



实时流计算服务

## 常见问题

文档版本 19

发布日期 2020-09-27

版权所有 © 华为技术有限公司 2020。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默示的声明或保证。

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。

# 目录

<b>1 通用问题</b>	<b>1</b>
1.1 什么是实时流计算?	1
1.2 什么是 CS?	1
1.3 CS 有哪些优势?	1
1.4 CS 有哪些产品特性?	2
1.5 CS 适用哪些场景?	3
1.6 CS 的用户需要什么技能?	5
1.7 什么是 SPU?	5
1.8 如何快速开始使用 CS?	5
1.9 什么是用户配额?	6
1.10 如何获取 AK/SK?	6
1.11 如何获取项目 ID?	6
1.12 如何获取用户支持?	6
1.13 CS 服务使用 SMN 主题, 提示 SMN 主题不存在, 怎么处理?	7
1.14 区域和可用区	7
<b>2 作业类</b>	<b>9</b>
2.1 CS 支持哪些数据源?	9
2.2 CS 支持输出到哪些目标?	9
2.3 CS 支持哪些数据格式?	10
2.4 支持何种代码开发的作业?	10
2.5 作业如何并发?	10
2.6 如何查看作业输出结果?	10
2.7 如何处理作业的 OBS Bucket 没有授权?	11
2.8 如何判断 CS 独享集群和指定地址是否连通?	11
2.9 如何解决 CS 作业提交失败, 日志显示连接超时?	11
2.10 作业语义检验时提示 DIS 通道不存在怎么处理?	12
2.11 CS 服务如何对接开启 sasl_ssl 认证的 DMS 服务?	12
2.12 Flink 自定义作业如何配置 checkpoint 保存到 OBS	13
<b>3 计费类</b>	<b>14</b>
3.1 CS 如何计费?	14
3.2 CS 使用其他云服务是如何收费的?	14
3.3 没有数据流入 CS 时用户是否需要付费?	14

---

<b>4 SQL 类</b> .....	<b>15</b>
4.1 CS SQL 支持配置表和流 JOIN 吗? .....	15

# 1 通用问题

## 1.1 什么是实时流计算？

“实时”指实时处理，计算框架支持按消息事件逐条处理；“流”指数据如水流，一个接着一个；“计算”指数学运算、数据分析、算法模型执行等。

“实时流计算”指实时处理当下正在发生的数据流，逐条大数据分析或算法运算。

## 1.2 什么是 CS？

实时流计算服务（Cloud Stream Service，简称CS），是运行在公有云上的实时流式大数据分析服务，全托管的方式用户无需感知计算集群，只需聚焦于Stream SQL业务，即时执行作业，完全兼容Apache Flink（1.5.3版本）API和Apache Spark（2.2.1版本）API。

实时流计算平台是公司在IT领域主推的低时延（ms级时延）、高吞吐、高可靠的分布式实时流计算服务。它以Flink为基础，加入公司沉淀的特性增强和安全增强，是批流合一的分布式计算服务，提供了数据处理所必须的Stream SQL特性，后续还会支持在Stream SQL上增加机器学习和图计算相关算法的功能。

## 1.3 CS 有哪些优势？

实时流计算服务主要具有以下优势。

- 分布式实时流计算  
支持大规模分布式集群，集群弹性伸缩，按作业使用的资源扩容和缩容集群，最大化节省成本。
- 简单易用  
在线SQL编辑平台编写Stream SQL，定义数据流入、数据处理、数据流出，快捷实现业务逻辑；用户无需关心计算集群，无需学习编程技能，降低流数据分析门槛。
- 独享集群  
全托管模式，自动弹性伸缩，用户完全不感知计算集群，无需关心大数据框架和资源调度框架，即时提交作业后，可视化查看作业的运行状态。提供共享集群和

独享集群两种模式，独享集群与其他租户和共享集群完全物理隔离，并支持配额限制管理。

- 安全隔离

租户安全机制保障，确保作业安全运行，且租户计算集群完全和其他租户物理隔离，独立的安防设置，确保计算集群的安全性。

- 按需计费

按资源使用量计费，用多少算多少，定价单位是SPU(Stream Processing Unit, 1 SPU=1核CPU+4G内存)，用户选定SPU数，按时长计费，精确到秒。

- 高吞吐低时延

采用Apache Flink的Dataflow模型，完全的实时计算框架。采用高性能计算资源，从用户自建的Kafka、MRS-Kafka、DMS-Kafka消费数据，单SPU每秒大约处理1万条消息。

## 1.4 CS 有哪些产品特性？

CS服务主要具有以下功能：

- 丰富的StreamSQL在线分析能力

支持Window、Join等聚合函数、地理函数、CEP函数等，用SQL表达业务逻辑，简便快捷实现业务。详细内容请参见《[实时流计算服务SQL语法参考](#)》。

- StreamingML

提供多种流式机器学习方法对数据进行实时分析与预测，用户仅需编写SQL调用相关函数便可实现数据统计，异常检测，实时聚类，时间序列分析等场景。详细内容请参见[StreamingML](#)。

- 地理位置分析

提供地理位置分析函数对地理空间数据进行实时分析，用户仅需编写SQL便可实现例如偏航检测，电子围栏等地理分析场景。详细内容请参见[地理函数](#)。

- CEP SQL

提供基于Match Recognize的模式匹配检测，帮助业务人员使用SQL实现基于复杂事件规则的异常检测业务。典型应用场景如欺诈检测、车辆异常行为检测、工业设备异常运行状态检测等。详细内容请参见[CEP模式匹配](#)。

- 数据可视化

提供多种图表类型实时展示作业数据输出，用户还可以通过API网关服务自由访问作业数据，接入自定义工作流中。详细内容请参见[数据可视化](#)。

- 可视化SQL编辑器

实时流计算服务针对不太熟悉SQL的用户，提供了可视化编辑器功能，它将实时流计算服务需要对接的上下游服务（如DIS、CloudTable等）和内部逻辑算子（如filter、window等）封装成可拖拽的组件，用户只需要将其拖入画布中并用连线将各个元素连接，就能轻松的创建一个作业的拓扑，再点击画布中的各个元素设置对应参数，就能轻松完成一个作业的编辑。详细内容请参见[可视化编辑器](#)。

- 支持独享集群及其作业的资源配额

允许租户创建独享集群，独享集群与其他租户和共享集群是完全物理隔离的，不会受其他作业影响。租户可配置独享集群最大可使用SPU配额，并为其子用户分配可用的集群以及可使用的SPU配额。详细内容请参见[集群管理](#)。

- 支持在线测试SQL作业  
作业调试功能可以帮助用户校验SQL语句逻辑是否正确，通过用户的样例数据输入（支持手动输入和OBS输入两种方式）结合SQL逻辑快速输出样例结果，确保在作业正式运行时，逻辑处理正确。详细内容请参见[调试作业](#)。
- 支持Flink和Spark自定义作业  
允许用户在独享集群上提交Flink和Spark自定义作业。
- 支持Spark streaming和Structured streaming  
允许用户在独享集群上提交Spark streaming自定义作业。
- 支持与多种云服务连通，形成丰富的流生态圈。  
实时流计算服务的生态分为云服务生态和开源生态：
  - 云服务生态：实时流计算服务在Stream SQL中支持与其他服务的连通。用户可以直接使用SQL从这些服务中读写数据，如DIS、OBS、CloudTable、MRS、RDS、SMN、DCS等。
  - 开源生态：通过对等连接建立与其他VPC的网络连接后，用户可以在实时流计算服务的租户独享集群中访问所有Flink和Spark支持的数据源与输出源，如Kafka、Hbase、ElasticSearch等。详细内容请参见[《实时流计算服务流生态开发指南》](#)。
- 支持智能边缘平台（Intelligent EdgeFabric）  
智能边缘平台（IEF）与实时流计算服务的结合，支持流处理云端管理，边缘侧运行，提供实时流处理能力。

## 1.5 CS 适用哪些场景？

实时流计算服务的使用，聚焦于互联网和物联网场景，适用于实时性要求高、吞吐量大的业务场景。主要应用在互联网行业中小企业、物联网、车联网、金融反欺诈等多种行业应用场景，如互联网汽车、日志在线分析、在线机器学习、在线图计算、在线推荐算法应用等。

### 实时流分析场景

提供易用、低时延、高吞吐的实时流分析服务。支持Stream SQL和用户自定义作业做流分析。

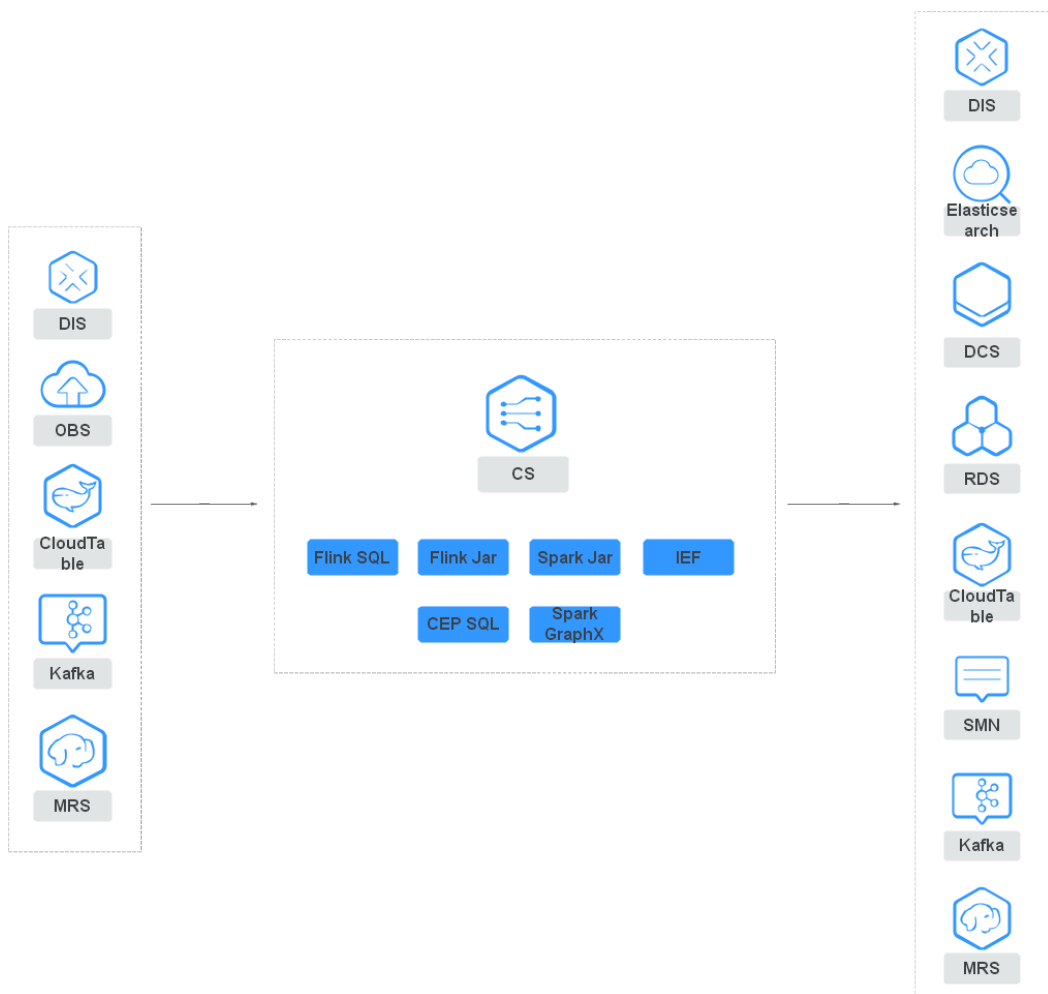
#### 场景优势

- 易用：在线编辑Stream SQL，丰富的SQL函数满足复杂业务需要。
- 完全托管：用户完全不感知计算集群，聚焦流分析本身。
- 按需计费：作业选定SPU资源量，按时长计费，精确到秒。

**场景特点：**面向流数据，支持Window、CEP、Join等复杂的流分析操作，毫秒级时延。

**适用场景：**实时日志分析，网络流量监控，实时风险管控，实时数据统计，实时数据ETL。

图 1-1 实时流分析场景



## 物联网 IoT 场景

物联网设备或边缘设备，上传数据到数据接入服务（DIS）或者其他云存储服务，实时流计算服务直接从DIS读取数据，实时分析数据流（故障检测、数据清洗、统计分析、指标预警等等），实时把流分析结果持久化或推送告警通知。

### 场景优势

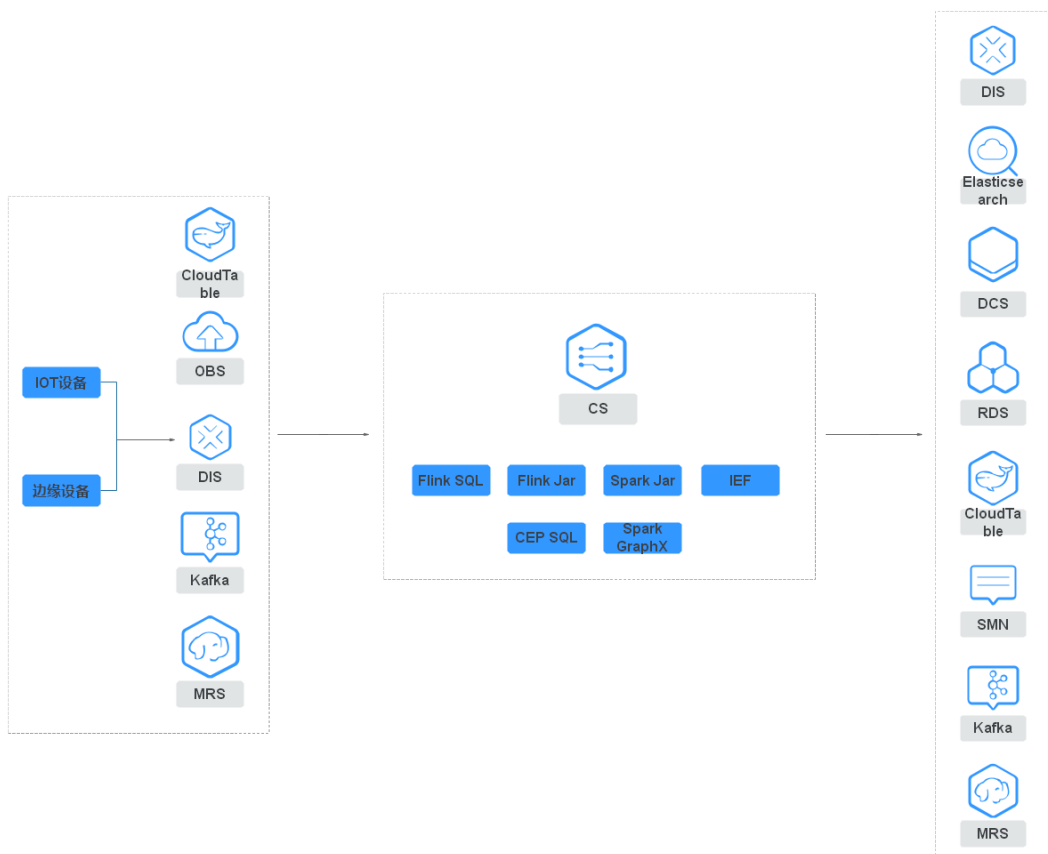
- 丰富的IoT SQL函数：区域检测函数、偏航检测函数、相对位置判断等常用的IoT函数。
- 高吞吐低时延：使用Apache Flink执行引擎，完全的实时计算框架。
- 安全隔离：租户之间完全隔离，确保数据安全。

**场景特点：**物联网IoT直接调用实时流计算服务API，实时流计算服务可以实时读取传感器信息并执行用户的分析逻辑，分析结果对接到DIS、RDS等服务，并用于可视化、持久化、告警或报表展示。

**适用场景：**电梯IoT、工业IoT、共享单车、互联网汽车和智能家居。



图 1-2 物联网 IoT 场景



## 1.6 CS 的用户需要什么技能?

CS用户需要具备一定的SQL编写技能，若要开发自定义作业还需要熟悉Flink API。

## 1.7 什么是 SPU?

流处理单元（Stream Process Units），简称SPU，是CS计费单位。1个SPU的标准配置为1核的CPU、4G的内存。用户作业可以设置多个SPU。

## 1.8 如何快速开始使用 CS?

快速使用CS服务基本流程如下：

1. 准备数据源和数据输出通道
2. 创建OBS桶保存输出数据
3. 申请实时流计算服务
4. 创建作业并提交
5. 发送数据到数据源
6. 查看作业信息及运行结果

详细操作步骤，请参见《[实时流计算服务快速入门](#)》。

## 1.9 什么是用户配额？

公有云对用户的资源数量和容量做了限制。如果资源配额限制满足不了用户的使用需求，可以通过工单系统来提交您的申请，并告知您申请提高配额的理由。在通过我们的审理之后，我们会更新您的配额并进行通知。关于配额的具体操作说明，请参见[关于配额](#)。

## 1.10 如何获取 AK/SK？

访问密钥即AK/SK（Access Key ID/Secret Access Key），表示一组密钥对，用于验证调用API发起请求的访问者身份，与密码的功能相似。用户通过调用API接口进行云资源管理（如创建集群）时，需要使用成对的AK/SK进行加密签名，确保请求的机密性、完整性和请求双方身份的正确性。获取AK/SK操作步骤如下：

1. 注册并登录[公有云管理控制台](#)。
2. 将鼠标移动到右上角用户名上，在下拉列表中单击“我的凭证”。
3. 单击“管理访问密钥”页签。
4. 单击“新增访问密钥”，进入“新增访问密钥”页面。
5. 输入当前用户的登录密码。
6. 通过邮箱或者手机进行验证，输入对应的验证码。

### 说明

在统一身份认证服务中创建的用户，如果创建时未填写邮箱或者手机号，则只需校验登录密码。

7. 单击“确定”，下载访问密钥。

### 说明

为防止访问密钥泄露，建议您将其保存到安全的位置。

## 1.11 如何获取项目 ID？

项目ID是系统所在区域的ID。用户在调用API接口进行云资源管理（如创建集群）时，需要提供项目ID。

查看项目ID步骤如下：

1. 登录[公有云管理控制台](#)。
2. 将鼠标移动到右上角用户名上，在下拉列表中单击“我的凭证”。

在“我的凭证”页面的项目列表中查看项目ID。例如  
project\_id:"5a3314075bfa49b9ae360f4ecd333695"。

## 1.12 如何获取用户支持？

为了更好的帮助和服务用户，我们提供多种方式的[用户支持](https://www.huaweicloud.com/about/contact.html)，详见<https://www.huaweicloud.com/about/contact.html>。

## 1.13 CS 服务使用 SMN 主题，提示 SMN 主题不存在，怎么处理？

进入统一身份认证服务（IAM），选择对应子账户所在的用户组，添加相应region的smn策略即可。

## 1.14 区域和可用区

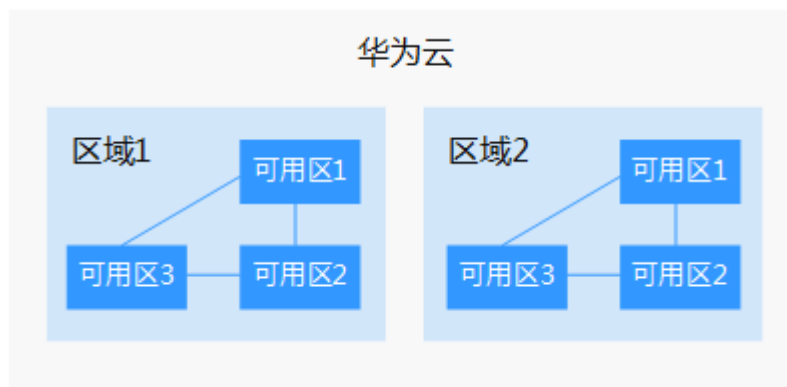
### 什么是区域、可用区？

我们用区域和可用区来描述数据中心的位置，您可以在特定的区域、可用区创建资源。

- 区域（Region）：从地理位置和网络时延维度划分，同一个Region内共享弹性计算、块存储、对象存储、VPC网络、弹性公网IP、镜像等公共服务。Region分为通用Region和专属Region，通用Region指面向公共租户提供通用云服务的Region；专属Region指只承载同一类业务或只面向特定租户提供业务服务的专用Region。
- 可用区（AZ，Availability Zone）：一个AZ是一个或多个物理数据中心的集合，有独立的风火水电，AZ内逻辑上再将计算、网络、存储等资源划分成多个集群。一个Region中的多个AZ间通过高速光纤相连，以满足用户跨AZ构建高可用性系统的需求。

图1-3阐明了区域和可用区之间的关系。

图 1-3 区域和可用区



目前，华为云已在全球多个地域开放云服务，您可以根据需求选择适合自己的区域和可用区。更多信息请参见[华为云全球站点](#)。

### 如何选择区域？

选择区域时，您需要考虑以下几个因素：

- 地理位置  
一般情况下，建议就近选择靠近您或者您的目标用户的区域，这样可以减少网络时延，提高访问速度。不过，在基础设施、BGP网络品质、资源的操作与配置等

方面，中国大陆各个区域间区别不大，如果您或者您的目标用户在中国大陆，可以不用考虑不同区域造成的网络时延问题。

香港、曼谷等其他地区和国家提供国际带宽，主要面向非中国大陆地区的用户。如果您或者您的目标用户在中国大陆，使用这些区域会有较长的访问时延，不建议使用。

- 在除中国大陆以外的亚太地区有业务的用户，可以选择“亚太-香港”、“亚太-曼谷”或“亚太-新加坡”区域。
- 在非洲地区有业务的用户，可以选择“南非-约翰内斯堡”区域。
- 在欧洲地区有业务的用户，可以选择“欧洲-巴黎”区域。
- 资源的价格  
不同区域的资源价格可能有差异，请参见[华为云服务价格详情](#)。

## 如何选择可用区？

是否将资源放在同一可用区内，主要取决于您对容灾能力和网络时延的要求。

- 如果您的应用需要较高的容灾能力，建议您将资源部署在同一区域的不同可用区内。
- 如果您的应用要求实例之间的网络延时较低，则建议您将资源创建在同一可用区内。

## 区域和终端节点

当您通过API使用资源时，您必须指定其区域终端节点。有关华为云的区域和终端节点的更多信息，请参阅[地区和终端节点](#)。

# 2 作业类

## 2.1 CS 支持哪些数据源？

CS可以处理来自如下服务的数据：

- 数据接入服务（DIS）
- 对象存储服务（OBS）
- MapReduce服务（MRS）
- Kafka集群
- 表格存储服务（CloudTable）

## 2.2 CS 支持输出到哪些目标？

CS支持将作业的结果数据输出到以下服务：

- 数据接入服务（DIS）
- 对象存储服务（OBS）
- 关系型数据库（RDS）
- 数据仓库服务（DWS）
- 文档数据库服务（DDS）
- 消息通知服务（SMN）
- Kafka集群
- 表格存储服务（CloudTable）
- 云搜索服务
- 分布式缓存服务（DCS）的Redis实例
- MapReduce服务（MRS）
- API网关（APIG）

创建如上输出流的语法请参见《[实时流计算服务流生态开发指南](#)》。

## 2.3 CS 支持哪些数据格式？

- CS支持从DIS上读取csv、json和email格式的数据。
- CS支持向DIS写入csv和json格式的数据。
- CS支持从OBS上读取csv格式的数据。
- CS支持向OBS写入csv、json和orc格式的数据。
- CS支持向SMN发送text格式的数据。
- CS支持从Kafka上读取和写入csv和json格式的数据。
- CS支持向CloudTable读取和存放table格式的数据。
- CS支持向云搜索服务发送json格式的数据。
- CS支持向DCS的Redis发送key-value格式的数据。
- CS支持向APIG写入csv和json格式的数据。

具体详情请参见《实时流计算服务流生态开发指南》中[数据格式](#)章节。

## 2.4 支持何种代码开发的作业？

CS支持使用SQL语句开发作业和使用jar包来自定义作业。

## 2.5 作业如何并发？

CS的SQL作业，不支持单个算子的并发，作业并发是指为流作业同时启动多个并发任务。

## 2.6 如何查看作业输出结果？

- CS支持将作业的结果数据输出到DIS上，因此用户需要到DIS上去查看输出结果，详细操作请参见《[数据接入服务用户指南](#)》中的“从DIS获取数据”章节。
- CS也支持将作业的结果数据输出到RDS上，因此用户需要到RDS上去查看输出结果，详细操作请参见《[关系型数据库快速入门](#)》。
- CS支持将作业的结果数据输出到SMN上，SMN服务会将输出结果反馈到用户绑定的终端设配上，详细操作请参见《[消息通知服务快速入门](#)》。
- CS支持将作业的结果数据输出到Kafka上，因此用户需要到Kafka集群上去查看输出结果，详细操作请参见[Kafka官网](#)。
- CS支持将作业的结果数据输出到CloudTable上，因此用户需要到CloudTable上去查看输出结果，详细操作请参见《[表格存储服务用户指南](#)》中的“入门”章节。
- CS支持将作业的结果数据输出到智能边缘平台上，因此用户需要到智能边缘平台上去查看输出结果，详细操作请参见《[智能边缘平台用户指南](#)》。
- CS支持将作业的结果数据输出到云搜索服务，因此用户需要到云搜索服务上去查看输出结果，详细操作请参见《[云搜索服务快速入门](#)》。
- CS支持将作业的结果数据输出到DCS上，因此用户需要到DCS上去查看输出结果，详细操作请参见《[分布式缓存服务快速入门](#)》。

## 2.7 如何处理作业的 OBS Bucket 没有授权？

用户在执行一个作业，提示OBS Bucket没有授权时，用户需要进行如下操作：

- 步骤1** 单击“作业管理”，找到操作的作业名称。
- 步骤2** 单击“操作”列表中的“编辑”，进入“作业编辑”页面。
- 步骤3** 用户需要在该作业的“作业编辑”页面进行“运行参数配置”。
  - 1. 选择“开启Checkpoint”或“保存作业日志”。
  - 2. 选择“OBS桶”。
  - 3. 单击“OBS授权”。

----结束

## 2.8 如何判断 CS 独享集群和指定地址是否连通？

CS独享集群创建完成后，用户可以测试当前集群和指定地址，如与CS对接的其他云服务地址，是否连通。

具体操作步骤如下：

- 步骤1** 登录CS管理控制台。
- 步骤2** 在CS管理控制台的左侧导航栏中，单击“集群管理”，进入“集群管理”页面。
- 步骤3** 在集群列表“名称”一列中，单击需要查看的集群名称，进入“集群详情”页面。
- 步骤4** 单击“测试地址连通性”，弹出“测试地址连通性”窗口。
- 步骤5** 输入待测试的地址，单击“测试”。

测试当前集群和指定地址是否连通，支持域名和IP，可指定端口。如果指定端口，请采用ip:port或者hostname:port的格式。

### 说明

如果指定地址没有开通icmp协议，建议优先使用指定端口测试。

如果CS独享集群与待对接的云服务网络不通，则可以通过建立VPC对等连接来实现网络互通。

如何建立VPC对等连接，请参考《实时流计算服务用户指南》中[对等连接](#)章节。

----结束

## 2.9 如何解决 CS 作业提交失败，日志显示连接超时？

原因：

运行CS作业的独享集群与待连接的地址网络不通。

解决方法：

通过建立VPC对等连接使作业所处的CS独享集群与待连接的地址网络互通。

如何建立VPC对等连接，请参考《实时流计算服务用户指南》中[对等连接](#)章节。

对等连接创建后，可以参考[如何判断CS独享集群和指定地址是否连通?](#) 测试地址连通性。

如果还是提示网络不通，请参考[为什么对等连接创建完成后不能互通?](#) 检查相关配置。

目前有如下场景的CS作业需要运行在CS的独享集群上：

- RDS输出流：CS将作业的输出数据输出到关系型数据库（RDS）中。
- DWS输出流（通过JDBC方式）：CS将作业的输出数据通过JDBC的方式输出到数据仓库服务（DWS）中。
- Elasticsearch输出流：CS将作业的输出数据输出到云搜索服务的Elasticsearch中。
- DCS输出流：CS将作业的输出数据输出到分布式缓存服务（DCS）的Redis中。
- Kafka输入流：创建source流从Kafka获取数据，作为CS作业的输入数据。
- Kafka输出流：CS将作业的输出数据输出到Kafka中。
- MRS HBase输出流：CS将作业的输出数据输出到MRS的HBase中。

运行以上场景的作业之前，需要确保作业所处的CS独享集群与待连接的地址已通过建立VPC对等连接实现网络互通。

## 2.10 作业语义检验时提示 DIS 通道不存在怎么处理？

1. 检查作业SQL语句中的DIS通道是否已创建。如果还未创建，请参见《[数据接入服务用户指南](#)》中“[开通DIS通道](#)”章节。
2. 检查并确保DIS通道和CS作业处于同一区域。

## 2.11 CS 服务如何对接开启 sasl\_ssl 认证的 DMS 服务？

CS服务支持创建输入流从DMS的Kafka专享版实例获取数据，作为作业的输入数据；也支持将作业的输出数据输出到DMS的Kafka专享版实例中。

DMS服务创建实例时开启SASL\_SSL访问，则数据加密传输，安全性更高。

CS服务采用命令行模式连接DMS实例，具体步骤如下：

**步骤1** 下载client.truststore.jks证书。

下载地址：<https://dms-demo.obs.myhuaweicloud.com/cert.zip>，下载压缩包后解压，获取压缩包中的客户端证书文件：client.truststore.jks。

**步骤2** 将client.truststore.jks上传至CS集群，例如，将文件存放在“/opt/cs/user\_files”目录下。

**步骤3** 在CS创建kafka输出流的SQL代码中，增加两个参数：kafka\_properties\_delimiter和kafka\_properties。

示例代码如下：

```
CREATE SINK STREAM kafka_sink (  
  id STRING  
)  
WITH (  
  type="kafka",
```



```
kafka_bootstrap_servers = "192.168.0.143:9093,192.168.0.113:9093,192.168.0.180:9093",
kafka_topic = "csdms",
encode = "csv",
kafka_properties_delimiter = ",",
kafka_properties = "sasljaa.config=org.apache.kafka.common.security.plain.PlainLoginModule required
username=\"test\" password=\"*****
\",saslmecanism=PLAIN,security.protocol=SASL_SSL,ssl.truststore.location=/opt/cs/user_files/
client.truststore.jks,ssl.truststore.password=dms@kafka"
);
```

其中，“username”和“password”是在DMS服务开启sasl\_ssl认证时配置的用户名和密码。

配置完成后，即可实现CS对开启SASL\_SSL认证的DMS服务的对接。

---结束

关于DMS服务连接已开启SASL的Kafka专享版实例的更多说明，请参考《[分布式消息服务用户指南](#)》。

## 2.12 Flink 自定义作业如何配置 checkpoint 保存到 OBS

CS提供将Flink的checkpoint数据保存到OBS的能力，配置的示例代码如下：

```
import org.apache.flink.streaming.api.environment.StreamExecutionEnvironment;
import org.apache.flink.runtime.state.filesystem.FsStateBackend;
import org.apache.flink.contrib.streaming.state.RocksDBStateBackend;

StreamExecutionEnvironment env = StreamExecutionEnvironment.getExecutionEnvironment();
RocksDBStateBackend rocksDbBackend = new RocksDBStateBackend(new FsStateBackend("obs路径"), true);
env.setStateBackend(rocksDbBackend);
```

其中，obs路径为obs://ak:sk@obs服务的endpoint/桶名/桶目录，ak、sk是需要上传的OBS桶所在的账号的AK/SK信息。

# 3 计费类

---

## 3.1 CS 如何计费？

CS的计费简单、易于预测，您既可以选择按照小时费率计费的按需计费方式，也可以选择更经济的按需套餐包计费方式。详细的服务资费费率标准请参见《[实时流计算服务价格说明](#)》。

## 3.2 CS 使用其他云服务是如何收费的？

使用的其他云服务按各服务收费标准收费，具体收费标准请参见[华为云计费详情页面](#)。

## 3.3 没有数据流入 CS 时用户是否需要付费？

作业一旦处于运行中即开始计费，即使没有数据流入也会计费。

# 4 SQL 类

---

## 4.1 CS SQL 支持配置表和流 JOIN 吗？

目前支持流和DCS服务的Redis表进行JOIN，具体可参见CS控制台“作业模板”中的[“云生态]DIS-CS=DCS-DIS样例模板”。流表JOIN的详细语法，请参见《实时流计算服务SQL语法参考》中[流表JOIN](#)。