的列名、类型、操作类型等信息。

图 6-40 添加信息



说明

- 需要添加规则的列为表中已有列，且不能为主键。
- 仅支持使用数据类型为时间戳类型（TIMESTAMP）的列作为规则列。

步骤4 检查无误后，单击“下一步”，进入下一步操作。

----结束

添加附加列

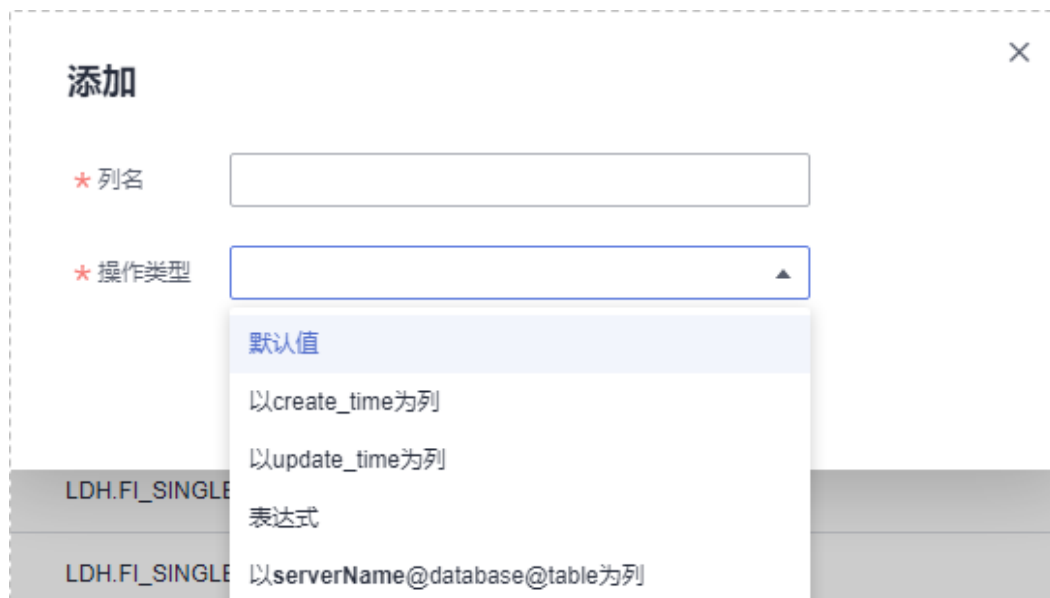
步骤1 在实时同步任务的“数据加工”页面选择“附加列”，对于需要加工的表对象，单击“操作”列的添加。

图 6-41 附加列



步骤2 在弹出的“添加”对话框中，填选列名、操作类型、字段类型等信息，选择的操作类型不同，需要填写的信息也不同。完成后单击“确定”。

图 6-42 操作类型



说明

- 多对一映射场景下，需要使用数据加工的附加列操作来避免数据冲突。
- 支持的操作类型有以下几种
 - “默认值”：使用输入的默认值填充新加的列。
 - “以create_time为列”和“以update_time为列”操作会分别使用数据产生时间和修改时间填充新加的列。
 - “表达式”：选择使用concat(_current_database, '@', _current_table) 表达式填充新加的列，暂不支持手动输入表达式。
 - “以serverName@database@table为列”：使用@符号分别拼接serverName、源库的库名、表名填充新加的列，其中serverName为手动输入。
 - “填充值”：选择填充值为新加的列，比如同步时间戳。
- 支持“批量设置”，将第一个可编辑表的附加列信息应用到此任务中所有可编辑的表。
- MySQL->GaussDB(for MySQL)同步，单表的列数超过500时，对该表添加附加列可能会超过列数上限，会导致任务失败。
- 如果使用serverName@database@table作为附加列，后续在进行数据对比时，目标端将默认使用该附加列作为行比对、内容比对的隐含过滤条件。

步骤3 检查无误后，单击“下一步”，进入下一步操作。

----结束

数据过滤

数据过滤规则添加后，在源库执行更新操作的处理原则，以数据一致为首要目标。即：

- 更新前不符合过滤条件，更新后符合。继续同步，在目标库执行相同的更新操作，如果匹配不到数据，则会忽略，会导致数据不一致。

- 更新前符合过滤条件，更新后不符合。继续同步，在目标库执行相同的更新操作。

步骤1 在“数据加工”页面选择“数据过滤”。

图 6-43 数据过滤



步骤2 在“对象选择”区域选择需要加工的表对象。

步骤3 在过滤条件区域，填写过滤条件（只需填写SQL语句的WHERE之后部分，例如id=1）后，单击“校验规则”。

说明

- 每张表仅支持添加一个校验规则。
- 数据过滤每次最多支持500张表。
- 过滤表达式不支持使用某种数据库引擎特有的package、函数、变量、常量等写法，须使用通用SQL标准。请直接输入SQL语句中WHERE之后的部分（不包含WHERE和分号，例如：`sid > 3 and sname like "G %"`），最多支持输入512个字符。
- 过滤条件填写的SQL语句中，关键字需要用反引号，datetime类型（包含日期和时间）和字符串类型需要用单引号，例如：``update` > '2022-07-13 00:00:00' and age >10、`update`='abc'`。
- 过滤条件不支持使用隐式转换规则，请填写正确的数据类型条件。例如Oracle的c列是字符类型varchar2，过滤条件应该设置为 `c > '10'`，而不是 `c > 10`。
- 不支持对LOB字段设置过滤条件，如CLOB、BLOB、BYTEA等大字段类型。
- 不支持库名、表名带有换行符的对象设置过滤规则。
- 不支持for update等行级锁的语法做过滤条件。
- 不支持对列名做函数操作，如果使用函数操作可能导致数据不一致。
- 建议不要对非精确类型字段设置过滤条件，如FLOAT、DECIMAL、DOUBLE等。
- 建议不要对带有特殊字符的字段设置过滤条件。
- 不建议使用非幂等表达式或函数作为数据加工条件，如SYSTIMESTAMP，SYSDATE等，因其每次调用返回的结果可能会有差异，导致达不到预期。
- 任务再编辑时，已经同步的表不支持修改过滤规则。

步骤4 校验通过后，单击“生成加工规则”，即可在加工规则表格中看到该规则。

步骤5 检查无误后，单击“下一步”，进入下一步操作。

----结束

数据过滤高级设置

如果需要使用关联表查询，可以使用数据加工的高级设置功能。

步骤1 在实时同步任务的“数据加工”页面选择“数据过滤”。

步骤2 在“对象选择”区域选择需要加工的表对象。

步骤3 在过滤条件区域，填写联表查询过滤条件，例如：id1 in (select id from db1.tab1 where id >=3 and id <10)，单击“校验”。

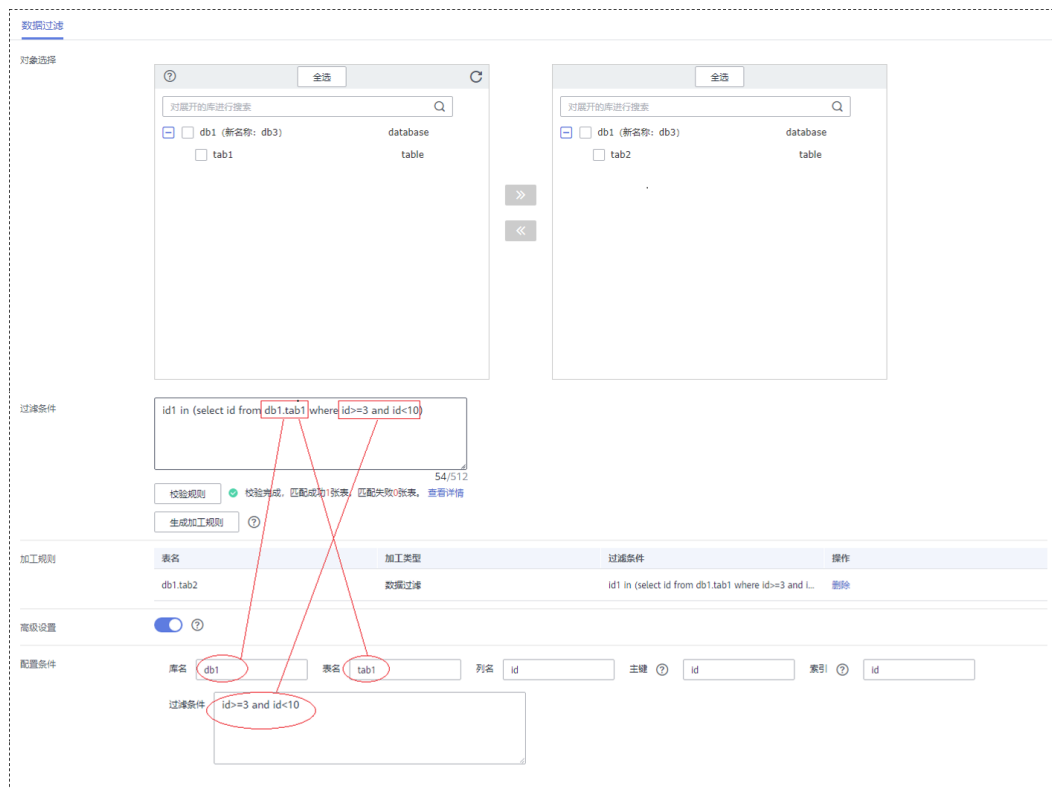
说明

- 每张表仅支持添加一个校验规则。
- 数据过滤每次最多支持500张表。
- 过滤表达式不支持使用某种数据库引擎特有的package、函数、变量、常量等写法，须使用通用SQL标准。请直接输入SQL语句中WHERE之后的部分（不包含WHERE和分号，例如：sid > 3 and sname like "G %"），最多支持输入512个字符。
- 过滤条件不支持使用隐式转换规则，请填写正确的数据类型条件。例如Oracle的c列是字符类型varchar2，过滤条件应该设置为 c > '10'，而不是c > 10。
- 不支持对LOB字段设置过滤条件，如CLOB、BLOB、BYTEA等大字段类型。
- 不支持库名、表名带有换行符的对象设置过滤规则。
- 不支持for update等行级锁的语法做过滤条件。
- 建议不要对非精确类型字段设置过滤条件，如FLOAT、DECIMAL、DOUBLE等。
- 建议不要对带有特殊字符的字段设置过滤条件。
- 不建议使用非幂等表达式或函数作为数据加工条件，如SYSTIMESTAMP，SYSDATE等，因其每次调用返回的结果可能会有差异，导致达不到预期。

步骤4 校验通过后，单击“生成加工规则”，即可在加工规则表格中看到该规则。

步骤5 在“高级设置”中设置关联表的“配置条件”和“配置规则”，辅助您进行数据过滤。

图 6-44 高级设置



1. 在“配置条件”中，根据在步骤3输入的过滤条件，填写关联表的相关信息。“库名”、“表名”、“列名”、“主键”、“索引”和“过滤条件”均为必填项。如果该表没有索引，则填写主键。“过滤条件”即在步骤3输入的该关联表的过滤条件。
2. 填写完成，单击“校验规则”。
3. 校验成功后，单击“生成配置规则”。即可在配置规则表格中看到该规则。如果需要进行多张关联表的数据过滤，重复执行步骤5即可。

说明

配置规则支持删除操作。

步骤6 检查无误后，单击“下一步”，进入下一步操作。

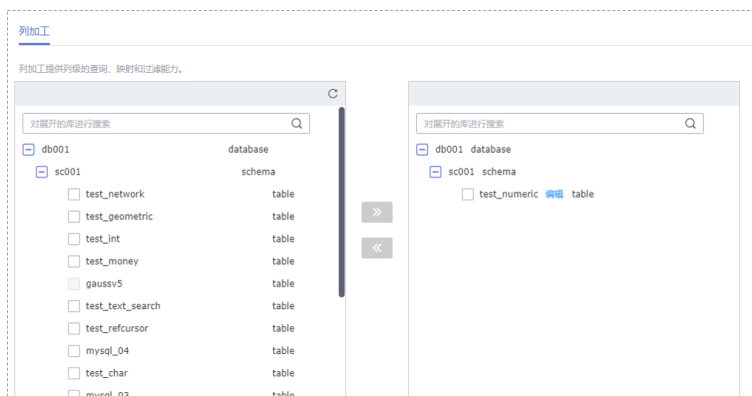
---结束

列加工

步骤1 在实时同步任务的“数据加工”页面选择“列加工”。

步骤2 在“对象选择”区域选择需要加工的对象。

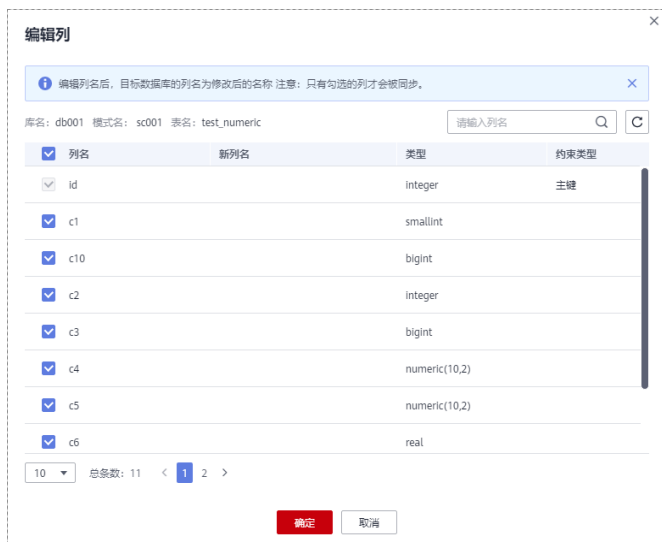
图 6-45 列加工



步骤3 单击右侧已选对象的“编辑”。

步骤4 在“编辑列”中，勾选需要映射的列，填写映射后的“新列名”。

图 6-46 编辑列



说明

- 列加工提供列级的查询、映射和过滤能力。
- 编辑列名后，目标数据库的列名为修改后的名称。
- 列映射名不能和原列名或已存在的映射名相同。
- 不支持库名、表名带有换行符的列进行映射。
- 任务再编辑时，已经同步的表不支持修改列信息。
- 只有勾选的列才会被同步，后续新增的列不在列加工范围内。
- MySQL->MySQL、MySQL->GaussDB(for MySQL)不支持分区表的分区字段进行列映射。
- GaussDB分区键不可以被过滤。
- 进行列过滤、列映射的表，增量阶段不支持对表中已过滤或映射的列进行DDL操作。


步骤5 填写完成后单击“确定”。

步骤6 检查无误后，单击“下一步”，进入下一步操作。

----结束

查看数据过滤

步骤1 在“实时同步管理”页面任务列表中，选择需要查看数据加工的任务，单击任务名称。

步骤2 选择“数据加工”页签，查看数据过滤记录。单击右上角的, 可刷新数据。

----结束

查看列加工

步骤1 在“实时同步管理”页面任务列表中，选择需要查看的任务，单击任务名称。


步骤2 选择“同步映射”页签，单击右上角选择列映射，查看列映射记录。单击右上角的, 可刷新数据。

图 6-47 查看列映射



源数据库	表名	目标数据库	表名	操作
id	MYSQL_02	h20010803	mysql_020803	编辑 删除列加工

----结束

6.6 参数管理

6.6.1 修改任务参数

为确保DRS任务发挥出最优性能，用户可根据业务需求对创建的DRS任务参数进行调整。

约束限制

- 支持任务状态为配置、全量同步、全量同步失败、增量同步、增量同步失败以及暂停时进行修改参数。
- 部分参数修改后，需要**重启同步任务**后才能生效。除此以外，重置、续传、任务重建、主备倒换、全量阶段修改限速操作也会重启任务进程使设置生效。
- 目前仅支持以下引擎的同步任务进行修改参数。常见的任务参数说明可参考[任务参数说明](#)。
 - MySQL->MySQL
 - MySQL->PostgreSQL
 - MySQL->GaussDB分布式版
 - MySQL->GaussDB主备版
 - MySQL->GaussDB(DWS)

- MySQL->GaussDB(for MySQL)
- MySQL->Kafka
- MySQL->CSS/ES
- MySQL->Oracle

操作步骤

步骤1 在“实时同步管理”界面，选择指定的同步任务。

步骤2 单击任务名称，页面跳转至“基本信息”页面。

步骤3 单击左侧“参数管理”，在“参数”页签可对当前同步任务的参数信息进行修改。

图 6-48 查看参数信息

参数名称	是否可配置	值	允许值	描述
increment_worker_num	是	16	1-32	增量阶段每日任务线程数
read_task_num	是	2	1-4	全量阶段从源数据库读取数据的线程数
datamove_source_socket_timeout	是	120000	120000-2147483647	全量阶段从源数据库读取数据超时时间
share_task_num	是	2	1-4	全量阶段从源数据库分片线程数
new_socket_timeout	是	30000	30000-2147483647	增量阶段从源数据库连接超时时间
supplier_thread_num	是	4	1-8	全量阶段写数据到目标数据库的线程数

步骤4 修改完成后，可选择“保存”或者“取消”设置。

- 单击“保存”，在弹出框中单击“是”，保存修改。
- 单击“取消”，在弹出框中单击“是”，放弃本次设置。

步骤5 部分参数完成修改后，需要重启任务后才生效，可单击右上角“重启”按钮使之生效，也可参考[重启同步任务](#)内容进行重启。

除此以外，重置、续传、任务重建、主备倒换、全量阶段修改限速操作也会重启任务进程使设置生效。

----结束

任务参数说明

表 6-11 任务参数说明

参数名	描述	取值范围	默认值	参数类型	是否需要重启
applier_thread_num	全量阶段写数据到目标数据库的线程数	与规格大小有关。 <ul style="list-style-type: none"> 极小规格：1-4 小规格：1-4 中规格：1-8 大规格：1-18 超大规格：1-24 	与规格大小有关。 <ul style="list-style-type: none"> 极小规格：2 小规格：2 中规格：4 大规格：6 超大规格：12 	int	是
read_task_num	全量阶段从源数据库读取数据的线程数	与规格大小有关。 <ul style="list-style-type: none"> 极小规格：1-2 小规格：1-2 中规格：1-4 大规格：1-8 超大规格：1-12 	与规格大小有关。 <ul style="list-style-type: none"> 极小规格：1 小规格：1 中规格：2 大规格：4 超大规格：6 	int	是
shard_task_num	全量阶段从源数据库分片线程数	与规格大小有关。 <ul style="list-style-type: none"> 极小规格：1-2 小规格：1-2 中规格：1-4 大规格：1-8 超大规格：1-8 	与规格大小有关。 <ul style="list-style-type: none"> 极小规格：1 小规格：1 中规格：2 大规格：4 超大规格：4 	int	是

参数名	描述	取值范围	默认值	参数类型	是否需要重启
increment_writer_num	增量阶段写目标库线程数	与规格大小有关。 <ul style="list-style-type: none"> 极小规格: 1-8 小规格: 1-16 中规格: 1-32 大规格: 1-128 超大规格: 1-128 	与规格大小有关。 <ul style="list-style-type: none"> 极小规格: 4 小规格: 8 中规格: 16 大规格: 64 超大规格: 64 	int	是
relay_socket_timeout	增量阶段从源数据库拉取日志的超时时间	30000-2147483647 毫秒	30000毫秒	int	是
datamove_source_socket_timeout	全量阶段从源数据库拉取数据的超时时间	120000-2147483647 毫秒	120000毫秒	int	是

6.6.2 查看参数修改历史

对于支持修改参数的DRS任务，您可以查看之前参数的修改历史，以满足业务需要。

操作步骤

- 步骤1** 在“实时同步管理”界面，选择指定的同步任务。
- 步骤2** 单击任务名称，页面跳转至“基本信息”页面。
- 步骤3** 单击左侧“参数管理”，在“参数修改历史”页签可对当前同步任务参数的修改历史进行查询。

默认查询7天内的参数修改历史。

图 6-49 查看修改历史



----结束

6.7 任务生命周期

6.7.1 查看任务详情

实时同步任务的配置信息，包括任务信息、同步实例信息和同步信息。同步任务创建成功后，可查看任务详细信息。

前提条件

已登录数据复制服务控制台。

操作步骤

说明

任务列表中，仅展示当前登录用户所创建的全部任务，同一租户下的不同用户之间的任务不可见。

步骤1 在“实时同步管理”界面，选择指定的同步任务，单击任务名称。

步骤2 在“基本信息”页签，查看当前同步任务的详细信息。

您可查看到当前同步任务下对应的任务信息、同步实例信息和同步信息。

---结束

6.7.2 修改任务信息

同步任务创建成功后，支持对部分任务基本信息进行修改，以方便区分和识别同步任务。

支持可修改的任务信息如下：

- 任务名称
- 描述
- SMN主题
- 时延阈值
- 任务异常结束天数
- 任务启动时间

前提条件

已登录数据复制服务控制台。

操作步骤

步骤1 在“实时同步管理”页面，选择指定的同步任务，单击任务名称。

步骤2 进入“基本信息”页签，在“任务信息”模块下，选择需要修改的任务信息。

- 对于任务名称、描述、SMN主题、时延阈值和任务异常结束天数，可以单击 ，进行修改。



- 单击 ，提交修改。
- 单击 ，取消修改。

表 6-12 实时同步任务信息

任务信息	描述
任务名称	任务名称在4到50位之间，必须以字母开头，可以包含字母、数字、中划线或下划线，不能包含其他的特殊字符。
描述	描述不能超过256位，且不能包含!<>&\"特殊字符。
SMN主题	需要先在SMN上申请主题并添加订阅。 SMN主题申请和订阅可参考《 消息通知服务用户指南 》。
时延阈值	时延阈值范围为0到3600s。 说明 当时延阈值设置为0时，不会发送通知给收件人。
任务异常结束天数（天）	输入值必须在14到100之间，默认值14。 说明 仅按需计费任务支持设置任务异常自动结束时间，包年/包月任务不支持。

- 对于任务启动时间，当同步任务为“等待启动”状态时，才可以进行修改。
需要在“任务信息”模块下的“计划启动时间”处，单击“修改”，选择指定时间，单击“确定”即可。

步骤3 修改完成后，可在“基本信息”页签下查看修改结果。

----结束

批量配置异常通知

步骤1 在“实时同步管理”页面的同步列表中，选择需要配置的任务。

步骤2 单击左上方的“批量操作”，选择“配置异常通知”。

图 6-50 批量操作



步骤3 在弹出的确认对话框中输入配置信息，确认后单击“是”，提交配置任务。

----结束

6.7.3 修改连接信息

数据同步过程中，您可能修改了源数据库或者目标数据库的密码信息，导致数据同步、数据对比、重置、续传、编辑同步对象、结束等操作失败，此时您需要通过数据复制服务控制台更新为正确的信息，然后续传任务。

对于如下的同步信息，您可以进行修改。

- 数据库密码
- 数据库IP
- 数据库端口
- 数据库用户名

约束限制

- 目前仅支持MySQL和GaussDB(for MySQL)为源且连接方式为输入IP地址的任务，在增量阶段修改IP、端口、用户名信息。当源库进行某些操作导致IP地址、端口、用户名变化时，可通过此功能更新为正确的信息。
- 修改IP功能适用于源库IP地址变化场景，修改前后的IP必须属于同一个数据实例，否则可能导致任务失败或者数据不一致。
- 连接信息修改后，将实时生效，不会清空目标数据库的数据。

操作步骤

步骤1 在“实时同步管理”界面，选中指定同步任务，单击任务名称。

步骤2 进入“基本信息”页签，在“连接信息”模块下，单击“修改连接信息”。

步骤3 在“修改连接信息”弹出框中对源库和目标库的密码等信息进行更新，更新完成后，单击“确认”即可。

----结束

6.7.4 修改流速模式

流速模式支持限速和不限速，默认为不限速。DRS提供创建任务后修改流速模式功能，目前DRS支持限速功能的同步链路可参考[支持限速的链路](#)。

约束限制

- 限速模式指的是限制秒级的流量上限，可能因为网络波动原因导致统计值的降低，实际统计的值可能会低于限速值大小。
- 限速模式只对全量同步阶段生效，增量同步阶段不生效。
- 增量阶段修改限速后，待任务再次进入全量阶段会生效。如增量中的任务修改限速，编辑任务新加同步对象后，限速会在同步任务的全量阶段生效。

前提条件

- 已登录数据复制服务控制台。
- 已创建同步任务。

方法一

步骤1 在“基本信息”页签的“限速信息”区域，单击“修改”。

步骤2 在弹出的对话框中修改流速模式。

图 6-51 修改流速模式



---结束

方法二

步骤1 在“实时同步管理”页面的同步列表中，选择需要修改流速模式的任务，单击操作列的“更多>限速”，或“限速”。

图 6-52 任务列表



步骤2 在弹出的对话框中修改流速模式。

图 6-53 修改流速模式

修改流速模式

流速模式 限速 不限速 ?

时区 GMT+08:00

流速设置 全天限流 时段限流 ?

时间范围 —

流速大小 MB/s(最大值为9999)

增加一个时间段 您还可以增加2个时间段

----结束

支持限速的链路

入云:

- MySQL->MySQL
- MySQL->GaussDB(for MySQL)
- MySQL->GaussDB主备版
- MySQL->GaussDB分布式版
- MySQL->PostgreSQL
- MySQL->GaussDB(DWS)
- MySQL->MariaDB
- Oracle->MySQL
- Oracle->GaussDB(for MySQL)
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式版
- Oracle->GaussDB(DWS)
- Oracle->DDM
- Oracle->PostgreSQL
- PostgreSQL->PostgreSQL
- PostgreSQL->GaussDB主备版
- PostgreSQL->GaussDB分布式版

- TiDB->GaussDB(for MySQL)
- Microsoft SQL Server->GaussDB(DWS)
- Microsoft SQL Server->GaussDB主备版
- Microsoft SQL Server->GaussDB分布式版
- Microsoft SQL Server->Microsoft SQL Server
- MongoDB->DDS
- MariaDB->MariaDB
- MariaDB->MySQL
- MariaDB->GaussDB(for MySQL)
- GaussDB(for MySQL)->GaussDB(for MySQL)

出云:

- MySQL->MySQL
- MySQL->CSS/ES
- MySQL->Oracle
- MySQL->Kafka
- MySQL->MariaDB
- GaussDB(for MySQL)->Oracle
- GaussDB(for MySQL)->Kafka
- GaussDB主备版->MySQL
- GaussDB主备版->Oracle
- GaussDB主备版->GaussDB(DWS)
- GaussDB主备版->GaussDB分布式版
- GaussDB主备版->GaussDB主备版
- GaussDB分布式版->MySQL
- GaussDB分布式版->Oracle
- GaussDB分布式版->GaussDB(DWS)
- GaussDB分布式版->GaussDB分布式版
- GaussDB分布式版->GaussDB主备版
- MariaDB->MariaDB
- PostgreSQL->PostgreSQL

自建-自建

- MySQL->CSS/ES
- MySQL->Kafka
- Oracle->Kafka
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式版
- GaussDB主备版->Oracle
- GaussDB主备版->GaussDB主备版

- GaussDB分布式版->Oracle
- GaussDB分布式版->GaussDB分布式版

6.7.5 编辑同步任务

已创建成功但是未确认启动的同步任务，数据复制服务支持编辑任务的配置信息，包括源库信息、目标库信息。创建实例后，对于如下状态的任务，您可进行编辑并提交同步任务。

- 创建中
- 配置

说明

对于增量同步中的任务，数据复制服务支持通过编辑同步对象功能进行对象修改，详细操作可参考[编辑同步对象](#)。

前提条件

已登录数据复制服务控制台。

方式一

- 步骤1** 在“实时同步管理”页面的任务列表中，选择要修改的任务，单击操作列“编辑”按钮。
- 步骤2** 进入“源库及目标库”页面，填写源库和目标库信息，单击“下一步”。
- 步骤3** 在“设定同步”页面，选择编辑同步对象，单击“下一步”。
- 步骤4** 在“预检查”页面，进行同步任务预校验，校验是否可进行实时同步。
- 步骤5** 在“任务确认”页面，设置同步任务的启动时间，并确认同步任务信息无误后，勾选协议，单击“启动任务”，提交同步任务。
- 步骤6** 同步任务提交后，您可在“实时同步管理”页面，查看并管理自己的任务。

----结束

方式二

- 步骤1** 在“实时同步管理”界面，选择指定同步任务，单击任务名称。
- 步骤2** 单击“进入”，进入“源库及目标库”页面。
- 步骤3** 执行[步骤2](#)到[步骤6](#)。

----结束

6.7.6 续传同步任务

实时同步中可能因外部因素导致同步暂时故障，用户根据同步日志信息解决问题后，可以通过续传功能继续同步，常见场景如存储空间不足。

对于以下状态的任务，可以通过续传功能继续同步任务。

- 同步失败

- 已暂停

📖 说明

- 如果因为非网络原因导致同步失败，系统默认进行三次自动续传，若三次续传后仍无法恢复，可手动进行续传。
- 如果因为网络原因导致同步失败，系统自动续传，直到恢复同步。

前提条件

已登录数据复制服务控制台。

方式一

在“实时同步管理”页面任务列表中，选择需要启动的任务，单击操作列“续传”，重新提交同步任务。

方式二

步骤1 在“实时同步管理”页面任务列表中，选择需要启动的任务，单击任务名称。

步骤2 跳转至“基本信息”页签，切换至“同步进度”页签，单击右上角“续传”，继续同步任务。

----结束

批量续传

步骤1 在“实时同步管理”页面的同步任务列表中，选择需要续传的任务。

步骤2 单击左上方的“批量操作”，选择“批量续传”。

图 6-54 批量操作



步骤3 在弹出的确认对话框中，确认任务信息后，单击“是”，提交续传任务。

----结束

6.7.7 跳跃续传同步任务

实时同步中可能因外部因素需要暂时停止同步，用户在解决问题后，可以通过跳跃续传功能跳过暂停时间段的数据继续同步。目前仅DDM->Oracle的增量同步阶段，支持跳跃续传功能。

前提条件

- 已登录数据复制服务控制台。

- 任务状态为暂停。

操作步骤

在“实时同步管理”页面任务列表中，选择需要启动的任务，单击操作列“跳跃续传”，重新提交同步任务。

6.7.8 暂停同步任务

DRS提供对实时同步中的任务进行暂停功能。目前支持暂停的同步链路可参考[支持暂停的链路](#)。

前提条件

- 已登录数据复制服务控制台。

暂停任务

步骤1 在“实时同步管理”页面的同步列表中，选择要暂停的任务，单击“操作 > 暂停”。

步骤2 在弹出的“暂停任务”对话框中，选择“暂停日志抓取”后，单击“是”。

📖 说明

- 暂停成功后，状态栏显示为“已暂停”。
- 勾选“暂停日志抓取”后，该功能会停止一切DRS和源库、目标库的连接，请注意暂停过长可能会因为源端所需日志过期而导致任务无法恢复续传。建议暂停时间不超过24小时，具体时间请排查相应日志配置。

----结束

批量暂停

步骤1 在“实时同步管理”页面的同步任务列表中，选择需要暂停的任务。

步骤2 单击左上方的“批量操作”，选择“批量暂停”。

图 6-55 批量操作



步骤3 在弹出的确认对话框中，确认任务信息后，单击“是”，提交暂停任务。

----结束

支持暂停的链路

以下链路支持增量同步阶段的暂停功能：

- 入云

- MySQL->MySQL
- MySQL->GaussDB(for MySQL)
- MySQL->GaussDB主备版
- MySQL->GaussDB分布式版
- MySQL->GaussDB(DWS)
- MySQL->PostgreSQL
- MySQL->MariaDB
- PostgreSQL->PostgreSQL
- PostgreSQL->GaussDB(DWS)
- PostgreSQL->GaussDB主备版
- PostgreSQL->GaussDB分布式版
- DDM->DDM
- DDM->GaussDB(DWS)
- Oracle->GaussDB(DWS)
- Oracle->PostgreSQL
- Oracle->GaussDB(for MySQL)
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式
- Oracle->MySQL
- DB2 for LUW->GaussDB主备版
- DB2 for LUW->GaussDB分布式版
- DB2 for LUW->GaussDB(DWS)
- TiDB->GaussDB(for MySQL)
- Microsoft SQL Server->GaussDB(DWS)
- Microsoft SQL Server->GaussDB主备版
- Microsoft SQL Server->GaussDB分布式版
- Microsoft SQL Server->Microsoft SQL Server
- MongoDB->DDS
- MariaDB->MariaDB
- MariaDB->MySQL
- MariaDB->GaussDB(for MySQL)
- GaussDB(for MySQL)->GaussDB(for MySQL)
- 出云
 - MySQL->MySQL
 - MySQL->CSS/ES
 - MySQL->Oracle
 - MySQL->Kafka
 - MySQL->MariaDB
 - DDM->Oracle
 - DDM->Kafka

- DDS->MongoDB
- DDS->Kafka
- GaussDB(for MySQL)->MySQL
- GaussDB(for MySQL)->Kafka
- GaussDB(for MySQL)->CSS/ES
- GaussDB(for MySQL)->Oracle
- GaussDB(for MySQL)->GaussDB(DWS)
- GaussDB主备版->MySQL
- GaussDB主备版->Oracle
- GaussDB主备版->Kafka
- GaussDB分布式版->MySQL
- GaussDB分布式版->Oracle
- GaussDB分布式版->Kafka
- PostgreSQL->PostgreSQL
- PostgreSQL->Kafka
- MariaDB->MariaDB
- Microsoft SQL Server->Kafka
- 自建-自建
 - MySQL->Kafka
 - MySQL->CSS/ES
 - MySQL->GaussDB主备版
 - MySQL->GaussDB分布式版
 - Oracle-> Kafka
 - Oracle->GaussDB主备版
 - Oracle->GaussDB分布式版
 - GaussDB主备版->Oracle
 - GaussDB主备版->Kafka
 - GaussDB分布式版->Oracle
 - GaussDB分布式版->Kafka
 - PostgreSQL->Kafka
 - DB2 for LUW->GaussDB主备版
 - DB2 for LUW->GaussDB分布式版
 - Microsoft SQL Server->Kafka

除此以外，以下链路还支持全量同步阶段的暂停功能：

- MySQL->MySQL
- MySQL->GaussDB(for MySQL)
- MySQL->GaussDB(DWS)
- MySQL->CSS/ES
- MySQL->Kafka

- GaussDB(for MySQL)->MySQL
- GaussDB(for MySQL)->GaussDB(for MySQL)
- GaussDB(for MySQL)->Kafka
- GaussDB(for MySQL)->GaussDB(DWS)
- Oracle->MySQL
- DDM->MySQL
- DDM->Oracle
- DDM->GaussDB(DWS)
- PostgreSQL->PostgreSQL
- PostgreSQL->GaussDB(DWS)
- MongoDB->DDS
- DDS->MongoDB

6.7.9 重置同步任务

在实时同步过程中，对于以下同步状态的任务，可通过重置功能，重新开始同步而不用再次配置任务。

- 同步暂停状态。
- 同步失败状态。

目前DRS支持重置的同步链路可参考[支持重置的链路](#)。

说明

DRS重置功能不会清空目标库，客户需要根据自己的需求选择是否清空目标库。

- 全量、全量+增量任务：为保证同步前后数据一致性，建议手动清空目标库后进行重置，任务重置后会重新进行全量同步，不需要再次配置任务。
- 单增量任务：仅同步增量数据，无需清空目标库，可直接进行重置。

前提条件

已登录数据复制服务控制台。

方法一

步骤1 在“实时同步管理”页面的任务列表中，选择要启动的任务，单击操作列“重置”按钮。

步骤2 在弹出的“重置任务”窗口，将重新进行同步任务预检查。

说明

重置多对一同步任务时，若预检查存在失败项，需要在查看失败详情中，单击对应失败的子任务名称，才能查看到具体任务的不通过原因。

步骤3 预检查完成后，且预检查通过率为100%时，单击“开始重置”按钮，重新提交同步任务。

----结束

方法二

步骤1 在“实时同步管理”页面，选择需要启动的任务，单击任务名称。

步骤2 页面跳转至“基本信息”页签，切换至“同步进度”页签，单击页面右上角“重置”按钮。

步骤3 继续执行方法一的步骤[步骤2](#)至[步骤3](#)。

----结束

支持重置的链路

入云：

- MySQL->MySQL
- MySQL->PostgreSQL
- MySQL->GaussDB(for MySQL)
- MySQL->GaussDB主备版
- MySQL->GaussDB分布式版
- MySQL->GaussDB(DWS)
- MySQL->MariaDB
- PostgreSQL->PostgreSQL
- PostgreSQL->GaussDB主备版
- PostgreSQL->GaussDB分布式版
- DDM->DDM
- Oracle->MySQL
- Oracle->GaussDB(for MySQL)
- Oracle->GaussDB(DWS)
- Oracle->PostgreSQL
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式
- DB2 for LUW->GaussDB主备版
- DB2 for LUW->GaussDB分布式版
- DB2 for LUW->GaussDB(DWS)
- TiDB->GaussDB(for MySQL)
- Microsoft SQL Server->GaussDB(DWS)
- Microsoft SQL Server->GaussDB主备版
- Microsoft SQL Server->GaussDB分布式版
- Microsoft SQL Server->Microsoft SQL Server
- MongoDB->DDS
- MariaDB->MariaDB
- MariaDB->MySQL
- MariaDB->GaussDB(for MySQL)

- GaussDB(for MySQL)->GaussDB(for MySQL)

出云:

- MySQL->MySQL
- MySQL->CSS/ES
- MySQL->Oracle
- MySQL->Kafka
- MySQL->MariaDB
- DDS->MongoDB
- DDS->Kafka
- GaussDB(for MySQL)->Kafka
- GaussDB(for MySQL)->CSS/ES
- GaussDB(for MySQL)->Oracle
- GaussDB(for MySQL)->MySQL
- GaussDB主备版->GaussDB主备版
- GaussDB主备版->GaussDB分布式版
- GaussDB主备版->GaussDB(DWS)
- GaussDB主备版->MySQL
- GaussDB主备版->Oracle
- GaussDB主备版->Kafka
- GaussDB分布式版->GaussDB分布式版
- GaussDB分布式版->GaussDB主备版版
- GaussDB分布式版->GaussDB(DWS)
- GaussDB分布式版->MySQL
- GaussDB分布式->Oracle
- GaussDB分布式->Kafka
- PostgreSQL->PostgreSQL
- PostgreSQL->Kafka
- MariaDB->MariaDB
- Microsoft SQL Server->Kafka

自建-自建:

- MySQL->Kafka
- MySQL->CSS/ES
- MySQL->GaussDB主备版
- MySQL->GaussDB分布式版
- Oracle->Kafka
- Oracle->GaussDB主备版
- Oracle->GaussDB分布式版
- GaussDB主备版->Oracle

- GaussDB主备版->Kafka
- GaussDB分布式版->Oracle
- GaussDB分布式版->Kafka
- PostgreSQL->Kafka
- DB2 for LUW->GaussDB主备版
- DB2 for LUW->GaussDB分布式版
- Microsoft SQL Server->Kafka

6.7.10 重置同步位点

对于MySQL->GaussDB(DWS)的实时同步链路，DRS支持增量位点重置功能，可根据用户选择的重置位点时间，将同步任务调整到问题发生之前，重新同步。

约束限制

- 目前仅支持引擎为MySQL->GaussDB(DWS)的同步任务支持重置位点。
- 任务在增量阶段时，需要先暂停任务，再进行重置位点。

操作步骤

步骤1 在“实时同步管理”页面的同步列表中，选择要已暂停的任务，单击任务“操作”列的“更多 > 重置位点”。

图 6-56 重置位点



步骤2 在重置位点对话框中，选择重置位点时间，单击“检查”。

图 6-57 确认信息



步骤3 检查重置位点无误后，单击“开始重置”，重新提交同步任务。

----结束

6.7.11 重启同步任务

DRS支持重启同步任务操作，方便用户在[修改任务参数](#)后，重启任务使设置生效。

约束限制

- 支持任务状态为全量同步、全量同步失败、增量同步、增量同步失败以及暂停时进行重启。
- 目前仅支持以下引擎的同步任务进行重启。
 - MySQL->MySQL
 - MySQL->PostgreSQL
 - MySQL->GaussDB分布式版
 - MySQL->GaussDB主备版
 - MySQL->GaussDB(DWS)
 - MySQL->GaussDB(for MySQL)
 - MySQL->Kafka
 - MySQL->CSS/ES
 - MySQL->Oracle

方法一

步骤1 在“实时同步管理”页面的任务列表中，选择要重启的任务，单击操作列“重启”按钮。

步骤2 在弹出的“重启任务”窗口，单击“是”，等待任务重启。

----结束

方法二

步骤1 在“实时同步管理”页面，选择需要重启的任务，单击任务名称。

步骤2 页面跳转至“基本信息”页签，单击页面右上角“重启”按钮。

步骤3 在弹出的“重启任务”窗口，单击“是”，等待任务重启。

----结束

6.7.12 跳过 DDL

对于目标数据库为GaussDB(DWS)的实时同步链路，增量阶段在源数据库执行的DDL操作时，DRS将原样在GaussDB(DWS)里执行。当目标数据库GaussDB(DWS)不支持该DDL语句时，会导致同步任务暂时故障。

针对这种情况，用户根据同步日志信息解决问题后，可以通过跳过DDL功能继续同步。

操作步骤

步骤1 联系GaussDB(DWS)技术支持，在目标数据库执行相同语义的DDL（当前提示显示的为源数据库的DDL，需在目标数据库执行符合目标库语法的语句）。

步骤2 执行完成后，在“实时同步管理”页面的同步列表中，选择要跳过的任务，单击任务“操作”列的“跳过DDL”。

图 6-58 跳过 DDL



步骤3 确认目标数据库执行完成后，单击“是”跳过报错，继续同步任务。

图 6-59 确认信息



----结束

6.7.13 双 AZ 任务倒换

DRS实时同步任务在创建时可选择任务类型为“双AZ”或“单AZ”。双AZ任务为一主一备的经典高可用架构，在任务创建后，DRS会同时创建一主一备两个子任务，当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步，提高同步任务的可靠性。

DRS目前支持以下链路在创建时进行任务类型的选择：

- 入云
 - MySQL->MySQL
 - MySQL->GaussDB(for MySQL)
 - MySQL->GaussDB(DWS)
 - MySQL->MariaDB
 - GaussDB(for MySQL)->GaussDB(for MySQL)
 - MariaDB->MySQL
 - MariaDB->GaussDB(for MySQL)
 - Oracle->GaussDB主备版
 - Oracle->GaussDB分布式版
- 出云
 - MySQL->Kafka
 - MySQL->MariaDB
 - GaussDB(for MySQL)->Kafka

- GaussDB(for MySQL)->CSS/ES
- GaussDB主备版->Kafka
- GaussDB分布式版->Kafka
- 自建-自建
 - MySQL->Kafka
 - Oracle->Kafka
 - Oracle->GaussDB主备版
 - Oracle->GaussDB分布式版
 - GaussDB主备版->Kafka
 - GaussDB分布式版->Kafka

前提条件

- 已登录数据复制服务控制台。

操作说明

步骤1 在“实时同步管理”页面，单击“创建同步任务”。

步骤2 在“同步实例”页面，填选任务名称、描述、同步实例等信息，其中DRS任务类型选择“双AZ”，完成后单击“下一步”。

图 6-60 同步实例信息



步骤3 返回“实时同步管理”页面可以看到，DRS会同时创建一主一备两个子任务。

图 6-61 主备任务



步骤4 配置并启动同步任务后，DRS会启动主AZ任务进行同步，此时备AZ任务处于暂停状态。

图 6-62 主备倒换前



步骤5 当主AZ任务发生异常，DRS会自动启动备AZ任务继续进行同步。

图 6-63 主备倒换后



----结束

6.7.14 克隆同步任务

DRS支持通过克隆功能，快速复制现有同步任务的配置。但是对于以下状态的任务，暂不支持进行克隆。

- 创建中
- 创建失败
- 配置
- 等待启动
- 启动中
- 删除

目前以下同步链路支持克隆功能：

- 入云
 - MySQL->MySQL
 - MySQL->MariaDB
 - PostgreSQL->PostgreSQL
- 出云
 - MySQL->MySQL
 - MySQL->CSS/ES
 - MySQL->MariaDB
 - DDS->MongoDB
- 自建-自建
 - MySQL->CSS/ES

📖 说明

- 任务克隆时，源和目标数据库密码信息不会被克隆，克隆后的新任务需再次输入。
- 克隆任务时，数据过滤中的高级设置信息不会被克隆，克隆后的新任务需重新进行设置。
- 不支持多对一任务的克隆。
- 对于变更中的任务进行克隆时，如果变更信息已存入数据库，克隆任务与变更后配置保持一致。
- 克隆任务创建成功后，新任务的弹性公网IP、私有IP与原任务不同，可能需要调整网络配置，确保克隆后的新任务和源库、目标库网络互通。

前提条件

- 已登录数据复制服务控制台。
- 已创建同步任务。

操作步骤

- 步骤1** 在“实时同步管理”页面的同步列表中，选择要克隆的任务，单击“操作 > 克隆任务”。
- 步骤2** 在弹出的“克隆任务”对话框中，确认新任务名称，单击“克隆”提交任务。
- 步骤3** 任务提交后，待任务克隆完成，状态为“配置”，您可单击“操作 > 编辑”，重新输入源和目标数据库密码信息，编辑和启动任务。
- 结束


6.7.15 同步规格变更

当DRS任务规格无法满足业务需要时，可参考以下内容升级DRS任务规格，变更开始后，任务状态为规格变更中，数据同步处于暂停阶段，变更完成后，会自动续传。该功能目前仅支持白名单用户，需要提交工单申请才能使用。

约束限制

- 账户余额大于等于0美元，才可变更规格。
- 目前同步仅支持任务类型为单AZ的同步任务升级实例规格，不支持父子任务升级实例规格，不支持降低规格。
- 目前仅支持任务状态为全量中、全量失败、增量中、增量失败的同步任务进行规格变更。
- 建议在业务低峰期操作规格变更。
- 变更开始后，任务处于暂停阶段，变更完成后，会自动续传。
- 任务规格变更时间大概需要5-10分钟左右，请耐心等待！

操作步骤

- 步骤1** 登录管理控制台。
- 步骤2** 单击管理控制台左上角的，选择区域和项目。
- 步骤3** 选择“数据库 > 数据复制服务 DRS”。进入数据复制服务页面。
- 步骤4** 在“实时同步管理”页面的任务列表中，选择操作列的“更多” > “规格变更”。
- 步骤5** 在“规格变更”页面，选择所需修改的性能规格并进行预检查后，单击“下一步”。
- 步骤6** 进行规格确认。
- 如需重新选择，单击“上一步”，回到上个页面，修改规格。
 - 按需计费模式的实例，单击“启动变更”，提交变更。
由规格变更产生的费用，您可在“费用中心 > 费用账单”中查看费用详情。
 - 包年/包月模式的实例，单击“启动变更”，进入付费确定页面，单击“确认付款”，支付成功后，才可进行规格变更。

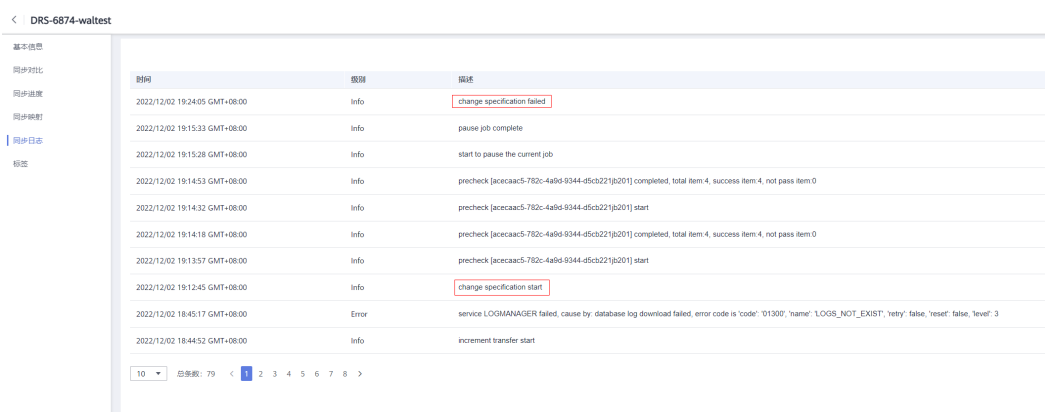
步骤7 查看变更结果。

提交成功后，单击“返回数据复制服务DRS列表”，在实时同步实例管理页面，可以看到实例状态为“规格变更中”。

稍后等到任务状态从规格变更中变成其他状态后，可在实例“基本信息”页面，查看实例规格，检查变更是否成功。或者在“同步日志”页面查看变更日志。

- 规格变更开始：change specification start
- 规格变更成功：change specification success
- 规格变更失败：change specification failed

图 6-64 规格变更日志



时间	级别	描述
2022/12/02 19:24:05 GMT+08:00	Info	change specification failed
2022/12/02 19:15:33 GMT+08:00	Info	pause job complete
2022/12/02 19:15:28 GMT+08:00	Info	start to pause the current job
2022/12/02 19:14:53 GMT+08:00	Info	precheck [jaceaa5-782c-4a9d-9344-d5cb2219201] completed, total item:4, success item:4, not pass item:0
2022/12/02 19:14:32 GMT+08:00	Info	precheck [jaceaa5-782c-4a9d-9344-d5cb2219201] start
2022/12/02 19:14:18 GMT+08:00	Info	precheck [jaceaa5-782c-4a9d-9344-d5cb2219201] completed, total item:4, success item:4, not pass item:0
2022/12/02 19:13:57 GMT+08:00	Info	precheck [jaceaa5-782c-4a9d-9344-d5cb2219201] start
2022/12/02 19:12:45 GMT+08:00	Info	change specification start
2022/12/02 18:45:17 GMT+08:00	Error	service LOGMANAGER failed, cause by database log download failed, error code is '01300', 'name': 'LOGS_NOT_EXIST', 'retry': false, 'reset': false, 'level': 3
2022/12/02 18:44:52 GMT+08:00	Info	increment transfer start

----结束

6.7.16 退订包周期任务

对于“包年/包月”模式的DRS任务，您需要先退订订单，结束任务，从而删除DRS实例资源。

前提条件

- 已登录数据复制服务控制台。
- 当前DRS实例计费模式为“包年/包月”。

方法一

您可在DRS“实时同步管理”页面的实例列表中，退订包周期任务。

步骤1 登录管理控制台。

步骤2 单击管理控制台左上角的📍，选择区域和项目。

步骤3 选择“数据库 > 数据复制服务 DRS”。进入数据复制服务页面。

步骤4 在“实时同步管理”页面的任务列表中，选择操作列的“更多” > “退订”。

步骤5 在退订任务提示窗中，单击“是”。系统自动跳转至“费用中心-退订资源”页面。

图 6-65 退订任务



步骤6 在“退订资源”页面，核实待退订实例信息，填写退订原因，勾选我已确认不再使用，单击下方“退订”。

说明

资源退订后，DRS任务将会立即结束，请确认已完成数据同步或不再使用。


步骤7 在退订弹窗中，单击“退订”，即可退订成功。

----结束

方法二

您可前往“费用中心”，退订包周期任务。

步骤1 登录管理控制台。

步骤2 单击管理控制台左上角的 ，选择区域和项目。

步骤3 选择“数据库 > 数据复制服务 DRS”。进入数据复制服务页面。

步骤4 单击页面右上角的“费用与成本”，进入费用中心页面。

步骤5 在左侧导航栏中选择“订单管理 > 云服务退订”，进入“云服务退订”页面。

步骤6 在“云服务退订”页面，勾选需要退订的实例订单，单击“操作”列的“退订资源”。

- 您可通过产品类型筛选出账号下所有的DRS订单。

图 6-66 筛选服务



- 您还可以在订单列表上方，通过实例名称、订单号或实例ID搜索资源。

步骤7 在“退订资源”页面，确认待退订实例信息，并选择退订原因，单击“退订”。

资源退订相关信息，请参考[退订规则说明](#)。

步骤8 在弹出框中确认是否退订该资源，单击“是”，提交退订申请。

说明

资源退订后，DRS任务将会立即结束，请确认已完成数据同步或不再使用。

----结束

6.7.17 结束同步任务

业务系统和数据库切换至目标数据库后，可选择结束同步任务。对于需要恢复目标数据库或停止同步的任务，您可选择结束任务，避免源数据库的操作继续同步到目标数据库，造成数据覆盖问题。

如下状态下的任务可以结束：

- 创建中
- 配置
- 等待启动
- 全量同步
- 全量同步失败
- 增量同步
- 增量同步失败
- 已暂停
- 故障恢复

须知

- 建议您先结束任务，再做断开源库与同步实例的网络等其他操作，避免产生无法连接源库的告警。
 - 对于“配置”状态的任务，配置失败的任务无法结束。
 - 对于“故障恢复”状态的任务，正在进行故障恢复的任务无法结束。
 - 任务结束后无法重试。
-

操作步骤

步骤1 在“实时同步管理”页面的任务列表中，选择要结束的任务，单击“结束”。

步骤2 在弹出框中单击“确定”，提交结束任务。

说明

- 当任务状态异常时（例如任务失败、网络异常），DRS会勾选“强制结束任务”优先结束任务，减少等待时间。
- “强制结束任务”会直接释放DRS资源，请检查是否存在影响，确认同步情况。
- 如果需要正常结束任务，请先修复DRS任务，待任务状态正常后，再单击“结束”，正常结束任务。

----结束

批量结束

步骤1 在“实时同步管理”页面的同步任务列表中，勾选需要结束的任务。

步骤2 单击左上方的“批量操作”，选择“批量结束”。

图 6-67 批量操作



步骤3 在弹出的确认对话框中，确认任务信息后，单击“是”，提交结束任务。

----结束

6.7.18 删除同步任务

对于已结束或者配置失败的实时同步任务，您可选择删除同步任务。被删除后的任务将不会再出现在任务列表中，请谨慎操作。

前提条件

已登录数据复制服务控制台。

删除任务

步骤1 在“实时同步管理”页面的任务列表中，选择要删除的任务，单击“删除”。

步骤2 单击“确定”，提交删除任务。

----结束

批量删除

步骤1 在“实时同步管理”页面的同步任务列表中，选择需要删除的任务。

步骤2 单击左上方的“批量操作”，选择“批量删除”。

图 6-68 批量操作



步骤3 在弹出的确认对话框中，确认任务信息后，单击“是”，提交删除任务。

---结束

6.7.19 批量导入任务

DRS支持通过Excel表格导入的形式创建同步任务，用户可以将任务基本信息、源库和目标库信息、以及待同步的对象等信息填写到模板中，再进行导入，批量创建同步任务。

约束限制

- 目前仅支持引擎为MySQL->MySQL、MySQL->GaussDB主备版、MySQL->GaussDB分布式版、Oracle->GaussDB主备版、Oracle->GaussDB分布式版、GaussDB主备版->GaussDB主备版、GaussDB主备版->GaussDB分布式版、GaussDB主备版->Kafka、GaussDB分布式版->GaussDB分布式版、GaussDB分布式版->GaussDB主备版、GaussDB分布式版->Kafka的任务进行导入。
- 修改导入文件时模板时，示例内容需要清除，填写为正确的值。
- 在填写源和目标库实例时，可以只填必须参数，但需要用户自行保证源和目标库的子网是打通的。
- 导入模板是结合API和界面信息来设计的，请根据以下参数说明严格填写，并自行保证参数准确性。
- 导入模板上传成功后即下发申请创建DRS实例，故部分参数不可修改，请仔细确认。
- 导入模板上传成功后任务会自动保存所有参数，并自动进行测试连接、对象选择、预检查等步骤。预检查成功后任务为配置中状态，可单击“立即启动”按钮直接启动任务，或单击“编辑”按钮进行其他参数修改。
- 导入模板上传任务不支持在模板中配置定时启动时间，待导入任务自动运行至预检查成功状态后，可以通过界面“编辑”进入到启动页面自行设置定时启动时间。
- 导入模板中的参数校验不能校验所有参数的正确性，建议单击“编辑”按钮进入页面，确认每一步保存的数据正确性后再进行启动。
- 导入模板为通用模板，一个sheet创建一个任务，若需要批量创建任务，请复制多个sheet，进行参数配置。

操作步骤

步骤1 在“实时同步管理”页面，单击左上角“批量导入任务”。

步骤2 在“批量导入任务”对话框中，单击“下载模板”。

图 6-69 批量导入任务



步骤3 编辑下载的Excel模板，按照要求填写任务信息。

步骤4 填写完成后，单击“添加文件”，完成模板上传。

步骤5 待上传并校验成功后，单击“确定”提交创建任务。

图 6-70 导入成功



说明

- DRS不会保存上传的Excel文件，只针对内容进行解析。
- 任务创建后，数据库的IP地址或域名、数据库用户名和密码，会被系统加密暂存，直至删除该任务后自动清除。

----结束

6.7.20 任务状态说明

实时同步提供了多种任务状态，以便区分不同阶段的同步任务。

实时同步任务的状态和说明，请参见表6-13。

表 6-13 实时同步任务状态和说明

状态	说明
创建中	正在创建数据复制服务需要用到的同步实例。

状态	说明
创建任务失败	创建实时同步实例失败。
配置	同步实例创建成功，但还没有启动任务，可以继续配置任务。
冻结	账户余额小于或等于0美元，系统对该用户下的实例进行冻结。
等待启动	已经下发了定时启动同步任务到同步实例上，等待同步实例启动任务。
启动中	正在启动实时同步任务。
启动失败	实时同步任务启动失败。
全量同步	正在进行源数据库到目标数据库的全量同步任务。
全量同步失败	实时同步任务全量同步失败。
增量同步	持续进行从源数据库到目标数据库的增量实时同步。
增量同步失败	实时同步任务增量同步失败。
任务变更中	编辑同步对象后，正在进行同步任务变更。
任务变更失败	编辑同步对象后，同步任务变更失败。
故障恢复	同步实例发生故障，系统自动恢复同步任务。
已暂停	实时同步任务已暂停。
克隆中	正在进行同步任务克隆。
克隆失败	克隆同步任务失败。
规格变更中	正在进行同步实例规格变更，变更结束任务自动恢复任务。
规格变更失败	同步实例规格变更失败。
AZ倒换失败	实时同步任务进行AZ倒换失败。
结束任务	释放执行同步任务所使用的同步实例和资源。
结束任务中	正在释放同步任务所使用的同步实例和资源。
结束任务失败	释放同步任务所使用的同步实例和资源失败。
已结束	释放同步任务所占同步实例成功。

📖 说明

- 对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。
- 对于未启动、状态为配置中的任务，DRS默认保留3天，超过3天DRS会自动删除后台资源，当前任务状态不变。当用户再次配置时，DRS会重新申请资源。
- 已删除的同步任务在状态列表中不显示。

7 标签管理

操作场景

标签管理服务（Tag Management Service, TMS）用于用户在云平台，通过统一的tag标签管理各种资源。TMS服务与各服务共同实现标签管理能力，TMS提供全局标签管理能力，各服务维护自身标签管理。当用户资源较多时，可通过不同的标签标识和查找任务。

- 建议您先在TMS系统中设置预定义标签。
- 标签由“键”和“值”组成，每个标签中的一个“键”只能对应一个“值”。
- 每个实例最多支持20个标签配额。

添加标签

- 步骤1** 在“实时同步管理”页面，选择指定的同步任务，单击任务名称进入“基本信息”页签。
- 步骤2** 在左侧导航栏，单击“标签”。
- 步骤3** 在“标签”页签单击“添加/编辑标签”，在“添加/编辑标签”弹出框中，输入标签的键和值，单击“添加”，然后单击“确定”。

添加/编辑标签

×

如果您需要使用同一标签标识多种云资源，即所有服务均可在标签输入框下拉选择同一标签，建议在TMS中创建预定义标签。[查看预定义标签](#) 

在下方键/值输入框输入内容后单击“添加”，即可将标签加入此处



您还可以添加19个标签。

- 输入标签的键和值时，系统会自动联想当前用户的所有实例（除当前实例外）的所有关联的预定义标签。
- 标签的键不能为空且必须唯一，长度为1~128个字符，可以包含任意语种字母、数字、空格和_./:=+@字符，但首尾不能含有空格，不能以_sys_开头。
- 标签的值可以为空字符串，长度为0~255个字符，可以包含任意语种字母、数字、空格和_./:=+@字符，但首尾不能含有空格。

步骤4 添加成功后，您可在当前任务的所有关联的标签集合中，查询并管理自己的标签。

----结束

编辑标签

步骤1 在“实时同步管理”页面，选择指定的同步任务，单击任务名称进入“基本信息”页签。

步骤2 在左侧导航栏，单击“标签”。

步骤3 在“标签”页签，单击“添加/编辑标签”，在弹出框中修改标签，单击“确定”。

----结束

删除标签

步骤1 在“实时同步管理”页面，选择指定的同步任务，单击任务名称进入“基本信息”页签。

步骤2 在左侧导航栏，单击“标签”。

步骤3 在“标签”页签，选择需要删除的标签，单击操作列的“删除”，在“删除标签”弹出框中单击“是”。

步骤4 删除成功后，该标签将不再显示在任务的所有关联的标签集合中。

----结束

8 连接诊断

当DRS任务与源库或目标库进行测试连接失败时，DRS提供一键诊断功能，可对源库或者目标库当前网络连接情况进行诊断并返回结果。

- 仅支持通过输入IP地址或者界面选择方式，获取数据库信息的任务节点进行网络诊断，暂不支持GaussDB的DN诊断。
- 集群或多AZ任务场景，只支持在主任务的节点进行诊断。

前提条件

- 已登录数据复制服务控制台。
- 已创建任务。

操作步骤

步骤1 在任务管理界面，选择指定的任务，单击任务名称。

步骤2 在“源库及目标库”页面，填选源库信息和目标库信息，单击“源库和目标库”处的“测试连接”，分别测试并确定与源库、目标库连通。

当测试连接失败时，可单击失败信息右侧的“一键诊断”按钮进行诊断。

图 8-1 一键诊断

源库信息

不支持数据库参数和系统数据库同步，源数据库参数设定和用户将不会同步至目标数据库中，请在目标数据库中使用参数组修改参数，手工创建用户。

IP地址或域名	<input type="text"/>
端口	<input type="text"/>
数据库用户名	<input type="text" value="root"/>
数据库密码	<input type="password" value="*****"/> 
SSL安全连接	<input type="checkbox"/>
<input type="button" value="测试连接"/> ! DRS.KE0015: 迁移实例和数据库的网络连接异常, 查看详情 	
<input type="button" value="一键诊断"/>	

步骤3 等待诊断完成后，可在弹出的“诊断详情”窗口查看结果，包括丢包率和端口检查结果。

图 8-2 诊断详情



----结束

9 查看异常数据

同步过程中可能因数据冲突、数据加工、异构类型转化、对象缺失等因素导致数据异常，您可以通过“异常数据”页签查看，以帮助异常写入的排查。

说明

目前仅DDM->GaussDB(DWS)、DDM->Oracle、Oracle->GaussDB(DWS)、Oracle->RDS for MySQL、Oracle->GaussDB(for MySQL)、Oracle->PostgreSQL、MySQL->GaussDB(DWS)、MySQL->GaussDB和MySQL->PostgreSQL支持该功能。

操作步骤

步骤1 在“实时同步管理”界面，选择指定的同步任务，单击任务名称。

步骤2 在“异常数据”页签，您可以通过时间段筛选查看当前同步任务的异常数据，包括执行语句和异常描述等信息。

---结束

10 对接云审计服务

10.1 支持审计操作的关键列表

云审计服务是安全解决方案中专业的日志审计服务，记录了数据复制服务的相关操作事件，方便您日后的查询、审计和回溯。

表 10-1 云审计服务支持的数据复制服务操作列表

操作名称	资源类型	事件名称
创建任务	job	createJob
编辑任务	job	modifyJob
删除任务	job	deleteJob
启动任务	job	startJob
续传任务	job	retryJob

10.2 如何查看审计日志

在您开启了云审计后，系统开始记录云服务资源的操作。云审计服务管理控制台保存最近7天的操作记录。


本节介绍如何在云审计管理控制台查看最近7天的操作记录。

前提条件

已开启云审计服务。

操作步骤

步骤1 登录管理控制台。

步骤2 在管理控制台左上角单击  图标，选择区域和项目。


步骤3 单击“服务列表”，选择“管理与监管 > 云审计服务 CTS”，进入云审计服务信息页面。

步骤4 单击左侧导航树的“事件列表”，进入事件列表信息页面。

步骤5 事件列表支持通过高级搜索来查询对应的操作事件，详细信息如下：

- 时间范围：可在页面右上角选择查询最近1小时、最近1天、最近1周及自定义时间段的操作事件。
- 事件类型、事件来源、资源类型和筛选类型：在下拉框中选择查询条件。其中筛选类型选择资源ID时，还需选择或者手动输入某个具体的资源ID。若事件类型选择了数据事件，则可根据追踪器来过滤，其他过滤条件不支持。
- 操作用户：在下拉框中选择某一具体的操作用户，此操作用户指用户级别，而非租户级别。
- 事件级别：可选项为“所有事件级别”、“normal”、“warning”、“incident”，只可选择其中一项。

步骤6 选择查询条件后，单击“查询”。

步骤7 在需要查看的记录左侧，单击  展开该记录的详细信息。

步骤8 在需要查看的记录右侧，单击“查看事件”，在弹出框中显示该操作事件结构的详细信息。

----结束

11 对接云监控服务

11.1 支持的监控指标

功能说明

本节定义了数据复制服务上报云监控的监控指标的命名空间，监控指标列表和维度定义，用户可以通过云监控提供的API接口来检索数据复制服务产生的监控指标和告警信息。

命名空间

SYS.DRS

实例监控指标

数据复制服务性能监控指标，如表11-1所示。

表 11-1 数据复制服务支持的监控指标

指标ID	指标名称	指标含义	取值范围	测量对象	监控周期 (原始指标)
cpu_util	CPU使用率	该指标用于统计测量对象的CPU使用率，以比率为单位。	0-100%	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟
mem_util	内存使用率	该指标用于统计测量对象的内存使用率，以比率为单位。	0-100%	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟

指标ID	指标名称	指标含义	取值范围	测量对象	监控周期 (原始指标)
network_incoming_bytes_rate	网络输入吞吐量	该指标用于统计平均每秒从测量对象的所有网络适配器输入的流量,以字节/秒为单位。	≥ 0 bytes/s	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
network_outgoing_bytes_rate	网络输出吞吐量	该指标用于统计平均每秒从测量对象的所有网络适配器输出的流量,以字节/秒为单位。	≥ 0 bytes/s	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
disk_read_bytes_rate	磁盘读吞吐量	该指标用于统计每秒从磁盘读取的字节数,以字节/秒为单位。	≥ 0 bytes/s	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
disk_write_bytes_rate	磁盘写吞吐量	该指标用于统计每秒写入磁盘的字节数,以字节/秒为单位。	≥ 0 bytes/s	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
disk_util	磁盘利用率	该指标用于统计测量对象的磁盘利用率,以比率为单位。	0-100%	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
extract_bytes_rate	读源库吞吐量	该指标用于统计每秒从源库读取表数据或者WAL的字节数。	≥ 0 bytes/s	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
extract_rows_rate	读源库频率	该指标用于统计每秒从源库读取表数据或者WAL的行数,以行/秒为单位。	≥ 0 row/s	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟

指标ID	指标名称	指标含义	取值范围	测量对象	监控周期 (原始指标)
extract_latency	源库WAL抽取延迟	该指标用于统计当前从源库抽取WAL的延迟，以毫秒为单位。	≥ms	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟
apply_bytes_rate	写目标库吞吐量	该指标用于统计每秒向目标库写入数据的字节数，以字节/秒为单位。	≥ 0 bytes/s	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟
apply_rows_rate	写目标库频率	该指标用于统计每秒向目标库写入数据的行数，以行/秒为单位。	≥ 0 row/s	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟
apply_transactions_rate	DML TPS	该指标用于统计每秒向目标库写入的DML事务数，以事务/秒为单位。	≥ 0 transaction/s	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟
apply_ddls_number 或 apply_ddls_rate 说明 apply_ddls_rate在2022年12月之后替换为apply_ddls_number。	DDL TPS	该指标用于统计向目标库写入的DDL事务总数。	≥ 0 transaction	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟
apply_latency	数据同步延迟	该指标用于统计当前数据向目标库回放延迟，以毫秒为单位。	≥ 0 ms	测量对象：弹性云服务器 监控实例类型：实时迁移、实时同步和实时灾备实例	1分钟

指标ID	指标名称	指标含义	取值范围	测量对象	监控周期 (原始指标)
apply_average_execute_time	事务平均执行时间	该指标用于统计当前一个事务在目标库的平均执行时间(一个事务的RT=执行时间+提交时间),以毫秒为单位。	≥ 0 ms	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
apply_average_commit_time	事务平均提交时间	该指标用于统计当前一个事务在目标库的平均提交时间(一个事务的RT=执行时间+提交时间),以毫秒为单位。	≥ 0 ms	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
apply_current_state	同步状态	该指标用于表示当前内核数据的同步状态(10: 异常, 1: 空闲, 2: 执行DML, 3: 执行DDL),不代表任务状态。	10: 异常 1: 空闲 2: 执行DML 3: 执行DDL	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
apply_thread_workers	同步线程数量	该指标用于表示当前数据同步的工作线程数量。	≥ 0	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟
apply_job_status	任务状态	该指标用于表示当前任务所处的状态(0: 正常, 1: 异常, 2: 暂停)。	0: 正常 1: 异常 2: 暂停	测量对象: 弹性云服务器 监控实例类型: 实时迁移、实时同步和实时灾备实例	1分钟

维度

Key	Value
instance_id	数据复制服务实例ID

11.2 设置告警规则

操作场景

通过设置数据复制服务告警规则，用户可自定义监控目标与通知策略，及时了解数据复制服务运行状况，从而起到预警作用。

设置关数据复制服务的告警规则包括设置告警规则名称、服务、维度、监控范围、模板和是否发送通知等参数。

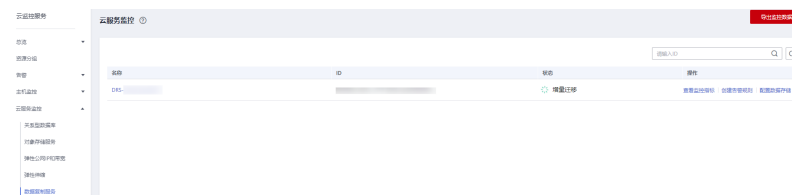
操作步骤

步骤1 登录管理控制台。

步骤2 在“所有服务”或“服务列表”中选择“管理与监管 > 云监控服务 CES”，进入“云监控服务”信息页面。

步骤3 在左侧导航栏选择“云服务监控 > 数据复制服务”。

图 11-1 选择监控对象




步骤4 选择需要添加告警规则的实例，单击操作列的“创建告警规则”。

步骤5 在“创建告警规则”页面，填选相关信息。

图 11-2 填选告警规则信息

告警规则	告警级别	操作
CPU使用率 原始值 > 90% 持续5个周期 则告警 每天告警一次	重要	删除
复制延迟 原始值 >= 300 秒 持续2个周期 则告警 每天告警一次	提示	删除
连接数使用率 原始值 >= 70% 持续2个周期 则告警 每天告警一次	重要	删除
磁盘利用率 原始值 > 75% 持续2个周期 则告警 每天告警一次	重要	删除
慢日志个数统计 原始值 >= 600 个/分钟	提示	删除

- 输入告警“名称”和“描述”。
- “选择类型”建议“从模板导入”，模板中，已经包含CPU使用率、内存使用率、磁盘利用率三个常用告警指标。
- 单击  开启“发送通知”，生效时间默认为全天，若没有您想要选择的主题，可以单击下一行的“创建主题”进行添加，“触发条件”勾选“出现告警”和“恢复正常”。

说明

该告警规则仅在生效时间段内发送通知消息。

步骤6 单击“立即创建”，告警规则创建完成。

关于告警参数的配置，请参见《[云监控用户指南](#)》。

----结束

11.3 查看监控指标

操作场景

云服务平台提供的云监控，可以对数据复制服务的实时迁移、实时同步和实时灾备实例（下文简称实例）的运行状态进行日常监控。您可以通过管理控制台，直观地查看实例的各项监控指标。由于监控数据的获取与传输会花费一定时间，因此，云监控显示的是当前时间5~10分钟前的实例状态。如果您的实例刚创建完成，请等待5~10分钟后查看监控数据。


前提条件

实例正常运行，即：

- 实时迁移：状态为全量迁移、增量迁移。
- 实时同步：状态为全量同步、增量同步。
- 实时灾备：状态为灾备中。

查看实例监控

步骤1 登录管理控制台。

步骤2 单击管理控制台左上角的 ，选择区域和项目。

步骤3 选择“数据库 > 数据复制服务 DRS”。进入数据复制服务页面。

步骤4 单击左侧目录的“实时迁移管理”，在“实时迁移管理”页面，单击目标实例名称，进入基本信息页面（此处以实时迁移为例，实时同步和实时灾备操作类似）。

步骤5 在实例的基本信息页面，单击右上角的“查看监控”，跳转到云监控页面。

该页面默认展示数据复制服务实例监控信息。

步骤6 在云监控页面，可以查看实例监控信息。

- 通过“设置监控指标”框可选择您在页面中要展示的指标名称并排序。
- 您也可根据业务需求，拖动其中的监控视图，调整监控视图的顺序。
- 云监控支持的性能指标监控时间窗包括：近1小时、近3小时、近12小时、近24小时、近7天和近6个月。

图 11-3 查看实例监控信息



----结束

12 对接云日志服务

12.1 日志配置管理

操作场景


配置访问日志后，DRS实例（包含实时迁移、备份迁移、实时同步、实时灾备和录制回放实例）新生成的所有日志记录会上传到云日志服务（Log Tank Service，简称LTS）进行管理。

注意事项

- 进行LTS日志配置后，会默认上传该任务的所有日志。
- 配置完成后不会立即生效，存在10分钟左右的时延。
- 配置完成后，会产生一定费用，费用情况请参考LTS的[定价详情](#)。
- 确保与DRS任务相同Region下的LTS服务已有日志组和日志流。
更多日志组和日志流信息，请参见[日志管理](#)。
- 关闭LTS日志配置后，将不再进行收费。

配置访问日志


步骤1 登录管理控制台。

步骤2 单击管理控制台左上角的，选择区域和项目。

步骤3 选择“数据库 > 数据复制服务 DRS”。进入数据复制服务页面。

步骤4 单击左侧目录的“实时迁移管理”，在“实时迁移管理”页面，单击目标实例名称，进入基本信息页面（此处以实时迁移为例，实时同步、实时灾备、录制回放操作类似）。

步骤5 在实例的基本信息页面，单击左侧“迁移日志”，进入日志页面。

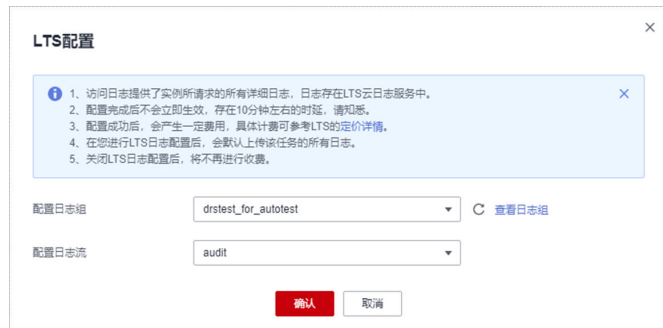
步骤6 单击页面上方“LTS配置”旁边的进行LTS日志配置。

步骤7 在下拉框分别选择LTS日志组和日志流，单击“确定”。

说明

配置完成后不会立即生效，存在10分钟左右的时延，请知悉。

图 12-1 LTS 配置



步骤8 如需解除或修改LTS配置，单击“LTS配置”开关或者旁边的修改进行LTS日志配置。

- **修改LTS配置：**单击“LTS配置”开关旁边的“修改”，在弹框中重新选择LTS日志组和日志流后，单击“确定”。
- **解除LTS配置：**单击“LTS配置”开关，在弹框中单击“确定”。

图 12-2 解除配置



----结束

12.2 查看或下载日志

操作场景

如果DRS任务已经配置了访问日志，操作详情请参见[日志配置管理](#)，就可以在LTS服务分析日志、搜索日志、日志可视化、下载日志和查看实时日志。

查看 LTS 日志

步骤1 登录管理控制台。

步骤2 单击管理控制台左上角的📍，选择区域和项目。

步骤3 在“所有服务”或“服务列表”中选择“管理与监管 > 云日志服务”，进入“云日志服务”信息页面。

步骤4 在“日志组列表”区域，选择目标日志组，单击日志组名称。更多关于LTS日志详细操作可参考《云日志服务LTS用户指南》。

图 12-3 查看日志详情

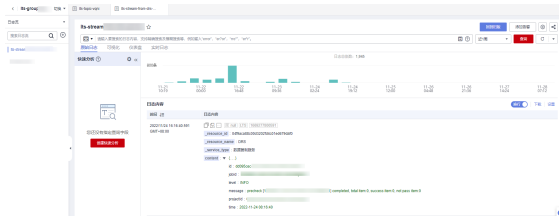


表 12-1 日志字段说明

名称	参数类型	说明
_resource_id	String	资源ID，DRS固定为projectId。
_resource_name	String	资源名称，固定为“DRS”。
_service_type	String	服务类型，固定为“数据复制服务”。

----结束

下载 LTS 日志文件

步骤1 登录管理控制台。

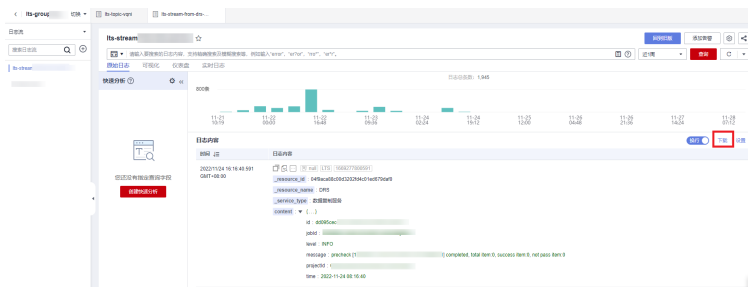
步骤2 单击管理控制台左上角的📍，选择区域和项目。

步骤3 在“所有服务”或“服务列表”中选择“管理与监管 > 云日志服务”，进入“云日志服务”信息页面。

步骤4 在“日志组列表”区域，选择目标日志组，单击日志组名称。

步骤5 单击右侧“下载”，进行日志下载。更多关于LTS日志详细操作可参考《云日志服务LTS用户指南》。

图 12-4 下载日志



----结束

13 同步场景操作参考

13.1 Kafka 消息格式

同步到Kafka集群中的数据以Avro、JSON和JSON-C格式存储。不同链路支持的数据格式可参考[表13-1](#)：

表 13-1 支持的数据格式

数据流向	Avro	JSON	JSON-C
MySQL->Kafka	支持	支持	支持
Oracle->Kafka	支持	支持	暂不支持
DDS->Kafka	暂不支持	支持	暂不支持
PostgreSQL->Kafka	支持	支持	暂不支持
GaussDB(for MySQL)->Kafka	支持	支持	支持
GaussDB主备版->Kafka	支持	支持	暂不支持
GaussDB分布式版->Kafka	支持	支持	暂不支持
Microsoft SQL Server->Kafka	支持	支持	暂不支持

Avro 格式

Avro格式的schema定义详情请参见[record.rar](#)。在实时同步到Kafka集群后，您需要根据schema定义进行数据解析。

JSON 格式

MySQL、GaussDB(MySQL)到Kafka的JSON格式定义详情参考[表13-2](#)，DDS到Kafka的JSON格式定义详情参考[表13-3](#)，Oracle、PostgreSQL、GaussDB、Microsoft SQL Server到Kafka的JSON格式定义详情参考[表13-4](#)。

表 13-2 MySQL 到 Kafka 的参数说明

参数名称	说明
mysqlType	源端表字段名称和类型。
id	DRS内部定义的事件操作的序列号，单调递增。
es	源库产生这一条记录的时间，13位Unix时间戳，单位为毫秒。
ts	写入到目标kafka的时间，13位Unix时间戳，单位为毫秒。
database	数据库名称。
table	表名。
type	操作类型，比如DELETE，UPDATE，INSERT，DDL，全量同步为INIT和INIT_DDL。
isDdl	是否是DDL操作。
sql	DDL的SQL语句，在DML操作中，取值为""。
sqlType	源端表字段的jdbc类型。
data	最新的数据，为JSON数组，如果type参数是插入则表示最新插入的数据，如果是更新，则表示更新后的最新数据。
old	旧数据，如果type参数是更新，则表示更新前的数据；如果是删除，则表示被删除的数据；如果是插入，取值为null。
pkNames	主键名称。

```
{
  "mysqlType":{
    "c11":"binary",
    "c10":"varchar",
    "c13":"text",
    "c12":"varbinary",
    "c14":"blob",
    "c1":"varchar",
    "c2":"varbinary",
    "c3":"int",
    "c4":"datetime",
    "c5":"timestamp",
    "c6":"char",
    "c7":"float",
    "c8":"double",
    "c9":"decimal",
    "id":"int"
  },
  "id":27677,
  "es":1624614713000,
  "ts":1625058726990,
  "database":"test01",
  "table":"test ",
  "type":"UPDATE",
  "isDdl":false,
  "sql":"",
  "sqlType":{
```

```
"c11":-2,
"c10":12,
"c13":-1,
"c12":-3,
"c14":2004,
"c1":12,
"c2":-3,
"c3":4,
"c4":94,
"c5":93,
"c6":1,
"c7":6,
"c8":8,
"c9":3,
"id":4
},
"data":[
  {
    "c11": "",
    "c10": "华为云huaweicloud",
    "c13": "asfiahfiaf939-0239uoituqorjqoirfojdjfqrniowejoiwjqroqwjrowqjoiqgoiegknkgoi23roiugou
ofdug9u90weurtg103",
    "c12": "[106, 103, 111, 106, 103, 111, 105, 100, 115, 106, 103, 111, 106, 111, 115, 111,
103, 57, 51, 52, 48, 57, 52, 51, 48, 57, 116, 106, 104, 114, 103, 106, 101, 119, 57, 116, 117, 48,
57, 51, 52, 48, 116, 101, 114, 111, 101, 106, 103, 57, 56, 51, 48, 52, 105, 101, 117, 114, 103, 57,
101, 119, 117, 114, 103, 48, 119, 101, 117, 116, 57, 114, 48, 52, 117, 48, 57, 53, 116, 117, 51, 48,
57, 50, 117, 116, 48, 57, 51, 117, 116, 48, 119, 57, 101]",
    "c14": "[106, 103, 111, 106, 103, 111, 105, 100, 115, 106, 103, 111, 106, 111, 115, 111,
103, 57, 51, 52, 48, 57, 52, 51, 48, 57, 116, 106, 104, 114, 103, 106, 101, 119, 57, 116, 117, 48,
57, 51, 52, 48, 116, 101, 114, 111, 101, 106, 103, 57, 56, 51, 48, 52, 105, 55, 57, 56, 52, 54, 53,
52, 54, 54, 54, 49, 52, 54, 53, 33, 64, 35, 36, 37, 94, 42, 40, 41, 95, 41, 43, 95, 43, 124, 125, 34,
63, 62, 58, 58, 101, 117, 114, 103, 57, 101, 119, 117, 114, 103, 48, 119, 101, 117, 116, 57, 114,
48, 52, 117, 48, 57, 53, 116, 117, 51, 48, 57, 50, 117, 116, 48, 57, 51, 117, 116, 48, 119, 57, 101]",
    "c1": "cf3f70a7-7565-44b0-ae3c-83bec549ea8e:104",
    "c2": "",
    "c3": "103",
    "c4": "2021-06-25 17:51:53",
    "c5": "1624614713.201",
    "c6": "!@#$$%90weurtg103",
    "c7": "10357.0",
    "c8": "1.2510357E7",
    "c9": "9874510357",
    "id": "104"
  }
],
"old":[
  {
    "c11": "",
    "c10": "华为云huaweicloud",
    "c13": "asfiahfiaf939-0239",
    "c12": "[106, 103, 111, 106, 103, 111, 105, 100, 115, 106, 103, 111, 106, 111, 115, 111,
103, 57, 51, 52, 48, 57, 52, 51, 48, 57, 116, 106, 104, 114, 103, 106, 101, 119, 57, 116, 117, 48,
57, 51, 52, 48, 116, 101, 114, 111, 101, 106, 103, 57, 56, 51, 48, 52, 105, 101, 117, 114, 103, 57,
101, 119, 117, 114, 103, 48, 119, 101, 117, 116, 57, 114, 48, 52, 117, 48, 57, 53, 116, 117, 51, 48,
57, 50, 117, 116, 48, 57, 51, 117, 116, 48, 119, 57, 101]",
    "c14": "[106, 103, 111, 106, 103, 111, 105, 100, 115, 106, 103, 111, 106, 111, 115, 111,
103, 57, 51, 52, 48, 57, 52, 51, 48, 57, 116, 106, 104, 114, 103, 106, 101, 119, 57, 116, 117, 48,
57, 51, 52, 48, 116, 101, 114, 111, 101, 106, 103, 57, 56, 51, 48, 52, 105, 55, 57, 56, 52, 54, 53,
52, 54, 54, 54, 49, 52, 54, 53, 33, 64, 35, 36, 37, 94, 42, 40, 41, 95, 41, 43, 95, 43, 124, 125, 34,
63, 62, 58, 58, 101, 117, 114, 103, 57, 101, 119, 117, 114, 103, 48, 119, 101, 117, 116, 57, 114,
48, 52, 117, 48, 57, 53, 116, 117, 51, 48, 57, 50, 117, 116, 48, 57, 51, 117, 116, 48, 119, 57, 101]"
  }
]
```

```

    "c1":"cf3f70a7-7565-44b0-ae3c-83bec549ea8e:104",
    "c2": "",
    "c3":"103",
    "c4":"2021-06-25 17:51:53",
    "c5":"1624614713.201",
    "c6":"!@#%$%90weurtg103",
    "c7":"10357.0",
    "c8":"1.2510357E7",
    "c9":"9874510357",
    "id":"103"
  }
],
"pkNames":[
  "id"
]
}

```

表 13-3 DDS 到 Kafka 的参数说明

参数名称	说明
id	DRS内部定义的事件操作的序列号，单调递增。
op	操作类型，比如DELETE，UPDATE，INSERT，DDL。
dbType	源库类型：Mongo。
db	数据库名称。
coll	集合名称。
value	这一条记录的变更值。
where	这一条记录的变更条件。
recordType	具体的记录类型，比如insert、update、replace、doc。其中，update和replace表示op中的UPDATE具体操作。doc表示op中的DELETE删除的是文档数据而非视图数据。
extra	拓展字段，一般和recordType保持一致，作为扩展oplog记录使用。
es	这一条记录的commit时间，13位Unix时间戳，单位为毫秒。
ts	写入到目标kafka的时间，13位Unix时间戳，单位为毫秒。
clusterTime	与事件关联的oplog条目的时间戳，格式为timestamp:inre（timestamp是10位unix时间戳，单位为秒；inre代表当前命令在同一秒内执行的次序）。

```

// insert操作
{
  "id": 256,
  "op": "INSERT",
  "dbType": "MongoDB",
  "db": "ljsx",
  "coll": "ljsx",
  "value": "{\"_id\": ObjectId(\"64650cf67dc36a464e76e583\"), \"c1\": \"baz\", \"tags\": [\"mongodb\", \"database\", \"NoSQL\"]}",

```

```

"where": null,
"recordType": "insert",
"extra": "insert",
"es": 1684315111439,
"ts": 1684315111576,
"clusterTime": "1684344064:1"
}

// replace操作
{
  "id": 340,
  "op": "UPDATE",
  "dbType": "MongoDB",
  "db": "ljx",
  "coll": "ljx",
  "value": "{\"_id\": ObjectId(\"64650cf67dc36a464e76e583\"), \"c1\": \"sss\"}",
  "where": "{\"_id\": ObjectId(\"64650cf67dc36a464e76e583\")}",
  "recordType": "replace",
  "extra": "replace",
  "es": 1684315951831,
  "ts": 1684315951961,
  "clusterTime": "1684344904:9"
}

// update 更新值操作
{
  "id": 386,
  "op": "UPDATE",
  "dbType": "MongoDB",
  "db": "ljx",
  "coll": "ljx",
  "value": "{\"$set\": {\"c1\": \"aaa\"}}",
  "where": "{\"_id\": ObjectId(\"64650cf67dc36a464e76e583\")}",
  "recordType": "update",
  "extra": "update",
  "es": 1684316412008,
  "ts": 1684316412146,
  "clusterTime": "1684345365:1"
}

// update 更新键操作
{
  "id": 414,
  "op": "UPDATE",
  "dbType": "MongoDB",
  "db": "ljx",
  "coll": "ljx",
  "value": "{\"$unset\": {\"c1\": true}, \"$set\": {\"column1\": \"aaa\"}}",
  "where": "{\"_id\": ObjectId(\"64650cf67dc36a464e76e583\")}",
  "recordType": "update",
  "extra": "update",
  "es": 1684316692054,
  "ts": 1684316692184,
  "clusterTime": "1684345648:1"
}

// remove 操作
{
  "id": 471,
  "op": "DELETE",
  "dbType": "MongoDB",
  "db": "ljx",

```

```

"coll": "lxx",
"value": "{\"_id\": ObjectId(\"64650cf67dc36a464e76e583\")}",
"where": null,
"recordType": "doc",
"extra": "doc",
"es": 1684317252747,
"ts": 1684317252869,
"clusterTime": "1684346209:1"
}

```

表 13-4 其他数据库到 Kafka 的参数说明

参数名称	说明
columnType	源端表字段名称和数据类型。 说明 <ul style="list-style-type: none"> 数据类型不带长度、精度等。 dbType为Oracle、Microsoft SQL Server时暂为空。
dbType	源库类型。
schema	schema名称。
opType	操作类型，比如DELETE，UPDATE，INSERT，DDL。
id	DRS内部定义的事件操作的序列号，单调递增。
es	源库不同引擎对应类型如下： GaussDB主备版：这一条记录上一个事务的commit时间，13位Unix时间戳，单位为毫秒。 GaussDB分布式：这一条记录上一个事务的commit时间，13位Unix时间戳，单位为毫秒。 PostgreSQL：这一条记录上一个事务的commit时间，13位Unix时间戳，单位为毫秒。 Oracle：这一条记录的commit时间，13位Unix时间戳，单位为毫秒。 Microsoft SQL Server：这一条记录的commit时间，13位Unix时间戳，单位为毫秒。
ts	写入到目标kafka的时间，13位Unix时间戳，单位为毫秒。
database	数据库名称，dbType为Oracle时暂时为空。
table	表名。
type	操作类型，比如DELETE，UPDATE，INSERT，DDL。
isDdl	是否是DDL操作。
sql	DDL的SQL语句，在DML操作中，取值为""。
sqlType	源端表字段的jdbc类型。
data	最新的数据，为JSON数组，如果type参数是插入则表示最新插入的数据，如果是更新，则表示更新后的最新数据。

参数名称	说明
old	旧数据，如果type参数是更新，则表示更新前的数据；如果是删除，则表示被删除的数据；如果是插入，取值为null。
pkNames	主键名称。

```
{
  "columnType": {
    "timestamp_column": "timestamp without time zone",
    "tstzrange_column": "tstzrange",
    "int4range_column": "int4range",
    "char_column": "character",
    "jsonb_column": "json",
    "boolean_column": "boolean",
    "bit_column": "bit",
    "smallint_column": "smallint",
    "bytea_column": "bytea"
  },
  "dbType": "GaussDB Primary/Standby",
  "schema": "schema01",
  "opType": "UPDATE",
  "id": 332,
  "es": 1639626187000,
  "ts": 1639629261915,
  "database": "database01",
  "table": "table01",
  "type": "UPDATE",
  "isDdl": false,
  "sql": "",
  "sqlType": {
    "timestamp_column": 16,
    "tstzrange_column": 46,
    "int4range_column": 42,
    "char_column": 9,
    "jsonb_column": 22,
    "boolean_column": 8,
    "bit_column": 20,
    "smallint_column": 2,
    "bytea_column": 15
  },
  "data": [
    {
      "timestamp_column": "2021-12-16 12:31:49.344365",
      "tstzrange_column": "(\"2010-01-01 14:30:00+08\", \"2010-01-01 15:30:00+08\")",
      "int4range_column": "[11,20)",
      "char_column": "g",
      "jsonb_column": "{\"key1\": \"value1\", \"key2\": \"value2\"}",
      "boolean_column": "false",
      "bit_column": "1",
      "smallint_column": "12",
      "bytea_column": "62797465615f64617461"
    }
  ],
  "old": [
    {
      "timestamp_column": "2014-07-02 06:14:00.742",
      "tstzrange_column": "(\"2010-01-01 14:30:00+08\", \"2010-01-01 15:30:00+08\")",
      "int4range_column": "[11,20)",

```



```

        "char_column": "g",
        "jsonb_column": "{\"key1\": \"value1\", \"key2\": \"value2\"}",
        "boolean_column": "true",
        "bit_column": "1",
        "smallint_column": "12",
        "bytea_column": "62797465615f64617461"
    }
],
    "pkNames": null
}

```

JSON-C 格式

JSON-C格式与JSON格式类似，区别是对于删除操作，JSON数据放在old上，JSON-C放在data上。对于timestamp类型数据转换成yyyy-mm-dd hh:mm:ss的字符串。

JSON-C定义详情参考[表13-5](#)：

表 13-5 JSON-C 参数说明

参数名称	说明
mysqlType	源端表字段名称和类型。
id	DRS内部定义的事件操作的序列号，单调递增。
es	源库产生这一条记录的时间，13位Unix时间戳，单位为毫秒。
ts	写入到目标kafka的时间，13位Unix时间戳，单位为毫秒。
database	数据库名称（Oracle数据库填写schema）。
table	表名。
type	操作类型，比如DELETE，UPDATE，INSERT，DDL。
isDdl	是否是DDL操作。
sql	DDL的SQL语句，在DML操作中，取值为""。
sqlType	源端表字段的jdbc类型。
data	最新的数据，为JSON数组，如果type参数是插入则表示最新插入的数据，如果是更新，则表示更新后的最新数据；如果是删除，则表示被删除的数据。
old	旧数据，如果type参数是更新，则表示更新前的数据；如果是插入，取值为null。
pkNames	主键名称。

JSON 格式数据中常见的转义字符

表 13-6 转义字符

字符	转义字符
<	\u003c
=	\u003d
>	\u003e
&	\u0026
'	\u0027

13.2 Kafka 认证方式

PLAINTEXT 方式

无安全认证方式，仅需输入IP和端口进行连接。

图 13-1 PLAINTEXT

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

SASL_PLAINTEXT 认证

使用SASL机制连接Kafka，需要设置SASL相关配置。

图 13-2 SASL_PLAINTEXT

目标库信息

IP地址或域名 ?

请确保所填写的多个IP地址或域名属于同一个实例

安全协议

SASL机制

用户名

密码

表 13-7 参数信息

参数	描述
SASL机制	用于客户端连接的SASL机制，支持以下四项，Kafka server默认是 GSSAPI 机制，更多说明可参考 SASL机制 。 <ul style="list-style-type: none"> • GSSAPI • PLAIN • SCRAM-SHA-256 • SCRAM-SHA-512
令牌委托	是否为委托令牌鉴权，SASL机制选择“SCRAM-SHA-256”或者“SCRAM-SHA-512”时可见。
用户名	登录使用的用户名。
密码	登录使用的密码

SSL 认证

使用SSL加密方式连接Kafka，需要设置SSL相关配置。

图 13-3 SSL

目标库信息

IP地址或域名 ?
请确保所填写的多个IP地址或域名属于同一个实例

安全协议

Truststore证书 选择文件

Truststore证书密码

主机名端点识别算法

SSL双向认证

Keystore证书 选择文件

Keystore证书密码

Keystore私钥密码

测试连接

表 13-8 参数信息

参数	描述
Truststore证书	后缀名为jks的SSL证书，证书文件生成可参考 SSL证书 。
Truststore证书密码	证书对应的密钥。
主机名端点识别算法	指定通过服务端证书验证服务端主机名的端点识别算法，选填，不填表示禁用主机名验证。
SSL双向认证	是否开启SSL双向认证。
Keystore证书	SSL双向认证开启可见，需要上传后缀名为jks的双向认证证书。
Keystore证书密码	SSL双向认证开启可见，SSL双向认证证书对应的密钥。
Keystore私钥密码	选填，Keystore证书中私钥的密码。

SASL_SSL 认证

使用SASL及SSL加密认证方式，需要设置SSL及SASL相关参数配置信息，可参考[SASL_PLAINTEXT认证](#)和[SSL认证](#)。

图 13-4 SASL_SSL

目标库信息

IP地址或域名 ?
请确保所填写的多个IP地址或域名属于同一个实例

安全协议

SASL机制

用户名

密码

Truststore证书

Truststore证书密码

主机名端点识别算法

SSL双向认证

Keystore证书

Keystore证书密码

Keystore私钥密码

13.3 GaussDB 分布式版为源强制结束任务

本小节介绍GaussDB分布式版为源的增量、全量+增量同步链路在强制结束任务后如何清理源库流复制槽。

复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务[同步日志](#)页面查看“task node id is ***”日志找到。

前提条件

由于普通用户没有执行execute direct操作的权限，删除流复制槽需要联系GaussDB运维人员执行以下操作。

操作步骤

步骤1 使用DRS同步任务测试连接时的用户登录GaussDB分布式版的每个主DN节点。

步骤2 执行如下语句，查询同步任务选择的database对象所对应的流复制槽名称。

```
select slot_name from pg_replication_slots where database = 'database';
```

须知

其中 *database* 为DRS同步任务中选择同步的database。

步骤3 执行如下语句，删除对应的流复制槽。

```
select * from pg_drop_replication_slot('slot_name');
```

须知

其中 *slot_name* 为**步骤2**中查询的流复制槽名称。

步骤4 执行如下语句，查询流复制槽是否成功删除

```
select slot_name from pg_replication_slots where database = 'database';
```

查询结果为空表示DRS同步任务对应的流复制槽已成功删除。

步骤5 重复执行以上操作，确保每个主DN都删除成功。

----结束

13.4 GaussDB 主备版为源强制结束任务

本小节介绍GaussDB主备版为源的同步链路在强制结束任务后如何清理源库流复制槽。

复制槽命名规则取决于数据库内核版本（可通过select working_version_num();命令查看），版本号低于92556默认使用串行解码，复制槽命名规则为“replication_slot_唯一性标识”；版本号大于等于92556默认使用并行解码，复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务node id中的“-”替换为“_”获取，node id可在DRS任务[同步日志](#)页面查看“task node id is ***”日志找到。

前提条件

由于普通用户没有执行execute direct操作的权限，删除流复制槽需要联系GaussDB运维人员执行以下操作。

操作步骤

步骤1 使用DRS同步任务测试连接时的用户登录GaussDB主备版实例。

步骤2 执行如下语句，查询同步任务选择的database对象所对应的流复制槽名称。

```
select slot_name from pg_replication_slots where database = 'database';
```

须知

其中 *database* 为 DRS 同步任务中选择同步的 database。

步骤3 执行如下语句，删除对应的流复制槽

```
select * from pg_drop_replication_slot('slot_name');
```

须知

其中 *slot_name* 为 **步骤2** 中查询的流复制槽名称。

步骤4 执行如下语句，查询流复制槽是否成功删除

```
select slot_name from pg_replication_slots where database = 'database';
```

查询结果为空表示 DRS 同步任务对应的流复制槽已成功删除。

----结束

13.5 PostgreSQL 为源强制结束任务

本小节介绍 PostgreSQL 为源的同步链路在强制结束任务后，如何清理源库逻辑复制槽、如何同步序列值以及如何源数据库已经无法连接的情况下，重置目标库中自增或自减列关联序列的序列值。

复制槽命名规则为“drs_唯一性标识”。其中唯一性标识需要将任务 node id 中的“-”替换为“_”获取，node id 可在 DRS 任务 [同步日志](#) 页面查看“task node id is ***”日志找到。

清理源库逻辑复制槽

步骤1 使用对应 DRS 任务的源数据库连接用户，登录该同步任务的源数据库。**步骤2** 查询同步任务选择的 database 对象所对应的流复制槽名称。

```
select slot_name from pg_replication_slots where database = 'database';
```

须知

其中 *database* 为 DRS 同步任务中选择同步的 database。

步骤3 执行如下语句，删除对应的流复制槽。

```
select * from pg_drop_replication_slot('slot_name');
```

须知

其中 *slot_name* 为 **步骤2** 中查询的流复制槽名称。

步骤4 执行如下语句，查询流复制槽是否成功删除。

```
select slot_name from pg_replication_slots where slot_name = 'slot_name';
```

查询结果为空表示DRS同步任务对应的流复制槽已成功删除。

---结束

同步序列值

如果未同步序列对象或者目标库为GaussDB，可忽略此节。

- 步骤1** 使用高权限账号（需要具有所有序列的USAGE权限）连接对应DRS任务同步的源数据库，执行如下语句。

```
select 'SELECT pg_catalog.setval('||quote_literal(quote_ident(n.nspname))||','||quote_ident(c.relname))||','||nextval(c.oid)||');' as sqls from pg_class c join pg_namespace n on c.relnamespace=n.oid where c.relkind = 'S' and n.nspname !~'^pg_' and n.nspname <>'information_schema' and not (c.relname='hwdrs_ddl_info_id_seq' and n.nspname='public') order by n.nspname, c.relname;
```

查询结果为需要在目标数据库中执行的sql语句。

- 步骤2** 使用对应DRS任务的目标数据库链接用户连接该任务同步的目标数据库，在目标库中执行步骤**步骤1**中查询出的sql语句。

- 步骤3** 在目标库中执行如下语句，检查序列值同步结果。

```
SELECT n.nspname, c.relname, nextval(c.oid) from pg_class c join pg_namespace n on c.relnamespace=n.oid where c.relkind = 'S' and n.nspname !~'^pg_' and n.nspname <>'information_schema' order by 1,2;
```

---结束

源库无法连接时，重置目标库中的序列值

在某些极端场景下，源数据库可能已经被损坏而无法连接，此时仍需将目标库中与自增或自减列相关联的序列值进行重置。如果源数据可以连接，请忽略此节，参考上节操作即可。

- 步骤1** 使用对应DRS任务的目标数据库测试链接用户，登录该同步任务的目标数据库。

- 步骤2** 使用如下语句，查询出将nextval作为表列默认值的序列对应的序列值重置sql语句。

```
set search_path to "; select 'SELECT pg_catalog.setval('||quote_literal(quote_ident(s.sequence_schema))||','||quote_ident(s.sequence_name))||','|| (SELECT '||case when s.increment::int<0 then 'min(' else 'max(' end||quote_ident(c.column_name))||')||case when s.increment::int<0 then '-1' else '+1' end||' FROM '||quote_ident(c.table_schema)||','||quote_ident(c.table_name)||');' as sqls from information_schema.columns c join information_schema.sequences s on (position(quote_literal(quote_ident(s.sequence_schema))||','||quote_ident(s.sequence_name))||':regclass' in c.column_default > 0) where c.data_type in ('bigint', 'int', 'integer', 'smallint', 'numeric', 'real', 'double precision', 'double') and c.column_default like 'nextval(%%)' order by s.sequence_schema, s.sequence_name;
```

查询结果为需要在目标数据库中执行的sql语句。

- 步骤3** 如果源库的版本小于10.0，请忽略此步。如果源库的版本不小于10.0，请在目标库中执行如下语句查询出重置表标识列附加序列的对应序列值的sql语句。

```
set search_path to "; select 'SELECT pg_catalog.setval('||quote_literal(seqname))||','|| (SELECT '||case when increment::int<0 then 'min(' else 'max(' end||colname||')||case when increment::int<0 then '-1' else '+1' end||' FROM '||tablename||');' as sqls from (select objid::regclass::text, refobjid::regclass::text, (pg_identify_object(refclassid,refobjid,refobjsubid)).identity, (pg_sequence_parameters(objid)).increment from pg_depend where deptype='i' and refobjsubid>0 and objid in (select c.oid from pg_class c join pg_namespace n on c.relnamespace=n.oid where c.relkind='S' and n.nspname !~ '^pg_' and n.nspname <>'information_schema')) p(seqname,tablename,colname,increment);
```

查询结果为需要在目标数据库中执行的sql语句。

- 步骤4** 在目标库中执行步骤**步骤2**和步骤**步骤3**中查询出的sql语句。

- 步骤5** 在目标库中执行如下语句，检查序列值同步结果。


```
SELECT n.nspname, c.relname, nextval(c.oid) from pg_class c join pg_namespace n on c.relnamespace=n.oid
where c.relkind = 'S' and n.nspname !~'^pg_' and n.nspname <> 'information_schema' order by 1,2;
```

---结束

13.6 通过创建事件触发器和函数实现 PostgreSQL 增量 DDL 同步

本小结介绍PostgreSQL->RDS for PostgreSQL实时同步，通过在源库创建事件触发器和函数获取源库的DDL信息，然后在DRS增量实时同步阶段实现DDL操作的同步。

前提条件

- 当前支持的DDL操作包含如下：
 - 表级同步支持：TRUNCATE（仅PostgreSQL 11及以上版本支持）、DROP TABLE、ALTER TABLE（包含ADD COLUMN、DROP COLUMN、ALTER COLUMN、RENAME COLUMN、ADD CONSTRAINT、DROP CONSTRAINT、RENAME）、COMMENT ON COLUMN、COMMENT ON TABLE。
 - 库级同步支持：TRUNCATE（仅PostgreSQL 11及以上版本支持）、CREATE SCHEMA/TABLE、DROP TABLE、ALTER TABLE（包含ADD COLUMN、DROP COLUMN、ALTER COLUMN、RENAME COLUMN、ADD CONSTRAINT、DROP CONSTRAINT、RENAME）、CREATE SEQUENCE、DROP SEQUENCE、ALTER SEQUENCE、CREATE INDEX、ALTER INDEX、DROP INDEX、CREATE VIEW、ALTER VIEW、COMMENT ON COLUMN、COMMENT ON TABLE、COMMENT ON SCHEMA、COMMENT ON SEQUENCE、COMMENT ON INDEX、COMMENT ON VIEW。

注意

- 表级同步：RENAME表名之后，向更改名称后的表插入新的数据时，DRS不会同步新的数据到目标库。
 - 库级同步：源库使用非CREATE TABLE方式创建的表不会同步到目标库。常见地如：使用CREATE TABLE AS创建表、调用函数创建表。
 - 暂不支持以注释开头的DDL语句的同步，以注释开头的DDL语句将被忽略。
 - 不支持函数和存储过程中DDL语句的同步，函数和存储过程中执行的DDL语句将被忽略。
-
- 源库和目标库版本不同时，请使用源库和目标库都兼容的SQL语句执行DDL操作。例如：源库为pg11，目标库为pg12，要将源库表的列类型从char修改为int时，请使用如下语句：

```
alter table tablename alter column columnname type int USING columnname::int;
```
 - 执行如下操作步骤前，请检查待同步的源数据库public模式下，是否存在名为hwdrs_ddl_info的表、名为hwdrs_ddl_function()的函数、名为hwdrs_ddl_event的触发器。如存在，请将其删除。
 - 库级同步时，如创建无主键表，请执行如下命令，将无主键表复制属性设置为full。

```
alter table tablename replica identity full;
```

操作步骤

- 如果源库为其他云上或自建PostgreSQL，执行以下步骤：

- a. 使用拥有创建事件触发器权限的用户连接要同步的数据库。
- b. 执行如下语句，创建存储DDL信息的表。

```
DROP TABLE IF EXISTS public.hwdrs_ddl_info;
DROP SEQUENCE IF EXISTS public.hwdrs_ddl_info_id_seq;
CREATE TABLE public.hwdrs_ddl_info(
  id          bigserial primary key,
  ddl         text,
  username    varchar(64) default current_user,
  txid        varchar(16) default txid_current()::varchar(16),
  tag         varchar(64),
  database    varchar(64) default current_database(),
  schema      varchar(64) default current_schema,
  client_address varchar(64) default inet_client_addr(),
  client_port integer default inet_client_port(),
  event_time  timestamp default current_timestamp
);
```

- c. 执行如下语句，创建函数。

```
CREATE OR REPLACE FUNCTION public.hwdrs_ddl_function()
  RETURNS event_trigger
  LANGUAGE plpgsql
  SECURITY INVOKER
AS $$
  declare ddl text;
  declare real_num int;
  declare max_num int := 50000;
begin
  if (tg_tag in ('CREATE TABLE','ALTER TABLE','DROP TABLE','CREATE SCHEMA','CREATE SEQUENCE','ALTER SEQUENCE','DROP SEQUENCE','CREATE VIEW','ALTER VIEW','DROP VIEW','CREATE INDEX','ALTER INDEX','DROP INDEX','COMMENT')) then
    select current_query() into ddl;
    insert into public.hwdrs_ddl_info(ddl, username, txid, tag, database, schema, client_address, client_port, event_time)
      values (ddl, current_user, cast(txid_current() as varchar(16)), tg_tag, current_database(), current_schema, inet_client_addr(), inet_client_port(), current_timestamp);
    select count(id) into real_num from public.hwdrs_ddl_info;
    if real_num > max_num then
      if current_setting('server_version_num')::int<100000 then
        delete from public.hwdrs_ddl_info where id<(select min(id)+1000 from public.hwdrs_ddl_info) and not exists (select 0 from pg_locks l join pg_database d on l.database=d.oid where d.datname=current_catalog and pid<>pg_backend_pid() and locktype='relation' and relation=to_regclass('public.hwdrs_ddl_info_pkey')::oid and mode='RowExclusiveLock');
      else
        delete from public.hwdrs_ddl_info where id<(select min(id)+1000 from public.hwdrs_ddl_info) and (xmax=0 or coalesce(txid_status(xmax::text::bigint, '')<>'in progress'));
      end if;
    end if;
  end if;
end;
$$;
```

- d. 执行以下语句，为**b**和**c**中创建的对象赋予必要权限。

```
GRANT USAGE ON SCHEMA public TO public;
GRANT SELECT,INSERT,DELETE ON TABLE public.hwdrs_ddl_info TO public;
GRANT SELECT,USAGE ON SEQUENCE public.hwdrs_ddl_info_id_seq TO public;
GRANT EXECUTE ON FUNCTION public.hwdrs_ddl_function() TO public;
```

- e. 执行以下语句，创建DDL事件触发器。

```
CREATE EVENT TRIGGER hwdrs_ddl_event ON ddl_command_end EXECUTE PROCEDURE public.hwdrs_ddl_function();
```

- f. 执行以下语句，将创建的事件触发器设置为enable。

```
ALTER EVENT TRIGGER hwdrs_ddl_event ENABLE ALWAYS;
```

- g. 返回数据复制服务控制台，创建PostgreSQL->RDS for PostgreSQL的同步任务。
- h. 待同步任务结束后，请执行下语句删除创建的表、函数、触发器。

```
DROP EVENT trigger hwdrs_ddl_event;
DROP FUNCTION public.hwdrs_ddl_function();
DROP TABLE public.hwdrs_ddl_info;
```
- 如果源库为RDS for PostgreSQL，执行以下步骤：
 - a. 执行如下语句，清理已经创建过的对象。

```
DROP EVENT TRIGGER IF EXISTS hwdrs_ddl_event;
DROP FUNCTION IF EXISTS public.hwdrs_ddl_function();
DROP TABLE IF EXISTS public.hwdrs_ddl_info;
```
 - b. 使用root用户执行如下语句，创建DDL插件。

```
SELECT CONTROL_EXTENSION('create', 'rds_hwdrs_ddl');
```
 - c. 执行如下语句，更新函数。

```
CREATE OR REPLACE FUNCTION public.hwdrs_ddl_function()
  RETURNS event_trigger
  LANGUAGE plpgsql
  SECURITY INVOKER
AS $BODY$
  declare ddl text;
  declare real_num int;
  declare max_num int := 50000;
begin
  if (tg_tag in ('CREATE TABLE','ALTER TABLE','DROP TABLE','CREATE SCHEMA','CREATE SEQUENCE','ALTER SEQUENCE','DROP SEQUENCE','CREATE VIEW','ALTER VIEW','DROP VIEW','CREATE INDEX','ALTER INDEX','DROP INDEX','COMMENT')) then
    select current_query() into ddl;
    insert into public.hwdrs_ddl_info(ddl, username, txid, tag, database, schema, client_address, client_port, event_time)
      values (ddl, current_user, cast(txid_current() as varchar(16)), tg_tag, current_database(), current_schema, inet_client_addr(), inet_client_port(), current_timestamp);
    select count(id) into real_num from public.hwdrs_ddl_info;
    if real_num > max_num then
      if current_setting('server_version_num')::int<100000 then
        delete from public.hwdrs_ddl_info where id<(select min(id)+1000 from public.hwdrs_ddl_info) and not exists (select 0 from pg_locks l join pg_database d on l.database=d.oid where d.datname=current_catalog and pid<>pg_backend_pid() and locktype='relation' and relation=to_regclass('public.hwdrs_ddl_info_pkey')::oid and mode='RowExclusiveLock');
      else
        delete from public.hwdrs_ddl_info where id<(select min(id)+1000 from public.hwdrs_ddl_info) and (xmax=0 or coalesce(txid_status(xmax::text::bigint), '')<>'in progress');
      end if;
    end if;
  end if;
end;
$BODY$;
```
 - d. 返回数据复制服务控制台，创建PostgreSQL->RDS for PostgreSQL的同步任务。
 - e. 待同步任务结束后，请执行下语句删除创建的表、函数、触发器。

```
SELECT CONTROL_EXTENSION('drop', 'rds_hwdrs_ddl');
```

14 附录

14.1 通用使用建议

在使用DRS前，您需要了解：

- DRS支持从数据库同版本或低版本到高版本的同步，不支持从高版本同步到低版本。
- 数据库同步与环境多样性和人为操作均有密切关系，为了确保同步的平顺，建议您在进行正式的数据库同步之前进行一次演练，可以帮助您提前发现问题并解决问题。
- 如果您使用的是全量同步模式，确保源和目标数据库无业务写入，保证同步前后数据一致。如果您使用的是全量+增量同步模式，支持在源数据库有业务数据写入的情况下进行同步。
- DRS任务启动和全量数据同步阶段，请不要在源数据库执行DDL操作，否则可能导致任务异常。
- 全量阶段读取表数据时，特别是大表的读取，可能会阻塞业务上对大表的独占锁操作。
- 同步无主键表时，为了确保数据一致性，会存在3s以内的单表级锁定。
- 正在同步的数据被其他事务长时间锁死，可能导致读数据超时。
- 由于MySQL固有特点限制，CPU资源紧张时，存储引擎为Tokudb的表，读取速度可能下降至10%。
- DRS并发读取数据库时，会占用大约6-10个session连接数，需要考虑该连接数对业务的影响。
- 在网络无瓶颈的情况下，全量同步会对源数据库增加约50MB/s的查询压力，以及占用2~4个CPU。
- 更多DRS对数据库的影响，可参考[DRS对源数据库和目标数据库有什么影响](#)。
- 建议您在启动任务时选择“稍后启动”功能，将启动时间设置在业务低峰期，相对静止的数据可以有效提升一次性同步成功率，避免同步对业务造成性能影响。如果同步不可避免业务高峰期，推荐使用同步限速功能，即“流速模式”选择“限速”。
- 数据对比

建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

- **如果涉及多对一同步场景，您需要了解：**
 - 如果涉及多对一场景的同步任务，可参考[多对一的场景约束及操作建议](#)。
 - 如果涉及表级汇集的多对一同步任务，则不支持DDL，否则会导致同步全部失败。
- **数据同步完成后，您需要了解**
 - 同步完成后，建议您结合[数据对比](#)的“稍后启动”功能，选择业务低峰期进行数据对比，以便得到更为具有参考性的对比结果。由于同步具有轻微的时差，在数据持续操作过程中进行对比任务，可能会出现少量数据不一致对比结果，从而失去参考意义。

14.2 DRS 预检查项一览表

DRS同步过程一般包含四个阶段：任务启动阶段、全量阶段、增量阶段、结束任务阶段（单全量任务则包含三个阶段）。为了确保同步各个阶段的平顺，DRS在任务正式启动前，会对任务源数据库、目标数据库中的参数、对象等信息进行自动化的扫描和检测，帮助您提高数据同步的成功率。

概览

根据同步任务的源数据库类型，查看预检查信息：

- 源数据库为MySQL的同步链路预检查一览表：
 - [MySQL->MySQL](#)
 - [MySQL->PostgreSQL](#)
 - [MySQL->MariaDB](#)

MySQL->MySQL

表 14-1 预检查一览表

分类	预检查项	检查项详情	不通过解决方法
权限类	源数据库权限	<ul style="list-style-type: none"> ● 全量同步需要具备如下最小权限： SELECT、SHOW VIEW、EVENT。 ● 全量+增量、增量同步需要具备如下最小权限： SELECT、SHOW VIEW、EVENT、 LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 	详见 MySQL权限授权方法 。

分类	预检查项	检查项详情	不通过解决方法
	目标数据库权限	<ul style="list-style-type: none"> 提供的目标数据库账号必须拥有如下权限： SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES。 RDS for MySQL实例的root账户默认已具备上述权限。 当目标库为8.0.14-8.0.18版本时，还需要有SESSION_VARIABLES_ADMIN权限。 	详见 MySQL权限授权方法 。
版本类	源数据库版本	支持5.5、5.6、5.7、8.0版本。	详见 支持的数据库 。
	目标数据库版本	支持5.5、5.6、5.7、8.0版本。	详见 支持的数据库 。
	同步版本检查	仅支持目标数据库版本等于或高于源数据库版本。	详见 版本是否符合从低到高或者同版本迁移 。
参数类	GTID状态	源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。	详见 源库的GTID功能检查 。
	性能参数	源数据库log_slave_updates参数需设置为开启状态，否则会导致同步失败。	详见 校验源数据库参数log_slave_updates 。
		源数据库的binlog_row_image参数需设置为full，否则会导致同步失败。	详见 binlog_row_image参数是否为full 。
	最大允许传输包的大小	DRS在同步数据量大或同步大字段情况下，源数据库max_allowed_packet参数过小可能会导致任务失败。	详见 检查源库的max_allowed_packet参数 。
		DRS在同步数据量大或同步大字段情况下，目标数据库的max_allowed_packet参数值过小导致目标库数据无法写入造成全量同步失败。	详见 检查目标库的max_allowed_packet参数 。
sql_mode取值检查	同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。	详见 源数据库是否存在不允许使用的sql_mode值 。	
增量同步类	索引列长度检查	源数据库索引列长度必须符合要求。	详见 源数据库索引列长度检查 。

分类	预检查项	检查项详情	不通过解决方法
	Binlog开启	增量同步时，源数据库的Binlog日志必须打开。	详见 源数据库binlog日志是否开启 。
	Binlog格式	源数据库的Binlog格式必须为行格式。	详见 源数据库binlog格式检查 。
	Binlog保留时长	在磁盘空间允许的情况下，建议源数据库Binlog保存时间越长越好，建议为3天，设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> 源数据库为自建MySQL时，通过设置expire_logs_days参数设置Binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的Binlog尚未过期，以保证任务中断后的顺利恢复。 源数据库为RDS for MySQL时，设置Binlog保留时间可参考设置RDS for MySQL本地Binlog日志清理。 	详见 源数据库binlog保留时间检查 。
	server_id值设置	增量同步时，必须设置MySQL源数据库的server_id。 <ul style="list-style-type: none"> 如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间。 如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 	详见 源数据库参数server_id是否符合增量迁移要求 。
	表字段检查	同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。	-
	附加列检查	单增量同步任务中如果目标数据库不存在新增附加列，会导致任务失败。	详见 目标数据库附加列检查 。
目标数据库检查	磁盘空间检查	目标数据库实例必须有足够的磁盘空间。	详见 目标数据库磁盘可用空间是否足够 。
	状态检查	目标数据库实例实例状态必须正常。	-
一致性检查	字符集	目标库和源库的字符集需要一致。	详见 源数据库和目标数据库的字符集是否一致 。
	字符序	目标库和源库的collation_server需要一致。	详见 collation_server的一致性检查 。
	时钟	目标库和源库的时钟需要一致。	-

分类	预检查项	检查项详情	不通过解决方法
	时区	目标库和源库的time_zone需要一致。	详见 time_zone的一致性检查 。
	大小写敏感	目标库和源库的lower_case_table_names参数设置需要一致。	详见 源数据库和目标数据库表名大小写敏感性检查 。
	group_concat函数计算结果字符的最大长度	目标库和源库的group_concat_max_len参数需要一致。	详见 数据库参数group_concat_max_len一致性检查 。
	InnoDB检查模式	目标库和源库innodb_strict_mode需要一致。	详见 数据库参数innodb_strict_mode一致性检查 。
	数据块加密参数	目标库和源库block_encryption_mode需要一致。	-
	SQL模式	目标库和源库sql_mode需要一致。	详见 数据库参数sql_mode的一致性检查 。
同步对象类	选择对象检查	<ul style="list-style-type: none"> 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 仅支持MyISAM和InnoDB表的同步。 不支持事件（EVENT）和触发器（TRIGGER）的同步。 不支持数据库参数和系统数据库同步。 	-
		库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。	详见 映射数据库对象检查 。
	无主键表检查	由于无主键表的性能低于主键表的性能，建议将无主键表修改为主键表。	详见 源库无主键表检查 。
	关联对象检查	相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。	详见 迁移对象未选择外键依赖的表 。
	外键引用操作检查	不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。	详见 源端存在不支持的外键引用操作 。

分类	预检查项	检查项详情	不通过解决方法
	同名检查	除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。	详见 目标数据库是否存在和源数据库同名的数据库 。
	库表名规范检查	<ul style="list-style-type: none"> 源数据库中的库名不能包含：'<>\'以及非ASCII字符。 源数据库中的表名、视图名不能包含：'<>\'以及非ASCII字符。 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“i”。 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 	-
SSL连接检查	SSL安全连接检查	选择SSL安全连接时，源数据库和目标数据库的SSL安全连接成功。	详见 SSL安全连接检查 。
网络情况	网络检查	目标数据库的IP，端口，用户名和密码连接成功。	详见 目标数据库连接是否成功 。
		源数据库的IP，端口，用户名和密码连接成功。	详见 源数据库连接是否成功 。

MySQL->PostgreSQL

表 14-2 预检查一览表

分类	预检查项	检查项详情	不通过解决方法
权限类	源数据库权限	<ul style="list-style-type: none"> 全量同步需要具备如下最小权限：SELECT。 全量+增量同步需要具备如下最小权限：SELECT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 	详见 MySQL权限授权方法 。
	目标数据库权限	RDS for PostgreSQL实例的账号默认已具有权限。	详见 MySQL权限授权方法 。

分类	预检查项	检查项详情	不通过解决方法
版本类	源数据库版本	支持5.5、5.6、5.7、8.0版本。	详见 支持的数据库 。
	目标数据库版本	支持9.5、9.6、10、11版本。	详见 支持的数据库 。
参数类	GTID状态	源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。	详见 源库的GTID功能检查 。
	性能参数	源数据库的binlog_row_image参数需设置为full，否则会导致同步失败。	详见 binlog_row_image参数是否为full 。
	最大允许传输包的大小	DRS在同步数据量大或同步大字段情况下，源数据库max_allowed_packet参数过小可能会导致任务失败。	详见 检查源库的max_allowed_packet参数 。
增量同步类	Binlog开启	增量同步时，源数据库的Binlog日志必须打开。	详见 源数据库binlog日志是否开启 。
	Binlog格式	源数据库的Binlog格式必须为行格式。	详见 源数据库binlog格式检查 。
	server_id值设置	增量同步时，必须设置MySQL源数据库的server_id。 <ul style="list-style-type: none"> 如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间。 如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 	详见 源数据库参数server_id是否符合增量迁移要求 。
	表字段检查	<ul style="list-style-type: none"> 同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。 源数据库中不支持的表字段类型有：xml、geometry、point、lineString、polygon、geometrycollection、multipoint、multilinestring、multipolygon、json。 	-
目标数据库检查	磁盘空间检查	目标数据库实例必须有足够的磁盘空间。	详见 目标数据库磁盘可用空间是否足够 。
	状态检查	目标数据库实例实例状态必须正常。	-

分类	预检查项	检查项详情	不通过解决方法
一致性检查	字符集	目标库和源库的字符集需要一致。	详见 源数据库和目标数据库的字符集是否一致 。
	时区	目标库和源库的time_zone需要一致。	详见 time_zone的一致性检查 。
	表结构检查	目标库和源库的表结构需要一致。	详见 表结构一致性检查 。
同步对象类	选择对象检查	<ul style="list-style-type: none"> 仅支持同步表结构、表数据、索引信息。 仅支持同步MyISAM和InnoDB表。 不支持同步存储过程等其他数据库对象。 	-
	无主键表检查	由于无主键表的性能低于主键表的性能，建议将无主键表修改为主键表。	详见 源库无主键表检查 。
	有主键表正确性检查	不建议在数据库中使用非精确数值类型做主键，该特性影响 DRS 增量场景下对 UPDATE、DELETE语句的同步，导致任务失败。	-
	外键引用操作检查	不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。	详见 源端存在不支持的外键引用操作 。
	同名检查	除了MySQL系统数据库之外，目标数据库不能包含与源数据库同名的数据库。	详见 目标数据库是否存在和源数据库同名的数据库 。
	库表名规范检查	源数据库中的库名和表名不能包含： '<'>/'\''以及非ASCII字符。	-
	对象名长度检查	目标数据库对象名最大支持63个字符。	详见 源数据库对象名长度检查 。
	源数据库检查	预检查阶段进行同步对象检查时，源数据库连接正常。	详见 源库选择对象预检查 。
	SSL连接检查	SSL安全连接检查	选择SSL安全连接时，源数据库和目标数据库的SSL安全连接成功。
网络情况	网络检查	目标数据库的IP，端口，用户名和密码连接成功。	详见 目标数据库连接是否成功 。
		源数据库的IP，端口，用户名和密码连接成功。	详见 源数据库连接是否成功 。

MySQL->MariaDB

表 14-3 预检查一览表

分类	预检查项	检查项详情	不通过解决方法
权限类	源数据库权限	<ul style="list-style-type: none"> 全量同步需要具备如下最小权限：SELECT、SHOW VIEW、EVENT。 全量+增量、增量同步需要具备如下最小权限：SELECT、SHOW VIEW、EVENT、LOCK TABLES、REPLICATION SLAVE、REPLICATION CLIENT。 	详见 MySQL权限授权方法 。
	目标数据库权限	<ul style="list-style-type: none"> 提供的目标数据库账号必须拥有如下权限：SELECT、CREATE、DROP、DELETE、INSERT、UPDATE、ALTER、CREATE VIEW、CREATE ROUTINE、REFERENCES、INDEX。 RDS for MariaDB实例的root账户默认已具备上述权限。 	详见 MySQL权限授权方法 。
版本类	源数据库版本	支持5.5、5.6、5.7、8.0版本。	详见 支持的数据库 。
	目标数据库版本	支持10.5版本。	详见 支持的数据库 。
参数类	GTID状态	源数据库GTID状态建议为开启状态，源数据库实例没有开启GTID的情况下不支持主备倒换，DRS任务会因为位点不续接而中断导致无法恢复。	详见 源库的GTID功能检查 。
	性能参数	源数据库log_slave_updates参数需设置为开启状态，否则会导致同步失败。	详见 校验源数据库参数log_slave_updates 。
		源数据库的binlog_row_image参数需设置为full，否则会导致同步失败。	详见 binlog_row_image参数是否为full 。
	sql_mode取值检查	同步的对象中包含引擎为MyISAM的表，则目标数据库sql_mode不能包含no_engine_substitution参数，否则可能会导致同步失败。	详见 源数据库是否存在不允许使用的sql_mode值 。
增量同步类	索引列长度检查	源数据库索引列长度必须符合要求。	详见 源数据库索引列长度检查 。
	Binlog开启	增量同步时，源数据库的Binlog日志必须打开。	详见 源数据库binlog日志是否开启 。

分类	预检查项	检查项详情	不通过解决方法
	Binlog格式	源数据库的Binlog格式必须为行格式。	详见 源数据库binlog格式检查 。
	Binlog保留时长	在磁盘空间允许的情况下，建议源数据库Binlog保存时间越长越好，建议为3天，设置为0，可能会导致同步失败。 <ul style="list-style-type: none"> 源数据库为自建MySQL时，通过设置expire_logs_days参数设置Binlog保留时间。建议将expire_logs_day参数设置在合理的范围，确保恢复时断点处的Binlog尚未过期，以保证任务中断后的顺利恢复。 源数据库为RDS for MySQL时，设置Binlog保留时间可参考设置RDS for MySQL本地Binlog日志清理。 	详见 源数据库binlog保留时间检查 。
	server_id值设置	增量同步时，必须设置MySQL源数据库的server_id。 <ul style="list-style-type: none"> 如果源数据库版本小于或等于MySQL5.6，server_id的取值范围在2 - 4294967296之间。 如果源数据库版本大于或等于MySQL5.7，server_id的取值范围在1 - 4294967296之间。 	详见 源数据库参数server_id是否符合增量迁移要求 。
	表字段检查	同步对象中如果存在包含longtext、longblob类型字段的表，建议创建大规格及以上规格的DRS任务进行同步，否则可能会导致capture OOM。	-
	附加列检查	单增量同步任务中如果目标数据库不存在新增附加列，会导致任务失败。	详见 目标数据库附加列检查 。
	目标数据库检查	磁盘空间检查	目标数据库实例必须有足够的磁盘空间。
状态检查		目标数据库实例实例状态必须正常。	-
一致性检查	字符集	目标库和源库的字符集需要一致。	详见 源数据库和目标数据库的字符集是否一致 。
	字符序	目标库和源库的collation_server需要一致。	详见 collation_server的一致性检查 。
	时钟	目标库和源库的时钟需要一致。	-
	大小写敏感	目标库和源库的lower_case_table_names参数设置需要一致。	详见 源数据库和目标数据库表名大小写敏感性检查 。

分类	预检查项	检查项详情	不通过解决方法
	group_concat函数计算结果字符的最大长度	目标库和源库的group_concat_max_len参数需要一致。	详见 数据库参数group_concat_max_len一致性检查 。
	InnoDB检查模式	目标库和源库innodb_strict_mode需要一致。	详见 数据库参数innodb_strict_mode一致性检查 。
	SQL模式	目标库和源库sql_mode需要一致。	详见 数据库参数sql_mode的一致性检查 。
同步对象类	选择对象检查	<ul style="list-style-type: none"> 支持表、主键索引、唯一索引、普通索引、存储过程、视图、函数的同步。 仅支持MyISAM和InnoDB表的同步。 不支持事件（EVENT）和触发器（TRIGGER）的同步。 不支持数据库参数和系统数据库同步。 	-
		库名映射时，同步的对象中如果存在存储过程、视图、函数对象，全量阶段这些对象不会同步，会导致对象对比不一致。	详见 映射数据库对象检查 。
	无主键表检查	由于无主键表的性能低于主键表的性能，建议将无主键表修改为主键表。	详见 源库无主键表检查 。
	关联对象检查	相互关联的数据对象要确保同时同步，避免因关联对象缺失，导致同步失败。常见的关联关系：视图引用表、视图引用视图、存储过程/函数/触发器引用视图/表、主外键关联表等。	详见 迁移对象未选择外键依赖的表 。
	外键引用操作检查	不支持外键级联操作。当外键是普通索引的时候，可能会导致表结构创建失败，建议改成唯一索引。	详见 源端存在不支持的外键引用操作 。
	同名检查	除了MySQL系统数据库之外，当目标库和源库同名时，目标数据库中若存在与源库同名的表，则表结构必须与源库保持一致。	详见 目标数据库是否存在和源数据库同名的数据库 。

分类	预检查项	检查项详情	不通过解决方法
	库表名规范检查	<ul style="list-style-type: none"> 源数据库中的库名不能包含：'<>\'以及非ASCII字符。 源数据库中的表名、视图名不能包含：'<>\'以及非ASCII字符。 源数据库中的库名或映射后的名称不允许以ib_logfile开头，也不能为ib_buffer_pool、ib_doublewrite、ibdata1、ibtmp1。 映射到目标库中的库名不能包含：“.”、“<”、“>”、“”、和“i”。 使用非字母、非数字、非下划线等字符作为库表名称，或库表映射后名称包含中划线、井号时，名称的长度请不要超过42个字符。 	-
SSL连接检查	SSL安全连接检查	选择SSL安全连接时，源数据库和目标数据库的SSL安全连接成功。	详见 SSL安全连接检查 。
网络情况	网络检查	目标数据库的IP，端口，用户名和密码连接成功。	详见 目标数据库连接是否成功 。
		源数据库的IP，端口，用户名和密码连接成功。	详见 源数据库连接是否成功 。

A 修订记录

发布日期	修改说明
2024-02-28	<p>第五十七次正式发布，新增如下内容：</p> <ul style="list-style-type: none">• Microsoft SQL Server->GaussDB的实时同步支持单全量、单增量模式。• GaussDB主备版->Oracle的实时同步支持选择DDL同步类型。• GaussDB主备为源的自建链路支持源数据库为容灾集群。• Oracle为源链路支持选择分片方式。• Oracle->GaussDB内容对比支持对LOB字段的长度进行比对。
2023-12-30	<p>第五十六就次正式发布，新增如下内容：</p> <ul style="list-style-type: none">• DRS实时同步支持修改任务参数。• DRS支持MySQL->MariaDB的实时同步。• DRS支持MariaDB->MySQL的实时同步。• DRS支持MariaDB->GaussDB(for MySQL)的实时同步。• PostgreSQL->GaussDB的实时同步支持文件导入、限速、暂停和重置。

发布日期	修改说明
2023-11-30	<p>第五十五次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS实时同步支持超大规格。 • DRS实时同步支持查看进度明细。 • DRS实时同步支持PostgreSQL 15版本。 • DRS支持Microsoft SQL Server->Kafka的实时同步。 • DRS支持MySQL兼容模式的OceanBase->GaussDB(for MySQL)的实时同步。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> • 以下链路转为商用支持标准。 <ul style="list-style-type: none"> - MySQL->GaussDB主备版的实时同步。 - MySQL->GaussDB分布式版的实时同步。 - DDS->Kafka的实时同步。 - GaussDB主备版->MySQL的实时同步。 - GaussDB主备版->GaussDB分布式版的实时同步。 - GaussDB分布式版->MySQL的实时同步 - GaussDB分布式版->GaussDB主备版的实时同步。
2023-10-30	<p>第五十四次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DB2 for LUW为源的实时同步增量支持分区表。 • DB2 for LUW为源的实时同步测试连接时需手动上传对应的JDBC驱动包。 • DRS多规格任务支持升级实例规格。
2023-09-30	<p>第五十三次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL->GaussDB(DWS)支持重置同步位点。
2023-08-30	<p>第五十二次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL->MySQL增量DDL同步支持自定义设置DROP PARTITION、RENAME COLUMN。 • PostgreSQL->PostgreSQL增量DDL同步支持COMMENT ON。 • DRS任务列表支持按照数据库实例ID、数据库IP进行筛选过滤。
2023-07-30	<p>第五十一次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS同步任务支持可用区的选择。

发布日期	修改说明
2023-06-30	第五十次正式发布，新增如下内容： <ul style="list-style-type: none"> ● DDM->MySQL、DDM->Oracle的实时同步支持单全量、单增量模式。 ● MySQL->GaussDB的实时同步支持同步对象类型的选择。 ● GaussDB->GaussDB的实时同步支持对象名映射。 ● GaussDB->Kafka的实时同步支持库级和schema级同步。 ● DB2 for LUW->GaussDB的实时同步支持数据过滤。 ● DRS实时同步支持Oracle 21c版本。
2023-05-30	第四十九次正式发布，新增如下内容： <ul style="list-style-type: none"> ● DRS支持通过导入方式创建同步任务。 ● DRS支持DDS->Kafka的出云同步。 ● DRS支持GaussDB(DWS) 8.2.0版本。 ● Oracle->GaussDB支持选择DDL同步类型。

发布日期	修改说明
2023-04-30	<p>第四十八次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● DRS支持DDS 4.2版本的同步。 ● DRS支持MySQL->Kafka单机类型同步任务升级实例规格。 ● MySQL->Kafka支持同步对象类型选择。 ● GaussDB(for MySQL)->GaussDB(for MySQL)的实时同步支持从备库读取进行全量数据同步。 ● MySQL->MySQL的入云实时同步支持单全量模式。 ● Oracle->MySQL、Oracle->GaussDB(DWS)的实时同步支持单增量模式。 ● GaussDB分布式版->MySQL的实时同步支持全量、全量+增量模式。 ● Oracle->PostgreSQL的入云实时同步支持PostgreSQL 13、14版本。 ● DRS测试连接失败时，支持一键网络诊断功能。 ● 支持MariaDB的入云和出云实时同步。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● 以下链路转为商用支持标准。 <ul style="list-style-type: none"> - Oracle->GaussDB主备版的实时同步。 - Oracle->GaussDB分布式版的实时同步。 - GaussDB(for MySQL)->GaussDB(for MySQL)的实时同步。 - GaussDB主备版->GaussDB主备版的实时同步。 - GaussDB分布式版->GaussDB分布式版的实时同步。 - GaussDB(for MySQL)->GaussDB(for MySQL)的实时同步。 ● 以下链路由白名单转为公测支持标准。 <ul style="list-style-type: none"> - MySQL->GaussDB主备的实时同步。 - DB2 for LUW->GaussDB主备的实时同步。 - DB2 for LUW->GaussDB分布式的实时同步。 - Microsoft SQL Server->GaussDB主备的实时同步。 - Microsoft SQL Server->GaussDB分布式的实时同步。 - Microsoft SQL Server->GaussDB(DWS)的实时同步。
2023-03-30	<p>第四十七次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● DRS支持GaussDB(for MySQL)->GaussDB(for MySQL)的入云同步。 ● DRS支持PostgreSQL->PostgreSQL的出云同步。 ● GaussDB(for MySQL)->CSS/ES同步任务支持选择任务类型。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● DRS创建任务界面，任务类型由“单机”、“主备”修改为“单AZ”、“双AZ”。

发布日期	修改说明
2023-02-28	第四十六次正式发布，新增如下内容： <ul style="list-style-type: none"> ● 行对比结果支持按照“源库表行数”、“目标库表行数”升序或者降序进行排序。 ● DRS支持MySQL->GaussDB(DWS)单机类型同步任务升级实例规格。
2023-01-30	第四十五次正式发布，新增如下内容： <ul style="list-style-type: none"> ● PostgreSQL为源的实时同步支持PostgreSQL 14版本。 ● DB2 for LUW为源同步链路支持DB2SECURITYLABEL类型。 ● MySQL->CSS/ES、GaussDB(for MySQL)->CSS/ES实时同步支持ElasticSearch 7.10版本。 ● MySQL->Kafka、GaussDB(for MySQL)->Kafka的单增量同步任务支持指定启动位点。 ● 支持Microsoft SQL Server->Microsoft SQL Server的入云实时同步。
2022-12-30	第四十四次正式发布，新增如下内容： <ul style="list-style-type: none"> ● MySQL->CSS/ES实时同步支持多规格。 ● Oracle->Kafka同步支持设置Topic、partition策略。 ● GaussDB分布式自建为源支持指定DN连接信息创建同步任务。 ● MySQL->Kafka、GaussDB(for MySQL)->Kafka的实时同步支持全量+增量模式。 ● MySQL->Kafka、GaussDB->Kafka的实时同步支持查看链路监控。

发布日期	修改说明
2022-11-30	<p>第四十三次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS对接云日志服务，配置访问日志后，DRS实例新生成的所有日志记录会上传到云日志服务进行管理。 • 支持Microsoft SQL Server->GaussDB主备版的入云实时同步。 • 支持Microsoft SQL Server->GaussDB分布式版的入云实时同步。 • DRS支持MySQL单机类型同步任务升级实例规格。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> • 以下链路转为商用支持标准。 <ul style="list-style-type: none"> - Oracle->GaussDB主备版的实时同步。 - Oracle->GaussDB分布式版的实时同步。 - MongoDB->DDS的实时同步。 - GaussDB主备版->Oracle的实时同步。 - GaussDB主备版->GaussDB(DWS)的实时同步。 - GaussDB主备版->Kafka的实时同步。 - GaussDB分布式版->Oracle的实时同步。 - GaussDB分布式版->GaussDB(DWS)的实时同步。 - GaussDB分布式版->Kafka的实时同步。
2022-10-30	<p>第四十二次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • Microsoft SQL Server->GaussDB(DWS)实时同步支持选择同步对象类型。 • MySQL为源同步链路，增加源库GTID状态预检查。
2022-08-30	<p>第四十一次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • Oracle为源同步链路支持WE8ISO8859P15字符集。 • PostgreSQL->Kafka同步支持多规格。 • MySQL->GaussDB(DWS)同步支持设置异常数据策略。
2022-07-30	<p>第四十次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • PostgreSQL为源同步链路增加主键列为行外存储模式的校验。 • Oracle-Kafka同步对象选择导入方式时，支持设置不同的表对象同步到不同Topic。 • DRS支持MySQL JSON数据格式同步GaussDB(DWS)。 • 对于创建失败的任务，DRS默认保留3天，超过3天将会自动结束任务。
2022-06-30	<p>第三十九次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL->GaussDB(DWS)的实时同步支持设置周期性行对比。 • MySQL->GaussDB(DWS)的实时同步支持跳过DDL功能。

发布日期	修改说明
2022-05-30	<p>第三十八次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● 支持MongoDB->DDS的入云实时同步。 ● 支持Microsoft SQL Server->GaussDB(DWS)的入云实时同步。 ● 支持TiDB->GaussDB(for MySQL)的入云实时同步。 ● 支持DB2 for LUW->GaussDB(DWS)的入云实时同步。 ● GaussDB主备版->Kafka、GaussDB分布式->Kafka实时同步支持多规格。 ● MySQL->CSS/ES实时同步支持列映射。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● 以下链路转为商用支持标准。 <ul style="list-style-type: none"> - DDM->Oracle的实时同步。 - MySQL->Oracle的实时同步。 - GaussDB(for MySQL)->MySQL的实时同步。 - GaussDB(for MySQL)->GaussDB(DWS)的实时同步。 - GaussDB(for MySQL)->Kafka的实时同步。 - GaussDB(for MySQL)->Oracle的实时同步。 - GaussDB(for MySQL)->CSS/ES的实时同步。 - PostgreSQL->Kafka的实时同步。
2022-04-30	<p>第三十七次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● MySQL->CSS/ES实时同步支持任务克隆功能。 ● PostgreSQL->PostgreSQL实时同步支持任务克隆功能。 ● DB2 for LUW 10.5及以上版本为源的同步链路支持SSL安全连接。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● 任务标签的键和值支持长度和字符范围调整。
2022-03-30	<p>第三十六次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● MySQL->CSS/ES、GaussDB(for MySQL)->CSS/ES实时同步支持数据过滤和列过滤。 ● MySQL->GaussDB(DWS)实时同步支持单增量模式。 ● DB2 for LUW到GaussDB的实时同步支持DB2 for LUW 10.1, 11.1版本。 ● Oracle为源实时同步链路，进度支持查看显示每张表的累积操作次数。 ● DDS->MongoDB实时同步支持暂停重置和任务克隆功能。 ● DRS支持GaussDB分布式版到GaussDB主备版的出云实时同步。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● 数据过滤where条件支持输入字符长度修改为512。 ● DRS任务时延阈值异常的通知支持关闭。

发布日期	修改说明
2022-02-28	<p>第三十五次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● GaussDB->Kafka、PostgreSQL->Kafka的实时同步支持指定复制槽名。 ● PostgreSQL同步支持快照导出。 ● DRS实时同步部分链路支持多规格。 ● DRS支持批量结束任务。
2022-01-30	<p>第三十四次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● DB2 for LUW到GaussDB的实时同步支持DB2 for LUW 10.5、11.5版本，且支持大字段。 ● 支持PostgreSQL->Kafka的出云和自建-自建实时同步。 ● PostgreSQL->PostgreSQL同步支持文件导入对象。 ● MySQL->CSS/ES、GaussDB(for MySQL)->CSS/ES实时同步支持SSL安全连接。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● MySQL->CSS/ES的实时同步由公测支持标准变更为商用支持标准。
2021-12-31	<p>第三十三次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● 新增DRS对数据库影响说明。 ● PostgreSQL增量同步支持DDL增加。 ● GaussDB同步到Kafka支持JSON消息格式。 ● MySQL->GaussDB(for MySQL)的实时同步支持限速。 ● 支持MySQL->GaussDB主备版的入云实时同步。 ● 支持GaussDB主备版->MySQL的出云实时同步。 ● 实时同步链路支持包周期。 ● Oracle->DDM同步支持目标端比源端多列场景。 ● MySQL->GaussDB(DWS)同步支持任务暂停和重置。 ● PostgreSQL->PostgreSQL同步支持任务暂停和重置。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● 任务异常同步设置从任务实例信息页面调整到任务确认页面。 ● GaussDB为源的实时同步链路不再支持库级同步。

发布日期	修改说明
2021-11-30	<p>第三十二次正式发布，新增如下内容：</p> <ul style="list-style-type: none">● 支持GaussDB主备版->GaussDB(DWS)的出云实时同步。● 支持GaussDB(for MySQL)->MySQL的出云实时同步。● 支持GaussDB(for MySQL)->GaussDB(DWS)的出云实时同步。● 支持DB2 for LUW->GaussDB主备版的入云实时同步。● 支持DB2 for LUW->GaussDB分布式版的入云实时同步。● PostgreSQL->GaussDB(DWS)的实时同步支持任务再编辑功能和暂停功能。● GaussDB分布式版->GaussDB(DWS)的实时同步支持全量+增量模式。● MySQL->GaussDB(for MySQL)数据加工附加列支持字段类型增加。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none">● 以下链路由白名单转为公测支持标准。<ul style="list-style-type: none">- PostgreSQL->GaussDB(DWS)的实时同步。- DDS->MongoDB的实时同步。- GaussDB主备版->Oracle的实时同步。- GaussDB主备版->Kafka的实时同步。- GaussDB主备版->GaussDB分布式版的实时同步。- GaussDB主备版->GaussDB主备版的实时同步。- GaussDB分布式版->MySQL的实时同步。- GaussDB(for MySQL)->Kafka的实时同步。- GaussDB(for MySQL)->Oracle的实时同步。- GaussDB(for MySQL)->CSS/ES的实时同步。- MySQL自建->CSS/ES的实时同步。

发布日期	修改说明
2021-09-30	<p>第三十一次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS增加产品架构和功能原理说明。 • 支持DDM->DDM的入云实时同步。 • 支持MySQL->Oracle的出云实时同步。 • 支持GaussDB(for MySQL)->Oracle的出云实时同步。 • 支持GaussDB(for MySQL)->ElasticSearch的出云实时同步。 • 支持MongoDB出云的单增量实时同步。 • 支持MySQL自建->ElasticSearch的实时同步。 • MySQL>GaussDB分布式实时同步支持对象级对比和行对比。 • MySQL->MySQL、MySQL->GaussDB(for MySQL)、MySQL->GaussDB(DWS)、MySQL->PostgreSQL同步场景支持DDL类型过滤。 • MySQL实时同步新增任务克隆功能。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> • 以下链路由公测支持标准变更为商用支持标准。 <ul style="list-style-type: none"> - Oracle->DDM的实时同步。 - Oracle->PostgreSQL的实时同步。 - DDM->kafka的实时同步。 - DDM->GaussDB(DWS)的实时同步。 - PostgreSQL->PostgreSQL的实时同步。
2021-08-30	<p>第三十次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • 支持MySQL->ElasticSearch的出云实时同步。 • 支持GaussDB主备版->GaussDB主备版的实时同步。 • 支持PostgreSQL 13版本的实时同步。 • MySQL->Kafka实时同步支持JSON-C消息格式。 • DRS数据对比支持导出对比结果。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> • 以下链路由白名单转为公测支持标准。 <ul style="list-style-type: none"> - MySQL->GaussDB分布式版 - Oracle->GaussDB分布式版 - GaussDB分布式版->Oracle - GaussDB分布式版->GaussDB(DWS) - GaussDB分布式版->Kafka - GaussDB分布式版->GaussDB分布式版 • DRS任务异常通知支持从SMN主题选取。

发布日期	修改说明
2021-08-03	<p>第二十九次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● 支持GaussDB(for MySQL)->Kafka的出云实时同步。 ● 支持DDM->MySQL的出云实时同步。 ● 支持PostgreSQL->GaussDB分布式版的入云实时同步。 ● 支持PostgreSQL->GaussDB主备版的入云实时同步。 ● MySQL->Kafka的出云实时同步支持过滤Delete操作。 ● MySQL->MySQL实时同步的数据加工支持列加工。 ● PostgreSQL->PostgreSQL实时同步支持同步用户和权限，并且支持用户对比。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● Oracle->GaussDB(DWS)由公测支持标准变更为商用支持标准。
2021-07-05	<p>第二十八次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● 支持DDM->Kafka的出云实时同步。 ● 支持MySQL->GaussDB(DWS)同步任务限速。 ● DRS新增权限，支持除删除实例外的所有权限。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● DRS公网网络入云开始收取数据传输费。 ● DRS出云费用调整。 ● 部分链路由公测支持标准变更为商用支持标准，将按需收取一定费用。 <ul style="list-style-type: none"> - DDM->RDS for MySQL的实时同步。 - MySQL->RDS for PostgreSQL的实时同步。 - MySQL->GaussDB(for MySQL)的实时同步。
2021-05-31	<p>第二十七次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> ● MySQL->GaussDB(DWS)支持无主键表的同步。 ● MySQL->GaussDB(DWS)的增量同步支持部分DDL ● PostgreSQL->PostgreSQL增量同步支持部分DDL。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> ● DRS任务异常通知支持从SMN主题选取。

发布日期	修改说明
2021-04-30	<p>第二十六次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS支持实时同步场景增加 <ul style="list-style-type: none"> - Oracle数据库到DDM。 - DDM到GaussDB(DWS)。 - GaussDB分布式版到GaussDB(DWS)。 <p>相较于上一个版本有如下变更：</p> <ul style="list-style-type: none"> • 部分链路由公测支持标准变更为商用支持标准，将按需收取一定费用。 <ul style="list-style-type: none"> - MySQL数据库->GaussDB(for MySQL)的实时同步。 - Oracle数据库->GaussDB(for MySQL)的实时同步。 - MySQL数据库->GaussDB(DWS)的实时同步。 - Oracle数据库->RDS for MySQL的实时同步。
2021-03-30	<p>第二十五次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS支持Oracle数据库到GaussDB分布式的实时同步。 • DRS支持MySQL数据库到GaussDB(for MySQL)的实时同步。 • DRS支持PostgreSQL数据库到GaussDB(DWS)的实时同步。 • DRS支持MySQL数据库的单增量同步。 • DRS支持部分同步链路正常情况下的暂停和重置功能。
2021-01-30	<p>第二十四次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS支持Oracle数据实时同步到GaussDB(DWS)的文件导入，编辑和数据加工。
2020-12-30	<p>第二十三次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS支持MySQL到GaussDB(DWS)的实时同步。
2020-11-30	<p>第二十二次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS支持MySQL到GaussDB同步增加再编辑和对象映射功能。 • DRS对象选择全面支持搜索框。 • DRS支持设置任务异常自动结束天数。
2020-10-31	<p>第二十一次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS增加同步场景的时延说明。
2020-09-30	<p>第二十次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS同步场景MySQL数据库->GaussDB数据库目标库最小权限优化。 • DRS同步场景支持GaussDB数据库->MySQL数据库增量启动。
2020-08-31	<p>第十九次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DDM到Oracle同步场景首次支持跳跃续传同步任务和暂停同步任务。 • 创建同步任务时，支持指定实例所在子网。

发布日期	修改说明
2020-07-31	<p>第十八次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS同步场景中MySQL支持多表行过滤。 • DRS支持同一租户下的不同用户可以各自管理DRS任务，且各用户之间任务互不可见。
2020-04-30	<p>第十七次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • 异构同步场景增加异常诊断功能。 • Oracle到kafka同步支持创建指定任务续传。
2020-03-31	<p>第十六次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS首次上线DDM到Oracle同步场景。 • MySQL到PostgreSQL同步支持公网网络。 • DRS提供任务暂停功能。 • DRS首次上线MySQL到GaussDB(DWS)同步场景。
2020-02-29	<p>第十五次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • 数据同步新增映射的分页和查询功能。 • DRS新增任务启动后修改流速模式的功能。 • DRS支持密码重置功能。
2020-01-30	<p>第十四次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • DRS首次支持MySQL 8.0。 • DRS提供强制结束任务功能，可优先结束任务。
2019-12-30	<p>第十三次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL支持自建库到自建库的同步。
2019-11-30	<p>第十二次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • Mysql同步场景：增加出云场景。
2019-10-30	<p>第十一次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • 支持标签系统。
2019-09-30	<p>第十次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL引擎的内容对比功能增加支持数据库测原地运算模式。 • MySQL引擎的同步场景开始收费。
2019-08-30	<p>第九正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL同步支持冲突处理的“覆盖”策略。 • 支持VPN和专线场景的数据库复制进行限速。
2019-07-30	<p>第八次正式发布，新增如下内容：</p> <ul style="list-style-type: none"> • MySQL同步场景取消重试功能，支持重置和续传功能。

发布日期	修改说明
2019-06-30	第七次正式发布，新增如下内容： <ul style="list-style-type: none"> MySQL同步场景增加重置流程。
2019-04-30	第六次正式发布，新增如下内容： <ul style="list-style-type: none"> 数据同步场景支持VPN，专线网络。
2019-02-28	第五次正式发布，新增如下内容： <ul style="list-style-type: none"> MySQL->MySQL同步场景支持过滤删除库操作。 数据同步场景基本信息页签新增查看映射。
2019-01-19	第四次正式发布，新增如下内容： <ul style="list-style-type: none"> MySQL->MySQL数据同步任务支持可编辑。 支持MySQL->PostgreSQL的数据同步。 数据同步场景数据冲突可视化。 MySQL->PostgreSQL的数据同步支持表映射。
2018-12-30	第三次正式发布，新增如下内容： <ul style="list-style-type: none"> 同步模式和同步对象优化。
2018-11-30	第二次正式发布，新增如下内容： <ul style="list-style-type: none"> 支持数据同步场景。 支持预检查项分类。 支持批量删除任务。
2018-10-31	第一次正式发布。