

ALIBABA CLOUD

Alibaba Cloud

消息队列Kafka版
用户指南

文档版本：20210702

 阿里云

法律声明

阿里云提醒您阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击 确定 。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.实例	08
1.1. 查看接入点	08
1.2. 升级实例配置	08
1.3. 变更消息配置	10
1.4. 升级实例版本	11
1.5. 配置白名单	13
1.6. Topic引流	13
1.7. 查看任务执行记录	15
1.8. 实例问题	16
1.8.1. 实例支持哪些开源版本?	16
1.8.2. 如何选择实例规格?	16
1.8.3. 如何获取部署实例时需要的VPC信息?	17
1.8.4. 升级Broker可能产生哪些影响?	17
1.8.5. 实例的地域或网络属性无法变更	17
1.8.6. 消息队列Kafka版何时删除旧消息?	18
1.8.7. 如何快速测试消息队列Kafka版服务端是否正常?	18
1.8.8. 是否支持延迟消息?	22
1.8.9. 能否通过公网访问消息队列Kafka版实例?	22
1.8.10. 不同VPC之间是否支持互连?	22
1.8.11. 经典网络是否必须要选择172.16.0.0/12网段?	24
1.8.12. 是否支持压缩消息?	24
2.Topic	25
2.1. 查看分区状态	25
2.2. Topic问题	25
2.2.1. 为什么限制Topic总数(分区总数)?	25
2.2.2. 为什么Topic不能减分区?	26

2.2.3. 新增Topic时对应的分区数量额度	26
2.2.4. Topic的存储有哪些类型?	26
2.2.5. 专业版实例附赠1倍的Topic	26
2.2.6. 为什么Topic有分区时间显示1970?	26
2.2.7. Topic删除后不能减少分区?	26
2.2.8. 是否支持Compact的日志清理策略?	27
3.Consumer Group	28
3.1. 重置消费位点	28
3.2. 查看消费状态	28
3.3. Consumer Group问题	30
3.3.1. Consumer Group规格是多少?	30
3.3.2. 为什么有Consumer Group对应的Topic分区显示1970?	31
3.3.3. Consumer Group删除后依然可以收到消息?	31
3.3.4. 为什么分区消费时间相差较大、滞后明显或乱序?	31
3.3.5. 控制台看到的最近消费时间是什么意思?	32
3.3.6. 如何查看哪些IP在消费消息?	32
3.3.7. 哪里可以找到消费最佳实践?	32
3.3.8. 消费端挂载NFS云盘是否会影响消费端处理速度?	32
3.3.9. 如何在控制台管理Consumer的offset?	33
4.标签	34
4.1. 标签管理	34
5.监控报警	37
5.1. 监控报警	37
5.2. 监控报警问题	39
5.2.1. 为什么需要升级实例来支持监报告警功能?	39
5.2.2. 报警项状态显示数据不足怎么办?	39
5.2.3. 如何监控开源Apache Kafka?	40
5.2.4. 父账号能看到监控信息, 子账户看不到监控信息	40

5.2.5. 为什么不能登录部署消息队列Kafka版的机器?	40
5.2.6. 删除了Consumer Group还会收到堆积告警?	40
5.2.7. 什么是死分区和假堆积?	40
5.2.8. 为什么告警堆积量和控制台堆积量不一致?	41
6.Connector	42
6.1. Connector概述	42
6.2. 开启Connector	43
6.3. 创建Connector	44
6.3.1. 创建FC Sink Connector	44
6.3.2. 创建MaxCompute Sink Connector	57
6.3.3. 创建OSS Sink Connector	68
6.3.4. 创建Elasticsearch Sink Connector	73
6.3.5. 创建MySQL Source Connector	79
6.4. 查看Connector任务配置	85
6.5. 查看Connector日志	86
6.6. 暂停Connector	86
6.7. 恢复Connector	87
6.8. 删除Connector	87
6.9. 修改Connector配置	88
6.10. 测试Connector	90
6.11. 为Connector开启公网访问	90
7.迁移	92
7.1. 迁移概述	92
7.2. 评估规格	95
7.3. 迁移Topic	97
7.3.1. 迁移Topic上云	97
7.3.2. 云上迁移Topic	98
7.4. 迁移Consumer Group	100

7.4.1. 迁移Consumer Group上云	100
7.4.2. 云上迁移Consumer Group	102
7.5. 迁移数据	103
7.5.1. 迁移数据上云	104
7.6. 查看迁移进度	106
8.查询消息	107

1. 实例

1.1. 查看接入点

如需使用SDK接入消息队列Kafka版收发消息，您需要根据实例的网络环境类型来配置接入点。您可以在消息队列Kafka版控制台查看您的实例的接入点。

背景信息

消息队列Kafka版提供以下类型的接入点：

- 默认接入点：适用于在VPC环境收发消息，但不支持SASL校验。
- SASL接入点：适用于在VPC环境收发消息，且支持SASL校验。
- SSL接入点：适用于在公网环境收发消息，且支持SASL校验。

关于以上接入点的差异，请参见[接入点对比](#)。

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在实例详情页面的接入点信息区域，查看接入点。

说明

- 如果实例类型为VPC实例，则仅显示默认接入点。
- 如果实例类型为公网/VPC实例，则同时显示默认接入点和SSL接入点。
- 实例默认不开启SASL接入点，因此默认不显示SASL接入点。如需使用SASL接入点，您需要申请开启。详情信息，请参见[SASL用户授权](#)。

后续步骤

获取实例的接入点后，您可以使用SDK接入消息队列Kafka版并收发消息。详细信息，请参见[SDK概述](#)。

1.2. 升级实例配置

本文介绍如何在消息队列Kafka版控制台完成实例升配，升级的配置项包括规格类型、流量规格、磁盘容量及Topic规格等。

应用场景

- 消息队列Kafka版实例的磁盘使用率偏高，即将影响正常业务进行。
- 消息队列Kafka版实例的流量规格持续超过已购买的规格，导致无法满足业务需求。

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。

4. 在实例详情页面，单击概览区域右上角的升配。
5. 在实例升级风险说明对话框，阅读并确认实例升级风险说明，然后单击确认。
6. 在变配页面，变更配置。

参数	说明
规格类型	<ul style="list-style-type: none"> ○ 支持将标准版升级为专业版。 ○ 支持将专业版的流量规格升级为更高流量规格。 ○ 不支持将专业版降低为标准版。 <p>关于消息队列Kafka版实例的规格类型和定价信息，请参见计费说明。</p>
公网流量	<p>公网流量分为双通道，读写一致。购买时请按照读流量峰值或写流量峰值的最大值购买公网流量带宽。该计费项目仅适用于公网/VPC实例。</p>
流量规格	<p>您在升级流量规格时需要关注以下配置：</p> <ul style="list-style-type: none"> ○ 规格类型 <ul style="list-style-type: none"> ■ 标准版：支持最高升级至120 MB/s。如需超过120 MB/s，请先将实例的规格类型升级至专业版，然后再升级流量规格。 ■ 专业版（高写版）：支持最高升级至2000 MB/s。 ■ 专业版（高读版）：支持读最高升级至150 MB/s，支持写最高升级至30 MB/s。 ○ 磁盘类型 <ul style="list-style-type: none"> ■ 高效云盘：超过120 MB/s，会触发集群横向扩容。升级完成后，需要进行Topic引流。具体操作，请参见Topic引流。 ■ SSD：超过300 MB/s，会触发集群横向扩容。升级完成后，需要进行Topic引流。具体操作，请参见Topic引流。
磁盘容量	<p>流量规格有默认推荐的磁盘容量，调整流量规格会相应调整磁盘容量，您也可以根据业务需求自行设置磁盘容量。</p>
Topic规格	<ul style="list-style-type: none"> ○ 每新加购1个Topic，相应增加16个分区的数量额度。 ○ 专业版实例附赠1倍的Topic，即实际可用Topic数为购买Topic数的2倍。

 注意

- 仅支持升级实例配置，不支持降级实例配置。
- 升级实例配置会引发集群逐台重启，可能会出现以下风险：
 - 客户端会短暂断开连接并重连，可能会造成少量报错。
 - 已经发送成功的消息，升级之后不会丢失。升级期间发送失败的消息，建议重试发送，可在客户端配置重试机制。
 - 升级预计持续约30分钟，升级的磁盘容量跨度越大，耗时越长。服务不会中断，可能会导致消费的分区消息发生乱序，请谨慎评估业务影响，建议您在业务低峰期升级实例配置。

7. 阅读并选中服务协议，然后单击去支付。

 说明 升级配置后，订单生效时间以升级订单页描述为准。

执行结果

实例详情页面显示升级后的配置。

1.3. 变更消息配置

您可以按照业务需求调整消息保留时长和最大消息大小。

前提条件

您已购买消息队列Kafka版实例，且实例处于服务中状态。

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在实例详情页面的配置信息区域，选择如下任意一种方式调整消息保留时长与最大消息大小：
 - 单击消息保留时长右侧的编辑，在编辑实例配置信息面板的消息保留时长文本框修改消息保留时长，在最大消息大小文本框修改收发消息的最大值，在消费位点保留时间文本框修改消息消费位点的保留时间，单击确定。具体调整数值，请参见[参数列表](#)。
 - 单击最大消息大小右侧的编辑，在编辑实例配置信息面板的消息保留时长文本框修改消息保留时长，在最大消息大小文本框修改收发消息的最大值，在消费位点保留时间文本框修改消息消费位点的保留时间，单击确定。具体调整数值，请参见[参数列表](#)。
 - 单击消费位点保留时间右侧的编辑，在编辑实例配置信息面板的消息保留时长文本框修改消息保留时长，在最大消息大小文本框修改收发消息的最大值，在消费位点保留时间文本框修改消息消费位点的保留时间，单击确定。具体调整数值，请参见[参数列表](#)。

参数列表

参数	描述
----	----

参数	描述
消息保留时长	在磁盘容量充足的情况下，消息的最长保留时间。 <ul style="list-style-type: none"> ■ 磁盘容量不足（即磁盘水位达到85%）时，将提前删除旧的消息，以确保服务可用性。 ■ 默认值为72小时，取值范围为24小时~480小时。
最大消息大小	消息队列Kafka版能收发的消息的最大值。 <ul style="list-style-type: none"> ■ 消息的最大值上限为10 MB，不区分标准版实例和专业版实例。 ■ 修改该配置前，请确认修改值是否匹配生产和消费客户端相应配置。
消费位点保留时间	消息消费位点的保留时间。默认保留时间为7天，即10080分钟。取值范围为1440分钟~43200分钟。

5. 在修改配置风险说明对话框，阅读并确认修改配置风险，然后单击**确认**。



后续步骤

[查看任务执行记录](#)

1.4. 升级实例版本

本文说明如何升级消息队列Kafka版实例的版本。

前提条件

消息队列Kafka版实例的状态为**服务中**。

背景信息

- 升级大版本

升级大版本是指跨版本升级，例如将消息队列Kafka版实例的版本从0.10.x版本升级至2.x版本。

说明

- 消息队列Kafka版实例支持0.10.x和2.x大版本。其中，0.10.x大版本提供0.10版本和0.10.2版本，2.x大版本只提供2.2.0版本。
- 由于2.x版本的消息存储格式与0.10.x版本不同，您在完成升级之后，将无法回退到0.10.x版本，请您谨慎操作。

升级小版本

升级小版本是指非跨版本升级，例如将消息队列Kafka版实例的版本从0.10版本升级至0.10.2版本，或者从0.10.2版本升级至0.10.2内核优化版。

说明 0.10版本有一定概率触发死锁、频繁Rebalance等问题，建议您将0.10版本升级至0.10.2版本。如何升级，请参见[升级小版本](#)。

注意事项

- 升级不会收取任何费用，且该升级兼容现运行的SDK和OpenAPI。
- 为保证升级过程中业务不受影响，请确保客户端支持重连机制（默认支持）及对断开连接的异常处理。
- 建议您先购买测试实例进行升级验证，再到生产实例进行操作。
- 建议您在业务低峰期执行升级操作。
- 建议您在升级后更新客户端版本，保持客户端和服务端版本一致。
- 升级大版本预计持续约25分钟，升级过程中服务不会中断，正常情况下不会影响业务。
- 升级小版本预计持续约15分钟，升级过程中服务不会中断，正常情况下不会影响业务。

升级大版本

- 登录[消息队列Kafka版控制台](#)。
- 在概览页面的资源分布区域，选择地域。
- 在实例列表页面，单击目标实例名称。
- 在实例详情页面的基础信息区域，单击大版本右侧的升级大版本。
- 在升级实例大版本面板，完成以下操作：
 - 认真阅读升级必读！中的内容。
 - 在紧急联系人文本框，输入您的姓名。
 - 在紧急联系电话文本框，输入您的紧急联系电话。
 - 单击确定。

升级小版本

- 登录[消息队列Kafka版控制台](#)。
- 在概览页面的资源分布区域，选择地域。
- 在实例列表页面，单击目标实例名称。
- 在实例详情页面的基础信息区域，单击小版本右侧的升级小版本。
- 在升级实例小版本面板，完成以下操作：
 - 认真阅读升级必读！中的内容。

- ii. 在紧急联系人文本框，输入您的姓名。
- iii. 在紧急联系电话文本框，输入您的紧急联系电话。
- iv. 单击确定。

1.5. 配置白名单

您可以编辑访问消息队列Kafka版的白名单，配置在白名单中的IP地址与端口才允许访问消息队列Kafka版实例。

前提条件

您已购买消息队列Kafka版实例，且该实例处于服务中的状态。

注意事项

- 白名单最多支持100条。
- 增加白名单时，每一条白名单中可以添加多个IP地址和IP地址段，需使用英文逗号进行分隔。
- 支持删除或增加单条白名单。
- 允许删除最后一条白名单，风险是消息队列Kafka版集群此端口范围不可访问，请谨慎操作。

操作步骤

请按以下步骤添加白名单IP地址或地址段。

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在实例详情页面的接入点信息区域，选择需要配置白名单的接入点，在其操作列，单击编辑白名单。
5. 在编辑接入点白名单面板，单击添加白名单 IP，输入IP地址或地址段，然后单击确定。

相关操作

如需删除白名单配置，可在编辑接入点白名单面板，找到需删除的IP地址或地址段，单击删除。

1.6. Topic引流

您在升级消息队列Kafka版实例的流量规格时，可能会触发集群横向扩容。集群横向扩容完成后，您需要进行Topic引流，使Topic流量重新均匀分布到扩容后的集群上。否则原有的Topic流量还是打在扩容前的集群节点上，原有的Topic的峰值流量会受限于扩容前的峰值流量。新增的Topic不受限于扩容前的流量规格。

前提条件

您的消息队列Kafka版实例处于**服务中（Topic待引流）**状态。

 **说明** 升级实例的流量规格操作以及集群横向扩容触发规则，请参见[升级实例配置](#)。

注意事项

消息队列Kafka版实例处于**服务中（Topic待引流）**状态时，您可以正常使用该实例收发消息，但不能在该实例下创建Topic、Consumer Group等资源。您必须完成Topic引流或者选择不引流，才能重新创建资源。

引流方式

消息队列Kafka版支持的引流方式如下。

引流方式	原理	影响	适用场景	持续时间
新增分区	为原集群节点上的所有Topic在扩容后的新节点中增加分区。	<ul style="list-style-type: none"> 分区消息乱序。 分区数量改变。如果您的客户端无法自动感知到新分区（例如：指定分区发送消息以及某些流计算场景），您可能需要重启或者修改客户端代码。 	<ul style="list-style-type: none"> 不要求分区顺序。 不指定分区发送。 消费方式采取订阅。 	秒级。
迁移分区（推荐）（推荐）	<ul style="list-style-type: none"> Local存储：使用kafka-reassign-partitions工具迁移分区数据。 云存储：修改映射关系，不迁移分区数据。 	<ul style="list-style-type: none"> Local存储：临时性的内部流量。 云存储：无临时性的内部流量。 	任何集群扩容场景。	<ul style="list-style-type: none"> Local存储：分钟级或小时级。取决于要迁移的Local存储数据量。如果数据量较大，可能持续几小时甚至更久，您需要谨慎评估。建议您在业务流量低峰期执行迁移操作。 云存储：秒级。迁移1个Topic大约需要30秒。
不引流（不推荐）	不进行任何操作，即原有的Topic依旧分布在扩容前的集群节点上，新增的Topic均衡分布到扩容后的所有集群节点上。	<ul style="list-style-type: none"> 原有的Topic峰值流量会受限于扩容前的流量规格。 如果原有的Topic流量较大，可能会出现集群节点之间流量不均衡。 	<ul style="list-style-type: none"> 原有的Topic流量非常小，并且集群扩容后原有的Topic流量没有较大提升。 集群扩容后会新建Topic，并且绝大部分流量会打在新建的Topic上。 	立即生效。

操作步骤

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在实例详情页面的概览区域的右上角，单击Topic引流。
5. 在实例的Topic引流对面板，选择引流方式。

- 新增分区
 - 选择新增分区，然后单击确定。
- 迁移分区
 - a. 提交工单联系消息队列Kafka版技术人员将服务端升级至最新版本。
 - b. 选择迁移分区（推荐），然后单击确定。
- 不引流
 - 选择不引流（不推荐），然后单击确定。

执行结果

Topic引流完成后，实例运行状态显示服务中。

1.7. 查看任务执行记录

在某个实例下发起重启类任务（包括配置变更、版本升级、开启ACL、实例升配）后，您可以在消息队列Kafka版控制台查看该实例的重启类任务的执行记录，获取任务的类型、开始执行时间、结束执行时间、状态等信息。

前提条件

创建并部署消息队列Kafka版实例，且实例处于服务中状态。

操作步骤

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在实例详情页面，单击任务执行记录页签。
在任务执行记录页签下方显示重启类任务的执行记录。

参数	描述	示例值
任务类型	执行的任务的类型。取值： <ul style="list-style-type: none"> ○ 配置变更：变更实例的消息配置，包括消息保留时长和最大消息大小。更多信息，请参见变更消息配置。 ○ 实例升配：升级实例的配置，包括规格类型、实例类型、流量规格、磁盘容量、Topic规格。更多信息，请参见升级实例配置。 ○ 版本升级：升级实例的版本，包括大版本和小版本。更多信息，请参见升级实例版本。 ○ 开启ACL：ACL是消息队列Kafka版提供的用于管理SASL用户及其资源访问权限的服务。更多信息，请参见SASL用户授权。 	实例升配

参数	描述	示例值
状态	任务的当前状态。取值： <ul style="list-style-type: none"> 未执行 执行中 已执行 已取消 	已执行
开始执行时间	任务开始执行的时间。	2021年5月26日20:51:11
结束执行时间	任务结束执行的时间。	2021年5月26日20:59:32
耗时	任务执行所消耗的时长。	8分20秒

1.8. 实例问题

1.8.1. 实例支持哪些开源版本？

本文介绍消息队列Kafka版实例支持哪些开源版本。

- 消息队列Kafka版标准版实例：仅支持0.10.x，默认部署0.10.x，如需更高版本请先将实例升级为专业版，然后升级实例版本，请参见[升级实例配置](#)和[升级实例版本](#)。
- 消息队列Kafka版专业版实例：支持0.10.x~2.x，默认部署0.10.x，如需升级实例版本，请参见[升级实例版本](#)。

说明

- 2.x向前兼容0.10.x、0.9.0。
- 0.10.x向前兼容0.9.0。

1.8.2. 如何选择实例规格？

消息队列Kafka版提供两种实例规格：标准版和专业版。您可以根据自建迁移情况选择实例规格。

- 标准版
 - 流量规格：集群流量的总量÷3（优化）
 - 磁盘大小：流量均值×存储时长×3（备份）
 - Topic数量：根据实际业务需求

说明 建议您在迁移上云过程中优化Topic以降低成本。

- 专业版
 - 流量规格：集群流量的总量÷3（优化）

- 磁盘大小：流量均值×存储时长×n（备份）

说明 在创建Topic时，选择云存储则n=1，选择Local存储则n=3。云存储和Local存储的对比，请参见[存储引擎对比](#)。

- Topic数量：根据实际业务需求

说明 建议您在迁移上云过程中优化Topic以降低成本。

1.8.3. 如何获取部署实例时需要的VPC信息？

您可以在VPC管理控制台获取VPC信息。

获取VPC信息的步骤如下：

1. 登录[VPC管理控制台](#)。
2. 在左侧导航栏，单击[交换机](#)。
3. 在[交换机](#)页面，查看以下信息：
 - 交换机实例ID/名称
 - 专有网络
 - 可用区

实例ID/名称	专有网络	标签	状态	IPv4网段	可用IP数	IPv6网段	默认交换机	可用区	路由表	路由表类型
实例ID/名称	defaultvpc		可用		4091	开通IPv6	是	可用区		系统

说明 请根据该页面的可用区（A~G）在消息队列Kafka版控制台中选择对应的可用区（A~G）。例如，某VPC交换机（VSwitch）显示在可用区B，那么在消息队列Kafka版控制台中就相应地选择可用区B。

1.8.4. 升级Broker可能产生哪些影响？

升级Broker可能产生消息乱序、客户端连接中断、消息量不均衡等影响。

升级Broker包含以下影响：

- 升级过程中，会逐个重启消息队列Kafka版集群中所有的Broker。在重启Broker的过程中服务不会中断，但是从每个Broker重启完成之后的5分钟内消费的分区消息可能会发生乱序（设置为顺序消息的，则不会乱序，而是短暂的不可用）。
- 重启过程中已有的客户端连接可能会中断。需要您的客户端有自动重连功能，服务端的其他Broker会自动接替服务。
- 此外，升级和重启Broker期间，各个分区处理的消息量也会出现一定的不均衡，需要您评估一下升级变更对您业务可能产生的影响。

升级所有Broker大概需要5分钟~15分钟。如果有多个实例，可以考虑先升级测试集群，验证通过后再升级生产集群。

1.8.5. 实例的地域或网络属性无法变更

如需变更实例的地域或网络属性，请释放实例，并重新购买。

实例购买部署之后，其地域和网络属性与物理资源紧密结合，无法变更。如需变更实例的地域或网络属性，请释放实例，并重新购买。

1.8.6. 消息队列Kafka版何时删除旧消息？

为避免因磁盘容量不足而导致机器宕机，进而影响服务可用性，消息队列Kafka版对磁盘使用率进行了动态控制。

- 磁盘使用率低于85%：每天凌晨4点集中删除超过存储时间的消息。
- 磁盘使用率达到85%：立即清除超过存储时间的消息。
- 磁盘使用率达到90%：无论消息是否超过存储时间，按服务端存储消息的时间先后顺序清除消息。

 **说明** 一般情况下，为了保证您的业务健康性（拥有充足的消息回溯能力），建议您的磁盘使用率不要超过70%。

1.8.7. 如何快速测试消息队列Kafka版服务端是否正常？

在创建并部署消息队列Kafka版实例后，您可以使用消息队列Kafka版控制台直接发送消息，快速测试服务端是否正常。

前提条件

您已创建并部署消息队列Kafka版实例，且实例处于**服务中**状态。

操作流程

快速测试消息队列Kafka版服务端的流程如下：

1. [创建Topic](#)
2. [发送消息](#)
3. [查看分区状态](#)
- 4.

您可以多次重复步骤2到步骤4，如果多次操作正常，则说明服务端正常。

 **说明** 如果服务端正常，但发送消息依然失败，建议您去调用方（例如，原生客户端、生态组件端等）排查问题。

创建Topic

创建用于接收消息的Topic。

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击**Topic 管理**。
5. 在**Topic 管理**页面，单击**创建 Topic**。

6. 在创建 Topic 面板，设置Topic属性，然后单击确定。

创建 Topic
✕

*** 名称**

长度限制为 3 ~ 64 个字符，只能包含英文、数字、短横线 (-) 以及下划线 (_)，且至少包含一个英文或数字。

*** 描述** 9/64

*** 分区数**

建议分区数是6的倍数，减少数据倾斜风险，分区数限制 (1 ~ 360)，特殊需求请提交工单。

存储引擎 云存储

i 底层接入阿里云云盘，具有低时延、高性能、持久性、高可靠等特点，采用分布式3副本机制。

消息类型 普通消息

i 默认情况下，保证相同 Key 的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，可能会造成消息乱序。

标签

参数	说明	示例
名称	Topic名称。	demo
描述	Topic的简单描述。	demo test
分区数	Topic的分区数量。	12
存储引擎	Topic消息的存储引擎。 消息队列Kafka版支持以下两种存储引擎。您可以了解界面的解释信息，选择满足业务需求的方式。 <ul style="list-style-type: none"> ◦ 云存储 ◦ Local 存储 	云存储
消息类型	Topic消息的类型。 <ul style="list-style-type: none"> ◦ 当存储引擎选择云存储时，默认选择普通消息。 ◦ 当存储引擎选择Local 存储时，默认选择分区顺序消息。 	普通消息

参数	说明	示例
日志清理策略	<p>Topic日志的清理策略。</p> <p>当存储引擎选择Local 存储时，需要配置日志清理策略。</p> <p>消息队列Kafka版支持以下两种日志清理策略。您可以了解界面的解释信息，选择满足业务需求的方式。</p> <ul style="list-style-type: none"> ◦ Delete ◦ Compact 	Compact
标签	Topic的标签。	demo

发送消息

向已创建的Topic发送消息。

1. 在Topic 管理页面，找到目标Topic，在其操作列中，选择更多 > 体验发送消息。
2. 在快速体验消息收发面板，发送测试消息。
 - 发送方式选择控制台。
 - a. 在消息 Key文本框中输入消息的Key值，例如demo。
 - b. 在消息内容文本框输入测试的消息内容，例如 {"key": "test"}。
 - c. 设置发送到指定分区，选择是否指定分区。
 - a. 单击是，在分区 ID文本框中输入分区的ID，例如0。如果您需查询分区的ID，请参见[查看分区状态](#)。
 - b. 单击否，不指定分区。
 - d. 根据界面提示信息，通过SDK订阅消息，或者执行Docker命令订阅消息。
 - 发送方式选择Docker，运行Docker容器。
 - a. 执行运行 Docker 容器生产示例消息区域的Docker命令，发送消息。
 - b. 执行发送后如何消费消息？区域的Docker命令，订阅消息。
 - 发送方式选择 SDK，根据您的业务需求，选择需要的语言或者框架的SDK以及接入方式，通过SDK体验消息收发。

查看分区状态

往指定分区发送消息后，查看该分区的状态。

1. 在Topic 管理页面，找到目标Topic，在其操作列中，选择更多 > 分区状态。分区状态信息

参数	说明
分区ID	该Topic分区的ID号。
最小位点	该Topic在当前分区下的最小消费位点。
最大位点	该Topic在当前分区下的最大消费位点。

参数	说明
最近更新时间	本分区中最近一条消息的存储时间。

配置信息	订阅关系	分区状态	云监控	消息查询
当前 Topic 中每个分区的具体状态				
分区 ID	最小位点	最大位点	最近更新时间	
0	0	5	2021年5月8日14:09:01	
1	0	3	2021年5月8日14:09:01	
2	0	1	2021年5月7日18:11:31	
3	0	8	2021年5月11日22:28:10	
4	0	0	--	
5	0	0	--	
6	0	0	--	
7	0	0	--	
8	0	7	2021年5月8日14:09:01	
9	0	0	--	
10	0	3	2021年5月7日18:11:32	
11	0	2	2021年5月7日18:11:31	

按位点查询消息

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击消息查询。
5. 在消息查询页面的查询方式列表中，选择按位点查询。
6. 在Topic列表中，选择消息所属Topic名称；在分区列表中，选择消息所属的分区；在起始位点文本框，输入消息所在分区的位点，然后单击查询。

展示该查询位点及以后连续的消息。例如，指定的分区和位点都为“5”，那么返回的结果从位点“5”开始。

查询结果参数解释

参数	描述
分区	消息的Topic分区。
位点	消息的所在的位点。
Key	消息的键（已强制转化为String类型）。
Value	消息的值（已强制转化为String类型），即消息的具体内容。

参数	描述
消息创建时间	<p>发送消息时，客户端自带的或是您指定的 <code>ProducerRecord</code> 中的消息创建时间。</p> <p>说明</p> <ul style="list-style-type: none"> 如果配置了该字段，则按配置值显示。 如果未配置该字段，则默认取消息发送时的系统时间。 如果显示值为1970/x/x x:x:x，则说明发送时间配置为0或其他有误的值。 0.9及以前版本的消息队列Kafka版客户端不支持配置该时间。
操作	<ul style="list-style-type: none"> 单击下载 Key：下载消息的键值。 单击下载 Value：下载消息的具体内容。 <p>注意</p> <ul style="list-style-type: none"> 查询到的每条消息在控制台上最多显示1 KB的内容，超过1 KB的部分将自动截断。如需查看完整的消息内容，请下载相应的消息。 仅专业版支持下载消息。 下载的消息最大为10 MB。如果消息超过10 MB，则只下载10 MB的内容。

1.8.8. 是否支持延迟消息？

和开源Apache Kafka一样，消息队列Kafka版同样不支持延迟消息。

1.8.9. 能否通过公网访问消息队列Kafka版实例？

网络类型为公网/VPC的消息队列Kafka版实例支持通过公网访问。

消息队列Kafka版提供VPC和公网/VPC两种网络类型的实例。如需通过公网访问，请购买公网/VPC实例。中国内地与非中国内地通过公网连接可能会出现异常。

注意 实例类型只能在购买时决定，购买之后无法变更。

1.8.10. 不同VPC之间是否支持互连？

不同VPC之间支持互连，您可以使用云企业网或VPN网关连接不同的VPC。

云企业网

云企业网CEN（Cloud Enterprise Network）帮助您在VPC间搭建私网通信通道，通过自动路由分发及学习，提高网络的快速收敛和跨网络通信的质量和安全性，实现全网资源的互通。详细信息，请参见[云企业网](#)。

您可以通过云企业网实现同账号VPC互通和跨账号VPC互通，互通场景如下表。

场景	配置方法
同账号VPC互通	同账号同地域VPC互连
	同账号跨地域VPC互连
跨账号VPC互通	跨账号同地域VPC互连
	跨账号跨地域VPC互连

云企业网的优势如下：

- **一网通天下**
云企业网打造的是一张能够实现阿里云全球网络资源互联、并能够与接入阿里云的网络资源互联的企业级网络。全网通过IP地址唯一性管理，避免地址冲突问题。用户不需要额外配置，网络通过控制器实现多节点、多级路由的自动转发与学习，实现全网的路由快速收敛。
- **低时延高速率**
云企业网提供低延迟、高速率的网络传输能力。本地互通最大速率可达到设备端口转发速率。在全球互通的时延中，整体时延较公网互通时延有很大提升。
- **就近接入与最短链路互通**
云企业网在全球超过60个地域部署了接入及转发节点，方便全球用户就近接入阿里云，避免绕行公网带来的时延及业务受损。
- **链路冗余及容灾**
云企业网具有高可用及网络冗余性，全网任意两点之间至少存在4组独立冗余的链路。即使部分链路中断，云企业网也可以保证客户的业务正常运行，不会发生抖动及中断。
- **系统化管理**
云企业网具有系统化的网络监控能力，自动检测由于系统变更而导致的路由冲突，保证网络运行的稳定性。

VPN网关

VPN网关是一款基于Internet的网络连接服务，支持基于路由的IPsec-VPN功能。您可以使用IPsec-VPN将不同的VPC进行安全可靠的连接。详细信息，请参见[建立VPC到VPC的连接](#)。

VPN网关的优势如下：

- **安全**
使用IKE和IPsec协议对传输数据进行加密，保证数据安全可靠。
- **高可用**
采用双机热备架构，故障时秒级切换，保证会话不中断，业务无感知。
- **成本低**
基于Internet建立加密通道，比建立专线的成本更低。
- **配置简单**

开通即用，配置实时生效，快速完成部署。

1.8.11. 经典网络是否必须要选择172.16.0.0/12网段？

不是必须，但使用172.16.0.0/12网段的配置最简单。

如需选择其他网段，请参见[ClassicLink概述](#)进行配置。

1.8.12. 是否支持压缩消息？

消息队列Kafka版服务端支持收发压缩消息。

如需使用压缩消息，您需要在消息队列Kafka版的客户端进行设置。在消息队列Kafka版客户端进行消息压缩的说明如下：

- 压缩格式：支持Snappy、LZ4、GZIP等压缩格式。其中，GZIP对CPU的消耗较高，因此不建议您选择GZIP，建议您选择Snappy或LZ4。
- 适用场景：一般来说，CPU的价格比流量和存储要高。对于日志类等压缩比较高的场景，您可以考虑使用压缩。其余场景，不建议您使用压缩。
- CPU消耗：压缩会消耗额外的CPU，平均在20%以上。具体额外CPU消耗，需要您根据实际场景进行测试。

2.Topic

2.1. 查看分区状态

您可以通过查看分区状态来了解服务端的消息总量或各个分区的消费进度。

前提条件

[创建Topic](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Topic 管理。
5. 在Topic 管理页面，找到目标Topic，在其操作列中，选择更多 > 分区状态。分区状态信息

参数	说明
分区ID	该Topic分区的ID号。
最小位点	该Topic在当前分区下的最小消费位点。
最大位点	该Topic在当前分区下的最大消费位点。
最近更新	本分区中最近一条消息的存储时间。

配置信息	订阅关系	分区状态	云监控	消息查询
当前 Topic 中每个分区的具体状态				
分区 ID	最小位点	最大位点	最近更新	
0	0	5	2021年5月8日 14:09:01	
1	0	3	2021年5月8日 14:09:01	
2	0	1	2021年5月7日 18:11:31	
3	0	8	2021年5月11日 22:28:10	
4	0	0	--	
5	0	0	--	
6	0	0	--	
7	0	0	--	
8	0	7	2021年5月8日 14:09:01	
9	0	0	--	
10	0	3	2021年5月7日 18:11:32	
11	0	2	2021年5月7日 18:11:31	

2.2. Topic问题

2.2.1. 为什么限制Topic总数（分区总数）？

Topic总数（分区总数）太多会使集群性能和稳定性能急剧下降。

消息队列Kafka版的存储和协调机制是以分区为粒度的，分区数太多，会导致存储碎片化严重，集群性能和稳定性都会急剧下降。

2.2.2. 为什么Topic不能减分区？

Topic减分区会造成数据丢失。

这是Apache Kafka自身设计所限制的。

2.2.3. 新增Topic时对应的分区数量额度

每新增1个Topic，会相应地增加16个分区数量额度。

在默认分区数量额度的基础上，每增购1个Topic，对应增加16个分区数量额度。例如，您购买了Topic数量为50、流量规格为20 MB/s、默认分区数为400的实例，那么当您增购10个Topic后，该实例对应增加160个分区数量额度，分区数量额度上升至560个。

2.2.4. Topic的存储有哪些类型？

云存储和Local存储。

消息队列Kafka版实例支持云存储和Local存储。详情请参见[存储引擎对比](#)。

2.2.5. 专业版实例附赠1倍的Topic

专业版实例的可用Topic数为购买Topic数的两倍。

专业版实例的可用Topic数=购买Topic数×2。例如，您购买了Topic数量为50的专业版实例，那么该实例实际可用的Topic数为100。

2.2.6. 为什么Topic有分区时间显示1970？

该Topic分区没有被发送过消息。

如果没有发送过消息到某Topic分区，那么对应Topic分区的最近更新时间就是显示为1970。

2.2.7. Topic删除后不能减少分区？

彻底清除路由后，可以重新指定Topic的分区数量。

Condition

您在删除某个分区数为XX的Topic后，再次创建该Topic并将分区数配置为小于XX，系统提示**创建topic成功**，分区数不能小于曾经配置过的分区数，已重置为XX。

Cause

老版本删除Topic时，不会彻底清除路由，导致再次创建Topic时无法减少分区。升级到新版本后，那些曾经在老版本删除过的Topic，其路由仍然保留着，为了对其进行彻底清除，需要“创建->删除->再创建”Topic，待该流程完成之后，就可以随意指定分区数量了。

Remedy

操作步骤

1. 确保实例的内部版本为最新版。在[消息队列Kafka版控制台](#)，进入实例详情的基础信息区域，在小版本右侧查看实例版本。
 - 如果显示为当前版本为最新版本，则无需处理。
 - 如果显示为升级小版本，请单击升级小版本，完成版本升级。
2. 创建->删除->再创建Topic。进入Topic 管理页面，创建之前删除的Topic，删除该Topic，然后再次创建该Topic并配置分区数。

2.2.8. 是否支持Compact的日志清理策略？

开源版本为2.2.0或以上的消息队列Kafka版实例支持Compact的日志清理策略。

开源版本为2.2.0或以上的消息队列Kafka版实例支持Compact的日志清理策略。如需升级实例的开源版本以支持Compact的日志清理策略，请参见[升级大版本](#)。

3.Consumer Group

3.1. 重置消费位点

重置消费位点是指改变订阅者当前的消费位置。您可通过重置消费位点，按需清除堆积的或不想消费的这部分消息，或直接跳转到某个时间点，从指定时间点的位点开始消费消息。

前提条件

已停止所有Consumer客户端（消息队列Kafka版不支持在线重置消费位点）。

 **注意** 在停止Consumer客户端后，需要经过 `ConsumerConfig.SESSION_TIMEOUT_MS_CONFIG` 配置的时间（默认10000 ms），服务端才认为Consumer真正下线。

背景信息

消息队列Kafka版支持以下重置消费位点方式：

- 消息清除：因为某种原因，订阅者不准备再继续消费Broker上堆积的消息，此时通过清除消息，可以将消费位点指定到最新的位置。

 **注意** 堆积的消息本身并不会因此被删除，改变的只是消费位点。

- 指定时间点开始消费：将某个Consumer Group的位点重置到过去或者将来的某个时间点（该时间点以Topic的消息存储时间为准，假设为“t”）。重置过后，Consumer Group将从“t”时间点的位点开始消费。

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Group 管理。
5. 在Group 管理页面，找到目标Group，在其右侧操作列，选择更多 > 重置消费位点。
6. 在重置Group的消费位点面板，了解其前提条件，设置重置策略。
 - i. 设置重置所有 Topic。
 - 单击是，重置所有Topic的消费位点。
 - 单击否，在Topic文本框输入需要重置Topic的名称。
 - ii. 设置重置方式。
 - 单击从最新位点开始消费，将消费位点指定到最新的位置，单击确定。
 - 单击从指定时间点的位点开始消费，在时间点文本框，单击，从指定时间点的位点开启消费功能，单击确定。
7. 在提示对话框，确认提示信息，单击确认。

3.2. 查看消费状态

当消费堆积或倾斜时，您可以查看Consumer Group和Topic之间的订阅关系，查看Topic各个分区的消费进度，了解消息的堆积总量，及时调整业务，预防风险。

查看Topic被订阅的Consumer Group

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Topic管理。
5. 在Topic管理页面，找到目标Topic，在其操作列，选择更多 > 订阅关系。在Topic 详情页面的订阅关系页签，显示所有订阅该Topic的Group。



6. 在订阅当前 Topic 的在线 Group 的列表中，找到目标Group，在其右侧操作列，单击消费详情。在消费详情列表，显示该Topic各个分区的消费详情。

分区 ID	客户端	最大位点	消费位点	堆积量	最近消费时间
0	consumer-4_172.16.10.128	6	6	0	2021年5月27日10:58:59
1	consumer-4_172.16.10.128	3	3	0	2021年5月8日14:09:01
2	consumer-4_172.16.10.128	3	3	0	今天 00:08:54
3	consumer-4_172.16.10.128	8	8	0	2021年5月11日22:28:10
4	consumer-4_172.16.10.128	2	2	0	2021年5月27日17:48:28
5	consumer-4_172.16.10.128	6	6	0	今天 00:38:59
6	consumer-4_172.16.10.128	2	2	0	2021年5月27日17:48:28

参数	说明
分区 ID	该Topic分区的ID。
客户端	实时订阅该Topic的在线Consumer。 <div style="border: 1px solid #ccc; padding: 5px; margin-top: 5px;"> <p>说明</p> <ul style="list-style-type: none"> 取值格式为 <Client ID>_<IP>。 非在线消费端则无法查看客户端信息。 </div>
最大位点	该Topic在当前分区下的最大消息消费位点。
消费位点	该Topic在当前分区下的消息消费位点。

参数	说明
堆积量	<p>当前分区下的消息堆积总量，即最大位点减去消费位点的值。</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p>注意 堆积总量是一个关键指标，如果发现堆积总量较大，则应用消费端可能产生了阻塞，或者消费速度跟不上生产速度。此时需要分析消费端的运行状况，提升消费速度。更多信息，请参见重置消费位点。</p> </div>
最近消费时间	该Consumer Group消费的最近一条消息被发送到Broker存储的时间。

查看Consumer Group订阅的Topic

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Group 管理。
5. 在Group 管理页面，找到目标Group，在其操作列，选择更多 > 消费状态。
在Group 详情的消费状态页签，显示该Group订阅的所有Topic以及各个Topic的堆积量和最近消费时间。

6. 在Topic列表，找到目标Topic，在其右侧操作列，单击消费详情。
在消费详情列表，显示该Group在相应Topic的各个分区的消费详情。

分区 ID	客户端	最大位点	消费位点	堆积量	最近消费时间
0	consumer-1_1719.0.0.0:9092	6	6	0	2021年5月27日10:58:59
1	consumer-1_1719.0.0.0:9092	3	3	0	2021年5月8日14:09:01
2	consumer-1_1719.0.0.0:9092	3	3	0	今天 00:08:54
3	consumer-1_1719.0.0.0:9092	8	8	0	2021年5月11日22:28:10
4	consumer-1_1719.0.0.0:9092	2	2	0	2021年5月27日17:48:28
5	consumer-1_1719.0.0.0:9092	6	6	0	今天 00:38:59
6	consumer-1_1719.0.0.0:9092	2	2	0	2021年5月27日17:48:28

3.3. Consumer Group问题

3.3.1. Consumer Group规格是多少？

Consumer Group规格默认为Topic规格的2倍。

Consumer Group规格默认为Topic规格的2倍。例如，某个实例的Topic规格为50个，则该实例下最多能创建100个Consumer Group。如需升级Consumer Group规格，您可升级Topic规格，每增加1个Topic配额，对应增加2个Consumer Group配额，详情请参见[升级实例配置](#)。

3.3.2. 为什么有Consumer Group对应的Topic分区显示1970?

Consumer Group没有消费过该Topic分区的信息。

如果Consumer Group没有消费过某Topic分区的信息，那么Consumer Group对应的Topic分区的最近消费时间就是显示为1970。

3.3.3. Consumer Group删除后依然可以收到消息?

老版本删除Consumer Group时，不会彻底清除路由。

Condition

某个Consumer Group被删除后，依然可以收到消息。

Cause

老版本删除Consumer Group时，不会彻底清除路由，导致被删除的Consumer Group依然可以收到消息。升级到新版本后，那些曾经在老版删除过的Consumer Group，其路由仍然保留着，为了对其进行彻底清除，需要“创建->删除”Consumer Group，待该流程完成之后，您可以再次创建Consumer Group。

Remedy

操作步骤

1. 将服务版本升级为最新版。在[消息队列Kafka版控制台](#)，进入实例详情的基础信息区域，在小版本右侧查看实例版本。
 - 如果显示为当前版本为最新版本，则无需处理。
 - 如果显示为升级小版本，请单击升级小版本，完成版本升级。
2. “创建->删除->再创建”Consumer Group。进入Group管理页面，创建之前删除的Consumer Group，删除该Consumer Group，然后再次创建Consumer Group。

3.3.4. 为什么分区消费时间相差较大、滞后明显或乱序?

分区分配的消费端不均。

现象

在消息队列Kafka版控制台的Consumer Group页签，单击查看消息堆积，显示某分区的最近消费时间比其他分区早，说明分区先收到来自Producer的消息。

原因

同一个Consumer Group的所有消费端实例，会按分区均匀分配。在此条件下，消费的消息是否均匀，取决于客户的Producer在发送消息时，发送消息对各个分区是否均匀：

- 如果消费端个数N可以整除24（默认分区个数），那么分区对消费端的分配就是均匀的。
- 如果消费端个数N不能整除24（默认分区个数），那么有的消费端处理的分区个数会比其他消费端大。

假设有5个消费端和24个分区，那么会有4个消费端每个消费5个分区，而余下的1个消费端只消费4个分区。消费的速度主要取决于消费端（您的客户端）的处理性能。如果处理性能完全一样，在此条件下分配到的分区个数多1的消费端，消费可能会稍慢。

解决方案

保证消费端个数N整除24。

3.3.5. 控制台看到的最近消费时间是什么意思？

消息发布到消息队列Kafka版服务器的存储时间。

控制台看到的最近消费时间是指消息发布到消息队列Kafka版服务器的存储时间。如果消费没有堆积，那么这个时间接近发送时间。

3.3.6. 如何查看哪些IP在消费消息？

在控制台根据Consumer Group查看消费状态。

在控制台根据Consumer Group查看消费状态，单击详情，可以看到各个分区的owner，进而看到对应的IP。

如果看到owner为空，说明客户端不在线或者处于Rebalance中。

3.3.7. 哪里可以找到消费最佳实践？

订阅者最佳实践。

请参见[订阅者最佳实践](#)。

3.3.8. 消费端挂载NFS云盘是否会影响消费端处理速度？

如果客户端的存储影响消费速度，说明存储是在处理消息的主线程里同步操作的，这会阻塞拉取和处理消息。

建议考虑把处理消息和存储处理结果用独立的且不同的线程，拉取消息只管消费，把消息转给缓存处理后就继续消费消息，这样可以保证消费较快。

需要注意，NFS有两方面影响性能：

- NFS本身速度就不太理想。
- NFS是网络共享存储，虽然有多机器共享访问的能力优势，但是多台机器和进程访问是争抢的，效率会降低。这解释了为什么消费者个数多了，性能反而下降。这一方面，可以考虑采用云盘，给每台消费端处理机挂载自己的云盘，独立存储，这样增加消费端不会因为争抢NFS而降低性能。

挂载云盘的存储是各自独立的。如果您需要把最终的处理结果集中到同一个NFS上保存，仍然可以通过一个异步的工具或者线程，把云盘上的结果再转发到NFS上，而不要让同步存储NFS阻塞消息处理。

总之，对于资源访问造成的处理低效，总是可以用异步处理的方式解决。

3.3.9. 如何在控制台管理Consumer的offset?

本文说明Consumer在读取消息异常中断后，如何在控制台管理Consumer的offset。

消费消息并不保证会提交消费位点，Broker记录的是客户端提交的消费位点。

提交消费位点的机制取决于您使用的客户端SDK，一般支持以下两种机制：

- 自动提交：按照时间间隔，SDK把消费过的最新消息的位点+1提交上去。
- 手动提交：应用程序里，把消费过的最新消息的位点+1提交上去。

您可在控制台的**Group 管理**页面，找到目标Group，在其操作列，选择**更多 > 消费状态**，查看Consumer提交的**消费位点**。Consumer消费时，从这个位点继续消费。具体操作，请参见[查看消费状态](#)。

在管理控制台上操作，可以人工移动Broker记录的消费位点。可以往前移动，重复消费，也可以往后移动，跳过消费。



注意 在控制台上**重置消费位点**，需要先停止消费客户端。否则，重置的结果很可能被消费端覆盖掉。

4. 标签

4.1. 标签管理

本文介绍如何为消息队列Kafka版的实例、Topic、Consumer Group等资源绑定标签，从而帮助您根据标签给资源归类，高效管理资源。

背景信息

标签由标签键和标签值组成，可以标记消息队列Kafka版的资源。如果您的账号下有多种消息队列Kafka版资源，且不同的资源类型之间有多种关联，您可以通过标签将作用相同或者相关联的资源进行分组归类。日常运维或者定位问题时，您便可以快速检索资源，进行批量操作，高效修复故障。

使用说明

- 标签都由一对键值对（Key-Value）组成。
- 资源的任一标签的标签键（Key）必须唯一。

例如，消息队列Kafka版实例先绑定了 `city:shanghai`，后续如果绑定 `city:newyork`，则 `city:shanghai` 自动被解绑。

- 不同地域中的标签信息不互通。例如在华东1（杭州）地域创建的标签在华东2（上海）地域不可见。
- 解绑标签时，如果解绑之后该标签不再绑定任何资源，则该标签会自动被删除。
- 如何设计标签键值，请参见[标签设计最佳实践](#)。

使用限制

- 每个资源最多可绑定20个标签。
- 每个标签最多可绑定50个资源。
- 单次操作绑定或解绑标签的数量分别不能超过20个。

绑定标签

为消息队列Kafka版的实例、Topic、Consumer Group等资源绑定标签，方便资源归类。

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在左侧导航栏，单击[标签管理](#)。
4. 在[标签管理](#)页面，单击[创建标签](#)。
5. 在[创建标签配置向导](#)的[选择标签](#)页面，创建标签，然后单击[下一步](#)。

 **注意** 标签的数量最多不超过20个，且不能重复。

- i. 在[标签键](#)文本框输入标签键，或者在下拉列表中选择已存在的标签键。
 - ii. 在[标签值](#)文本框输入标签值，或者在下拉列表中选择已存在的标签值。
6. 在[创建标签配置向导](#)的[绑定资源](#)页面，给标签绑定资源，单击[绑定](#)。

 **注意** 您最多可选择50个同类型的资源。

- 单击实例列表页签，选中需要绑定的实例。
- 单击Topic 列表页签，选中需要绑定的Topic。
- 单击Group 列表页签，选中需要绑定的Group。

在标签管理页面，可以看到已经创建的标签，以及标签绑定的资源。单击其操作列的绑定资源，还可修改绑定的资源。

编辑标签

当某个实例添加的标签不便于资源的归类或检索，需要修改时，您可以给该实例绑定其他已有的标签或者重新创建合适的标签，高效管理资源。

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 根据实际需求，编辑实例标签、Topic标签或者Consumer Group的标签。
 - 编辑实例的标签
 - a. 在实例列表页面，找到需要编辑标签的实例，在其标签列，将鼠标悬停在图标之上，然后单击编辑。
 - b. 在编辑实例的标签面板，找到需修改的标签键。
 - c. 在标签键文本框重新编辑标签键，在标签值文本框重新编辑标签值，然后单击确定。
 - 编辑Topic的标签
 - a. 在实例列表页面，单击目标实例名称。
 - b. 在左侧导航栏，单击Topic 管理。
 - c. 在Topic 管理页面，找到需修改标签的Topic，在其标签列，将鼠标悬停在图标之上，然后单击编辑。
 - d. 在编辑Topic的标签面板，找到需修改的标签键。
 - e. 在标签键文本框重新编辑标签键，在标签值文本框重新编辑标签值，然后单击确定。
 - 编辑Group的标签
 - a. 在实例列表页面，单击目标实例名称。
 - b. 在左侧导航栏，单击Group 管理。
 - c. 在Group 管理页面，找到需修改标签的Group，在其标签列，将鼠标悬停在图标之上，然后单击编辑。
 - d. 在编辑Group的标签面板，找到需修改的标签键。
 - e. 在标签键文本框重新编辑标签键，在标签值文本框重新编辑标签值，然后单击确定。

标签检索

为资源绑定标签后，您可以使用标签对资源进行检索。

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在左侧导航栏，单击标签管理。
4. 在标签管理页面的选择标签键的下拉列表中，选择目标标签键，页面显示此标签键对应的标签值，在

搜索框中模糊搜索标签值，找到目标标签值之后，单击其右侧操作列的详情。

 说明 若无标签值，则显示为空。

5. 在**标签详情**页面，分别在**实例列表**、**Topic 列表**或**Group 列表**页签，可查看该标签绑定的资源。

解绑标签

当标签被绑定到不适用的资源，或者不再适用于管理和检索资源，您可以批量解除标签和资源之间的绑定。

 注意 单次最多可以解绑20个标签。解除绑定后，如果标签绑定的资源数量为零，标签会被自动删除。

1. 登录**消息队列Kafka版控制台**。
2. 在**概览**页面的**资源分布**区域，选择地域。
3. 在左侧导航栏，单击**标签管理**。
4. 在**标签管理**页面的**选择标签键**的下拉列表中，选择需解绑的标签键，页面显示此标签键对应的标签值，找到需解绑的标签值，单击其右侧**操作列的详情**。
5. 在**标签详情**页面，分别在**实例列表**、**Topic 列表**或**Group 列表**页签，选择一个或者多个需解绑的资源。
 - 单个解绑：在资源列表，找到需解绑的资源，单击其**操作列的解绑资源**。
 - 批量解绑：在资源列表，选中需解绑的所有资源，单击资源列表下方的**解绑资源**。
6. 在提示对话框，认真阅读提示信息，单击**确认**。

5. 监控报警

5.1. 监控报警

消息队列Kafka版支持监控您账户下创建的资源，包括实例、Topic、Consumer Group，帮助您实时掌握资源状态。您还可以为监控项设置报警规则，当监控项数据超过设定的报警阈值时，云监控可以通过电话、短信、邮件、钉钉机器人等方式通知您，帮助您及时应对异常情况。

监控项

说明

- 监控项的数据聚合周期为1分钟，即1分钟计算一次，计算出来每秒多少字节。您可以将该数据理解为一分钟内的平均值。
- 监控项的数据延时1分钟。

针对各资源类型提供的监控项如下：

- 实例监控项
 - 实例消息生产流量 (bytes/s)
 - 实例消息消费流量 (bytes/s)
 - 实例磁盘使用率 (%)

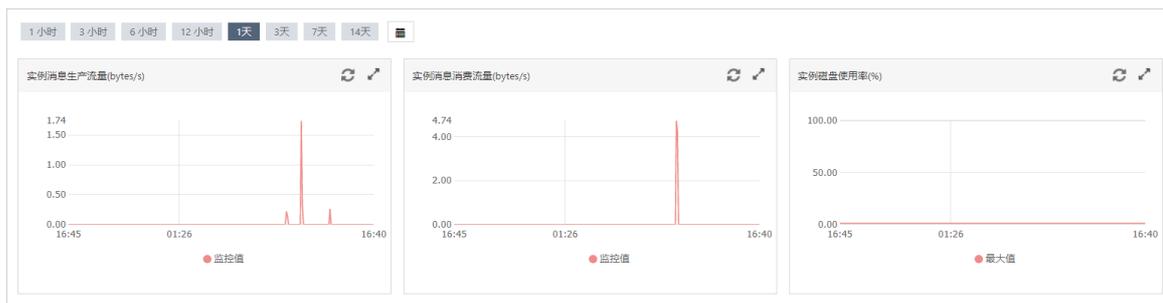
说明 实例磁盘使用率 (%) 显示的是实例各节点中磁盘使用率的最大值。

- Topic监控项
 - Topic消息生产流量 (bytes/s)
 - Topic消息消费流量 (bytes/s)
- Consumer Group监控项
 - Consumer Group未消费消息总数 (个)

查看监控数据

查看监控数据的步骤如下：

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击[监控报警](#)。
5. 在[监控报警](#)页面，单击要查看监控数据的资源页签，找到要查看监控数据的资源，单击其右侧操作列的云监控，设置时间范围，查看监控数据。



设置报警规则

设置报警的操作步骤如下：

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击[监控报警](#)。
5. 在[监控报警](#)页面，选择资源页签，设置报警规则。
 - 单击实例页签，然后单击[创建报警规则](#)，设置报警规则和通知方式，单击[确认](#)。如果您需要对已创建的规则进行修改，请执行如下步骤。
 - a. 在实例页签，单击操作列的[报警规则](#)。
 - b. 在实例关联的报警规则面板，选择目标报警规则，单击其操作列的[编辑规则](#)，对报警规则进行修改。
 - 单击Topic页签，然后单击[创建报警规则](#)，设置报警规则和通知方式，单击[确认](#)。如果您需要对已创建的规则进行修改，请执行如下步骤。
 - a. 在Topic页签，选择需要修改报警规则的Topic，单击其操作列的[报警规则](#)。
 - b. 在Topic关联的报警规则面板，选择目标报警规则，单击其操作列的[编辑规则](#)，对报警规则进行修改。
 - 单击Group页签，然后单击[创建报警规则](#)，设置报警规则和通知方式，单击[确认](#)。如果您需要对已创建的规则进行修改，请执行如下步骤。
 - a. 在Group页签，选择需要修改报警规则的Group，单击其操作列的[报警规则](#)。
 - b. 在Group关联的报警规则面板，选择目标报警规则，单击其操作列的[编辑规则](#)，对报警规则进行修改。

查看报警信息

查看报警信息的步骤如下：

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击[监控报警](#)。
5. 在[监控报警](#)页面，单击要查看报警信息的资源页签。
 - 查看实例的报警信息
 - a. 单击实例页签，然后单击操作列的[报警规则](#)。

- b. 在实例关联的报警规则面板，选择目标报警规则，单击其操作列的详情，您可以查看报警规则、禁用或启用报警规则、删除报警规则、查看报警历史。
- o 查看Topic的报警信息
 - a. 单击Topic页签，选择需要查看报警信息的Topic，然后单击操作列的报警规则。
 - b. 在Topic关联的报警规则面板，选择目标报警规则，单击其操作列的详情，您可以查看报警规则、禁用或启用报警规则、删除报警规则、查看报警历史。
- o 查看Group的报警信息
 - a. 在Group页签，选择需要查看报警信息的Group，单击其操作列的报警规则。
 - b. 在Group关联的报警规则面板，选择目标报警规则，单击其操作列的详情，您可以查看报警规则、禁用或启用报警规则、删除报警规则、查看报警历史。

5.2. 监控报警问题

5.2.1. 为什么需要升级实例来支持监报告警功能？

2018年11月以前部署的消息队列Kafka版集群不支持上报监报告警数据，需要升级来支持监报告警功能。

在控制台上可以查看和设置监报告警功能，但是在2018年11月以前部署的消息队列Kafka版集群不支持上报监报告警数据，因此需要升级旧版本的消息队列Kafka版实例集群，让它可以上报监报告警数据。

若您要升级您的消息队列Kafka版实例集群，请参见[升级实例版本](#)。

5.2.2. 报警项状态显示数据不足怎么办？

将实例的小版本升级至最新版本。

Condition

配置了报警规则，但在资源关联的报警规则面板的状态列显示数据不足。

Cause

老版消息队列Kafka版实例不支持报警数据上报。

Remedy

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在实例详情页面，单击配置信息页签，然后单击小版本右侧的升级小版本。
5. 在升级实例小版本对话框，完成以下操作：
 - i. 认真阅读升级必读！中的内容。
 - ii. 在紧急联系人文本框，输入您的姓名。
 - iii. 在紧急联系电话文本框，输入您的紧急联系电话。
 - iv. 单击确定。

5.2.3. 如何监控开源Apache Kafka?

本文为您介绍监控开源Apache Kafka的解决方案。

如果您需要监控开源Apache Kafka，请参见以下解决方案：

- [How to Monitor Kafka](#)
- [Monitoring Kafka performance metrics](#)

5.2.4. 父账号能看到监控信息，子账户看不到监控信息

子账号没有被授予查看监控信息的权限。

父账号给予账号授权AliyunCloudMonitorReadOnlyAccess权限后，子账号即可看到监控信息。操作步骤，请参见[步骤二：为RAM用户添加权限](#)。

5.2.5. 为什么不能登录部署消息队列Kafka版的机器?

消息队列Kafka版提供全托管免运维服务，您无需登录机器。

消息队列Kafka版是全托管的，由消息队列Kafka版专业团队进行运维保障，您无需登录机器。集群的一些基础信息会通过监控告警进行透传。

5.2.6. 删除了Consumer Group还会收到堆积告警?

删除Consumer Group并不会删除服务端消费位点等信息。

删除了Consumer Group，只是从控制台逻辑删除，并不会实际删除服务端消费位点等信息。堆积告警是根据位点处理的，因此您还会收到堆积告警。

删除了Consumer Group后，如果您不希望继续收到堆积告警，您可以选择：

- 禁用堆积告警
- 等待位点过期

消息队列Kafka版的消费位点保存在一个内部Topic里面，无法直接删除。位点在超过消息保留期限后仍然没有任何更新，就会被过期清理。

 **注意** 较早购买的用户，没有开启这个清理机制，请先在实例详情页，升级服务端版本到最新。

5.2.7. 什么是死分区和假堆积?

死分区指长期没有数据发送进来的分区，死分区存在没有被完全删除的消息就会出现假堆积。

当一个分区，长期没有数据发送进来时，就慢慢变成了死分区。死分区，通常不会造成使用问题，但却给监控报警等问题带来干扰。常见的问题就是查看堆积状态显示堆积，且堆积一直存在。这是因为死分区长期没有数据进来，消费端也不会再提交位点，超过保留时间后，位点会过期。当一个分区没有位点在服务端时，默认取最大位点-最小位点作为堆积值。此时，如果分区已死，但里面的消息并没有被完全删除时，就会出现假堆积。

一般来说，如果数据量足够多时，保留时间一到，消息就都被删除了。但如果数据量少时，消息可能要很久才会真正被删除，因为服务端存储是1G为Segment，至少保留一个Segment导致的。

解决办法就是，消息尽量发送均匀，且保证每个分区有数据。如果业务数据比较少，可以考虑定时发一些心跳数据到各个分区里面去。

5.2.8. 为什么告警堆积量和控制台堆积量不一致？

一般是因为死分区。

一般是因为死分区。当一个分区在服务端没有位点时，控制台默认取最大位点-最小位点作为堆积值，堆积告警默认取最大位点作为堆积值。小版本为最新版的实例优化了堆积告警，改为默认取最大位点-最小位点作为堆积值。如果您发现告警堆积量和控制台堆积量不一致，您可以将实例的小版本升级至最新版，详情请参见[升级小版本](#)。

6.Connector

6.1. Connector概述

消息队列Kafka版提供全托管、免运维的Connector，用于消息队列Kafka版和其他阿里云服务之间的数据同步。本文介绍Connector支持的数据同步任务的类型、使用流程、使用限制以及跨地域数据同步。

注意 消息队列Kafka版的Connector组件处于公测阶段，且独立于消息队列Kafka版实例，因此不会在消息队列Kafka版侧产生费用。同时，阿里云不承诺Connector的SLA，使用Connector所依赖的其他产品的SLA和费用说明请以对应产品为准。

Connector类型

消息队列Kafka版支持以下两大类的Connector：

- Sink Connector: Sink代表数据向外流转，即消息队列Kafka版为数据源，其他产品为数据目标。

Connector	描述	文档
FC Sink Connector	将数据从消息队列Kafka版导出至函数计算。	创建FC Sink Connector
MaxCompute Sink Connector	将数据从消息队列Kafka版导出至大数据计算服务MaxCompute。	创建MaxCompute Sink Connector
OSS Sink Connector	将数据从消息队列Kafka版导出至对象存储OSS。	创建OSS Sink Connector
Elasticsearch Sink Connector	将数据从消息队列Kafka版导出至阿里云Elasticsearch。	创建Elasticsearch Sink Connector

- Source Connector: Source代表数据向内流转，即消息队列Kafka版为数据目标，其他产品为数据源。

Connector	描述	文档
MySQL Source Connector	将数据从阿里云数据库RDS MySQL版导出至消息队列Kafka版。	创建MySQL Source Connector

使用流程

Connector的使用流程如下：

1. [开启Connector](#)
2. [创建Connector](#)
 - [创建FC Sink Connector](#)
 - [创建MaxCompute Sink Connector](#)
 - [创建OSS Sink Connector](#)
 - [创建Elasticsearch Sink Connector](#)
 - [创建MySQL Source Connector](#)

3. 更多Connector的相关操作

- [查看Connector任务配置](#)
- [查看Connector日志](#)
- [暂停Connector](#)
- [恢复Connector](#)
- [删除Connector](#)
- [修改Connector配置](#)
- [测试Connector](#)
- [为Connector开启公网访问](#)

使用限制

消息队列Kafka版对Connector的限制如下：

项目	限制值
数量	单实例最多创建3个
地域	<ul style="list-style-type: none"> • 华东1（杭州） • 华东2（上海） • 华北2（北京） • 华北3（张家口） • 华北5（呼和浩特） • 华南1（深圳） • 西南1（成都） • 中国香港 • 新加坡（新加坡） • 日本（东京）

 **说明** 如果您需要提升您的实例的Connector的数量限制或者需要更多的地域，请提交工单[工单](#)联系消息队列Kafka版技术人员。

跨地域数据同步

如果您需要将某个地域的数据，通过Connector同步到另一个地域的阿里云服务，您需要为该Connector开启公网访问，然后在公网进行数据同步。具体操作步骤，请参见[为Connector开启公网访问](#)。

 **说明** MySQL Source Connector的跨地域数据同步比较特殊，需要您自行开通企业网。更多信息，请参见[创建MySQL Source Connector](#)。

6.2. 开启Connector

本文说明如何为消息队列Kafka版实例开启Connector。

前提条件

您已购买并部署消息队列Kafka版实例，且该实例必须满足以下条件：

项目	说明
状态	服务中
版本	消息队列Kafka版实例的版本必须满足以下任意一种要求： <ul style="list-style-type: none"> 大版本为0.10.2，且小版本为最新版本。 大版本为2.2.0。

 **说明** 您可以在消息队列Kafka版控制台的实例详情页面的基础信息区域查看实例的运行状态和版本。

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在左侧导航栏，单击Connector 管理。
4. 在Connector 管理页面，单击创建 Connector。
5. 在您尚未开通当前实例的 Connector 功能提示对话框，单击确认。

后续步骤

为消息队列Kafka版实例开启Connector后，您可以创建Connector将消息队列Kafka版实例的数据同步到函数计算或大数据计算服务MaxCompute。

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)

6.3. 创建Connector

6.3.1. 创建FC Sink Connector

本文说明如何创建FC Sink Connector将数据从消息队列Kafka版实例的数据源Topic导出至函数计算的函数。

前提条件

在创建FC Sink Connector前，请确保您已完成以下操作：

- 为消息队列Kafka版实例开启Connector。更多信息，请参见[开启Connector](#)。
- 为消息队列Kafka版实例创建数据源Topic。更多信息，请参见[步骤一：创建Topic](#)。

本文以名称为fc-test-input的Topic为例。

- 在函数计算创建函数。更多信息，请参见[使用控制台创建函数](#)。

 **注意** 函数类型必须为事件函数。

本文以服务名称为guide-hello_world、函数名称为hello_world、运行环境为Python的事件函数为例。该示例函数的代码如下：

```
# -*- coding: utf-8 -*-
import logging
# To enable the initializer feature
# please implement the initializer function as below:
# def initializer(context):
#     logger = logging.getLogger()
#     logger.info('initializing')
def handler(event, context):
    logger = logging.getLogger()
    logger.info('hello world:' + bytes.decode(event))
    return 'hello world:' + bytes.decode(event)
```

操作流程

使用FC Sink Connector将数据从消息队列Kafka版实例的数据源Topic导出至函数计算的函数的操作流程如下：

1. (可选) 使FC Sink Connector跨地域访问函数计算

 **注意** 如果您不需要使FC Sink Connector跨地域访问函数计算，您可以直接跳过该步骤。

(可选) [为FC Sink Connector开启公网访问](#)

2. (可选) 使FC Sink Connector跨账号访问函数计算

 **注意** 如果您不需要使FC Sink Connector跨账号访问函数计算，您可以直接跳过该步骤。

- o [创建自定义权限策略](#)
- o [创建RAM角色](#)
- o [添加权限](#)

3. (可选) 创建FC Sink Connector依赖的Topic和Consumer Group

 **注意**

- o 如果您不需要自定义Topic和Consumer Group的名称，您可以直接跳过该步骤。
- o 部分FC Sink Connector依赖的Topic的存储引擎必须为Local存储，大版本为0.10.2的消息队列Kafka版实例不支持手动创建Local存储的Topic，只支持自动创建。

- i. [创建FC Sink Connector依赖的Topic](#)
- ii. [创建FC Sink Connector依赖的Consumer Group](#)

4. [创建并部署FC Sink Connector](#)

5. 结果验证

- i. [发送测试消息](#)
- ii. [查看函数日志](#)

(可选)

为FC Sink Connector开启公网访问

如需使FC Sink Connector跨地域访问其他阿里云服务，您需要为FC Sink Connector开启公网访问。具体操作，请参见[为Connector开启公网访问](#)。

(可选)

创建自定义权限策略

在目标账号下创建访问函数计算的自定义权限策略。

1. 登录[访问控制控制台](#)。
2. 在左侧导航栏，选择**权限管理 > 权限策略管理**。
3. 在权限策略管理页面，单击**创建权限策略**。
4. 在**新建自定义权限策略**页面，创建自定义权限策略。
 - i. 在**策略名称**文本框，输入 *KafkaConnectorFcAccess*。
 - ii. 在**配置模式**区域，选择**脚本配置**。
 - iii. 在**策略内容**区域，输入自定义权限策略脚本。访问函数计算的自定义权限策略脚本示例如下：

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": [
        "fc:InvokeFunction",
        "fc:GetFunction"
      ],
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

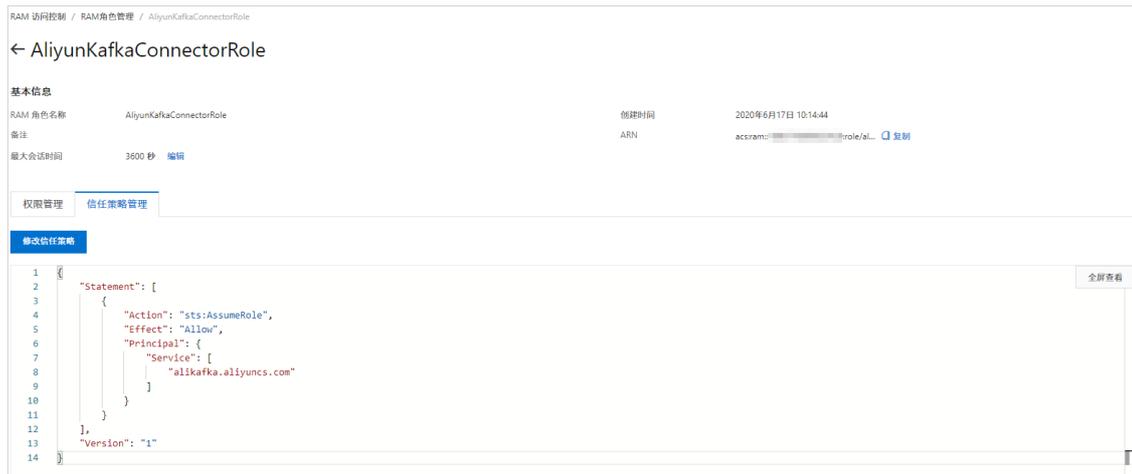
- iv. 单击**确定**。

(可选)

创建RAM角色

在目标账号下创建RAM角色。由于RAM角色不支持直接选择消息队列Kafka版作为受信服务，您在创建RAM角色时，需要选择任意支持的服务作为受信服务。RAM角色创建后，手工修改信任策略。

1. 在左侧导航栏，单击**RAM角色管理**。
2. 在**RAM角色管理**页面，单击**创建RAM角色**。
3. 在**创建RAM角色**面板，创建RAM角色。
 - i. 在**当前可信实体类型**区域，选择**阿里云服务**，单击**下一步**。
 - ii. 在**角色类型**区域，选择**普通服务角色**，在**角色名称**文本框，输入 *AliyunKafkaConnectorRole*，从**选择受信服务**列表，选择**函数计算**，然后单击**完成**。
4. 在**RAM角色管理**页面，找到 *AliyunKafkaConnectorRole*，单击 *AliyunKafkaConnectorRole*。
5. 在 *AliyunKafkaConnectorRole* 页面，单击**信任策略管理**页签，单击**修改信任策略**。
6. 在**修改信任策略**面板，将脚本中 *fc* 替换为 *alikafka*，单击**确定**。



(可选)

添加权限

在目标账号下为创建的RAM角色授予访问函数计算的权限。

1. 在左侧导航栏，单击RAM角色管理。
2. 在RAM角色管理页面，找到AliyunKafkaConnectorRole，在其右侧操作列，单击添加权限。
3. 在添加权限面板，添加KafkaConnectorFcAccess权限。
 - i. 在选择权限区域，选择自定义策略。
 - ii. 在权限策略名称列表，找到KafkaConnectorFcAccess，单击KafkaConnectorFcAccess。
 - iii. 单击确定。
 - iv. 单击完成。

(可选)

创建FC Sink Connector依赖的Topic

您可以在消息队列Kafka版控制台手动创建FC Sink Connector依赖的5个Topic，包括：任务位点Topic、任务配置Topic、任务状态Topic、死信队列Topic以及异常数据Topic。每个Topic所需要满足的分区数与存储引擎会有差异，具体信息，请参见配置源服务参数列表。

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。

注意 Topic需要在应用程序所在的地域（即所部署的ECS的所在地域）进行创建。Topic不能跨地域使用。例如Topic创建在华北2（北京）这个地域，那么消息生产端和消费端也必须运行在华北2（北京）的ECS。

3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Topic管理。
5. 在Topic管理页面，单击创建Topic。
6. 在创建Topic面板，设置Topic属性，然后单击确定。

创建 Topic
✕

*** 名称**

长度限制为 3 ~ 64 个字符，只能包含英文、数字、短横线 (-) 以及下划线 (_)，且至少包含一个英文或数字。

*** 描述** 9/64

*** 分区数**

建议分区数是6的倍数，减少数据倾斜风险，分区数限制（1~360），特殊需求请提交工单。

存储引擎 云存储

i 底层接入阿里云云盘，具有低时延、高性能、持久性、高可靠等特点，采用分布式3副本机制。

消息类型 普通消息

i 默认情况下，保证相同 Key 的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，可能会造成消息乱序。

标签

参数	说明	示例
名称	Topic名称。	demo
描述	Topic的简单描述。	demo test
分区数	Topic的分区数量。	12
存储引擎	<p>Topic消息的存储引擎。</p> <p>消息队列Kafka版支持以下两种存储引擎。</p> <ul style="list-style-type: none"> ◦ 云存储：底层接入阿里云云盘，具有低时延、高性能、持久性、高可靠等特点，采用分布式3副本机制。实例的规格类型为标准版（高写版）时，存储引擎只能为云存储。 ◦ Local 存储：使用原生Kafka的ISR复制算法，采用分布式3副本机制。 	云存储

参数	说明	示例
消息类型	<p>Topic消息的类型。</p> <ul style="list-style-type: none">普通消息：默认情况下，保证相同Key的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，可能会造成消息乱序。当存储引擎选择云存储时，默认选择普通消息。分区顺序消息：默认情况下，保证相同Key的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，仍然保证分区内按照发送顺序存储。但是会出现部分分区发送消息失败，等到分区恢复后即可恢复正常。当存储引擎选择Local存储时，默认选择分区顺序消息。	普通消息

参数	说明	示例
日志清理策略	<p>Topic日志的清理策略。</p> <p>当存储引擎选择Local 存储时，需要配置日志清理策略。</p> <p>消息队列Kafka版支持以下两种日志清理策略。</p> <ul style="list-style-type: none"> ◦ Delete：默认的消息清理策略。在磁盘容量充足的情况下，保留在最长保留时间范围内的消息；在磁盘容量不足时（一般磁盘使用率超过85%视为不足），将提前删除旧消息，以保证服务可用性。 ◦ Compact：使用Kafka Log Compaction日志清理策略。Log Compaction清理策略保证相同Key的消息，最新的value值一定会被保留。主要适用于系统宕机后恢复状态，系统重启后重新加载缓存等场景。例如，在使用Kafka Connect或Confluent Schema Registry时，需要使用Kafka Compact Topic存储系统状态信息或配置信息。 <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> <p> 注意 Compact Topic一般只用在某些生态组件中，例如Kafka Connect或Confluent Schema Registry，其他情况的消息收发请勿为Topic设置该属性。具体信息，请参见消息队列Kafka版Demo库。</p> </div>	Compact
标签	Topic的标签。	demo

创建完成后，在Topic 管理页面的列表中显示已创建的Topic。

(可选)

创建FC Sink Connector依赖的Consumer Group

您可以在消息队列Kafka版控制台手动创建FC Sink Connector数据同步任务使用的Consumer Group。该Consumer Group的名称必须为connect-任务名称，具体信息，请参见[配置源服务参数列表](#)。

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。

4. 在左侧导航栏，单击**Group 管理**。
5. 在**Group 管理**页面，单击**创建 Group**。
6. 在**创建 Group**面板的**Group ID**文本框输入Group的名称，在**描述**文本框简要描述Group，并给Group添加标签，单击**确定**。
创建完成后，在**Group 管理**页面的列表中显示已创建的Group。

创建并部署FC Sink Connector

创建并部署用于将数据从消息队列Kafka版同步至函数计算的FC Sink Connector。

1. 登录**消息队列Kafka版控制台**。
2. 在**概览**页面的**资源分布**区域，选择地域。
3. 在**实例列表**页面，单击目标实例名称。
4. 在左侧导航栏，单击**Connector 管理**。
5. 在**Connector 管理**页面，单击**创建 Connector**。
6. 在**创建 Connector**配置向导面页面，完成以下操作。
 - i. 在**配置基本信息**页签，按需配置以下参数，然后单击**下一步**。

参数	描述	示例值
名称	Connector的名称。命名规则： <ul style="list-style-type: none"> ▪ 可以包含数字、小写英文字母和短划线（-），但不能以短划线（-）开头，长度限制为48个字符。 ▪ 同一个消息队列Kafka版实例内保持唯一。 Connector的数据同步任务必须使用名称为connect-任务名称的Consumer Group。如果您未手动创建该Consumer Group，系统将为您自动创建。	kafka-fc-sink
实例	默认配置为实例的名称与实例ID。	demo alikafka_post-cn-st21p8vj****

- ii. 在**配置源服务**页签，选择数据源为**消息队列Kafka版**，并配置以下参数，然后单击**下一步**。

 **说明** 如果您已创建好Topic和Consumer Group，那么请选择**手动创建资源**，并填写已创建的资源信息。否则，请选择**自动创建资源**。

配置源服务参数列表

参数	描述	示例值
数据源 Topic	需要同步数据的Topic。	fc-test-input
消费线程并发数	数据源Topic的消费线程并发数。默认值为6。取值说明如下： <ul style="list-style-type: none"> ▪ 6 ▪ 12 	6

参数	描述	示例值
消费初始位置	<p>开始消费的位置。取值说明如下：</p> <ul style="list-style-type: none"> ■ 最早位点：从最初位点开始消费。 ■ 最近位点：从最新位点开始消费。 	最早位点
VPC ID	数据同步任务所在的VPC。单击 配置运行环境 显示该参数。默认为消息队列Kafka版实例所在的VPC，您无需填写。	vpc-bp1xpdnd3l***
VSwitch ID	数据同步任务所在的交换机。单击 配置运行环境 显示该参数。该交换机必须与消息队列Kafka版实例处于同一VPC。默认为部署消息队列Kafka版实例时填写的交换机。	vsw-bp1d2jgg81***
失败处理	<p>消息发送失败后，是否继续订阅出现错误的Topic的分区。单击配置运行环境显示该参数。取值说明如下。</p> <ul style="list-style-type: none"> ■ 继续订阅：继续订阅出现错误的Topic的分区，并打印错误日志。 ■ 停止订阅：停止订阅出现错误的Topic的分区，并打印错误日志 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> 说明</p> <ul style="list-style-type: none"> ■ 如何查看日志，请参见查看Connector日志。 ■ 如何根据错误码查找解决方案，请参见错误码。 ■ 如需恢复对出现错误的Topic的分区的订阅，您需要提交工单联系消息队列Kafka版技术人员。 </div>	继续订阅
创建资源方式	<p>选择创建Connector所依赖的Topic与Group的方式。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> ■ 自动创建 ■ 手动创建 	自动创建
Connector 消费组	Connector的数据同步任务使用的Consumer Group。单击 配置运行环境 显示该参数。该Consumer Group的名称必须为connect-任务名称。	connect-kafka-fc-sink
任务位点 Topic	<p>用于存储消费位点的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> ■ Topic：建议以connect-offset开头。 ■ 分区数：Topic的分区数量必须大于1。 ■ 存储引擎：Topic的存储引擎必须为Local存储。 ■ cleanup.policy：Topic的日志清理策略必须为compact。 	connect-offset-kafka-fc-sink

参数	描述	示例值
任务配置 Topic	用于存储任务配置的Topic。单击配置运行环境显示该参数。 <ul style="list-style-type: none"> Topic: 建议以connect-config开头。 分区数: Topic的分区数量必须为1。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-config-kafka-fc-sink
任务状态 Topic	用于存储任务状态的Topic。单击配置运行环境显示该参数。 <ul style="list-style-type: none"> Topic: 建议以connect-status开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-status-kafka-fc-sink
死信队列 Topic	用于存储Connect框架的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和异常数据Topic为同一个Topic, 以节省Topic资源。 <ul style="list-style-type: none"> Topic: 建议以connect-error开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-fc-sink
异常数据 Topic	用于存储Sink的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和死信队列Topic为同一个Topic, 以节省Topic资源。 <ul style="list-style-type: none"> Topic: 建议以connect-error开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-fc-sink

iii. 在配置目标服务页签, 选择目标服务为函数计算, 并配置以下参数, 然后单击创建。

参数	描述	示例值
是否跨账号/地域	FC Sink Connector是否跨账号/地域向函数计算服务同步数据。默认为否。取值: <ul style="list-style-type: none"> 否: 同地域同账号模式。 是: 跨地域同账号模式、同地域跨账号模式或跨地域跨账号模式。 	否
服务地域	函数计算服务的地域。默认为FC Sink Connector所在地域。如需跨地域, 您需要为Connector开启公网访问, 然后选择目标地域。更多信息, 请参见 为FC Sink Connector开启公网访问 。 <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;">  注意 是否跨账号/地域为是时, 显示服务地域。 </div>	cn-hangzhou

参数	描述	示例值
服务接入点	<p>函数计算服务的接入点。在函数计算控制台的概览页的常用信息区域获取。</p> <ul style="list-style-type: none"> 内网Endpoint：低延迟，推荐。适用于消息队列Kafka版实例和函数计算处于同一地域场景。 公网Endpoint：高延迟，不推荐。适用于消息队列Kafka版实例和函数计算处于不同地域的场景。如需使用公网Endpoint，您需要为Connector开启公网访问。更多信息，请参见为FC Sink Connector开启公网访问。 <p> 注意 是否跨账号/地域为是时，显示服务接入点。</p>	<p>http://188***.cn - hangzhou.fc.aliyuncs.com</p>
服务账号	<p>函数计算服务的阿里云账号ID。在函数计算控制台的概览页的常用信息区域获取。</p> <p> 注意 是否跨账号/地域为是时，显示服务账号。</p>	<p>188***</p>
授权角色名	<p>消息队列Kafka版访问函数计算服务的RAM角色。</p> <ul style="list-style-type: none"> 如不需跨账号，您需要在本账号下创建RAM角色并为其授权，然后输入该授权角色名。操作步骤，请参见创建自定义权限策略、创建RAM角色和添加权限。 如需跨账号，您需要在目标账号下创建RAM角色并为其授权，然后输入该授权角色名。操作步骤，请参见创建自定义权限策略、创建RAM角色和添加权限。 <p> 注意 是否跨账号/地域为是时，显示授权角色名。</p>	<p>AliyunKafkaConnectorRole</p>
服务名	<p>函数计算服务的名称。</p>	<p>guide-hello_world</p>
函数名	<p>函数计算服务的函数名称。</p>	<p>hello_world</p>
版本或别名	<p>函数计算服务的版本或别名。</p>	<p>LATEST</p>
发送模式	<p>消息发送模式。取值说明如下：</p> <ul style="list-style-type: none"> 异步：推荐。 同步：不推荐。同步发送模式下，如果函数计算的处理时间较长，消息队列Kafka版的处理时间也会较长。当同一批次消息的处理时间超过5分钟时，会触发消息队列Kafka版客户端Rebalance。 	<p>异步</p>
	<p>批量发送的消息条数。默认为20。Connector根据发送批次大小和请求大小限制（同步请求大小限制为6 MB，异步请求大小限制为128 KB）将多条消息聚合后发送。例如，发送模式为异步，发送批次大小为20，如果要发送18条消息，其中有17条消</p>	

参数	描述	示例值
发送批大小	<p>消息的总大小为127 KB，有1条消息的大小为200 KB，Connector 会将总大小不超过128 KB的17条消息聚合后发送，将大小超过128 KB的1条消息单独发送。</p> <p>说明 如果您在发送消息时将key设置为null，则请求中不包含key。如果将value设置为null，则请求中不包含value。</p> <ul style="list-style-type: none"> 如果批量发送的多条消息的大小不超过请求大小限制，则请求中包含消息内容。请求示例如下： <pre>[{ "key":"this is the message's key2", "offset":8, "overflowFlag":false, "partition":4, "timestamp":1603785325438, "topic":"Test", "value":"this is the message's value2", "valueSize":28 }, { "key":"this is the message's key9", "offset":9, "overflowFlag":false, "partition":4, "timestamp":1603785325440, "topic":"Test", "value":"this is the message's value9", "valueSize":28 }, { "key":"this is the message's key12", "offset":10, "overflowFlag":false, "partition":4, "timestamp":1603785325442, "topic":"Test", "value":"this is the message's value12", "valueSize":29 }, { "key":"this is the message's key38", "offset":11, "overflowFlag":false, "partition":4, "timestamp":1603785325464, "topic":"Test", "value":"this is the message's value38", "valueSize":29 }]</pre>	50

参数	描述	示例值
	<p>■ 如果发送的单条消息的大小超过请求大小限制，则请求中不包含消息内容。请求示例如下：</p> <pre data-bbox="564 293 1161 719">[{ "key":"123", "offset":4, "overflowFlag":true, "partition":0, "timestamp":1603779578478, "topic":"Test", "value":"1", "valueSize":272687 }]</pre> <p> 说明 如需获取消息内容，您需要根据位点主动拉取消息。</p>	
重试次数	<p>消息发送失败后的重试次数。默认为2。取值范围为1~3。部分导致消息发送失败的错误不支持重试。错误码与是否支持重试的对应关系如下：</p> <ul style="list-style-type: none"> ■ 4XX：除429支持重试外，其余错误码不支持重试。 ■ 5XX：支持重试。 <p> 说明 Connector调用InvokeFunction向函数计算发送消息。</p>	2

创建完成后，在Connector 管理页面，查看创建的Connector。

7. 创建完成后，在Connector 管理页面，找到创建的Connector，单击其操作列的部署。如需配置函数计算资源，单击其操作列的更多 > 配置函数，跳转至函数计算控制台完成操作。

发送测试消息

部署FC Sink Connector后，您可以向消息队列Kafka版的数据源Topic发送消息，测试数据能否被同步至函数计算。

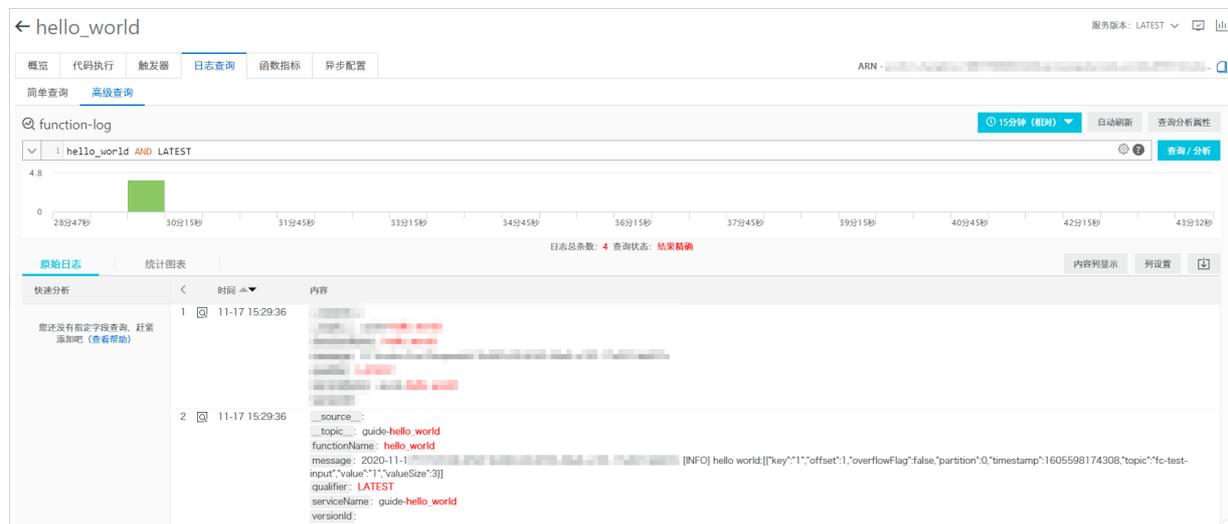
1. 在Connector 管理页面，找到目标Connector，在其右侧操作列，单击测试。
2. 在发送消息面板，发送测试消息。
 - 发送方式选择控制台。
 - a. 在消息 Key文本框中输入消息的Key值，例如demo。

- b. 在消息内容文本框输入测试的消息内容，例如 {"key": "test"}。
- c. 设置发送到指定分区，选择是否指定分区。
 - a. 单击是，在分区 ID 文本框中输入分区的ID，例如0。如果您需查询分区的ID，请参见[查看分区状态](#)。
 - b. 单击否，不指定分区。
- o 发送方式选择 Docker，执行运行 Docker 容器生产示例消息区域的 Docker 命令，发送消息。
- o 发送方式选择 SDK，根据您的业务需求，选择需要的语言或者框架的 SDK 以及接入方式，通过 SDK 发送消息。

查看函数日志

向消息队列 Kafka 版的数据源 Topic 发送消息后，查看函数日志，验证是否收到消息。更多信息，请参见[配置并查看函数日志](#)。

日志中显示发送的测试消息。



6.3.2. 创建 MaxCompute Sink Connector

本文说明如何创建 MaxCompute Sink Connector 将数据从消息队列 Kafka 版实例的数据源 Topic 导出至 MaxCompute 的表。

前提条件

在创建 MaxCompute Sink Connector 前，请确保您已完成以下操作：

- 为消息队列 Kafka 版实例开启 Connector。更多信息，请参见[开启 Connector](#)。
- 为消息队列 Kafka 版实例创建数据源 Topic。更多信息，请参见[步骤一：创建 Topic](#)。

本文以名称为 maxcompute-test-input 的 Topic 为例。

- 通过 MaxCompute 客户端创建表。更多信息，请参见[创建表](#)。

本文以名称为 connector_test 的项目下名称为 test_kafka 的表为例。该表的建表语句如下：

```
CREATE TABLE IF NOT EXISTS test_kafka(topic STRING,partition BIGINT,offset BIGINT,key STRING,value STRING) PARTITIONED BY (pt STRING);
```

操作流程

使用MaxCompute Sink Connector将数据从消息队列Kafka版实例的数据源Topic导出至MaxCompute的表操作流程如下：

1. 授予消息队列Kafka版访问MaxCompute的权限。
 - 创建RAM角色
 - 添加权限
2. （可选）创建MaxCompute Sink Connector依赖的Topic和Consumer Group

如果您不需要自定义Topic和Consumer Group的名称，您可以直接跳过该步骤，在下一步骤选择自动创建。

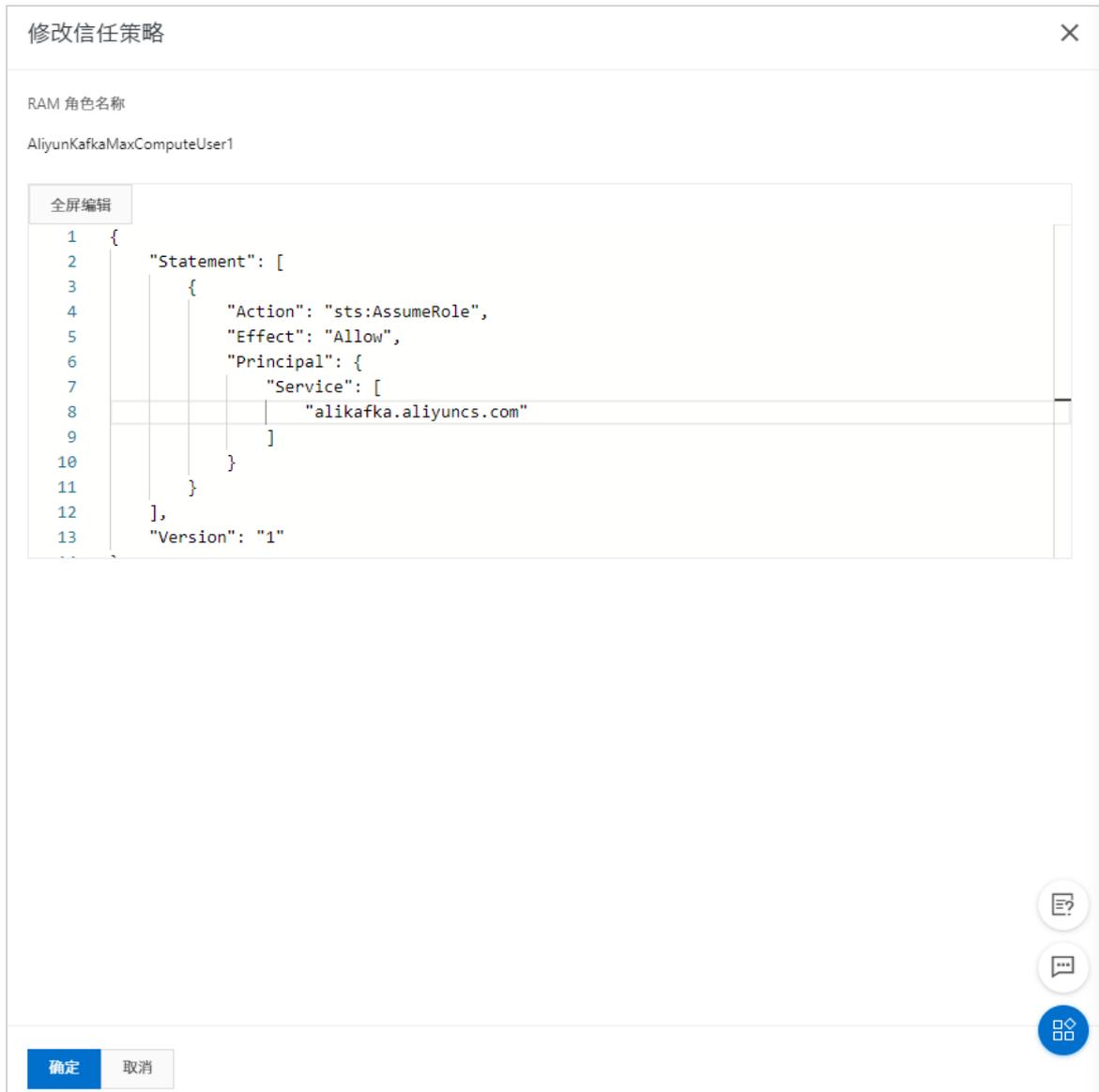
 **注意** 部分MaxCompute Sink Connector依赖的Topic的存储引擎必须为Local存储，大版本为0.10.2的消息队列Kafka版实例不支持手动创建Local存储的Topic，只支持自动创建。

- i. 创建MaxCompute Sink Connector依赖的Topic
 - ii. 创建MaxCompute Sink Connector依赖的Consumer Group
3. 创建并部署MaxCompute Sink Connector
4. 结果验证
 - i. 发送测试消息
 - ii. 查看表数据

创建RAM角色

由于RAM角色不支持直接选择消息队列Kafka版作为受信服务，您在创建RAM角色时，需要选择任意支持的服务作为受信服务。RAM角色创建后，手动修改信任策略。

1. 登录[访问控制控制台](#)。
2. 在左侧导航栏，单击RAM角色管理。
3. 在RAM角色管理页面，单击创建RAM角色。
4. 在创建RAM角色面板，执行以下操作。
 - i. 当前可信实体类型区域，选择阿里云服务，然后单击下一步。
 - ii. 在角色类型区域，选择普通服务角色，在角色名称文本框，输入AliyunKafkaMaxComputeUser1，从选择受信服务列表，选择大数据计算服务，然后单击完成。
5. 在RAM角色管理页面，找到AliyunKafkaMaxComputeUser1，单击AliyunKafkaMaxComputeUser1。
6. 在AliyunKafkaMaxComputeUser1页面，单击信任策略管理页签，单击修改信任策略。
7. 在修改信任策略面板，将脚本中fc替换为alikaafka，单击确定。



添加权限

为使Connector将消息同步到MaxCompute表，您需要为创建的RAM角色至少授予以下权限：

客体	操作	描述
Project	CreateInstance	在项目中创建实例。
Table	Describe	读取表的元信息。
Table	Alter	修改表的元信息或添加删除分区。
Table	Update	覆盖或添加表的数据。

关于以上权限的详细说明以及授权操作，请参见[授权](#)。

为本文创建的AliyunKafkaMaxComputeUser1添加权限的示例步骤如下：

1. 登录MaxCompute客户端。

2. 执行以下命令添加RAM角色为用户。

```
add user `RAM${<accountid>}:role/aliyunkafkamaxcomputeuser1`;
```

 **说明** 将<accountid>替换为您自己的阿里云账号ID。

3. 为RAM用户授予访问MaxCompute所需的最小权限。

- i. 执行以下命令为RAM用户授予项目相关权限。

```
grant CreateInstance on project connector_test to user `RAM${<accountid>}:role/aliyunkafkamaxcomputeuser1`;
```

 **说明** 将<accountid>替换为您自己的阿里云账号ID。

- ii. 执行以下命令为RAM用户授予表相关权限。

```
grant Describe, Alter, Update on table test_kafka to user `RAM${<accountid>}:role/aliyunkafkamaxcomputeuser1`;
```

 **说明** 将<accountid>替换为您自己的阿里云账号ID。

创建MaxCompute Sink Connector依赖的Topic

您可以在消息队列Kafka版控制台手动创建MaxCompute Sink Connector依赖的5个Topic，包括：任务位点Topic、任务配置Topic、任务状态Topic、死信队列Topic以及异常数据Topic。每个Topic所需要满足的分区数与存储引擎会有差异，具体信息，请参见[配置源服务参数列表](#)。

1. 登录[消息队列Kafka版控制台](#)。
2. 在**概览**页面的**资源分布**区域，选择地域。

 **注意** Topic需要在应用程序所在的地域（即所部署的ECS的所在地域）进行创建。Topic不能跨地域使用。例如Topic创建在华北2（北京）这个地域，那么消息生产端和消费端也必须运行在华北2（北京）的ECS。

3. 在**实例列表**页面，单击目标实例名称。
4. 在左侧导航栏，单击**Topic 管理**。
5. 在**Topic 管理**页面，单击**创建 Topic**。
6. 在**创建 Topic**面板，设置Topic属性，然后单击**确定**。

创建 Topic
✕

*** 名称**

长度限制为 3 ~ 64 个字符，只能包含英文、数字、短横线 (-) 以及下划线 (_)，且至少包含一个英文或数字。

*** 描述** 9/64

*** 分区数**

建议分区数是6的倍数，减少数据倾斜风险，分区数限制（1~360），特殊需求请提交工单。

存储引擎 云存储

i 底层接入阿里云云盘，具有低时延、高性能、持久性、高可靠等特点，采用分布式3副本机制。

消息类型 普通消息

i 默认情况下，保证相同 Key 的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，可能会造成消息乱序。

标签

参数	说明	示例
名称	Topic名称。	demo
描述	Topic的简单描述。	demo test
分区数	Topic的分区数量。	12
存储引擎	Topic消息的存储引擎。 消息队列Kafka版支持以下两种存储引擎。 <ul style="list-style-type: none"> ◦ 云存储：底层接入阿里云云盘，具有低时延、高性能、持久性、高可靠等特点，采用分布式3副本机制。实例的规格类型为标准版（高写版）时，存储引擎只能为云存储。 ◦ Local 存储：使用原生Kafka的ISR复制算法，采用分布式3副本机制。 	云存储

参数	说明	示例
消息类型	<p>Topic消息的类型。</p> <ul style="list-style-type: none">普通消息：默认情况下，保证相同Key的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，可能会造成消息乱序。当存储引擎选择云存储时，默认选择普通消息。分区顺序消息：默认情况下，保证相同Key的消息分布在同一个分区中，且分区内消息按照发送顺序存储。集群中出现机器宕机时，仍然保证分区内按照发送顺序存储。但是会出现部分分区发送消息失败，等到分区恢复后即可恢复正常。当存储引擎选择Local存储时，默认选择分区顺序消息。	普通消息

参数	说明	示例
日志清理策略	<p>Topic日志的清理策略。</p> <p>当存储引擎选择Local 存储时，需要配置日志清理策略。</p> <p>消息队列Kafka版支持以下两种日志清理策略。</p> <ul style="list-style-type: none"> ◦ Delete：默认的消息清理策略。在磁盘容量充足的情况下，保留在最长保留时间范围内的消息；在磁盘容量不足时（一般磁盘使用率超过85%视为不足），将提前删除旧消息，以保证服务可用性。 ◦ Compact：使用Kafka Log Compaction日志清理策略。Log Compaction清理策略保证相同Key的消息，最新的value值一定会被保留。主要适用于系统宕机后恢复状态，系统重启后重新加载缓存等场景。例如，在使用Kafka Connect或Confluent Schema Registry时，需要使用Kafka Compact Topic存储系统状态信息或配置信息。 <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> <p> 注意 Compact Topic一般只用在某些生态组件中，例如Kafka Connect或Confluent Schema Registry，其他情况的消息收发请勿为Topic设置该属性。具体信息，请参见消息队列Kafka版Demo库。</p> </div>	Compact
标签	Topic的标签。	demo

创建完成后，在Topic 管理页面的列表中显示已创建的Topic。

创建MaxCompute Sink Connector依赖的Consumer Group

您可以在消息队列Kafka版控制台手动创建MaxCompute Sink Connector数据同步任务使用的Consumer Group。该Consumer Group的名称必须为connect-任务名称，具体信息，请参见[配置源服务参数列表](#)。

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Group 管理。

5. 在Group 管理页面，单击创建 Group。
6. 在创建 Group 面板的Group ID 文本框输入Group的名称，在描述文本框简要描述Group，并给Group添加标签，单击确定。
创建完成后，在Group 管理页面的列表中显示已创建的Group。

创建并部署MaxCompute Sink Connector

创建并部署用于将数据从消息队列Kafka版同步至MaxCompute的MaxCompute Sink Connector。

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，单击创建 Connector。
6. 在创建 Connector配置向导页面，完成以下操作。
 - i. 在配置基本信息页签，按需配置以下参数，然后单击下一步。

参数	描述	示例值
名称	Connector的名称。命名规则： <ul style="list-style-type: none"> ▪ 可以包含数字、小写英文字母和短划线（-），但不能以短划线（-）开头，长度限制为48个字符。 ▪ 同一个消息队列Kafka版实例内保持唯一。 Connector的数据同步任务必须使用名称为connect-任务名称的Consumer Group。如果您未手动创建该Consumer Group，系统将为您自动创建。	kafka-maxcompute-sink
实例	默认配置为实例的名称与实例ID。	demoalikafka_post-cn-st21p8vj****

- ii. 在配置源服务页签，选择数据源为消息队列Kafka版，并配置以下参数，然后单击下一步。

 **说明** 如果您已创建好Topic和Consumer Group，那么请选择手动创建资源，并填写已创建的资源信息。否则，请选择自动创建资源。

配置源服务参数列表

参数	描述	示例值
数据源 Topic	需要同步数据的Topic。	maxcompute-test-input
消费线程并发数	数据源Topic的消费线程并发数。默认值为6。取值说明如下： <ul style="list-style-type: none"> ▪ 6 ▪ 12 	6

参数	描述	示例值
消费初始位置	<p>开始消费的位置。取值说明如下：</p> <ul style="list-style-type: none"> ■ 最早位点：从最初位点开始消费。 ■ 最近位点：从最新位点开始消费。 	最早位点
VPC ID	<p>数据同步任务所在的VPC。单击配置运行环境显示该参数。默认为消息队列Kafka版实例所在的VPC，您无需填写。</p>	vpc-bp1xpdnd3l***
VSwitch ID	<p>数据同步任务所在的交换机。单击配置运行环境显示该参数。该交换机必须与消息队列Kafka版实例处于同一VPC。默认为部署消息队列Kafka版实例时填写的交换机。</p>	vsw-bp1d2jgg81***
失败处理	<p>消息发送失败后，是否继续订阅出现错误的Topic的分区。单击配置运行环境显示该参数。取值说明如下。</p> <ul style="list-style-type: none"> ■ 继续订阅：继续订阅出现错误的Topic的分区，并打印错误日志。 ■ 停止订阅：停止订阅出现错误的Topic的分区，并打印错误日志 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> 说明</p> <ul style="list-style-type: none"> ■ 如何查看日志，请参见查看Connector日志。 ■ 如何根据错误码查找解决方案，请参见错误码。 ■ 如需恢复对出现错误的Topic的分区的订阅，您需要提交工单联系消息队列Kafka版技术人员。 </div>	继续订阅
创建资源方式	<p>选择创建Connector所依赖的Topic与Group的方式。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> ■ 自动创建 ■ 手动创建 	自动创建
Connector 消费组	<p>Connector的数据同步任务使用的Consumer Group。单击配置运行环境显示该参数。该Consumer Group的名称必须为connect-任务名称。</p>	connect-kafka-maxcompute-sink
任务位点 Topic	<p>用于存储消费位点的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> ■ Topic名称：建议以connect-offset开头。 ■ 分区数：Topic的分区数量必须大于1。 ■ 存储引擎：Topic的存储引擎必须为Local存储。 ■ cleanup.policy：Topic的日志清理策略必须为compact。 	connect-offset-kafka-maxcompute-sink

参数	描述	示例值
任务配置 Topic	<p>用于存储任务配置的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic名称：建议以connect-config开头。 分区数：Topic的分区数量必须为1。 存储引擎：Topic的存储引擎必须为Local存储。 cleanup.policy：Topic的日志清理策略必须为compact。 	connect-config-kafka-maxcompute-sink
任务状态 Topic	<p>用于存储任务状态的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic名称：建议以connect-status开头。 分区数：Topic的分区数量建议为6。 存储引擎：Topic的存储引擎必须为Local存储。 cleanup.policy：Topic的日志清理策略必须为compact。 	connect-status-kafka-maxcompute-sink
死信队列 Topic	<p>用于存储Connect框架的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和异常数据Topic为同一个Topic，以节省Topic资源。</p> <ul style="list-style-type: none"> Topic名称：建议以connect-error开头。 分区数：Topic的分区数量建议为6。 存储引擎：Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-maxcompute-sink
异常数据 Topic	<p>用于存储Sink的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和死信队列Topic为同一个Topic，以节省Topic资源。</p> <ul style="list-style-type: none"> Topic名称：建议以connect-error开头。 分区数：Topic的分区数量建议为6。 存储引擎：Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-maxcompute-sink

iii. 在配置目标服务页签，选择目标服务为大数据计算服务，并配置以下参数，然后单击创建。

参数	描述	示例值
连接地址	<p>MaxCompute的服务接入点。更多信息，请参见Endpoint。</p> <ul style="list-style-type: none"> VPC网络Endpoint：低延迟，推荐。适用于消息队列Kafka版实例和MaxCompute处于同一地域场景。 外网Endpoint：高延迟，不推荐。适用于消息队列Kafka版实例和MaxCompute处于不同地域的场景。如需使用公网Endpoint，您需要为Connector开启公网访问。更多信息，请参见为Connector开启公网访问。 	http://service.cn-hangzhou.maxcompute.aliyun-inc.com/api
工作空间	MaxCompute的工作空间。	connector_test
表	MaxCompute的表。	test_kafka
表地域	MaxCompute表所在地域。	华东1（杭州）

参数	描述	示例值
服务账号	MaxCompute的阿里云账号ID。	188***
授权角色名	消息队列Kafka版的RAM角色的名称。更多信息，请参见 创建RAM角色 。	AliyunKafkaMaxComputeUser1
模式	<p>消息同步到Connector的模式。默认为DEFAULT。取值说明如下：</p> <ul style="list-style-type: none"> ▪ KEY：只保留消息的Key，并将Key写入MaxCompute表的key列。 ▪ VALUE：只保留消息的Value，并将Value写入MaxCompute表的value列。 ▪ DEFAULT：同时保留消息的Key和Value，并将Key和Value分别写入MaxCompute表的key列和value列。 <p> 注意 DEFAULT模式下，不支持选择CSV格式，只支持TEXT格式和BINARY格式。</p>	DEFAULT
格式	<p>消息同步到Connector的格式。默认为TEXT。取值说明如下：</p> <ul style="list-style-type: none"> ▪ TEXT：消息的格式为字符串。 ▪ BINARY：消息的格式为字节数组。 ▪ CSV：消息的格式为逗号(,)分隔的字符串。 <p> 注意 CSV格式下，不支持DEFAULT模式，只支持KEY模式和VALUE模式：</p> <ul style="list-style-type: none"> ▪ KEY模式：只保留消息的Key，根据逗号(,)分隔Key字符串，并将分隔后的字符串按照索引顺序写入表。 ▪ VALUE模式：只保留消息的Value，根据逗号(,)分隔Value字符串，并将分隔后的字符串按照索引顺序写入表。 	TEXT
分区	<p>分区的粒度。默认为HOUR。取值说明如下：</p> <ul style="list-style-type: none"> ▪ DAY：每天将数据写入一个新分区。 ▪ HOUR：每小时将数据写入一个新分区。 ▪ MINUTE：每分钟将数据写入一个新分区。 	HOUR
时区	向Connector的数据源Topic发送消息的消息队列Kafka版生产者客户端所在时区。默认为GMT 08:00。	GMT 08:00

创建完成后，在Connector管理页面，查看创建的Connector。

7. 创建完成后，在Connector管理页面，找到创建的Connector，单击其操作列的部署。

发送测试消息

部署MaxCompute Sink Connector后，您可以向消息队列Kafka版的数据源Topic发送消息，测试数据能否被同步至MaxCompute。

1. 在Connector 管理页面，找到目标Connector，在其右侧操作列，单击测试。
2. 在发送消息面板，发送测试消息。
 - 发送方式选择控制台。
 - a. 在消息 Key文本框中输入消息的Key值，例如demo。
 - b. 在消息内容文本框输入测试的消息内容，例如 {"key": "test"}。
 - c. 设置发送到指定分区，选择是否指定分区。
 - a. 单击是，在分区 ID文本框中输入分区的ID，例如0。如果您需查询分区的ID，请参见[查看分区状态](#)。
 - b. 单击否，不指定分区。
 - 发送方式选择Docker，执行运行 Docker 容器生产示例消息区域的Docker命令，发送消息。
 - 发送方式选择 SDK，根据您的业务需求，选择需要的语言或者框架的SDK以及接入方式，通过SDK发送消息。

查看表数据

向消息队列Kafka版的数据源Topic发送消息后，在MaxCompute客户端查看表数据，验证是否收到消息。

查看本文写入的test_kafka的示例步骤如下：

1. 登录MaxCompute客户端。
2. 执行以下命令查看表的数据分区。

```
show partitions test_kafka;
```

返回结果示例如下：

```
pt=11-17-2020 15
OK
```

3. 执行以下命令查看分区的数据。

```
select * from test_kafka where pt="11-17-2020 14";
```

返回结果示例如下：

```
+-----+-----+-----+-----+-----+-----+
| topic      | partition | offset | key | value | pt      |
+-----+-----+-----+-----+-----+-----+
| maxcompute-test-input| 0      | 0      | 1 | 1 | 11-17-2020 14 |
+-----+-----+-----+-----+-----+-----+
```

6.3.3. 创建OSS Sink Connector

本文介绍如何创建OSS Sink Connector将数据从消息队列Kafka版实例的数据源Topic导出至对象存储OSS。

前提条件

在导出数据前，请确保您已完成以下操作：

- 为消息队列Kafka版实例开启Connector。更多信息，请参见[开启Connector](#)。
- 为消息队列Kafka版实例创建数据源Topic。更多信息，请参见[步骤一：创建Topic](#)。
- 在[OSS管理控制台](#)创建存储空间。更多信息，请参见[创建存储空间](#)。
- 开通函数计算服务。更多信息，请参见[开通函数计算服务](#)。

注意事项

- 仅支持在同地域内，将数据从消息队列Kafka版实例的数据源Topic导出至函数计算，再由函数计算导出至对象存储。Connector的限制说明，请参见[使用限制](#)。
- 该功能基于函数计算服务提供。函数计算为您提供了一定的免费额度，超额部分将产生费用，请以函数计算的计费规则为准。计费详情，请参见[计费概述](#)。
- 函数计算的函数调用支持日志查询。具体操作步骤，请参见[配置并查看函数日志](#)。
- 消息转储时，消息队列Kafka版中消息用UTF-8 String序列化，暂不支持二进制的数据格式。

创建并部署OSS Sink Connector

1. 登录[消息队列Kafka版控制台](#)。
2. 在[概览](#)页面的资源分布区域，选择地域。
3. 在[实例列表](#)页面，单击目标实例名称。
4. 在左侧导航栏，单击[Connector 管理](#)。
5. 在[Connector 管理](#)页面，单击[创建 Connector](#)。
6. 在[创建 Connector配置向导](#)页面，完成以下操作。
 - i. 在[配置基本信息](#)页签，按需配置以下参数，然后单击[下一步](#)。

 **注意** 消息队列Kafka版会为您自动选中授权创建服务关联角色。

- 如果未创建服务关联角色，消息队列Kafka版会为您自动创建一个服务关联角色，以便您使用消息队列Kafka版导出数据至对象存储的功能。
- 如果已创建服务关联角色，消息队列Kafka版不会重复创建。

关于该服务关联角色的更多信息，请参见[服务关联角色](#)。

参数	描述	示例值
名称	Connector的名称。命名规则： <ul style="list-style-type: none"> ■ 可以包含数字、小写英文字母和短划线（-），但不能以短划线（-）开头，长度限制为48个字符。 ■ 同一个消息队列Kafka版实例内保持唯一。 Connector的数据同步任务必须使用名称为connect-任务名称的Consumer Group。如果您未手动创建该Consumer Group，系统将为您自动创建。	kafka-oss-sink
实例	默认配置为实例的名称与实例ID。	demo alikafka_post-cn-st21p8vj****

- ii. 在[配置源服务](#)页签，选择数据源为[消息队列Kafka版](#)，并配置以下参数，然后单击[下一步](#)。

参数	描述	示例值
数据源 Topic	需要同步数据的Topic。	oss-test-input
消费线程并发数	数据源Topic的消费线程并发数。默认值为6。取值说明如下： <ul style="list-style-type: none"> 6 12 	6
消费初始位置	开始消费的位置。取值说明如下： <ul style="list-style-type: none"> 最早位点：从最初位点开始消费。 最近位点：从最新位点开始消费。 	最早位点
VPC ID	数据同步任务所在的VPC。单击配置运行环境显示该参数。默认为消息队列Kafka版实例所在的VPC，您无需填写。	vpc-bp1xpdnd3l***
VSwitch ID	数据同步任务所在的交换机。单击配置运行环境显示该参数。该交换机必须与消息队列Kafka版实例处于同一VPC。默认为部署消息队列Kafka版实例时填写的交换机。	vsw-bp1d2jgg81***
失败处理	消息发送失败后，是否继续订阅出现错误的Topic的分区。单击配置运行环境显示该参数。取值说明如下。 <ul style="list-style-type: none"> 继续订阅：继续订阅出现错误的Topic的分区，并打印错误日志。 停止订阅：停止订阅出现错误的Topic的分区，并打印错误日志 <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> 说明</p> <ul style="list-style-type: none"> 如何查看日志，请参见查看Connector日志。 如何根据错误码查找解决方案，请参见错误码。 如需恢复对出现错误的Topic的分区订阅，您需要提交工单联系消息队列Kafka版技术人员。 </div>	继续订阅
创建资源方式	选择创建Connector所依赖的Topic与Group的方式。单击配置运行环境显示该参数。 <ul style="list-style-type: none"> 自动创建 手动创建 	自动创建
Connector 消费组	Connector使用的Consumer Group。单击配置运行环境显示该参数。该Consumer Group的名称建议以connect-cluster开头。	connect-cluster-kafka-oss-sink

参数	描述	示例值
任务位点 Topic	<p>用于存储消费位点的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic: 建议以connect-offset开头。 分区数: Topic的分区数量必须大于1。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-offset-kafka-oss-sink
任务配置 Topic	<p>用于存储任务配置的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic: 建议以connect-config开头。 分区数: Topic的分区数量必须为1。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-config-kafka-oss-sink
任务状态 Topic	<p>用于存储任务状态的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic: 建议以connect-status开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-status-kafka-oss-sink
死信队列 Topic	<p>用于存储Connect框架的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和异常数据Topic为同一个Topic, 以节省Topic资源。</p> <ul style="list-style-type: none"> Topic: 建议以connect-error开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-oss-sink
异常数据 Topic	<p>用于存储Sink的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和死信队列Topic为同一个Topic, 以节省Topic资源。</p> <ul style="list-style-type: none"> Topic: 建议以connect-error开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-oss-sink

iii. 在配置目标服务页签，选择目标服务为对象存储，并配置以下参数，然后单击创建。

参数	描述	示例值
Bucket 名称	对象存储Bucket的名称。	bucket_test
Access Key	阿里云账号的AccessKey ID。	LTAl4GG2RGAjppjk*****
Secret Key	阿里云账号的AccessKey Secret。	WbGPVb5rrecVw3SQvEPw6R** *****

请确保您使用的AccessKey ID所对应的账号已被授予以下最小权限：

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": [
        "oss:GetObject",
        "oss:PutObject"
      ],
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

说明

AccessKey ID和AccessKey Secret是消息队列Kafka版创建任务时作为环境变量传递至对象存储的数据，任务创建成功后，消息队列Kafka版不保存AccessKey ID和AccessKey Secret信息。

创建完成后，在Connector 管理页面，查看创建的Connector。

7. 创建完成后，在Connector 管理页面，找到创建的Connector，单击其操作列的部署。

发送测试消息

您可以向消息队列Kafka版的数据源Topic发送消息，测试数据能否被导出至对象存储。

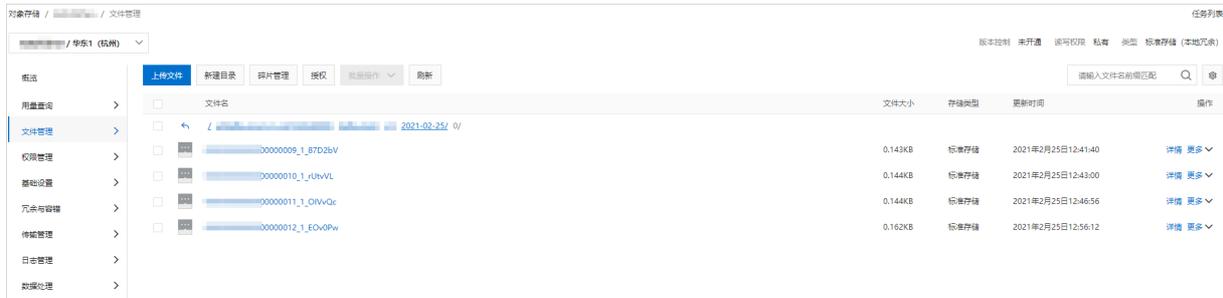
- 在Connector 管理页面，找到目标Connector，在其右侧操作列，单击测试。
- 在发送消息面板，发送测试消息。
 - 发送方式选择控制台。
 - 在消息 Key文本框中输入消息的Key值，例如demo。
 - 在消息内容文本框输入测试的消息内容，例如{"key": "test"}。
 - 设置发送到指定分区，选择是否指定分区。
 - 单击是，在分区 ID文本框中输入分区的ID，例如0。如果您需查询分区的ID，请参见[查看分区状态](#)。
 - 单击否，不指定分区。
 - 发送方式选择Docker，执行运行 Docker 容器生产示例消息区域的Docker命令，发送消息。

- 发送方式选择 SDK，根据您的业务需求，选择需要的语言或者框架的SDK以及接入方式，通过SDK发送消息。

验证结果

向消息队列Kafka版的数据源Topic发送消息后，查看OSS文件管理，验证数据导出结果。更多信息，请参见[文件概览](#)。

文件管理中显示新导出的文件。



消息队列Kafka版数据导出至对象存储的格式示例如下：

```
[
  {
    "key": "123",
    "offset": 4,
    "overflowFlag": true,
    "partition": 0,
    "timestamp": 1603779578478,
    "topic": "Test",
    "value": "1",
    "valueSize": 272687
  }
]
```

更多操作

您可以按需对该Connector所依赖的函数计算资源进行配置。

1. 在Connector 管理页面，找到创建的Connector，单击其操作列的更多 > 配置函数。页面跳转至函数计算控制台，您可以按需配置函数资源。

6.3.4. 创建Elasticsearch Sink Connector

本文介绍如何创建Elasticsearch Sink Connector将数据从消息队列Kafka版实例的数据源Topic导出至阿里云Elasticsearch的索引。

前提条件

在导出数据前，请确保您已完成以下操作：

- 为消息队列Kafka版实例开启Connector。更多信息，请参见[开启Connector](#)。
- 为消息队列Kafka版实例创建数据源Topic。更多信息，请参见[步骤一：创建Topic](#)。
- 在Elasticsearch管理控制台创建实例和索引。更多信息，请参见[快速开始](#)。

 **说明** 函数计算使用的Elasticsearch客户端版本为7.7.0，为保持兼容，您需创建7.0或以上版本的Elasticsearch实例。

- **开通函数计算服务。**

注意事项

- 仅支持在同地域内，将数据从消息队列Kafka版实例的数据源Topic导出至函数计算，再由函数计算导出至Elasticsearch。Connector的限制说明，请参见[使用限制](#)。
- 该功能基于函数计算服务提供。函数计算为您提供了一定的免费额度，超额部分将产生费用，请以函数计算的计费规则为准。计费详情，请参见[计费概述](#)。
- 函数计算的函数调用支持日志查询。具体操作步骤，请参见[配置并查看函数日志](#)。
- 消息转储时，消息队列Kafka版中消息用UTF-8 String序列化，暂不支持二进制的数据格式。

创建并部署Elasticsearch Sink Connector

1. 登录[消息队列Kafka版控制台](#)。
2. 在**概览**页面的**资源分布**区域，选择地域。
3. 在**实例列表**页面，单击目标实例名称。
4. 在左侧导航栏，单击**Connector 管理**。
5. 在**Connector 管理**页面，单击**创建 Connector**。
6. 在**创建 Connector**配置向导页面，完成以下操作。
 - i. 在**配置基本信息**页签，按需配置以下参数，然后单击**下一步**。

 **注意** 消息队列Kafka版会为您自动选中授权创建服务关联角色。

- 如果未创建服务关联角色，消息队列Kafka版会为您自动创建一个服务关联角色，以便您使用消息队列Kafka版导出数据至Elasticsearch的功能。
- 如果已创建服务关联角色，消息队列Kafka版不会重复创建。

关于该服务关联角色的更多信息，请参见[服务关联角色](#)。

参数	描述	示例值
名称	Connector的名称。命名规则： <ul style="list-style-type: none"> ■ 可以包含数字、小写英文字母和短划线（-），但不能以短划线（-）开头，长度限制为48个字符。 ■ 同一个消息队列Kafka版实例内保持唯一。 Connector的数据同步任务必须使用名称为connect-任务名称的Consumer Group。如果您未手动创建该Consumer Group，系统将为您自动创建。	kafka-elasticsearch-sink
实例	默认配置为实例的名称与实例ID。	demoalikafka_post-cn-st21p8vj****

- ii. 在**配置源服务**页签，选择数据源为**消息队列Kafka版**，并配置以下参数，然后单击**下一步**。

参数	描述	示例值
数据源 Topic	需要同步数据的Topic。	elasticsearch-test-input
消费线程并发数	数据源Topic的消费线程并发数。默认值为6。取值说明如下： <ul style="list-style-type: none"> 6 12 	6
消费初始位置	开始消费的位置。取值说明如下： <ul style="list-style-type: none"> 最早位点：从最初位点开始消费。 最近位点：从最新位点开始消费。 	最早位点
VPC ID	数据同步任务所在的VPC。单击配置运行环境显示该参数。默认为消息队列Kafka版实例所在的VPC，您无需填写。	vpc-bp1xpdnd3l***
VSwitch ID	数据同步任务所在的交换机。单击配置运行环境显示该参数。该交换机必须与消息队列Kafka版实例处于同一VPC。默认为部署消息队列Kafka版实例时填写的交换机。	vsw-bp1d2jgg81***
失败处理	消息发送失败后，是否继续订阅出现错误的Topic的分区。单击配置运行环境显示该参数。取值说明如下。 <ul style="list-style-type: none"> 继续订阅：继续订阅出现错误的Topic的分区，并打印错误日志。 停止订阅：停止订阅出现错误的Topic的分区，并打印错误日志 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p>? 说明</p> <ul style="list-style-type: none"> 如何查看日志，请参见查看Connector日志。 如何根据错误码查找解决方案，请参见错误码。 如需恢复对出现错误的Topic的分区的订阅，您需要提交工单联系消息队列Kafka版技术人员。 </div>	继续订阅
创建资源方式	选择创建Connector所依赖的Topic与Group的方式。单击配置运行环境显示该参数。 <ul style="list-style-type: none"> 自动创建 手动创建 	自动创建
Connector 消费组	Connector的数据同步任务使用的Consumer Group。单击配置运行环境显示该参数。该Consumer Group的名称必须为connect-任务名称。	connect-kafka-elasticsearch-sink

参数	描述	示例值
任务位点 Topic	<p>用于存储消费位点的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic: 建议以connect-offset开头。 分区数: Topic的分区数量必须大于1。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-offset-kafka-elasticsearch-sink
任务配置 Topic	<p>用于存储任务配置的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic: 建议以connect-config开头。 分区数: Topic的分区数量必须为1。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-config-kafka-elasticsearch-sink
任务状态 Topic	<p>用于存储任务状态的Topic。单击配置运行环境显示该参数。</p> <ul style="list-style-type: none"> Topic: 建议以connect-status开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎必须为Local存储。 cleanup.policy: Topic的日志清理策略必须为compact。 	connect-status-kafka-elasticsearch-sink
死信队列 Topic	<p>用于存储Connect框架的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和异常数据Topic为同一个Topic, 以节省Topic资源。</p> <ul style="list-style-type: none"> Topic: 建议以connect-error开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-elasticsearch-sink
异常数据 Topic	<p>用于存储Sink的异常数据的Topic。单击配置运行环境显示该参数。该Topic可以和死信队列Topic为同一个Topic, 以节省Topic资源。</p> <ul style="list-style-type: none"> Topic: 建议以connect-error开头。 分区数: Topic的分区数量建议为6。 存储引擎: Topic的存储引擎可以为Local存储或云存储。 	connect-error-kafka-elasticsearch-sink

iii. 在配置目标服务页签，选择目标服务为 *Elasticsearch*，并配置以下参数，然后单击创建。

参数	描述	示例值
ES 实例 ID	阿里云Elasticsearch实例ID。	es-cn-oew1o67x0000****
接入地址	阿里云Elasticsearch实例的公网或私网地址。更多信息，请参见 查看实例的基本信息 。	es-cn-oew1o67x0000****.elasticsearch.aliyuncs.com
接入端口	访问阿里云Elasticsearch的公网或私网端口，取值如下： <ul style="list-style-type: none"> 9200：基于HTTP或HTTPS。 9300：基于TCP。 更多信息，请参见 查看实例的基本信息 。	9300
用户名	登录Kibana控制台的用户名，默认为elastic。您也可以创建自定义用户，创建步骤，请参见 创建用户 。	elastic
用户密码	登录Kibana控制台的密码。elastic用户的密码在创建实例时设定，如果忘记可重置。如需重置，请参见 重置实例访问密码 。	*****
索引	阿里云Elasticsearch的索引名称。	elastic_test

说明

- 用户名和用户密码会被用来初始化Elasticsearch对象，通过bulk投递消息，请确认账号对索引有写权限。
- 用户名和用户密码是消息队列Kafka版创建任务时作为环境变量传递至函数计算的函数，任务创建成功后，消息队列Kafka版不保存相关信息。

创建完成后，在Connector 管理页面，查看创建的Connector。

7. 创建完成后，在Connector 管理页面，找到创建的Connector，单击其操作列的部署。

配置函数服务

您在消息队列Kafka版控制台成功创建并部署Elasticsearch Sink Connector后，函数计算会自动为您创建给该Connector使用的函数服务，服务命名格式为 `kafka-service-<connector_name>-<随机String>`。

- 在Connector 管理页面，找到目标Connector，在其右侧操作列，选择更多 > 查看日志。页面跳转至函数计算控制台。
- 在函数计算控制台，找到自动创建的函数服务，并配置其VPC和交换机信息。请确保该信息和您阿里云Elasticsearch相同。配置的具体步骤，请参见[更新服务](#)。

发送测试消息

您可以向消息队列Kafka版的数据源Topic发送消息，测试数据能否被导出至阿里云Elasticsearch。

1. 在**Connector** 管理页面，找到目标Connector，在其右侧操作列，单击**测试**。
2. 在**发送消息**面板，发送测试消息。
 - **发送方式选择控制台**。
 - a. 在**消息 Key**文本框中输入消息的Key值，例如demo。
 - b. 在**消息内容**文本框输入测试的消息内容，例如 {"key": "test"}。
 - c. 设置**发送到指定分区**，选择是否指定分区。
 - a. 单击**是**，在**分区 ID**文本框中输入分区的ID，例如0。如果您需查询分区的ID，请参见[查看分区状态](#)。
 - b. 单击**否**，不指定分区。
 - **发送方式选择Docker**，执行运行 **Docker 容器生产示例消息**区域的Docker命令，发送消息。
 - **发送方式选择 SDK**，根据您的业务需求，选择需要的语言或者框架的SDK以及接入方式，通过SDK发送消息。

验证结果

向消息队列Kafka版的数据源Topic发送消息后，[登录Kibana控制台](#)，执行 `GET /<index_name>/_search` 查看索引，验证数据导出结果。

消息队列Kafka版数据导出至Elasticsearch的格式示例如下：

```
{
  "took": 8,
  "timed_out": false,
  "_shards": {
    "total": 5,
    "successful": 5,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": {
      "value": 1,
      "relation": "eq"
    },
    "max_score": 1.0,
    "hits": [
      {
        "_index": "product_****",
        "_type": "_doc",
        "_id": "TX3TZHgBfHNEDGoZ****",
        "_score": 1.0,
        "_source": {
          "msg_body": {
            "key": "test",
            "offset": 2,
            "overflowFlag": false,
            "partition": 2,
            "timestamp": 1616599282417,
            "topic": "dv****",
            "value": "test1",
            "valueSize": 8
          },
          "doc_as_upsert": true
        }
      }
    ]
  }
}
```

6.3.5. 创建MySQL Source Connector

本文介绍如何创建MySQL Source Connector，通过DataWorks将数据从阿里云数据库RDS MySQL版导出至消息队列Kafka版实例的Topic。

前提条件

在导出数据前，请确保您已完成以下操作：

- 为消息队列Kafka版实例开启Connector。更多信息，请参见[开启Connector](#)。

 **注意** 请确保您的消息队列Kafka版实例部署在华南1（深圳）、西南1（成都）、华北2（北京）、华北3（张家口）、华东1（杭州）、华东2（上海）或新加坡地域。

- 创建RDS MySQL实例。
- 创建数据库和账号。
- 创建数据库表。常见的SQL语句，请参见[常用语句](#)。
- 阿里云账号和RAM用户均须授予DataWorks访问您弹性网卡ENI资源的权限。授予权限，请访问[云资源访问授权](#)。

注意 如果您使用的是RAM用户，请确保您的账号有以下权限：

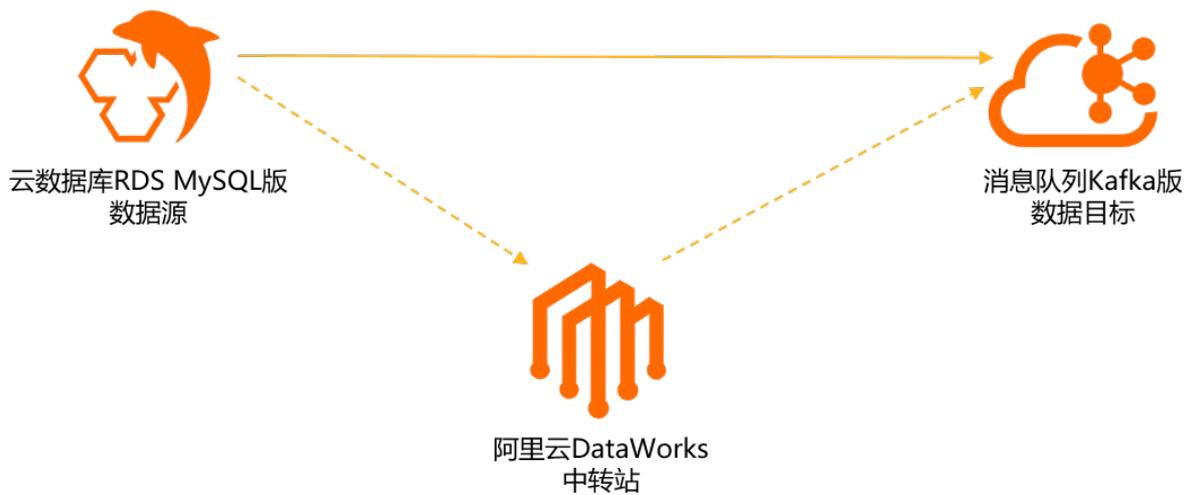
- AliyunDataWorksFullAccess：DataWorks所有资源的管理权限。
- AliyunBSSOrderAccess：购买阿里云产品的权限。

如何为RAM用户添加权限策略，请参见[步骤二：为RAM用户添加权限](#)。

- 请确保您是阿里云数据库RDS MySQL版实例（数据源）和消息队列Kafka版实例（数据目标）的所有者，即创建者。
- 请确保阿里云数据库RDS MySQL版实例（数据源）和消息队列Kafka版实例（数据目标）所在的VPC网段没有重合，否则无法成功创建同步任务。

背景信息

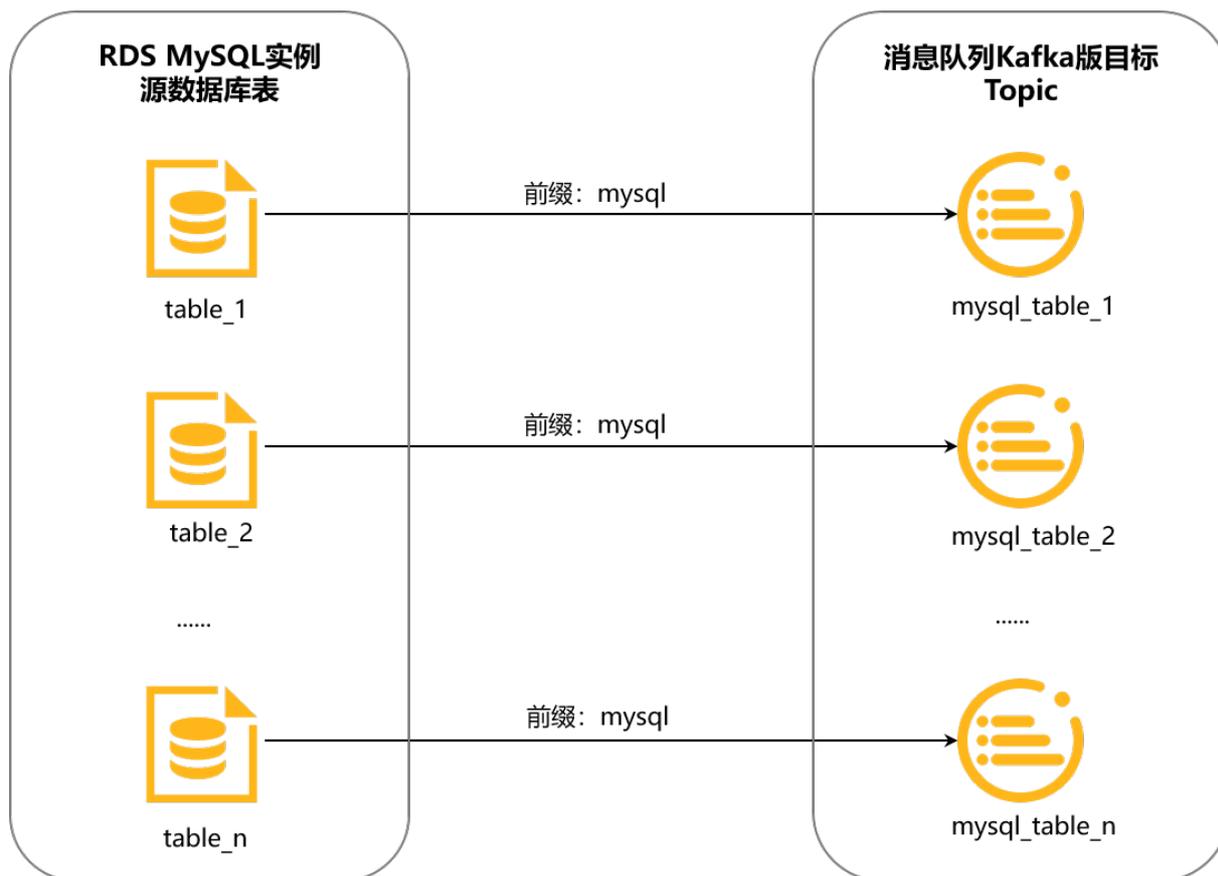
您可以在[消息队列Kafka版控制台](#)创建数据同步任务，将您在阿里云数据库RDS MySQL版数据库表中的数据同步至消息队列Kafka版的Topic。该同步任务将依赖阿里云DataWorks产品实现，流程图如下所示。



如果您在消息队列Kafka版控制台成功创建了数据同步任务，那么阿里云DataWorks会自动为您开通DataWorks产品基础版服务（免费）、新建DataWorks项目（免费）、并新建数据集成独享资源组（需付费），资源组规格为4c8g，购买模式为包年包月，时长为1个月并自动续费。阿里云DataWorks的计费详情，请参见[DataWorks计费概述](#)。

此外，DataWorks会根据您数据同步任务的配置，自动为您生成消息队列Kafka版的目标Topic。数据库表和Topic是一对一的关系，对于有主键的表，默认6分区；无主键的表，默认1分区。请确保实例剩余Topic数和分区数充足，不然任务会因为创建Topic失败而导致异常。

Topic的命名格式为 `<配置的前缀>_<数据库表名>`，下划线（`_`）为系统自动添加的字符。详情如下图所示。



例如，您将前缀配置为mysql，需同步的数据库表名为table_1，那么DataWorks会自动为您生成专用Topic，用来接收table_1同步过来的数据，该Topic的名称为mysql_table_1；table_2的专用Topic名称为mysql_table_2，以此类推。

注意事项

- 地域说明
 - 如果数据源和目标实例位于不同地域，请确保您使用的账号拥有云企业网实例，且云企业网实例已挂载数据源和目标实例所在的VPC，并配置好流量带宽完成网络打通。
 否则，可能会新建云企业网实例，并将目标实例和独享资源组ECS全部挂载到云企业网实例来打通网络。这样的云企业网实例没有配置带宽，所以带宽流量很小，可能导致创建同步任务过程中的网络访问出错，或者同步任务创建成功后，在运行过程中出错。
 - 如果数据源和目标实例位于同一地域，创建数据同步任务会自动在其中一个实例所在VPC创建ENI，并绑定到独享资源组ECS上，以打通网络。
- DataWorks独享资源组说明
 - DataWorks的每个独享资源组可以运行最多3个同步任务。创建数据同步任务时，如果DataWorks发现您的账号名下有资源组的历史购买记录，并且运行的同步任务少于3个，将使用已有资源组运行新建的同步任务。
 - DataWorks的每个独享资源组最多绑定两个VPC的ENI。如果DataWorks发现已购买的资源组绑定的ENI与需要新绑定的ENI有网段冲突，或者其他技术限制，导致使用已购买的资源组无法创建出同步任务，此时，即使已有的资源组运行的同步任务少于3个，也将新建资源组确保同步任务能够顺利创建。

创建并部署MySQL Source Connector

1. 登录消息队列Kafka版控制台。

2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，单击创建 Connector。
6. 在创建 Connector配置向导中，完成以下操作。
 - i. 在配置基本信息页签的名称文本框，输入Connector名称，然后单击下一步。

参数	描述	示例值
名称	<p>Connector的名称。命名规则：</p> <ul style="list-style-type: none"> ▪ 可以包含数字、小写英文字母和短划线（-），但不能以短划线（-）开头，长度限制为48个字符。 ▪ 同一个消息队列Kafka版实例内保持唯一。 <p>Connector的数据同步任务必须使用名称为connect-任务名称的Consumer Group。如果您未手动创建该Consumer Group，系统将为您自动创建。</p>	kafka-source-mysql
实例	默认配置为实例的名称与实例ID。	demo alikafka_post-cn-st21p8vj****

ii. 在配置源服务页签，选择数据源为云数据库RDS MySQL版，配置以下参数，然后单击下一步。

参数	描述	示例值
RDS 实例所在地域	从下拉列表中，选择阿里云数据库RDS MySQL版实例所在的地域。	华南1（深圳）
云数据库 RDS 实例 ID	需要同步数据的阿里云数据库RDS MySQL版的实例ID。	rm-wz91w3vk6owmz****
数据库名称	需要同步的阿里云数据库RDS MySQL版实例数据库的名称。	mysql-to-kafka
数据库账号	需要同步的阿里云数据库RDS MySQL版实例数据库账号。	mysql_to_kafka
数据库账号密码	需要同步的阿里云数据库RDS MySQL版实例数据库账号的密码。	无
数据库表	需要同步的阿里云数据库RDS MySQL版实例数据库表的名称，多个表名以英文逗号（,）分隔。 数据库表和目标Topic是一对一的关系。	mysql_tbl
Topic 前缀	阿里云数据库RDS MySQL版数据库表同步到消息队列Kafka版的Topic的命名前缀，请确保前缀全局唯一。	mysql

 注意

请确保阿里云数据库RDS MySQL版数据库账号有以下最小权限：

- SELECT
- REPLICATION SLAVE
- REPLICATION CLIENT

授权命令示例：

```
GRANT SELECT, REPLICATION SLAVE, REPLICATION CLIENT ON *.* TO '同步账号'@'%';//授予数据库账号的SELECT、REPLICATION SLAVE和REPLICATION CLIENT权限。
```

iii. 在配置目标服务页签，显示数据将同步到目标消息队列Kafka版实例，确认信息无误后，单击创建。

7. 创建完成后，在Connector 管理页面，找到创建的Connector，单击其操作列的部署。在Connector 管理页面，您可以看到创建的任务状态为运行中，则说明任务创建成功。

 说明 如果创建失败，请再次检查本文前提条件中的操作是否已全部完成。

验证结果

1. 向阿里云数据库RDS MySQL版数据库表插入数据。示例如下。

```
INSERT INTO mysql_tbl
(mysql_title, mysql_author, submission_date)
VALUES
("mysql2kafka", "tester", NOW())
```

更多SQL语句，请参见[常用语句](#)。

2. 使用消息队列Kafka版提供的消息查询功能，验证数据能否被导出至消息队列Kafka版目标Topic。查询的具体步骤，请参见[查询消息](#)。
云数据库RDS MySQL版数据库表导出至消息队列Kafka版Topic的数据示例如下。消息结构及各字段含义，请参见[附录：消息格式](#)。

```
{
  "schema":{
    "dataColumn":[
      {
        "name":"mysql_id",
        "type":"LONG"
      },
      {
        "name":"mysql_title",
        "type":"STRING"
      },
      {
        "name":"mysql_author",
        "type":"STRING"
      },
      {
        "name":"submission_date",
        "type":"DATE"
      }
    ],
    "primaryKey":[
      "mysql_id"
    ],
    "source":{
      "dbType":"MySQL",
      "dbName":"mysql_to_kafka",
      "tableName":"mysql_tbl"
    }
  },
  "payload":{
    "before":null,
    "after":{
      "dataColumn":{
        "mysql_title":"mysql2kafka",
        "mysql_author":"tester",
        "submission_date":1614700800000
      }
    }
  },
  "sequenceId":"1614748790461000000",
  "timestamp":{
    "eventTime":1614748870000,
    "systemTime":1614748870925,
    "checkpointTime":1614748870000
  },
  "op":"INSERT",
  "ddl":null
},
"version":"0.0.1"
}
```

6.4. 查看Connector任务配置

如果您需要查看已创建的Connector的基础信息、源实例配置、目标实例配置等信息，您可以在消息队列Kafka版控制台查看Connector的任务配置。

前提条件

您已创建以下任意一种Connector:

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)
- [创建OSS Sink Connector](#)
- [创建Elasticsearch Sink Connector](#)
- [创建MySQL Source Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，找到目标Connector，单击操作的详情。
在Connector详情页面，显示Connector的基础信息、源信息和目标信息。

6.5. 查看Connector日志

如果您的Connector出现异常，您可以在消息队列Kafka版控制台查看Connector的日志以进行异常排查。

前提条件

您已创建以下任意一种Connector:

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)
- [创建OSS Sink Connector](#)
- [创建Elasticsearch Sink Connector](#)
- [创建MySQL Source Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，找到目标Connector，在其操作列，选择更多 > 查看日志。
在查看 Connector日志页面，查看Connector的日志。

6.6. 暂停Connector

如果您需要暂时中止某个运行中的Connector，您可以在消息队列Kafka版控制台暂停该Connector。

前提条件

您已创建以下任意一种Connector，且该Connector处于运行中状态：

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)
- [创建OSS Sink Connector](#)
- [创建Elasticsearch Sink Connector](#)
- [创建MySQL Source Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，找到目标Connector，在其操作列，选择更多 > 暂停。在提示对话框，单击确认，暂停Connector。
在Connector 管理页面，找到暂停的Connector，状态已变更为已暂停。

6.7. 恢复Connector

如果您需要恢复某个暂停的Connector，您可以在消息队列Kafka版控制台恢复该Connector。

前提条件

您已创建以下任意一种Connector，且该Connector处于暂停状态：

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)
- [创建OSS Sink Connector](#)
- [创建Elasticsearch Sink Connector](#)
- [创建MySQL Source Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，找到目标Connector，单击其操作的启用。
在Connector 管理页面，找到启用的Connector，其状态已变更为进行中。

6.8. 删除Connector

消息队列Kafka版限制了每个实例的Connector数量。如果您不再需要某个Connector，您可以在消息队列Kafka版控制台删除该Connector。

前提条件

您已创建以下任意一种Connector：

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)
- [创建OSS Sink Connector](#)
- [创建Elasticsearch Sink Connector](#)
- [创建MySQL Source Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，找到目标Connector，在其操作列，选择更多 > 删除。
6. 在提示对话框，单击确认，删除Connector。

注意

- 如果MySQL Source Connector任务处于运行状态，在消息队列Kafka版控制台将无法直接删除，您需登录DataWorks控制台停止并下线Connector任务，然后联系消息队列Kafka版值班号清理消息队列Kafka版Connector任务的元信息。其他FC Sink Connector、MaxCompute Sink Connector、OSS Sink Connector以及Elasticsearch Sink Connector任务，均可在消息队列Kafka版控制台直接删除。
- 当删除Connector时，系统会同时删除该Connector依赖的5个Topic和2个Consumer Group，无论这些资源当时是自动创建的还是手动创建的。

6.9. 修改Connector配置

成功创建FC Sink Connector后，您可以在消息队列Kafka版控制台更新其配置。

前提条件

[创建FC Sink Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector 管理。
5. 在Connector 管理页面，找到目标Connector，在其操作列，选择更多 > 修改配置。
6. 在修改配置面板，按需修改以下参数，然后单击确定。

参数	描述
----	----

参数	描述
消费线程并发数	<p>数据源Topic的消费线程并发数。默认值为6。取值说明如下：</p> <ul style="list-style-type: none"> ◦ 1 ◦ 2 ◦ 3 ◦ 6 ◦ 12
失败处理	<p>消息发送失败后，是否继续订阅出现错误的Topic的分区。取值说明如下。</p> <ul style="list-style-type: none"> ◦ 继续订阅：继续订阅出现错误的Topic的分区，并打印错误日志。 ◦ 停止订阅：停止订阅出现错误的Topic的分区，并打印错误日志 <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> 说明</p> <ul style="list-style-type: none"> ◦ 如何查看日志，请参见查看Connector日志。 ◦ 如何根据错误码查找解决方案，请参见错误码。 ◦ 如需恢复对出现错误的Topic的分区的订阅，您需要提交工单联系消息队列Kafka版技术人员。 </div>
重试次数	<p>消息发送失败后的重试次数。默认为2。取值范围为1~3。部分导致消息发送失败的错误不支持重试。错误码与是否支持重试的对应关系如下：</p> <ul style="list-style-type: none"> ◦ 4XX：除429支持重试外，其余错误码不支持重试。 ◦ 5XX：支持重试。 <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> 说明 Connector调用InvokeFunction向函数计算发送消息。</p> </div>

执行结果

修改完成后，在Connector管理页面，找到目标Connector，单击其操作的详情。在Connector的详情页面，查看到更新后的Connector配置。

6.10. 测试Connector

如果您需要测试某个Connector，您可以在消息队列Kafka版控制台向Connector发送测试消息。

前提条件

您已创建以下任意一种Connector：

- [创建FC Sink Connector](#)
- [创建MaxCompute Sink Connector](#)
- [创建OSS Sink Connector](#)
- [创建Elasticsearch Sink Connector](#)
- [创建MySQL Source Connector](#)

操作步骤

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击Connector管理。
5. 在Connector管理页面，找到目标Connector，在其右侧操作列，单击测试。
6. 在发送消息面板，发送测试消息。
 - 发送方式选择控制台。
 - a. 在消息Key文本框中输入消息的Key值，例如demo。
 - b. 在消息内容文本框输入测试的消息内容，例如{"key": "test"}。
 - c. 设置发送到指定分区，选择是否指定分区。
 - a. 单击是，在分区ID文本框中输入分区的ID，例如0。如果您需查询分区的ID，请参见[查看分区状态](#)。
 - b. 单击否，不指定分区。
 - 发送方式选择Docker，执行运行Docker容器生产示例消息区域的Docker命令，发送消息。
 - 发送方式选择SDK，根据您的业务需求，选择需要的语言或者框架的SDK以及接入方式，通过SDK发送消息。

6.11. 为Connector开启公网访问

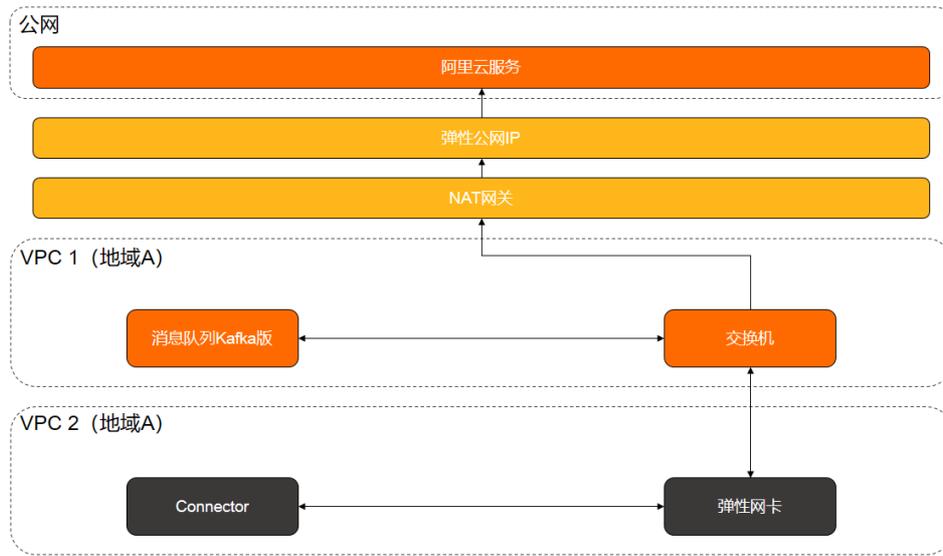
如需使Connector跨地域访问其他阿里云服务，您需要为Connector开启公网访问。本文介绍如何为Connector开启公网访问。

前提条件

为Connector开启公网访问前，请确保您的消息队列Kafka版实例已开启Connector。更多信息，请参见[开启Connector](#)。

开启公网访问

为Connector开启公网访问的方案如下：



为Connector开启公网访问的操作流程如下：

1. 为部署了消息队列Kafka版实例的VPC 1创建NAT网关。更多信息，请参见[创建NAT网关实例](#)。
2. 为创建的NAT网关绑定弹性公网IP。更多信息，请参见[绑定NAT网关](#)。
3. 为VPC 1下消息队列Kafka版实例使用的交换机创建SNAT条目。更多信息，请参见[创建SNAT实现访问公网服务](#)。

7. 迁移

7.1. 迁移概述

本文介绍将自建Kafka集群迁移到消息队列Kafka版实例的优势、原理和流程。

迁移优势

将自建Kafka集群迁移到消息队列Kafka版实例的优势，请参见[产品优势](#)。

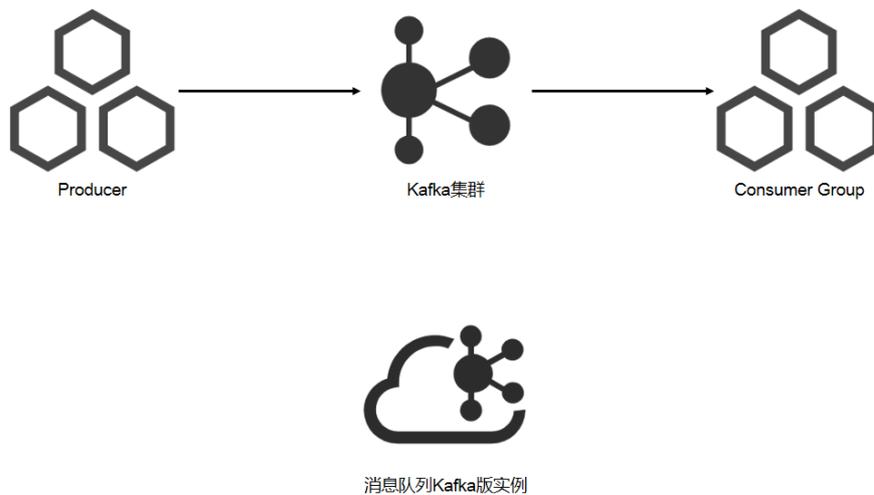
迁移原理

对于消息队列来说，如果要实现集群迁移，只需消费完旧集群的消息即可。由于Producer和Consumer Group都是集群化的，您可以通过一台一台操作的方式实现上层业务无感知。

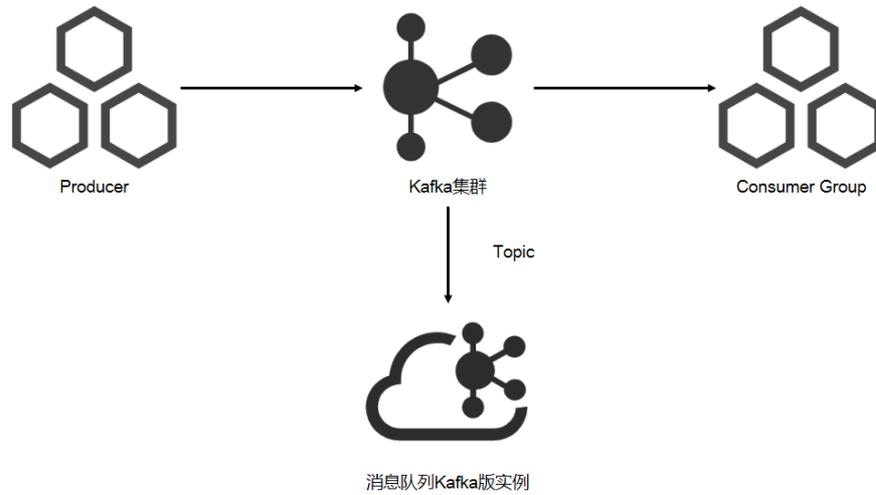
迁移流程

将自建Kafka集群迁移到消息队列Kafka版实例的操作流程如下：

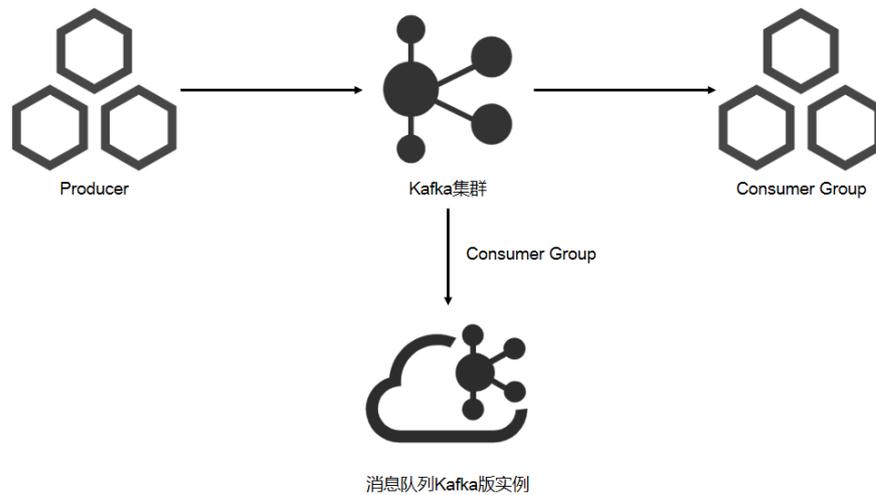
1. 评估需要购买的消息队列Kafka版实例规格。详情请参见[评估规格](#)。
2. 根据推荐意见，购买并部署消息队列Kafka版实例。



3. 迁移自建Kafka集群的Topic到消息队列Kafka版实例。详情请参见[迁移Topic上云](#)。



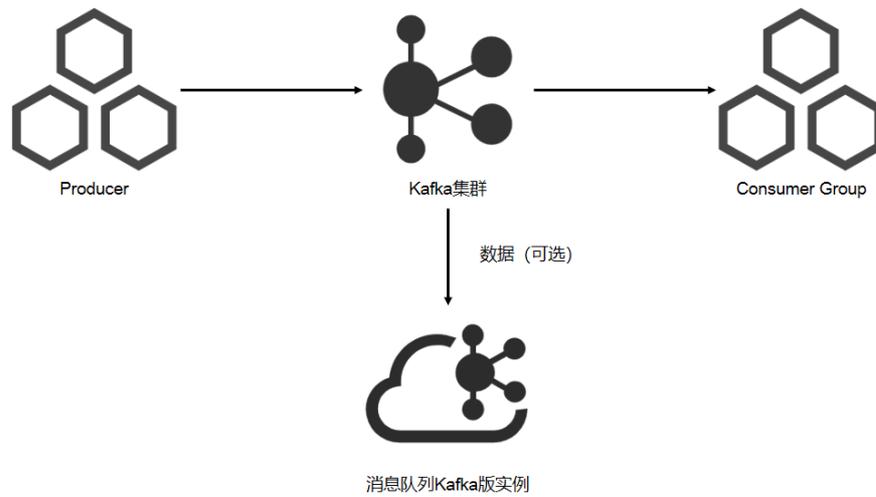
4. 迁移自建Kafka集群的Consumer Group到消息队列Kafka版实例。详情请参见[迁移Consumer Group上云](#)。



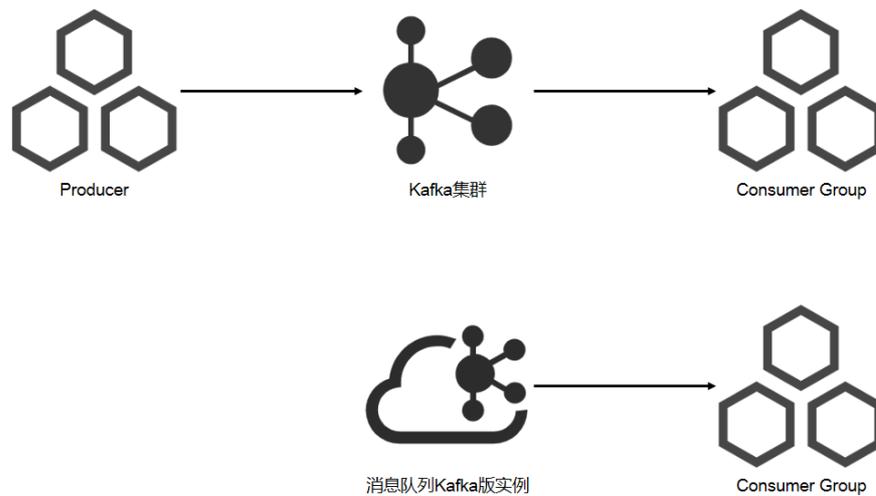
5. (可选) 迁移自建Kafka集群的数据到消息队列Kafka版实例。

注意 消息队列的特点是，数据一旦被消费，则已经完成使命。因此，除了需要将自建Kafka集群的数据备份消息队列Kafka版实例的情况外，一般情况下不推荐您迁移数据。

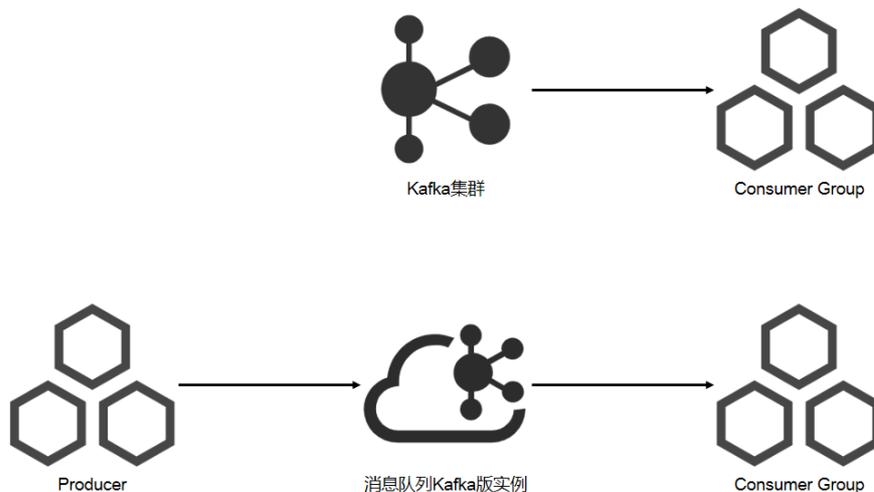
详情请参见[迁移数据上云](#)。



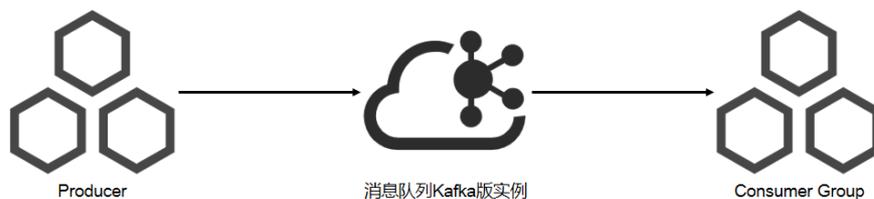
6. 为消息队列Kafka版实例开启新的Consumer Group，准备消费消息队列Kafka版实例的消息。



7. 为消息队列Kafka版实例开启新的Producer，下线旧的Producer，并使旧的Consumer Group继续消费自建Kafka集群的消息。



8. 待自建Kafka集群的消息全部被旧的Consumer Group消费后，下线旧的Consumer Group和自建Kafka集群。



7.2. 评估规格

本文说明如何使用消息队列Kafka版的规格评估功能为迁移上云的自建Kafka集群选择合适的实例规格。

前提条件

注册阿里云账号并完成实名认证。详情请参见[注册阿里云账号](#)。

评估规格

说明 消息队列Kafka版的规格评估功能遵循最小原则，即在满足业务正常需求的基础上，尽量推荐价格较低的消息队列Kafka版实例。

评估规格的操作步骤如下：

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在左侧导航栏，单击迁移上云。

4. 在迁移上云页面，单击规格评估。
5. 在规格评估面板，输入自建Kafka集群的规格信息，然后单击确定获取规格建议。

在获取推荐实例规格成功！之后，在规格评估面板显示系统为您推荐实例规格。

规格评估

获取推荐实例规格成功!
您可以在实例列表中点击“购买实例”，并根据如下建议购买 Kafka 实例。

推荐实例规格

规格类型	标准版	实例类型	公网/VPC实例
流量峰值	20 MB/s	磁盘类型	SSD
磁盘容量	900 GB	Topic 规格	50 个
服务端版本	0.10.2		

* 服务端版本 0.10.x 以下 0.10.x 1.x.x 2.x.x

* IDC 迁移 是 否

* 集群峰值流量 MB/s

* 公网流量 MB/s

* SSD 磁盘 是 否

* 磁盘容量 GB

* Topic 数量 ↑

* 分区总数 ↑

如需了解消息队列Kafka版实例规格详情，请参见[规格类型](#)。

后续步骤

购买并部署消息队列Kafka版实例：

- [购买并部署VPC实例](#)
- [购买并部署公网/VPC实例](#)

7.3. 迁移Topic

7.3.1. 迁移Topic上云

本文介绍如何使用消息队列Kafka版提供的迁移工具将自建Kafka集群的Topic迁移到消息队列Kafka版实例。

前提条件

您已完成以下操作：

- [下载JDK 8](#)
- [下载迁移工具JAR文件](#)

注意

- 迁移不会删除自建的源Kafka集群的Topic，只是在目标消息队列Kafka版实例创建相同配置的Topic。
- 迁移内容仅为Topic配置，不包含Topic中存储的消息。

操作步骤

1. 打开命令行工具。
2. 使用cd命令将路径切换到迁移工具所在目录。
3. 执行以下命令确认要迁移的Topic。 `java -jar kafka-migration.jar TopicMigrationFromZk --sourceZkConnect 192.168.XX.XX --destAk <yourdestAccessKeyId> --destSk <yourdestAccessKeySecret> --destRegionId <yourdestRegionId> --destInstanceId <yourdestInstanceId>`

参数	描述
sourceZkConnect	自建的源ZooKeeper集群的IP地址
destAk	目标消息队列Kafka版实例所属阿里云账号的AccessKey ID
destSk	目标消息队列Kafka版实例所属阿里云账号的AccessKey Secret
destRegionId	目标消息队列Kafka版实例的地域ID
destInstanceId	目标消息队列Kafka版实例的ID

待确认的返回结果示例如下：

```
13:40:08 INFO - Begin to migrate topics:[test]
13:40:08 INFO - Total topic number:1
13:40:08 INFO - Will create topic:test, isCompactTopic:false, partition number:1
```

4. 执行以下命令提交要迁移的Topic。 `java -jar kafka-migration.jar TopicMigrationFromZk --sourceZkConnect 192.168.XX.XX --destAk <yourAccessKeyId> --destSk <yourAccessKeySecret> --destRegionId <yourRegionId> --destInstanceId <yourInstanceId> --commit`

参数	描述
commit	提交迁移

提交迁移的返回结果示例如下：

```
13:51:12 INFO - Begin to migrate topics:[test]
13:51:12 INFO - Total topic number:1
13:51:13 INFO - cmd=TopicMigrationFromZk, request=null, response={"code":200,"requestId":"7F76C7D7-AAB5-4E29-B49B-CD6F1E0F508B","success":true,"message":"operation success"}
13:51:13 INFO - TopicCreate success, topic=test, partition number=1, isCompactTopic=false
```

5. 确认Topic迁移是否成功。
 - i. 登录[消息队列Kafka版控制台](#)。
 - ii. 在概览页面的资源分布区域，选择地域。
 - iii. 在实例列表页面，单击目标实例名称。
 - iv. 在左侧导航栏，单击Topic 管理。
 - v. 在Topic 管理页面的Topic列表中，显示成功迁移的Topic。

7.3.2. 云上迁移Topic

本文介绍如何使用消息队列Kafka版提供的迁移工具将某个消息队列Kafka版实例的Topic迁移到另一个消息队列Kafka版实例。

前提条件

您已完成以下操作：

- [下载JDK 8](#)
- [下载迁移工具JAR文件](#)

注意

- 迁移不会删除源消息队列Kafka版实例的Topic，只是在目标消息队列Kafka版实例创建相同配置的Topic。
- 迁移内容仅为Topic配置，不包含Topic中存储的数据。

操作步骤

1. 打开命令行工具。
2. 使用cd命令将路径切换到迁移工具所在目录。
3. 执行以下命令确认要迁移的Topic。


```
java -jar kafka-migration.jar TopicMigrationFromAlivun --sourceAk <voursourceAccessKeyId> --sourceSk <voursourceAccessKeySecret> --sourceRegionId <voursourceRegionId> --sourceInstanceId <voursourceInstanceId> --destAk <yourdestAccessKeyId> --destSk <yourdestAccessKeySecret> --destRegionId <yourdestRegionId> --destInstanceId <yourdestInstanceId>
```

参数	描述
sourceAk	源消息队列Kafka版实例所属阿里云账号的AccessKey ID

参数	描述
sourceSk	源消息队列Kafka版实例所属阿里云账号的AccessKey Secret
sourceRegionId	源消息队列Kafka版实例的地域ID
sourceInstanceId	源消息队列Kafka版实例的ID
destAk	目标消息队列Kafka版实例所属阿里云账号的AccessKey ID
destSk	目标消息队列Kafka版实例所属阿里云账号的AccessKey Secret
destRegionId	目标消息队列Kafka版实例的地域ID
destInstanceId	目标消息队列Kafka版实例的ID

待确认的返回结果示例如下：

```
15:13:12 INFO - cmd=TopicMigrationFromAliyun, request=null, response={"total":4,"code":200,"requestId":"1CBAB340-2146-43A3-8470-84D77DB8B43E","success":true,"pageSize":10000,"currentPage":1,"message":"operation success.", "topicList":[{"instanceId":"alikafka_pre-cn-0pp1cng20***", "localTopic":false, "createTime":1578558314000, "regionId":"cn-hangzhou", "statusName":"服务中", "topic":"agdagasd", "remark":"agdadgdasg", "partitionNum":12, "compactTopic":false, "status":0, "tags":[]}, {"instanceId":"alikafka_pre-cn-0pp1cng20***", "localTopic":false, "createTime":1578558294000, "regionId":"cn-hangzhou", "statusName":"服务中", "topic":"135215", "remark":"1315215", "partitionNum":12, "compactTopic":false, "status":0, "tags":[]}, {"instanceId":"alikafka_pre-cn-0pp1cng20***", "localTopic":false, "createTime":1578558214000, "regionId":"cn-hangzhou", "statusName":"服务中", "topic":"1332", "remark":"13414", "partitionNum":12, "compactTopic":false, "status":0, "tags":[]}, {"instanceId":"alikafka_pre-cn-0pp1cng20***", "localTopic":false, "createTime":1578558141000, "regionId":"cn-hangzhou", "statusName":"服务中", "topic":"aete", "remark":"est", "partitionNum":12, "compactTopic":false, "status":0, "tags":[]}]}
15:13:12 INFO - Will create topic:agdagasd, isCompactTopic:false, partition number:12
15:13:12 INFO - Will create topic:135215, isCompactTopic:false, partition number:12
15:13:12 INFO - Will create topic:1332, isCompactTopic:false, partition number:12
15:13:12 INFO - Will create topic:aete, isCompactTopic:false, partition number:12
```

4. 执行以下命令提交要迁移的Topic。 `java -jar kafka-migration.jar TopicMigrationFromAliyun --sourceAk <voursourceAccessKeyId> --sourceSk <voursourceAccessKeySecret> --sourceRegionId <voursourceRegionId> --sourceInstanceId <voursourceInstanceId> --destAk <vourdestAccessKeyId> --destSk <vourdestAccessKeySecret> --destRegionId <yourdestRegionId> --destInstanceId <yourdestInstanceId> --commit`

参数	描述
commit	提交迁移

提交迁移的返回结果示例如下：

```
16:38:30 INFO - cmd=TopicMigrationFromAliyun, request=null, response={"code":200,"requestId":"A0C
A4D70-46D4-45CF-B9E0-B117610A26DB"},"success":true,"message":"operation success"}
16:38:30 INFO - TopicCreate success, topic=agdagasd, partition number=12, isCompactTopic=false
16:38:36 INFO - cmd=TopicMigrationFromAliyun, request=null, response={"code":200,"requestId":"05E
88C75-64B6-4C87-B962-A63D906FD993"},"success":true,"message":"operation success"}
16:38:36 INFO - TopicCreate success, topic=135215, partition number=12, isCompactTopic=false
16:38:42 INFO - cmd=TopicMigrationFromAliyun, request=null, response={"code":200,"requestId":"9D5
4F6DB-6FA0-4F6D-B19A-09109F70BDDA"},"success":true,"message":"operation success"}
16:38:42 INFO - TopicCreate success, topic=1332, partition number=12, isCompactTopic=false
16:38:49 INFO - cmd=TopicMigrationFromAliyun, request=null, response={"code":200,"requestId":"6C2
65013-D15E-49AE-BE55-BF7657ADA1B7"},"success":true,"message":"operation success"}
16:38:49 INFO - TopicCreate success, topic=aete, partition number=12, isCompactTopic=false
```

5. 确认Topic迁移是否成功。
 - i. 登录[消息队列Kafka版控制台](#)。
 - ii. 在概览页面的资源分布区域，选择地域。
 - iii. 在实例列表页面，单击目标实例名称。
 - iv. 在左侧导航栏，单击Topic 管理。
 - v. 在Topic 管理页面的Topic列表中，显示成功迁移的Topic。

7.4. 迁移Consumer Group

7.4.1. 迁移Consumer Group上云

本文介绍如何使用消息队列Kafka版提供的迁移工具将自建Kafka集群的Consumer Group迁移到消息队列Kafka版实例。

前提条件

您已完成以下操作：

- [下载JDK 8](#)
- [下载迁移工具JAR文件](#)

注意

- 迁移不会删除源Kafka的Consumer Group，只是在目标消息队列Kafka版实例创建相同配置的Consumer Group。
- 迁移内容仅为Consumer Group配置，不包含消费的Topic及位点信息。

操作步骤

1. 打开命令行工具。
2. 使用cd命令将路径切换到迁移工具所在目录。
3. 创建配置文件*kafka.properties*。*kafka.properties*用于构造Kafka Consumer，从自建Kafka集群获取消费者位点信息。配置文件内容如下：

```
## 接入点。
bootstrap.servers=localhost:9092
## Consumer Group，注意该Consumer Group不能有消费者位点信息，以保证能从第一个消息开始消费。
group.id=XXX
## 如果无安全配置，可以不配置以下内容。
## SASL鉴权方式。
#sasl.mechanism=PLAIN
## 接入协议。
#security.protocol=SASL_SSL
## SSL根证书的路径。
#ssl.truststore.location=/Users/***/Documents/code/aliware-kafka-demos/main/resources/kafka.client.truststore.jks
## SSL密码。
#ssl.truststore.password=***
## SASL路径。
#java.security.auth.login.config=/Users/***/kafka-java-demo/vpc-ssl/src/main/resources/kafka_client_jaas.conf
```

4. 执行以下命令确认要迁移的Consumer Group。 `java -jar kafka-migration.jar ConsumerGroupMigrationFromTopic --propertiesPath /usr/local/kafka-2.12-2.4.0/config/kafka.properties --destAk <yourAccessKeyId> --destSk <yourAccessKeySecret> --destRegionId <yourRegionId> --destInstanceId <yourInstanceId>`

参数	描述
propertiesPath	配置文件 <code>kafka.properties</code> 的文件路径
destAk	目标消息队列Kafka版实例所属阿里云账号的AccessKey ID
destSk	目标消息队列Kafka版实例所属阿里云账号的AccessKey Secret
destRegionId	目标消息队列Kafka版实例的地域ID
destInstanceId	目标消息队列Kafka版实例的ID

待确认的返回结果示例如下：

```
15:29:45 INFO - Will create consumer groups:[XXX, test-consumer-group]
```

5. 执行以下命令提交要迁移的Consumer Group。 `java -jar kafka-migration.jar ConsumerGroupMigrationFromTopic --propertiesPath /usr/local/kafka-2.12-2.4.0/config/kafka.properties --destAk LTAI4FwO5aK1mFYCspJ1h*** --destSk wvDxiiRO1tHPil0oj7Y2Z7WDNkS*** --destRegionId cn-hangzhou --destInstanceId alika_fka_pre-cn-v0h1cng00*** --commit`

参数	描述
commit	提交迁移

提交迁移的返回结果示例如下：

```
15:35:51 INFO - cmd=ConsumerGroupMigrationFromTopic, request=null, response={"code":200,"requestId":"C9797848-FD4C-411F-966D-0D4AB5D12F55","success":true,"message":"operation success"}
15:35:51 INFO - ConsumerCreate success, consumer group=XXX
15:35:57 INFO - cmd=ConsumerGroupMigrationFromTopic, request=null, response={"code":200,"requestId":"3BCFDBF2-3CD9-4D48-92C3-385C8DBB9709","success":true,"message":"operation success"}
15:35:57 INFO - ConsumerCreate success, consumer group=test-consumer-group
```

6. 确认Group迁移是否成功。
 - i. 登录[消息队列Kafka版控制台](#)。
 - ii. 在概览页面的资源分布区域，选择地域。
 - iii. 在实例列表页面，单击目标实例名称。
 - iv. 在左侧导航栏，单击Group 管理。
 - v. 在Group 管理页面的Group列表显示成功迁移的Group。

7.4.2. 云上迁移Consumer Group

本文介绍如何使用消息队列Kafka版提供的迁移工具将某个消息队列Kafka版实例的Consumer Group迁移到另一个消息队列Kafka版实例。

前提条件

您已完成以下操作：

- [下载JDK 8](#)
- [下载迁移工具JAR文件](#)

注意

- 迁移不会删除源消息队列Kafka版实例的Consumer Group，只是在目标消息队列Kafka版实例创建相同配置的Consumer Group。
- 迁移内容仅为Consumer Group配置，不包含Consumer Group消费的Topic及位点信息。

操作步骤

1. 打开命令行工具。
2. 使用cd命令将路径切换到迁移工具所在目录。
3. 确认要迁移的Consumer Group。 `java -jar kafka-migration.jar ConsumerGroupMigrationFromAlivun --sourceAk <yoursourceAccessKeyId> --sourceSk <yoursourceAccessKeySecret> --sourceRegionId <yoursourceRegionId> --sourceInstanceId <yoursourceInstanceId> --destAk <yourdestAccessKeyId> --destSk <yourdestAccessKeySecret> --destRegionId <yourdestRegionId> --destInstanceId <yourdestInstanceId>`

参数	描述
sourceAk	源消息队列Kafka版实例所属阿里云账号的AccessKey ID
sourceSk	源消息队列Kafka版实例所属阿里云账号的AccessKey Secret
sourceRegionId	源消息队列Kafka版实例的地域ID

参数	描述
sourceInstanceld	源消息队列Kafka版实例的ID
destAk	目标消息队列Kafka版实例所属阿里云账号的AccessKey ID
destSk	目标消息队列Kafka版实例所属阿里云账号的AccessKey Secret
destRegionId	目标消息队列Kafka版实例的地域ID
destInstanceld	目标消息队列Kafka版实例的ID

待确认的返回结果示例如下：

```
10:54:26 INFO - cmd=ConsumerGroupMigrationFromAliyun, request=null, response={"code":200,"requestId":"9793DADB-55A5-4D4E-9E9C-D4DA8B35370C","success":true,"consumerList":[{"instanceld":"alika_fka_post-cn-0pp1h0uv6***","regionId":"cn-hangzhou","consumerId":"Demo","tags":[{"value":"","key":"migration"}]}],"message":"operation success."}
10:54:26 INFO - Will create consumer groups:[Demo]
```

4. 执行以下命令提交要迁移的Consumer Group。


```
java -jar kafka-migration.jar ConsumerGroupMigrationFromAliyun --sourceAk LTAI4FwO5aK1mFYCsnJ1h*** --sourceSk wvDxiiRO1tHPiL0oi7Y2Z7WDNkS*** --sourceRegionId cn-hangzhou --sourceInstanceld alikafka_post-cn-0pp1h0uv6*** --destAk LTAI4FwO5aK1mFYCsnJ1h*** --destSk wvDxiiRO1tHPiL0oj7Y2Z7*** --destRegionId cn-hangzhou --destInstanceld alikafka_pre-cn-v0h1cng00*** --commit
```

参数	说明
commit	提交迁移

提交迁移的返回结果示例如下：

```
10:54:40 INFO - cmd=ConsumerGroupMigrationFromAliyun, request=null, response={"code":200,"requestId":"49E53B79-3C2C-4BCF-8BC8-07B0BB14B52A","success":true,"consumerList":[{"instanceld":"alika_fka_post-cn-0pp1h0uv6***","regionId":"cn-hangzhou","consumerId":"Demo","tags":[{"value":"","key":"migration"}]}],"message":"operation success."}
10:54:41 INFO - cmd=ConsumerGroupMigrationFromAliyun, request=null, response={"code":200,"requestId":"5AEEFB13-2A6B-4265-97CB-902CFA483339","success":true,"message":"operation success"}
10:54:41 INFO - ConsumerCreate success, consumer group=Demo
```

5. 确认Group迁移是否成功。
 - i. 登录[消息队列Kafka版控制台](#)。
 - ii. 在概览页面的资源分布区域，选择地域。
 - iii. 在实例列表页面，单击目标实例名称。
 - iv. 在左侧导航栏，单击Group 管理。
 - v. 在Group 管理页面的Group列表显示成功迁移的Group。

7.5. 迁移数据

7.5.1. 迁移数据上云

本文介绍如何使用MirrorMaker将自建Kafka集群的数据迁移到消息队列Kafka版集群。

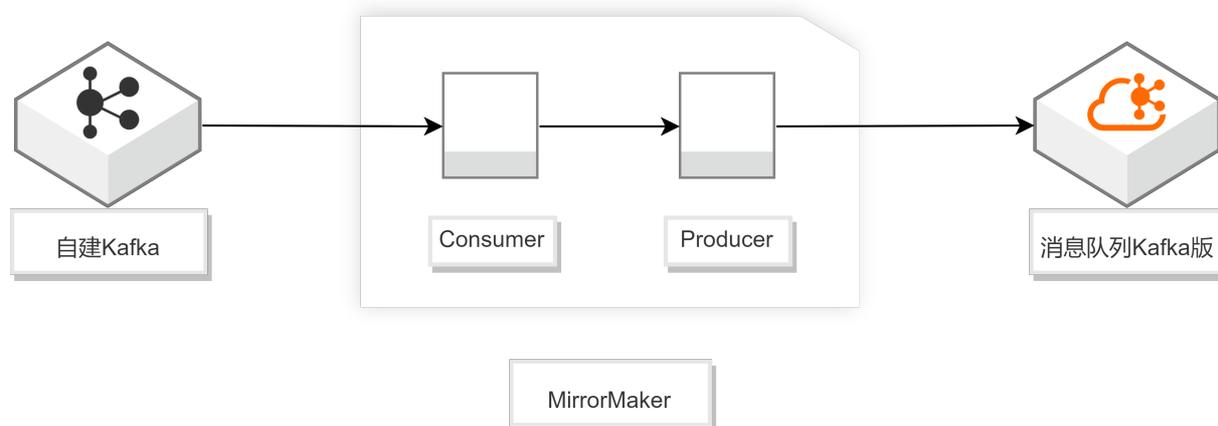
前提条件

您已完成以下操作：

- [下载迁移工具MirrorMaker](#)
- [迁移Topic上云](#)

背景信息

Kafka的镜像特性可实现Kafka集群的数据备份。实现这一特性的工具就是MirrorMaker。您可以使用MirrorMaker将源集群中的数据镜像拷贝到目标集群。如下图所示，MirrorMaker使用一个内置的Consumer从源自建Kafka集群消费消息，然后再使用一个内置的Producer将这些消息重新发送到目标消息队列Kafka版集群。



更多信息，请参见[Apache Kafka MirrorMaker](#)。

注意事项

- Topic名称必须一致。
- 分区数量可以不一致。
- 在同一个分区中的数据迁移后并不保证依旧在同一个分区中。
- 默认情况下，Key相同的消息会分布在同一分区中。
- 普通消息在宕机时可能会乱序，分区顺序消息在宕机时依然保持顺序。

VPC接入

1. 配置 *consumer.properties*。

```
## 自建Kafka集群的接入点
bootstrap.servers=XXX.XXX.XXX.XXX:9092
## 消费者分区分配策略
partition.assignment.strategy=org.apache.kafka.clients.consumer.RoundRobinAssignor
## Consumer Group的名称
group.id=test-consumer-group
```

2. 配置 *producer.properties*。

```
## 消息队列Kafka版集群的默认接入点（可在消息队列Kafka版控制台获取）
bootstrap.servers=XXX.XXX.XXX.XXX:9092
## 数据压缩方式
compression.type=none
```

3. 执行以下命令开启迁移进程。

```
sh bin/kafka-mirror-maker.sh --consumer.config config/consumer.properties --producer.config config/producer.properties --whitelist topicName
```

公网接入

1. 下载 `kafka.client.truststore.jks`。
2. 配置 `kafka_client_jaas.conf`。

```
KafkaClient {
  org.apache.kafka.common.security.plain.PlainLoginModule required
  username="your username"
  password="your password";
};
```

3. 配置 `consumer.properties`。

```
## 自建Kafka集群的接入点
bootstrap.servers=XXX.XXX.XXX.XXX:9092
## 消费者分区分配策略
partition.assignment.strategy=org.apache.kafka.clients.consumer.RoundRobinAssignor
## Consumer Group名称
group.id=test-consumer-group
```

4. 配置 `producer.properties`。

```
## 消息队列Kafka版集群的SSL接入点（可在消息队列Kafka版控制台获取）
bootstrap.servers=XXX.XXX.XXX.XXX:9093
## 数据压缩方式
compression.type=none
## truststore（使用步骤1下载的文件）
ssl.truststore.location=kafka.client.truststore.jks
ssl.truststore.password=KafkaOmsClient
security.protocol=SASL_SSL
sasl.mechanism=PLAIN
## 消息队列Kafka版2.X版本在配置SASL接入时需要做以下配置，2.X以下版本不需要配置。
ssl.endpoint.identification.algorithm=
```

5. 设置 `java.security.auth.login.config`。

```
export KAFKA_OPTS="-Djava.security.auth.login.config=kafka_client_jaas.conf"
```

6. 执行以下命令开启迁移进程。

```
sh bin/kafka-mirror-maker.sh --consumer.config config/consumer.properties --producer.config config/producer.properties --whitelist topicName
```

结果验证

您可以通过以下任一方法验证MirrorMaker是否运行成功。

- 通过 `kafka-consumer-groups.sh` 查看自建集群消费进度。

```
bin/kafka-consumer-groups.sh --new-consumer --describe --bootstrap-server自建集群接入点 --group test-consumer-group
```

- 往自建集群中发送消息，在消息队列Kafka版控制台中查看Topic的分区状态，确认当前服务器上消息总量是否正确。您还可以通过消息队列Kafka版控制台来查看具体消息内容。具体操作，请参见[查询消息](#)。

7.6. 查看迁移进度

本文说明如何查看自建Kafka集群迁移到消息队列Kafka版的进度。

前提条件

您已完成以下操作：

1. 购买并部署消息队列Kafka版实例。
2. 开始自建Kafka集群迁移上云。迁移类型包括：
 - 元数据迁移
 - [迁移Topic上云](#)
 - [迁移Consumer Group上云](#)
 - 数据迁移
 - [迁移数据上云](#)
 - 位点迁移

 **注意** 目前不支持将自建Kafka集群的位点迁移到消息队列Kafka版。

查看迁移进度

 **注意** 消息队列Kafka版目前不支持数据迁移和位点迁移信息上报，即如果您已经开始数据迁移或位点迁移，您目前无法在消息队列Kafka版控制台查看迁移进度。

查看迁移进度的操作步骤如下：

1. 登录[消息队列Kafka版控制台](#)。
2. 在概览页面的资源分布区域，选择地域。
3. 在左侧导航栏，单击迁移上云。
4. 在迁移上云页面，显示该地域下的所有迁移上云任务。选择要查看的迁移任务，查看Topic迁移进度和Group迁移进度，单击详情查看详细信息。

8. 查询消息

如遇消息消费有问题，您可通过查询消息排查问题。消息队列Kafka版控制台支持按位点查询和按时间查询。

背景信息

您可根据具体情况按以下推荐方式查询消息：

- 如果您确定消息所在Topic的分区ID以及消息位点，则推荐您使用。位点的相关概念，请参见[名词解释](#)。
- 如果您不确定消息的位置，但确定消息发送的时间段，则推荐您使用。

注意事项

- 查询到的每条消息在消息队列Kafka版控制台上最多显示1 KB的内容，超过1 KB的部分将自动截断。如需查看完整的消息内容，请下载相应的消息。

目前，仅专业版实例支持下载，且下载的消息最大为10 MB。

- 标准版实例查询结果最多显示10条消息，且总大小不能超过256 KB。
 - 如果10条消息的总大小已经超过256 KB，那么控制台上只显示256 KB以内的内容。
 - 如果10条消息的总大小还不到256 KB，在控制台上最多也只能看到10条消息。在这样的情况下，请以客户端的实际消费为准。
- 专业版实例查询结果最多显示30条消息，且总大小不能超过10 MB。
 - 如果30条消息的总大小已经超过10 MB，那么控制台上只显示10 MB以内的内容。
 - 如果30条消息的总大小还不到10 MB，在控制台上最多也只能看到30条消息。在这样的情况下，请以客户端的实际消费为准。

关于实例版本的更多信息，请参见[计费说明](#)。

- 能否查询到消息，与消息队列Kafka版的消息清除策略相关：
 - 磁盘使用率低于85%，每天凌晨4点集中删除超过消息保留时间的消息。
 - 磁盘使用率达到85%，立即清除超过消息保留时间的消息。
 - 磁盘使用率达到90%，无论消息是否超过消息保留时间，按服务端存储消息的时间先后顺序清除消息。

 **注意** 消息队列Kafka版在清除消息时，会至少保留一个存储文件。因此，您有可能查询到超过消息保留时间的消息。

按位点查询消息

1. 登录[消息队列Kafka版控制台](#)。
2. 在**概览**页面的**资源分布**区域，选择地域。
3. 在**实例列表**页面，单击目标实例名称。
4. 在左侧导航栏，单击**消息查询**。
5. 在**消息查询**页面的**查询方式**列表中，选择**按位点查询**。
6. 在**Topic**列表中，选择消息所属Topic名称；在**分区**列表中，选择消息所属的分区；在**起始位点**文本框，输入消息所在分区的位点，然后单击**查询**。

展示该查询位点及以后连续的消息。例如，指定的分区和位点都为“5”，那么返回的结果从位点“5”开始。

查询结果参数解释

参数	描述
分区	消息的Topic分区。
位点	消息的所在的位点。
Key	消息的键（已强制转化为String类型）。
Value	消息的值（已强制转化为String类型），即消息的具体内容。
消息创建时间	<p>发送消息时，客户端自带的或是您指定的 ProducerRecord 中的消息创建时间。</p> <div style="background-color: #e6f2ff; padding: 10px;"> <p>说明</p> <ul style="list-style-type: none"> 如果配置了该字段，则按配置值显示。 如果未配置该字段，则默认取消息发送时的系统时间。 如果显示值为1970/x/x x:x:x，则说明发送时间配置为0或其他有误的值。 0.9及以前版本的消息队列Kafka版客户端不支持配置该时间。 </div>
操作	<ul style="list-style-type: none"> 单击下载 Key：下载消息的键值。 单击下载 Value：下载消息的具体内容。 <div style="background-color: #e6f2ff; padding: 10px;"> <p>注意</p> <ul style="list-style-type: none"> 查询到的每条消息在控制台上最多显示1 KB的内容，超过1 KB的部分将自动截断。如需查看完整的消息内容，请下载相应的消息。 仅专业版支持下载消息。 下载的消息最大为10 MB。如果消息超过10 MB，则只下载10 MB的内容。 </div>

按时间查询消息

按时间查询允许查询全部分区的消息，不确定消息的位置，但确定消息发送的时间段，请指定该时间段中的某一个时间点查询该时间点附近的消息。

1. 登录消息队列Kafka版控制台。
2. 在概览页面的资源分布区域，选择地域。

3. 在实例列表页面，单击目标实例名称。
4. 在左侧导航栏，单击消息查询。
5. 在消息查询页面的查询方式列表中，选择按时间点查询。
6. 在Topic列表中，选择消息所属Topic名称；在分区列表中，选择消息所属的分区；在时间点中，单击选择需查询的时间点，然后单击查询。

展示该查询时间点及之后连续的消息。查询结果的参数，请参见。

分区的取值会影响查询结果：

- 如果选择全部分区，则显示全部分区。
- 如果指定具体分区，则显示指定分区的消息。