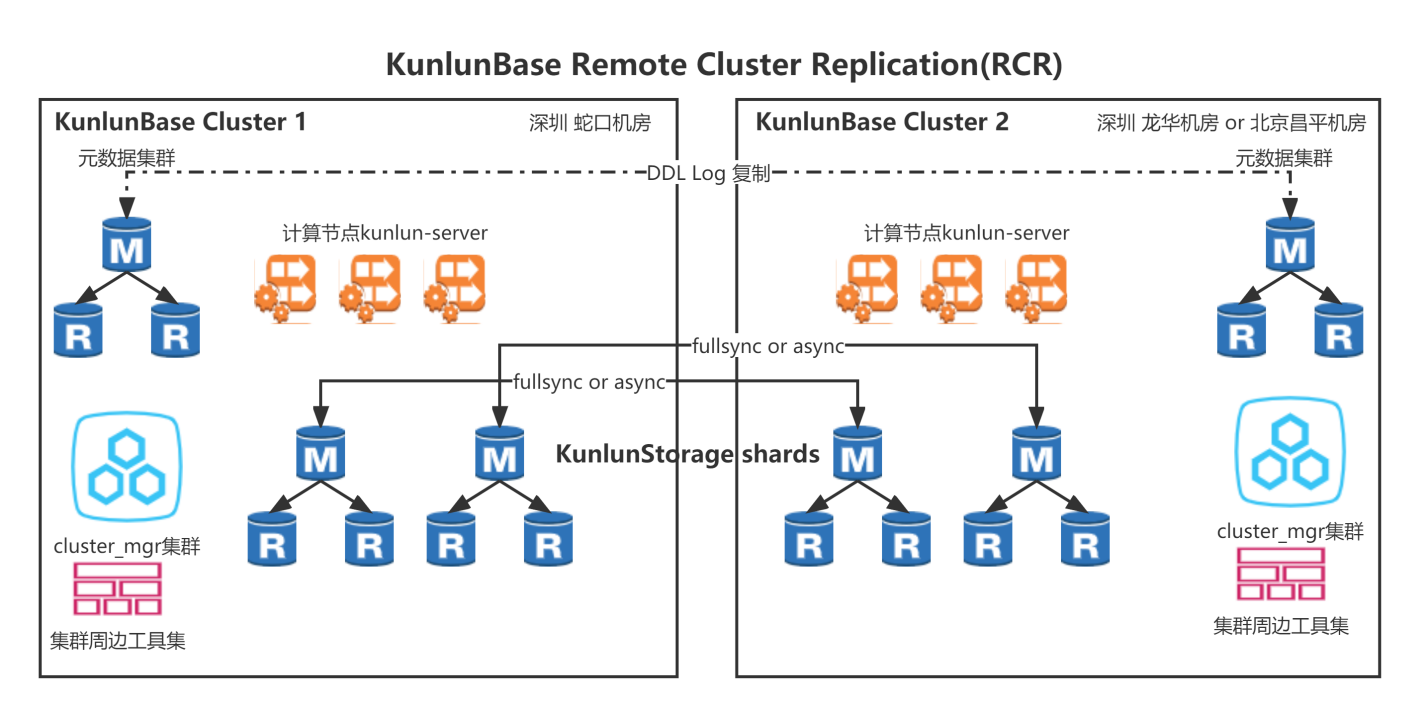
# klustron RCR 使用手册

## 一. 需求背景和功能概况

klustron分布式数据库为了支持集群同城/跨城实现高可用性，支持用户在同城/异地的备用机房建立另一个klustron集群（备集群）与主城klustron集群（主集群）建立RCR (remote cluster replication)关系（主集群和备集群合称为主备集群），实现klustron集群层面的数据更新实时同步。

当主机房整体不可用时，用户的业务系统可以自动切换到klustron在备机房的备份集群继续对外提供服务。如果备机房在同城，则DBA可以配置主备集群同步复制，因此切换后可以确保不丢失任何数据；如果备集群在异地机房，则考虑到网络延时较大，通常用户会设置为主备集群做异步复制。这样，一旦主机房失效，则切换到klustron备集群后，可能会丢失一些最新的数据更新。

## 二．部署架构



Klustron 企业版cluster\_mgr组件负责创建和维护RCR集群间复制关系。Klustron在cluster\_mgr提供API接口，并且在XPanel 提供GUI，供用户管理和维护RCR复制关系。用户DBA需要在备用机房创建备集群，并且使用RCR配置管理功能配置好RCR集群复制关系。用户的业务系统需要探测主集群或者主机房的可用性，在探测到主集群失效时自动切换到备集群。

主备集群必须具有各自的元数据集群和cluster\_mgr集群，不能共用，这主要是考虑主备集群之间的网络延时可能较大，共用的话会导致性能下降。

Klustron RCR要求主备集群具有相同数量的存储集群(storage shard, 简称shard)，cluster\_mgr 的RCR组件为主备集群的每一对shard建立复制关系，这样备集群每个shard就可以从主集群对应的shard复制数据更新。

备集群可以使用主集群的全量物理备份数据或者全量逻辑备份数据做回档来提供初始化数据。

Klustron RCR功能可以确保在主集群或者备集群发生存储节点故障而导致部分或者全部RCR连接失效时，自动恢复连接，并且不丢失数据更新。

## 三. 实现原理和功能用法

klustron集群间数据同步包括以下两个部分。

### shard之间用户数据同步

基于 binlog replication 主备同步来实现，其中如果是同步复制则使用KunlunBase特有的fullsync技术。

### klustron集群的元数据同步

需要同步元数据是因为计算节点需要复制执行DDL以便与源集群具有相同的元数据信息。其中库表分布信息不能直接使用，需要根据目标集群的情况转换映射。

该同步采用binlog\_sync工具来完成，binlog\_sync通过binlog\_dump方式连接主cluster所在元数据集群，捕获对目标集群的DDL log系统表的更新，捕获到binlog变化后，将binlog记录中的DDL SQL语句做必要的转换后再写入备cluster元数据表中。

目前必要的转换包括：shard id信息进行映射。需要映射转换是因为目标集群的shard id与源集群的对应的shard id未必相同。

### 功能实现和用法说明

#### 建立RCR数据同步

cluster\_mgr检查主备cluster的shard节点数，配置参数（参数为主备RCR cluster\_id以及对应cluster所在元数据集群ip:port）等是否一致。如果不一致则返回失败。如果一致则主备cluster 各个shard之间建立binlog同步链路。为了保证主备RCR shard

数据一致性，cluster\_mgr会清理掉备RCR shard中各个节点中数据，重新从对应主RCR shard节点上同步数据过来。建立RCR关系时，cluster\_mgr会自动将主备RCR 各个shard 一一对应起来。cluster\_mgr将备cluster中各个shard中主节点与主cluster中对应shard中一个延迟较小备节点建立数据同步。cluster\_mgr会根据建立好数据层同步，自动启动binlog\_sync工具来实时同步klustron主集群的元数据更新，经过转换后同步到备集群的元数据集群的DDL log表，让备集群的计算节点去执行。

为了确保数据一致性，Klustron RCR会把备集群每个shard主节点设置为只读，避免用户误写入数据导致其数据与主集群不一致。

特别说明：建立RCR时，cluster\_mgr自动清理掉备RCR shard中数据。

#### 删除RCR数据同步

cluster\_mgr根据输入参数（参数为主备RCR cluster\_id以及对应cluster所在元数据集群ip:port）校验是否合法，如果输入参数与cluster\_mgr中记录主备RCR cluster\_id以及元数据集群信息不一致则返回任务失败。校验通过后，cluster\_mgr连接备RCR cluster中各个shard主节点，断开binlog同步链路，断开klustron集群元数据同步关系。放开备cluster 各个shard主节点可读写等操作。

#### 手动切换RCR关系

cluster\_mgr根据输入参数（参数为主备RCR cluster\_id以及对应cluster所在元数据集群ip:port）校验是否合法，如果输入参数与cluster\_mgr中记录主备RCRcluster\_id以及元数据集群信息不一致则返回任务失败。校验通过后cluster\_mgr检查主备RCR同步延迟，如果延迟大约配置值（5s）则返回任务失败。cluster\_mgr将主RCR cluster各个shard主节点设置为只读，阻止数据继续写入；连接备RCR cluster各个shard主节点，检查主备数据是否同步（检查gtid）完成，该阶段cluster\_mgr超时时间为10s，如果超时则任务回滚并返回任务失败。断开klustron集群元数据同步关系。断开主备RCR shard之间原先数据同步链路，反向建立shard之间数据同步链路，即原主变为备，原备变为主。设置原备RCR cluster 各个shard主节点可读写。原主RCR cluster各个shard主节点只读。cluster\_mgr根据shard新同步关系，重新启动binlog\_sync工具建立klustron集群元数据同步。

手动切换场景：

1. 业务部署发生变化，从主RCR切换到备RCR所在地区。直接发起手动RCR切换即可。由于RCR手动切换保证数据一致性的，不必担心数据丢失问题。

2. 容灾演练。

#### 配置RCR同步关系

当前支持暂停RCR主备数据同步，恢复RCR主备数据同步以及设置主备RCR之间数据延迟复制。

#### 集群拓扑结构变化

如果在建立RCR关系cluster中任何一个cluster添加/删除shard，cluster\_mgr自动根据RCR关系到对端cluster\_mgr中添加/删除shard。目前cluster\_mgr自动会创建一个添加/删除shard任务发送对端cluster\_mgr，并等待查询该任务执行情况。如果执行失败，则用户发起的添加/删除shard任务失败。

1.添加shard，cluster\_mgr在主备RCR 都建立好shard后，发起对新增shard建立RCR关系。

2.删除shard，cluster\_mgr先断开删除shard RCR同步关系，再发起主备RCR shard删除任务。

## 四. 配置使用

1.使用cluster\_mgr api操作

1.1建立RCR关系

curl -d '

{

"version":"1.0",

"job\_id":"",

"job\_type":"create\_rcr",

"timestamp":"1435749309",

"user\_name":"kunlun\_test",

"paras":{

"master\_info": {

"meta\_db":"127.0.0.1:1001, 127.0.0.2:1001, 127.0.0.3:1001",

"cluster\_id":"2"

},

"cluster\_id":"3"

}

}

' -X POST http://127.0.0.1:58000/HttpService/Emit

1.2 删除RCR关系

curl -d '

{

"version":"1.0",

"job\_id":"",

"job\_type":"delete\_rcr",

"timestamp":"1435749309",

"user\_name":"kunlun\_test",

"paras":{

"master\_info": {

"meta\_db":"127.0.0.1:1001, 127.0.0.2:1001, 127.0.0.3:1001",

"cluster\_id":"2"

},

"cluster\_id":"3"

}

}

' -X POST http://127.0.0.1:58000/HttpService/Emit

1.3 手动RCR主备切换，支持在主备RCR所在cluster\_mgr发起

在备RCR所在cluster\_mgr发起

curl -d '

{

"version":"1.0",

"job\_id":"",

"job\_type":"manualsw\_rcr",

"timestamp":"1435749309",

"user\_name":"kunlun\_test",

"paras":{

"master\_info": {

"meta\_db":"127.0.0.1:1001, 127.0.0.2:1001, 127.0.0.3:1001",

"cluster\_id":"2"

},

"cluster\_id":"3"

}

}

' -X POST <http://127.0.0.1:58000/HttpService/Emit>

在主RCR所在cluster\_mgr发起

curl -d '

{

"version":"1.0",

"job\_id":"",

"job\_type":"manualsw\_rcr",

"timestamp":"1435749309",

"user\_name":"kunlun\_test",

"paras":{

"slave\_info": {

"meta\_db":"127.0.0.1:1001, 127.0.0.2:1001, 127.0.0.3:1001",

"cluster\_id":"2"

},

"cluster\_id":"3"

}

}

' -X POST <http://127.0.0.1:58000/HttpService/Emit>

1.4 配置主备RCR同步，支持在主备RCR所在cluster\_mgr发起

在备RCR所在cluster\_mgr发起

curl -d '

{

"version":"1.0",

"job\_id":"",

"job\_type":"modify\_rcr",

"timestamp":"1435749309",

"user\_name":"kunlun\_test",

"paras":{

"master\_info": {

"meta\_db":"127.0.0.1:1001, 127.0.0.2:1001, 127.0.0.3:1001",

"cluster\_id":"2"

},

"cluster\_id":"3",

"work\_mode":"modify\_sync\_delay|stop\_rcr|start\_rcr",

"sync\_delay":"10" //如果是modify\_sync\_delay，则需要配置sync\_delay

}

}

' -X POST <http://127.0.0.1:58000/HttpService/Emit>

在主RCR所在cluster\_mgr发起

curl -d '

{

"version":"1.0",

"job\_id":"",

"job\_type":"modify\_rcr",

"timestamp":"1435749309",

"user\_name":"kunlun\_test",

"paras":{

"slave\_info": {

"meta\_db":"127.0.0.1:1001, 127.0.0.2:1001, 127.0.0.3:1001",

"cluster\_id":"2"

},

"cluster\_id":"3"

"work\_mode":"modify\_sync\_delay|stop\_rcr|start\_rcr",

"sync\_delay":"10" //如果是modify\_sync\_delay，则需要配置sync\_delay

}

}

' -X POST <http://127.0.0.1:58000/HttpService/Emit>

2.使用xpanel页面操作

由于xpanel支持管理多个klustron集群，首先上报要建立RCR关系klustron集群元数据信息，如下图

A picture containing table

Description automatically generated

A picture containing chart

Description automatically generated

设置元数据集群名称，以及集群ip，port；如果多个，点击加号添加。

2.1 配置好元数据集群后，点击新增RCR

A picture containing table

Description automatically generated

Graphical user interface, application

Description automatically generated

点击确认发起建立RCR

Graphical user interface, text

Description automatically generated with medium confidence

Table, calendar

Description automatically generated

点击详情查看 元数据同步以及shard同步对应关系，如下图



2.2 删除RCR关系

Table

Description automatically generated

Graphical user interface, text, application

Description automatically generated

输入验证码，点击确认发起删除命令

Shape, rectangle

Description automatically generated

2.3 手动切换RCR关系

Table

Description automatically generated

Graphical user interface, text

Description automatically generated

输入验证码，点击确认发起手动切换命令

Shape, rectangle

Description automatically generated

2.4 配置RCR关系

* 设置延迟复制时间

Table

Description automatically generated

Icon

Description automatically generated

输入配置延迟时间，点击确认。

Shape, rectangle

Description automatically generated with medium confidence

* 暂停同步复制

Table

Description automatically generated

Graphical user interface, text, application

Description automatically generated

输入验证码，点击确认

Shape, rectangle

Description automatically generated

* 启动同步复制

Table

Description automatically generated

Graphical user interface, text, application

Description automatically generated

输入验证码，点击确认

Shape, rectangle

Description automatically generated