

表格存储服务

快速入门

文档版本

01

发布日期

2025-07-17



版权所有 © 华为云计算技术有限公司 2025。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云计算技术有限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，华为云计算技术有限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。

华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址：<https://www.huaweicloud.com/>

目 录

1 快速创建 HBase 集群并进行离线数据查询.....

1

2 快速创建 Doris 集群并进行报表分析.....

5

3 快速创建 ClickHouse 集群并进行统计分析.....

10

4 快速创建 StarRocks 集群并进行实时数据分析.....

15

5 入门实践.....

19

1 快速创建 HBase 集群并进行离线数据查询

HBase是一个高可靠、高性能、面向列、可伸缩的分布式存储系统。本章节提供从零开始使用HBase的操作指导：通过HBase Shell命令实现创建表，往表中插入数据，修改表，读取表数据，删除表中数据以及删除表的功能。

背景信息

假定用户开发一个应用程序，用于管理企业中的使用A业务的用户信息，使用HBase客户端实现A业务操作流程如下：

- 创建用户信息表user_info。
- 在用户信息中新增用户的学历、职称信息。
- 根据用户编号查询用户姓名和地址。
- 根据用户姓名进行查询。
- 用户销户，删除用户信息表中该用户的数据。
- A业务结束后，删除用户信息表。

表 1-1 用户信息

编号	姓名	性别	年龄	地址
12005000201	A	Male	19	IPA, IPB
12005000202	B	Female	23	IPC, IPD
12005000203	C	Male	26	IPE, IPF
12005000204	D	Male	18	IPG, IPH
12005000205	E	Female	21	IPI, IPJ
12005000206	F	Male	32	IPK, IPL
12005000207	G	Female	29	IPM, IPN
12005000208	H	Female	30	IPO, IPP
12005000209	I	Male	26	IPQ, IPR

编号	姓名	性别	年龄	地址
12005000210	J	Male	25	IPS, IPT

准备工作

- 已注册华为账号并开通华为云，具体请参见[注册华为账号并开通华为云](#)，且在使用CloudTable前检查账号状态，账号不能处于欠费或冻结状态。
- 已创建虚拟私有云和子网，参见创建[虚拟私有云和子网](#)。

步骤一：购买 HBase 集群

1. 登录表格存储服务控制台。
2. 控制台左上角选择区域。
3. 单击“集群管理”进入集群管理界面。
4. 单击集群管理界面右上角的“购买集群”，进入购买集群页面配置相关参数（安全组规则配置端口请参见[HBase安全规则组](#)）。
5. 单击“立即购买”，进入确认规格页面，确认无误，单击“完成”。
6. 返回集群列表查看创建集群进度，当集群状态为“服务中”时，集群创建成功，详细参见[创建HBase集群](#)。

表 1-2 HBase 安全规则组

方向	协议	端口/范围	源地址/安全组	用途
出方向	全部	全部	0.0.0.0/0	出方向放行
入方向	TCP	16000	CloudTable HBase集群所在的安全组	HMaster RPC端口
	TCP	16020		RegionServer RPC端口
	TCP	2181		监听ZooKeeper客户端连接 监听端口
	TCP	2888		Follower连接监听端口
	TCP	3888		ZooKeeper选举端口
	TCP	2000		Hagent访问端口

步骤二：准备弹性云服务器

1. 购买弹性云服务器，登录ECS控制台。
2. 控制台左上角选择区域。
3. 单击左侧的服务列表，选择“计算 > 弹性云服务器 ECS”，进入弹性云服务器管理界面。
4. 单击页面右上角“购买弹性云服务器”，进入配置界面。

- 配置弹性云服务器参数：基础配置、实例、操作系统、存储备份、网络、安全组、公网访问、云服务器管理、高级配置、购买量。
- 参数配置后，勾选“协议”并单击“立即购买”。ECS创建成功后会默认开机。具体请参见弹性云服务“[自定义购买ECS](#)”章节。

说明

为了确保集群连接成功VPC、安全组必须与ECS保持一致。

步骤三：添加安全组

在ECS安全组中添加本地主机的IP地址。

- 获取本地主机IP，按键盘中“win+R”，弹出运行界面。
- 在输入框中输入“cmd”，单击“确定”，进入cmd运行窗口，
- 在命令窗口输入“ipconfig”，单击“回车”，查询本机IP。
- 登录ECS服务控制台。
- 进入“弹性云服务器”列表页面，单击“弹性云服务名称 > 基本信息 > 安全组名称 > 入方向规则”，进入入方向规则页面。
- 单击页面右上角的“添加规则”，弹出“添加入方向规则”页面。
- 源地址中填写步骤3中获取的本地IP地址。单击“确定”安全组添加完成。

步骤四：安装客户端并校验客户端

安装客户端有一键部署客户端、[手动安装](#)部署两种方法，此处以一键部署为例。

- 下载客户端一键部署工具。使用SSH登录工具（如PuTTY）通过“弹性IP”远程登录到Linux弹性云服务器。

具体登录操作步骤请参见《弹性云服务器用户指南》中的“[登录Linux弹性云服务器 > SSH密码方式登录](#)”。

然后执行如下命令，获取客户端一键部署工具：

```
curl -O -k "https://cloudtable-publish.obs.myhuaweicloud.com/quick_start_hbase_shell.sh"
```

说明

此命令适用于HBase 1.x版本。

```
curl -O -k "https://cloudtable-publish.obs.myhuaweicloud.com/cloudtable-client/quick_start_hbase_shell.sh"
```

说明

- 此命令适用于HBase 2.x版本。
 - 一键部署包内置校验文件。
- 获取访问集群地址。
 - 登录表格存储服务管理控制台，进入集群管理界面。
 - 单击目标集群“集群名称 > 详情”，进入集群信息页面。
 - 获取相应的“ZK链接地址（内网）”，该参数值就是集群访问地址。
 - 使用工具部署客户端。请将以下命令中的"\$zookeeper_address"参数替换为2获取的ZK链接地址，然后在弹性云服务器命令行窗口，执行该命令一键部署客户端。
 - 普通集群一键部署命令。


```
source quick_start_hbase_shell.sh $zookeeper_address
```

- 安全集群一键部署命令。

```
source quick_start_hbase_shell $zookeeper_address enable
```

4. 启动Shell访问集群。

在上一步执行source命令自动部署客户端后，就已自动启动了HBase Shell。在后续使用过程中，您也可以执行“bin/hbase shell”命令启动Shell访问集群。

步骤五：运行 HBase 客户端命令，实现 A 业务

1. 根据表1-1创建用户信息表user_info并添加相关数据。

```
create 'user_info',{NAME => 'i'}
```

以增加编号12005000201的用户信息为例，其他用户信息参照如下命令依次添加：

```
put 'user_info','12005000201','i:name','A'
put 'user_info','12005000201','i:gender','Male'
put 'user_info','12005000201','i:age','19'
put 'user_info','12005000201','i:address','IPA, IPB'
```

2. 在用户信息表user_info中新增用户的学历、职称信息。

以增加编号为12005000201的用户的学历、职称信息为例，其他用户类似。

```
put 'user_info','12005000201','i:degree','master'
put 'user_info','12005000201','i:pose','manager'
```

3. 根据用户编号查询用户姓名和地址。

以查询编号为12005000201的用户姓名和地址为例，其他用户类似。

```
scan 'user_info',
{STARTROW=>'12005000201',STOPROW=>'12005000201',COLUMNS=>['i:name','i:address']}
```

4. 根据用户姓名进行查询。

以查询A用户信息为例，其他用户类似。

```
scan 'user_info',{FILTER=>"SingleColumnValueFilter('i','name',=,'binary:A')"}]
```

5. 删除用户信息表中该用户的数据。

所有用户的数据都需要删除，以删除编号为12005000201的用户数据为例，其他用户类似。

```
delete 'user_info','12005000201','i'
```

6. 删除用户信息表。

```
disable 'user_info';drop 'user_info'
```

2 快速创建 Doris 集群并进行报表分析

Doris是一个基于MPP架构的高性能、实时的分析型数据库，以极速易用的特点被人们所熟知，仅需亚秒级响应时间即可返回海量数据下的查询结果，不仅可以支持高并发的点查询场景，也能支持高吞吐的复杂分析场景。本章节提供从零开始使用Doris操作指导：通过MySQL命令实现创建表，往表中插入数据，修改表，读取表数据，删除表中数据以及删除表的功能。

背景信息

假定这是一张记录用户访问某商品页面行为的表信息，使用MySQL客户端实现业务操作：

- 创建用户信息表example_tbl。
- 在用户信息中新增访问时间、城市、性别、停留时长、花费。
- 根据用户user_id查询用户的基本信息。
- 业务结束后，删除用户信息表。

表 2-1 原始数据

user_id	date	city	age	sex	last_visit_date	cost	max_dwell_time	min_dwell_time
10000	2017-10-01	A	20	0	2017-10-01 06:00:00	20	10	10
10000	2017-10-01	A	20	0	2017-10-01 07:00:00	15	2	2
10001	2017-10-01	A	30	1	2017-10-01 17:05:45	2	22	22

user_id	date	city	age	sex	last_visit_date	cost	max_dwell_time	min_dwell_time
10002	2017-10-02	B	20	1	2017-10-02 12:59:12	200	5	5
10003	2017-10-02	C	32	0	2017-10-02 11:20:00	30	11	11
10004	2017-10-01	D	35	0	2017-10-01 10:00:15	100	3	3
10004	2017-10-03	D	35	0	2017-10-03 10:20:22	11	6	6

- 字段含义说明。

表 2-2 参数说明

参数	数据	说明
user_id	10000	用户id，每个用户唯一识别id。
date	2017-10-01	数据入库时间，精确到日期。
city	A	用户所在城市。
age	20	用户年龄。
sex	0	性别男（1代表女性）。
last_visit_date	2017-10-01 06:00:00	用户本次访问该页面的时间，精确到秒。
cost	20	用户本次访问产生的消费。
max_dwell_time	10	用户本次访问，驻留该页面的时间。
min_dwell_time	10	用户本次访问，驻留该页面的时间（冗余）。

准备工作

- 已注册华为账号并开通华为云，具体请参见[注册华为账号并开通华为云](#)，且在使用CloudTable前检查账号状态，账号不能处于欠费或冻结状态。

- 已创建虚拟私有云和子网，参见创建[虚拟私有云和子网](#)。

步骤一：购买 Doris 集群

1. 登录表格存储服务控制台。
2. 控制台左上角选择区域。
3. 单击“集群管理”进入集群管理界面。
4. 单击集群管理界面右上角的“购买集群”，进入购买集群页面配置相关参数（安全组规则配置端口请参见[Doris安全规则组](#)）。
5. 单击“立即购买”，进入确认规格页面，确认无误，单击“完成”。
6. 返回集群列表查看创建集群进度，当集群状态为“服务中”时，集群创建成功，详细参见[创建Doris集群](#)。

表 2-3 Doris 安全规则组

方向	策略	协议端口/范围	类型	目的/源地址	用途
出方向	允许	全部	IPv4/ IPv6	0.0.0.0/0	出方向放行。
入方向	允许	9030		CloudTable Doris集群所在的安全组。	FE上的mysql server端口。
	允许	8030			FE上的http server端口。
	允许	8040			BE上的http server的端口。
	允许	8050			FE上的https server端口

步骤二：准备弹性云服务器

1. 购买弹性云服务器，登录ECS控制台。
2. 控制台左上角选择区域。
3. 单击左侧的服务列表，选择“计算 > 弹性云服务器 ECS”，进入弹性云服务器管理界面。
4. 单击页面右上角“购买弹性云服务器”，进入配置界面。
5. 配置弹性云服务器参数：基础配置、实例、操作系统、存储备份、网络、安全组、公网访问、云服务器管理、高级配置、购买量。
6. 参数配置后，勾选“协议”并单击“立即购买”。ECS创建成功后会默认开机。具体请参见弹性云服务“[自定义购买ECS](#)”章节。

步骤三：添加安全组

在ECS安全组中添加本地主机的IP地址。

1. 获取本地主机IP，按键盘中“win+R”，弹出运行界面。

2. 在输入框中输入“cmd”，单击“确定”，进入cmd运行窗口，
3. 在命令窗口输入“ipconfig”，单击“回车”，查询本机IP。
4. 登录ECS服务控制台。
5. 进入“弹性云服务器”列表页面，单击“弹性云服务名称 > 基本信息 > 安全组名称 > 入方向规则”，进入入方向规则页面。
6. 单击页面右上角的“添加规则”，弹出“添加入方向规则”页面。
7. 源地址中填写步骤3中获取的本地IP地址。单击“确定”安全组添加完成。

步骤四：安装 MySQL 客户端

用户在弹性云服务器里可以通过手动安装客户端的方法安装客户端。

1. 使用SSH登录工具（如PuTTY）通过“弹性IP”远程登录到Linux弹性云服务器。具体登录操作步骤请参见《弹性云服务器用户指南》中的“[登录Linux弹性云服务器](#)> SSH密码方式登录”。
2. 将客户端传输到步骤二中创建的弹性云服务器。
3. 登录后执行以下命令解压安装包。

```
cd <客户端安装包存放路径>
tar xzvf 客户端压缩包名
```

说明

操作步骤提到的“<客户端安装包存放路径>”请替换为实际的客户端存放路径。

4. 进入bin目录。

```
cd mysql-5.7.22-linux-glibc2.12-x86_64/bin/
```
5. 然后执行下面的命令连接Doris集群。

```
./mysql -uadmin -ppassword -h集群内网地址 -P9030
```

说明

- 集群内网地址：集群详情页面中集群访问地址，这里替换成您自己购买的集群的访问地址（FE节点的访问地址，都可以访问集群）。
- `password`为您购买集群时设置的密码，如果有特殊符号，特殊符号前面需要\进行转义，如果密码整体用单引号括起来，则特殊字符不需要转义。
- 端口：FE上的mysql server端口9030。

步骤五：执行 MySQL 命令插入数据

1. 创建一个数据库。

```
CREATE DATABASE demo;
```
2. 创建数据表
 - 使用数据库。

```
USE demo;
```
 - 创建表。

```
CREATE TABLE IF NOT EXISTS demo.example_tbl
(
  `user_id` LARGEINT NOT NULL COMMENT "用户id",
  `date` DATE NOT NULL COMMENT "数据插入日期时间",
  `city` VARCHAR(20) COMMENT "用户所在城市",
  `age` SMALLINT COMMENT "用户年龄",
  `sex` TINYINT COMMENT "用户性别",
  `last_visit_date` DATETIME REPLACE DEFAULT "1970-01-01 00:00:00" COMMENT "用户最后一次访问时间",
  `cost` BIGINT SUM DEFAULT "0" COMMENT "用户总消费",
  `max_dwell_time` INT MAX DEFAULT "0" COMMENT "用户最大停留时间",
```

```
`min_dwell_time` INT MIN DEFAULT "99999" COMMENT "用户最小停留时间"
)
AGGREGATE KEY(`user_id`, `date`, `city`, `age`, `sex`)
DISTRIBUTED BY HASH(`user_id`) BUCKETS 1
PROPERTIES (
  "replication_allocation" = "tag.location.default: 3"
);
```

3. 插入数据。

```
INSERT INTO demo.example_tbl
(user_id,date,city,age,sex,last_visit_date,cost,max_dwell_time,min_dwell_time)
VALUES('10000','2017-10-01','A','20','0','2017-10-01 07:00:00','35','10','2'),
('10001','2017-10-01','A','30','1','2017-10-01 17:05:45','2','22','22'),
('10002','2017-10-02','B','20','1','2017-10-02 12:59:12','200','5','5'),
('10003','2017-10-02','C','32','0','2017-10-02 11:20:12','30','11','11'),
('10004','2017-10-01','D','35','0','2017-10-01 10:00:15','100','3','3'),
('10004','2017-10-03','D','35','0','2017-10-03 10:20:22','11','6','6');
```

4. 查询数据。

- 我们上面完成了建表，输数据导入，下面我们就可以体验Doris的数据快速查询分析能力。

```
mysql> SELECT * FROM
demo.example_tbl;
```

user_id	date	city	age	sex	last_visit_date	cost	max_dwell_time	min_dwell_time
10000	2017-10-01	A	20	0	2017-10-01 07:00:00	35	10	2
10001	2017-10-01	A	30	1	2017-10-01 17:05:45	2	22	22
10002	2017-10-02	B	20	1	2017-10-02 12:59:12	200	5	5
10003	2017-10-02	C	32	0	2017-10-02 11:20:12	30	11	11
10004	2017-10-01	D	35	0	2017-10-01 10:00:15	100	3	3
10004	2017-10-03	D	35	0	2017-10-03 10:20:22	11	6	6

6 rows in set (0.02 sec)

- 查看指定城市信息。

```
mysql> SELECT * FROM demo.example_tbl where city='B';
```

user_id	date	city	age	sex	last_visit_date	cost	max_dwell_time	min_dwell_time
10002	2017-10-02	B	20	1	2017-10-02 12:59:12	200	5	5

1 row in set (0.10 sec)

5. 删除数据。

- 删除指定行数据。

```
mysql> DELETE FROM demo.example_tbl WHERE user_id = 10003;
Query OK, 0 rows affected (0.04 sec)
{'label':'delete_77ed273a-a052-4d64-bac0-23916b698003', 'status':'VISIBLE', 'txnId':'39'}
```

- 删除表。

```
mysql> DROP TABLE demo.example_tbl;
Query OK, 0 rows affected (0.01 sec)
```

3 快速创建 ClickHouse 集群并进行统计分析

ClickHouse为您提供方便易用、灵活稳定的云端ClickHouse托管服务。只需要几分钟，便可完成海量数据查询数据仓库的搭建，简单轻松地完成对数据的实时查询分析，提升数据价值挖掘的整体效率。云数据仓库ClickHouse是一种基于MPP（大规模并行处理）架构的数仓服务，基于ClickHouse优异的查询性能，查询效率数倍于传统数据仓库。

背景信息

假定这是一张学生成绩表信息，使用客户端实现业务操作：

- 创建用户信息表demo_t。
- 在用户信息中新增访问性别、科目。
- 根据用户user_id查询用户的基本信息。
- 业务结束后，删除用户信息表。

表 3-1 成绩表

user_id	name	sex	subject	score	time
10000	A	1	语文	89	2023-07-01 09:00:00
10001	B	0	数学	132	2023-07-01 09:00:00
10002	C	0	数学	90	2023-07-02 09:00:00
10003	D	0	英语	120	2023-07-01 14:00:00
10004	E	1	语文	101	2023-07-01 09:00:00
10005	F	1	语文	110	2023-07-01 09:00:00

表 3-2 参数说明

参数	数据	说明
user_id	10000	用户id，每个用户唯一识别id。
name	2023-07-01 09:00:00	数据入库时间。
sex	A	学生姓名。
subject	1	性别女（0代表男生）。
score	语文	代表学科。
time	89	代表学成绩。

准备工作

- 已注册华为账号并开通华为云，具体请参见[注册华为账号并开通华为云](#)，且在使用CloudTable前检查账号状态，账号不能处于欠费或冻结状态。
- 已创建虚拟私有云和子网，参见创建[虚拟私有云和子网](#)。

步骤一：购买 ClickHouse 集群

1. 登录表格存储服务控制台。
2. 控制台左上角选择区域。
3. 单击“集群管理”进入集群管理界面。
4. 单击集群管理界面右上角的“购买集群”，进入购买集群页面配置相关参数（安全组规则配置端口请参见[ClickHouse安全规则组](#)）。
5. 单击“立即购买”，进入确认规格页面，确认无误，单击“完成”。
6. 返回集群列表查看创建集群进度，当集群状态为“服务中”时，集群创建成功，详细参见[创建ClickHouse集群](#)。

表 3-3 ClickHouse 安全规则组

方向	策略	协议端口/范围	类型	目的/源地址	说明
出方向	允许	全部	IPv4/IPv6	0.0.0.0/0	出方向放行。
入方向	允许	8123		CloudTable ClickHouse集群所在的安全组。	ClickHouse的Http协议端口。
	允许	9000			ClickHouse的TCP协议端口。
	允许	8443			ClickHouse的Https协议端口。
	允许	9440			ClickHouse的安全TCP安全端口。

方向	策略	协议端口/范围	类型	目的/源地址	说明
	允许	2181			监听ZooKeeper客户端连接监听端口。

步骤二：下载 ClickHouse 客户端和客户端校验文件

1. 登录表格存储服务控制台。
2. 在控制台左上角选择区域。
3. 单击“帮助”，进入帮助页面。
4. 单击帮助页面右侧“常用链接 > ClickHouse客户端下载”，下载客户端安装包。
5. 单击“客户端校验文件下载”，下载校验文件。

步骤三：准备弹性云服务器

1. 购买弹性云服务器，登录ECS控制台。
2. 控制台左上角选择区域。
3. 单击左侧的服务列表，选择“计算 > 弹性云服务器 ECS”，进入弹性云服务器管理界面。
4. 单击页面右上角“购买弹性云服务器”，进入配置界面。
5. 配置弹性云服务器参数：基础配置、实例、操作系统、存储备份、网络、安全组、公网访问、云服务器管理、高级配置、购买量。
6. 参数配置后，勾选“协议”并单击“立即购买”。ECS创建成功后会默认开机。具体请参见弹性云服务“[自定义购买ECS](#)”章节。

说明

为了确保集群连接成功VPC、安全组必须与ECS保持一致。

步骤四：添加安全组

在ECS安全组中添加本地主机的IP地址。

1. 获取本地主机IP，按键盘中“win+R”，弹出运行界面。
2. 在输入框中输入“cmd”，单击“确定”，进入cmd运行窗口，
3. 在命令窗口输入“ipconfig”，单击“回车”，查询本机IP。
4. 登录ECS服务控制台。
5. 进入“弹性云服务器”列表页面，单击“弹性云服务名称 > 基本信息 > 安全组名称 > 入方向规则”，进入入方向规则页面。
6. 单击页面右上角的“添加规则”，弹出“添加入方向规则”页面。
7. 源地址中填写步骤3中获取的本地IP地址。单击“确定”安全组添加完成。

步骤五：安装 ClickHouse 客户端并校验客户端

用户在弹性云服务器里可以通过手动安装客户端的方法安装客户端。

1. 使用SSH登录工具（如PuTTY）通过“弹性IP”远程登录到Linux弹性云服务器。

具体登录操作步骤请参见《弹性云服务器用户指南》中的“[登录Linux弹性云服务器 > SSH密码方式登录](#)”。

2. 将[步骤二](#)中下载的客户端传输到Linux弹性云服务器。

3. 安装客户端，连接集群。

a. 使用SSH登录工具，通过弹性IP远程登录到Linux弹性云服务器。

具体登录操作步骤请参见《弹性云服务器用户指南》中的“[SSH密码方式登录](#)”。

b. 进入SSH工具的根目录。

```
cd /
```

c. 在根目录中新建一个文件夹。

```
mkdir 文件夹名
```

d. 进入新建的目录中。

```
cd /文件夹名/
```

e. 将客户端放在新建的目录下。

f. 解压客户端。

```
tar -xzf 客户端包名
```

g. 解压“客户端校验文件”，与客户端解压在同一个目录下。

i. 解压“客户端校验文件”。

```
cd <客户端校验文件存放路径>
```

```
tar xzvf Client_sha256.tar.gz
```

ii. 获取客户端校验码。

```
sha256sum ClickHouse_Client_23.3.tar.gz
```

iii. 查看客户端校验文件中的校验码，并与客户端校验码进行对比。如果对比结果一致，代表客户端未被篡改，如果对比结果不一致，则代表客户端篡改。

```
less ClickHouse_Client_23.3.tar.gz.sha256
```

h. 加载so。

```
sh install.sh
```

i. 进入bin目录，加权限。

```
cd bin/
```

加权限。

```
chmod 700 clickhouse
```

j. 然后执行以下命令连接ClickHouse集群，端口

非安全集群连接命令

```
./clickhouse client --host 集群内网地址 --port 9000 --user admin --password password
```

安全集群连接命令，详细操作请参见[安全集群连接步骤](#)。

```
./clickhouse client --host 集群内网地址 --port 9440 --user admin --password password --secure --config-file /root/config.xml
```

说明

- 集群内网地址：集群详情页面中集群访问地址，这里替换成您自己购买的集群的访问地址。
- password为您购买集群时设置的密码，如果有特殊符号，特殊符号前面需要\进行转义，如果密码整体用单引号括起来，则特殊字符不需要转义。

步骤六：插入数据

在[步骤五](#)的命令窗口中执行以下命令，使用ClickHouse集群快速创建数据表，并对表数据进行查询。

1. 创建数据库。

```
create database DB_demo;
```

2. 使用数据库。

```
use DB_demo;
```

3. 创建表。

```
create table DB_demo_t(user_id Int32,name String,sex Tinyint ,subject String,score Int32,time  
datetime)engine=TinyLog;
```

4. 插入数据。

```
insert into DB_demo_t(user_id,name,sex,subject,score,time) values('10000','A','1','语文','89','2023-07-01  
09:00:00');  
insert into DB_demo_t(user_id,name,sex,subject,score,time) values('10001','B','0','数学  
,','132','2023-07-01 09:00:00');  
insert into DB_demo_t(user_id,name,sex,subject,score,time) values('10002','C','0','数学','90','2023-07-02  
09:00:00');  
insert into DB_demo_t(user_id,name,sex,subject,score,time) values('10003','D','0','英语  
,','120','2023-07-01 14:00:00');  
insert into DB_demo_t(user_id,name,sex,subject,score,time) values('10004','E','1','语文  
,','101','2023-07-01 09:00:00');  
insert into DB_demo_t(user_id,name,sex,subject,score,time) values('10005','F','1','语文  
,','110','2023-07-01 09:00:00');
```

5. 查询数据。

- 查询导入的数据。

```
host-172-16-13-95 :) select * from DB_demo_t;
```

```
SELECT *
```

```
FROM DB_demo_t
```

```
Query id: 4e119f77-0592-4131-bbe2-31f42bc069a1
```

user_id	name	sex	subject	score	time
10000	A	1	语文	89	2023-07-01 09:00:00
10001	B	0	数学	132	2023-07-01 09:00:00
10002	C	0	数学	90	2023-07-02 09:00:00
10003	D	0	英语	120	2023-07-01 14:00:00
10004	E	1	语文	101	2023-07-01 09:00:00
10005	F	1	语文	110	2023-07-01 09:00:00

```
6 rows in set. Elapsed: 0.004 sec.
```

6. 删除数据。

- 删除表。

```
drop table DB_demo_t;
```

- 删除数据库。

```
drop database DB_demo;
```

4

快速创建 StarRocks 集群并进行实时数据分析

析

StarRocks是新一代极速全场景MPP(Massively Parallel Processing)数据库。StarRocks的愿景是能够让用户的数据分析变得更加简单和敏捷。用户无需经过复杂的预处理，就可以用StarRocks来支持多种数据分析场景的极速分析。

操作场景

假设某电商平台，需要实时分析用户订单行为，包括：

- 记录用户订单数据（用户ID、订单时间、金额、商品类别）。
- 快速查询特定时间段内的订单金额、热门商品类别。

表 4-1 原始数据

user_id	order_time	amount	category
101	2025-02-15 10:00:00	50.99	Electronics
102	2025-02-18 10:30:00	99.99	Clothing
103	2025-02-28 17:30:06	159.99	Cosmetics
104	2025-02-17 19:30:06	20.00	Shoes
105	2025-01-01 11:30:06	99.00	Drinks
106	2025-01-20 19:30:06	199.00	Furniture
107	2025-01-03 09:30:06	50.99	Kitchen

准备工作

- 已注册华为账号并开通华为云，具体请参见[注册华为账号并开通华为云](#)，且在使用CloudTable前检查账号状态，账号不能处于欠费或冻结状态。
- 已创建虚拟私有云和子网，参见创建[虚拟私有云和子网](#)。

步骤一：购买 StarRocks 集群

1. 登录表格存储服务控制台。
2. 控制台左上角选择区域。
3. 单击“集群管理”进入集群管理界面。
4. 单击集群管理界面右上角的“购买集群”，进入购买集群页面配置相关参数（安全组规则配置端口请参见[StarRocks安全规则组](#)）。
5. 单击“立即购买”，进入确认规格页面，确认无误，单击“完成”。
6. 返回集群列表查看创建集群进度，当集群状态为“服务中”时，集群创建成功。

表 4-2 StarRocks 安全规则组

方向	策略	协议端口/范围	类型	目的/源地址	用途
出方向	允许	全部	IPv4/ IPv6	0.0.0.0/0	出方向放行。
入方向	允许	9030		StarRocks集群所在的安全组。	FE上的mysql server端口。
	允许	8030			FE上的http server端口。
	允许	8040			BE上的http server的端口。
	允许	8050			FE上的https server端口

步骤二：准备弹性云服务

1. 购买弹性云服务器，登录ECS控制台。
2. 控制台左上角选择区域。
3. 单击左侧的服务列表，选择“计算 > 弹性云服务器 ECS”，进入弹性云服务器管理界面。
4. 单击页面右上角“购买弹性云服务器”，进入配置界面。
5. 配置弹性云服务器参数：基础配置、实例、操作系统、存储备份、网络、安全组、公网访问、云服务器管理、高级配置、购买量。
6. 参数配置后，勾选“协议”并单击“立即购买”。ECS创建成功后会默认开机。具体请参见弹性云服务“[自定义购买ECS](#)”章节。

步骤三：添加安全组

在ECS安全组中添加本地主机的IP地址。

1. 获取本地主机IP，按键盘中“win+R”，弹出运行界面。
2. 在输入框中输入“cmd”，单击“确定”，进入cmd运行窗口，
3. 在命令窗口输入“ipconfig”，单击“回车”，查询本机IP。
4. 登录ECS服务控制台。
5. 进入“弹性云服务器”列表页面，单击“弹性云服务名称 > 基本信息 > 安全组名称 > 入方向规则”，进入入方向规则页面。
6. 单击页面右上角的“添加规则”，弹出“添加入方向规则”页面。
7. 源地址中填写步骤3中获取的本地IP地址。单击“确定”安全组添加完成。

步骤四：手动安装 MySQL 客户端

步骤1 使用SSH登录工具，通过弹性IP远程登录到Linux弹性云服务器。具体登录操作步骤请参见弹性云服务器《用户指南》中的“[SSH密码方式登录](#)”。

步骤2 进入SSH工具的根目录。

```
cd /
```

步骤3 在根目录中新建一个文件夹。

```
mkdir 文件夹名
```

步骤4 登录后执行以下命令解压安装包。

```
cd <客户端安装包存放路径>  
tar -xvzf 客户端压缩包名
```

说明

“<客户端安装包存放路径>”请替换为实际的客户端存放路径。

步骤5 进入bin目录。

```
cd mysql-5.7.22-linux-glibc2.12-x86_64/bin/
```

步骤6 然后执行下面的命令连接StarRocks集群。

```
./mysql -uadmin -ppassword -h集群内网地址 -P9030
```

说明

- 集群内网地址：集群详情页面中集群访问地址，这里替换成您自己购买的集群的访问地址（FE节点的访问地址，都可以访问集群）。
- `password`为您购买集群时设置的密码，如果有特殊符号，特殊符号前面需要\进行转义，如果密码整体用单引号括起来，则特殊字符不需要转义。
- 端口：FE上的mysql server端口。

----结束

步骤五：在 MySQL 客户端创建数据表进行数据分析

步骤1 创建数据库。

```
CREATE DATABASE ecommerce;
```

步骤2 使用数据库。

```
USE ecommerce;
```

步骤3 创建订单表user_orders（按天分区，按用户ID分桶）。

```
CREATE TABLE user_orders (
  user_id BIGINT,
  order_time DATETIME,
  amount DECIMAL(10,2),
  category VARCHAR(50)
)
PARTITION BY RANGE(order_time) (
  PARTITION p202301 VALUES LESS THAN ('2025-02-01'),
  PARTITION p202302 VALUES LESS THAN ('2025-03-01')
)
DISTRIBUTED BY HASH(user_id) BUCKETS 8;
```

步骤4 插入数据。

```
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('101','2025-02-15
10:00:00','50.99','Electronics');
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('102','2025-02-18
10:30:00','99.99','clothing');
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('103','2025-02-28
17:30:06','159.99','Cosmetics');
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('104','2025-02-17
19:30:06','20.00','Shoes');
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('105','2025-01-01
11:30:06','99.00','Drinks');
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('106','2025-01-20
19:30:06','199.00','Furniture');
INSERT INTO user_orders (user_id,order_time,amount,category) VALUES('107','2025-01-03
09:30:06','150.99','Kitchen');
```

步骤5 简单聚合查询2月订单总金额。

```
SELECT
  DATE(order_time) AS day,
  SUM(amount) AS total_amount,
  category
FROM user_orders
WHERE order_time >= '2025-02-01'
GROUP BY day, category
ORDER BY total_amount DESC;
```

结果示例：

```
+-----+-----+-----+
| day      | total_amount | category |
+-----+-----+-----+
| 2025-02-28 | 159.99 | Cosmetics |
| 2025-02-18 | 99.99 | clothing |
| 2025-02-15 | 50.99 | Electronics |
| 2025-02-17 | 20.00 | Shoes |
+-----+-----+-----+
4 rows in set (0.16 sec)
```

----结束

5 入门实践

当用户购买并连接到CloudTable集群后，可以根据自身的业务需求使用CloudTable提供的一系列常用实践。

表 5-1 常用最佳实践

实践		描述
数据导入	使用DLI Flink作业实时同步MRS Kafka数据至CloudTable HBase集群	本实践演示通过数据湖探索服务DLI Flink作业将分布式消息服务Kafka的消费数据实时同步至CloudTable HBase，实现Kafka实时入库到CloudTable HBase的过程。
	使用DLI Flink作业实时同步MRS Kafka数据至CloudTable ClickHouse集群	本实践演示通过数据湖探索服务DLI Flink作业将分布式消息服务Kafka的消费数据实时同步至CloudTable ClickHouse，实现Kafka实时入库到CloudTable ClickHouse的过程。