

云数据库 HBase

云HBase主备双活

云HBase主备双活

主备双活

背景

在大数据背景下每时每刻都会有大量的用户行为数据被收集、分析、以及应用, 典型的应用场景包括用户推荐、安全风控;这些场景的共同特点是数据量大、可用性要求高、同时对随机读请求的响应时间非常敏感;云HBase作为海量大数据存储天然适合数据量TB、PB级别的业务场景,同时提供主备容灾方案(详见)满足用户的高可用需求,再此基础上为了满足用户大数据量随机读低毛刺的要求新增追求极致稳定性的主备双活能力;

使用场景

对随机读响应有高要求(P999 小于50ms)的实时在线业务场景,例如用户推荐、安全风控等

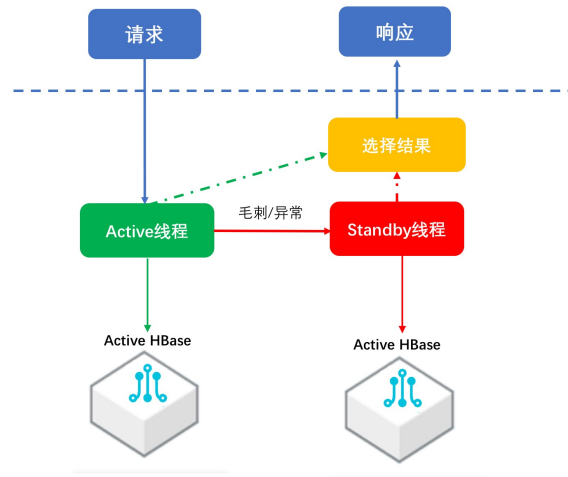
核心能力

- 请求低毛刺
- 自动容错
- 资源利用率高

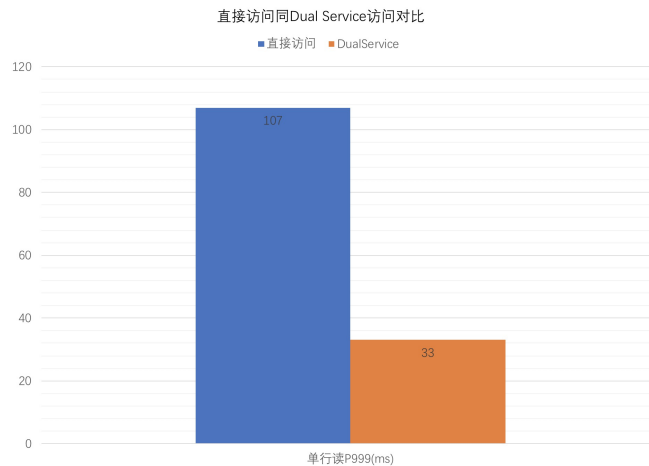
原理介绍

低毛刺

如果一个节点的稳定性是99.9%, 那么两个独立节点的组合的稳定性是 $1 - 0.1 * 0.1 = 99.99\%$ 。采用双节点服务同一份数据可以在最终一致性条件下提升一个数量级的稳定性。当用户请求发起时会首先向主库发起请求,在等待一段时间(Glitch Time)后如果主库仍没有返回结果,则并发向备库发起请求,最终取最快返回的值作为结果。

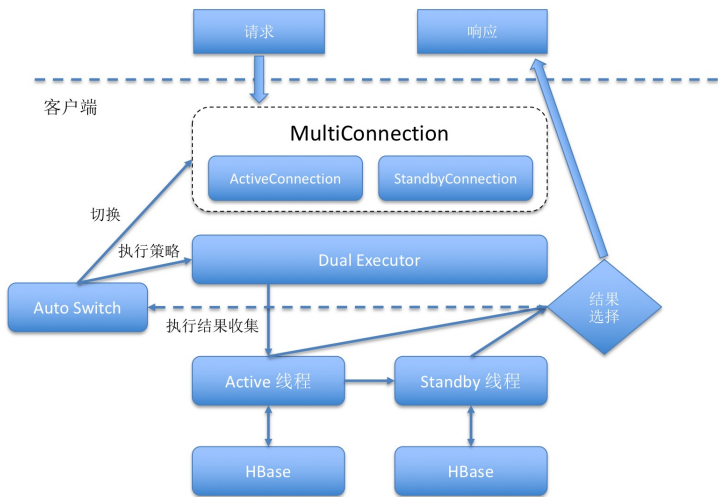


毛刺对比



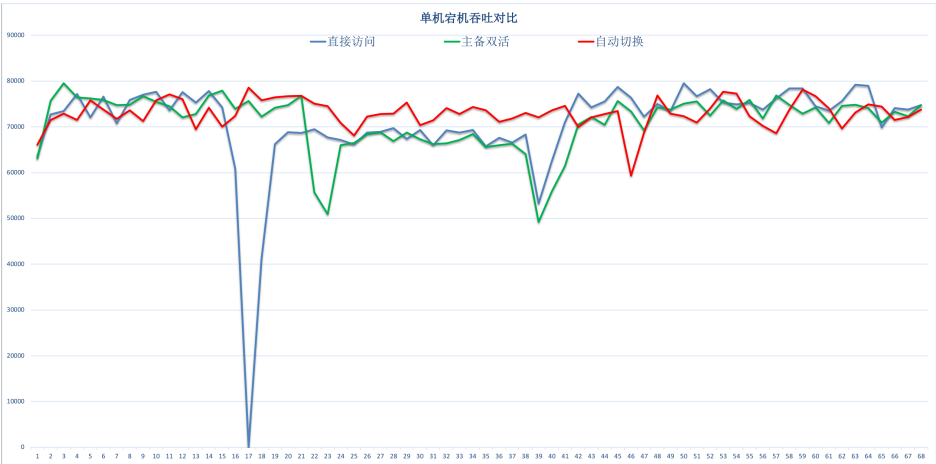
自动容错

主备双活中我们提供了自动容错能力,在故障场景下自动进行切换,做到故障对业务完全透明;通常需要进行主备集群切换的场景如下:1,断网断电等机房维度故障,无法正常链接主集群,请求全部抛错;2,由于软件BUG造成的全集群宕机;3,由于慢盘、坏盘造成的集群访问超时;从用户视角看就是主集群访问抛错或者主集群访问超时;自动容错会历史的执行结果进行收集,当主集群连续抛错或者连续超时超过用户指定次数时,即判定主集群存在故障需要进行“切换”,在切换状态下在主库服务恢复可以进行正常访问的情况会进行自动回切,对用户完全透明.

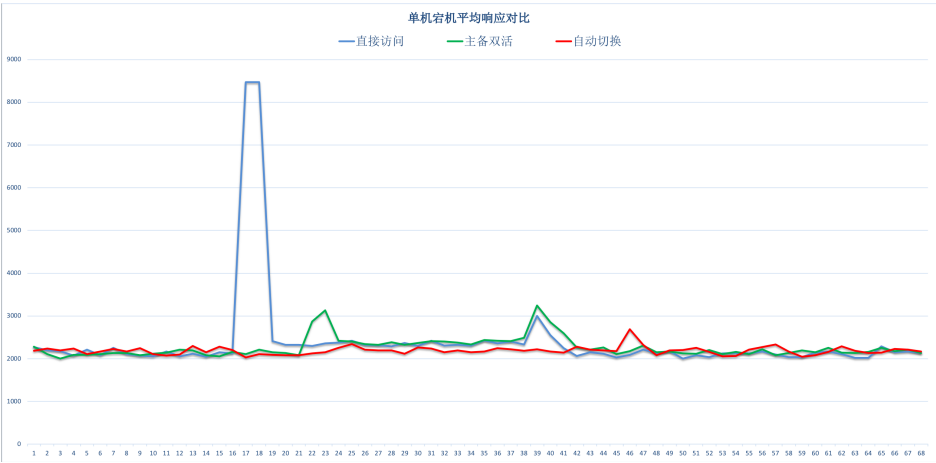


宕机测试

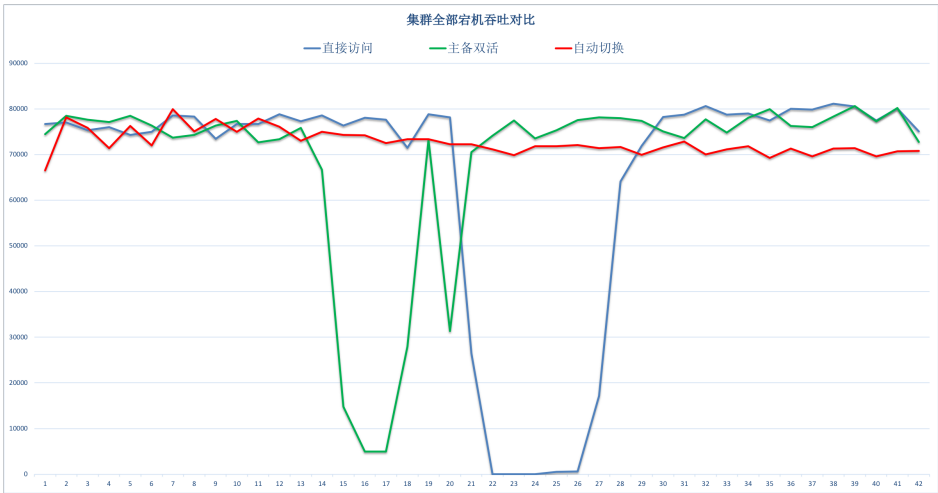
单机宕机吞吐对比



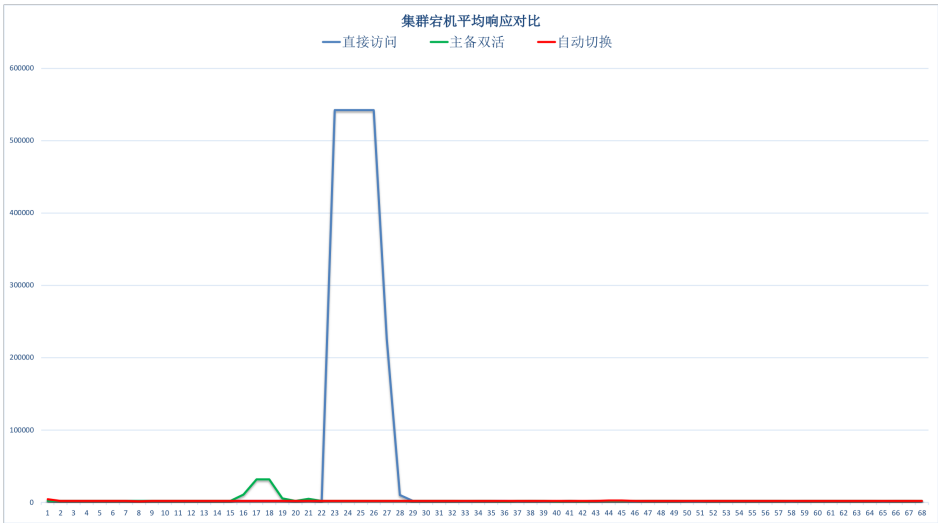
单机宕机平均响应对比



集群宕机吞吐对比



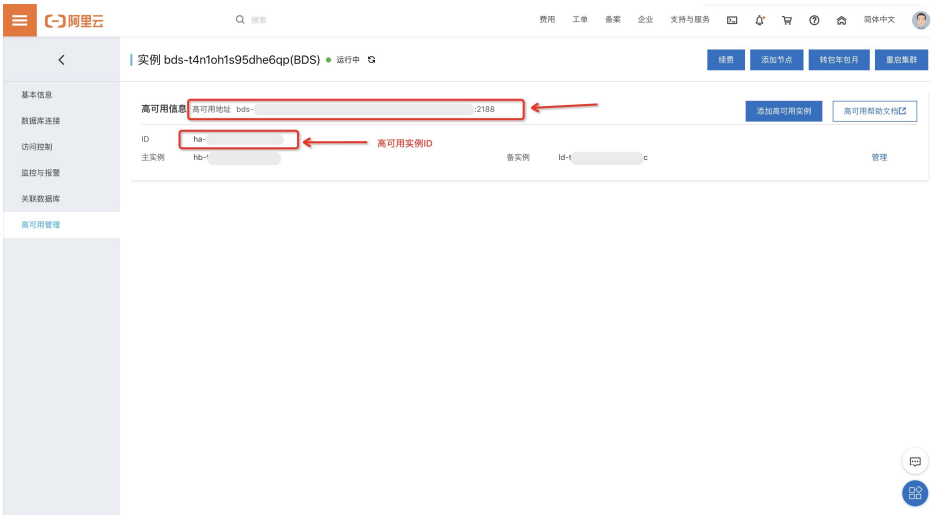
集群宕机平均响应对比



使用方法

准备

1. 需要使用云HBase主备容灾方案,(详见)
2. 在高可用管理页面获取高可用链接地址和高可用实例ID：



3. 客户端依赖版本

```
hbase1.X版本使用客户端
<dependency>
<groupId>com.aliyun.hbase</groupId>
<artifactId>alihbase-client</artifactId>
<version>1.1.12</version>
</dependency>
```

```
hbase2.X版本使用客户端
<dependency>
<groupId>com.aliyun.hbase</groupId>
<artifactId>alihbase-client</artifactId>
<version>2.0.7</version>
</dependency>
<dependency>
<groupId>com.aliyun.hbase</groupId>
<artifactId>alihbase-endpoint</artifactId>
<version>2.0.7</version>
</dependency>
```

配置客户端参数

您可以通过以下两种方式，来配置客户端参数：

方式一：配置文件

```
<configuration>
<!-- 高可用链接地址 -->
<property>
<name>hbase.zookeeper.quorum</name>
<value>HOST:PORT</value>
</property>
<!-- 高可用实例ID-->
<property>
<name>haclient.cluster.id</name>
<value>ha-xxxxxx</value>
```

```

</property>
<!--
如果需要连接HBase增强版，需要设置用户名密码，默认root:root，可根据实际情况调整
-->
<property>
<name>hbase.client.username</name>
<value>root</value>
</property>
<property>
<name>hbase.client.password</name>
<value>root</value>
</property>
<!--开启主备双活-->
<property>
<name>hbase.dualservice.enable</name>
<value>true</value>
</property>
<!--全部表使用主备双活访问-->
<property>
<name>hbase.dualservice.table.enable</name>
<value>true</value>
</property>
<!--可以表级别设置使用主备双活访问-->
<property>
<name>${tablename}.hbase.dualservice.enable</name>
<value>true</value>
</property>
<!--设置glitchtimeout(当访问主库超过这个时间后,dual service并发访问备库),单位ms-->
<property>
<name>hbase.dualservice.glitchtimeout</name>
<value>xxx</value>
</property>
<!--表级别设置glitchtimeout,单位ms-->
<property>
<name>${tablename}.hbase.dualservice.glitchtimeout</name>
<value>xxx</value>
</property>
<!--开启自动容错-->
<property>
<name>hbase.autoswitch.enable</name>
<value>true</value>
</property>
</configuration>

```

方式二：代码

通过代码Create Configuration，然后增加相关配置

```

// 新建一个Configuration
Configuration conf = HBaseConfiguration.create();
// 高可用链接地址
conf.set("hbase.zookeeper.quorum", "HOST:PORT");
// 高可用实例ID
conf.set("haclient.cluster.id", "ha-xxxxxx");
// 设置用户名密码，默认root:root，可根据实际情况调整

```

```
conf.set("hbase.client.username", "root");
conf.set("hbase.client.password", "root");
//开启Dual Service
conf.setBoolean("hbase.dualservice.enable", true);
//全部表使用DualService访问
conf.setBoolean("hbase.dualservice.table.enable", true);
//可以表级别设置使用DualService访问
conf.setBoolean("${tablename}.hbase.dualservice.enable", true);
//设置glitchtimeout(当访问主库超过这个时间后,dual service并发访问备库),单位ms
conf.setInt("hbase.dualservice.glitchtimeout", xxx);
//表级别设置glitchtimeout,单位ms
conf.setInt("${tablename}.hbase.dualservice.glitchtimeout", xxx);
//开启自动容错
conf.setBoolean("hbase.autoswitch.enable", true);
```

使用限制

Dual Service仅支持以下访问方式:

- get
- delete
- put
- batchGet
- batchDelete
- batchPut
- small scan(2.x客户端需要scan设置setLimit且limit < 500)

主备容灾

背景

云HBase通过master节点自动切换、数据两副本、快速自动failover等手段最大程度的保证了HBase实例服务和数据可用性，但是面对可用区级别的断电、断网及其他极端故障单个HBase集群无法满足用户的高可用要求，针对跨可用区灾备需求我们提供了跨可用区的HBase同城主备容灾，提供混合场景(云实例同云实例、云实例同自建实例以及自建与自建实例)下的故障一键、透明、快速切换能力。

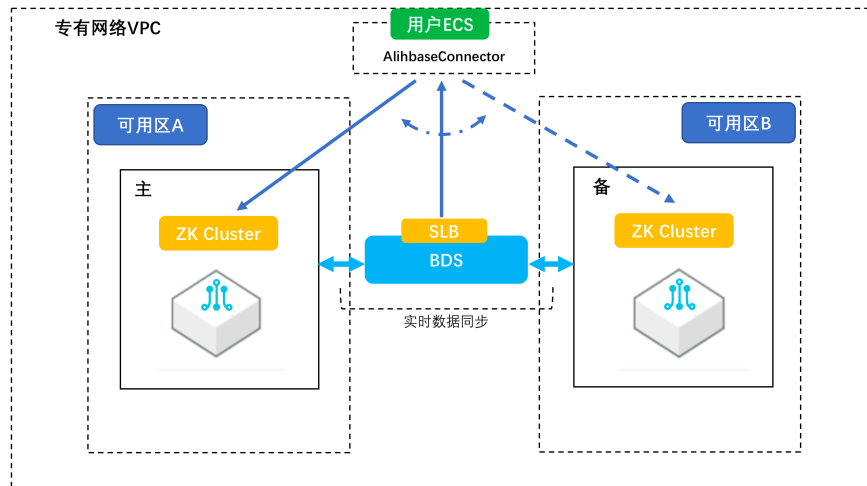
产品优势

- 支持云HBase、EMR HBase、自建HBase混合主备
- 零代码改造，支持1.x、2.x多个版本
- 自动数据同步管理，高效双向同步，延迟在百毫秒级
- 控制台一键切换

- 主备控制台，可显示同步延迟、客户端链接数、主备状态等重要信息

架构说明

用户主备HBase实例分别部署在不同可用区，使用BDS数据通道服务(服务介绍)完成主备实例之间实时增量数据双向同步，同时在BDS上进行主备切换操作，用户通过alihbase-connector访问HBase实例，alihbase-connector监听主备实例zk集群的切换事件完成实例切换，切换过程对上层业务透明。



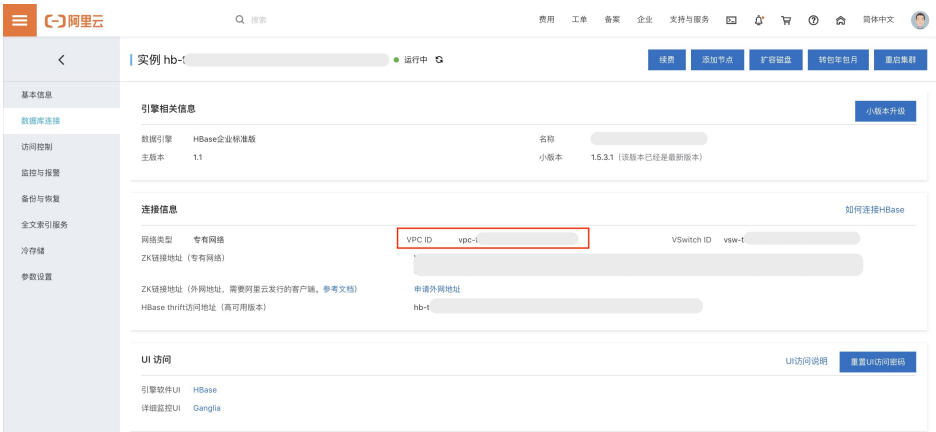
使用限制

- 当前仅支持相同VPC内跨可用区部署方式
- 2019年9月18日后购买的BDS服务支持该功能
- 主备实例数据同步目前使用最终一致性模型，暂时不支持强一致模型
- 目前只支持HBase服务切换，不支持phoenix，solr切换

使用方法

购买实例

HBase实例购买，如已经购买可以忽略，如未购买可以到 [购买链接](#) 页面购买 HBase 实例，购买好的 HBase 实



例如如下：
BDS实例购买，如已经购买可以忽略，如未购买可以到 购买链接 页面购买BDS实例，BDS实例需要同HBase在相同VPC下，并且同主备HBase实例不在相同可用区，如果没有多余可用区BDS建议购买在备实例所在可用区，购买好的 BDS实



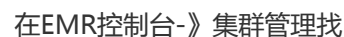
例如如下：

添加白名单

主备HBase实例以及BDS实例都要添加客户端的白名单，否则客户端无法访问相应的实例，添加白名单方法详见

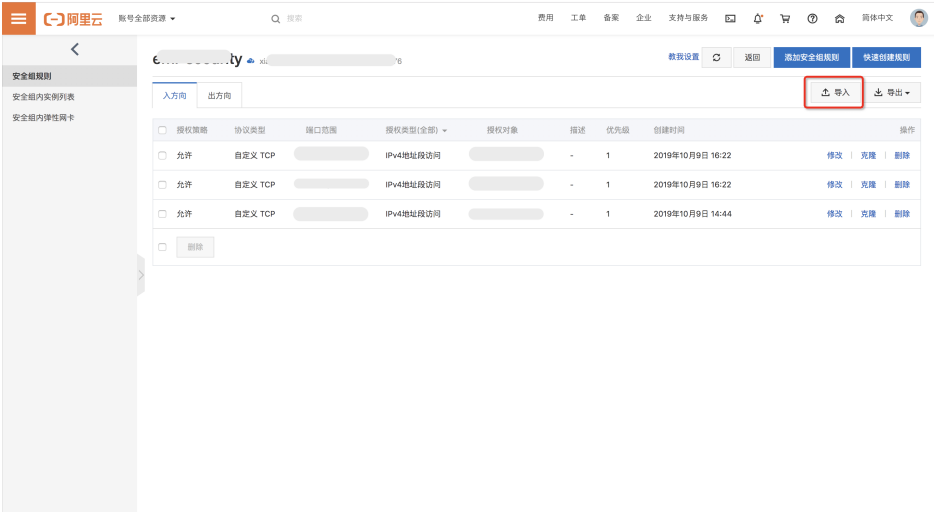
自建实例网络打通

当关联自建实例时，BDS需要能够访问自建实例，这里以EMR HBase实例为例(用户自己购买ECS自建实例情况时一致的)，需要在部署HBase的ECS实例安全组中为BDS的节点授权；进入bds页面，选择配置管理，点击生成安全组规则，会下载xxxx_securityrule.json的文件到本地：



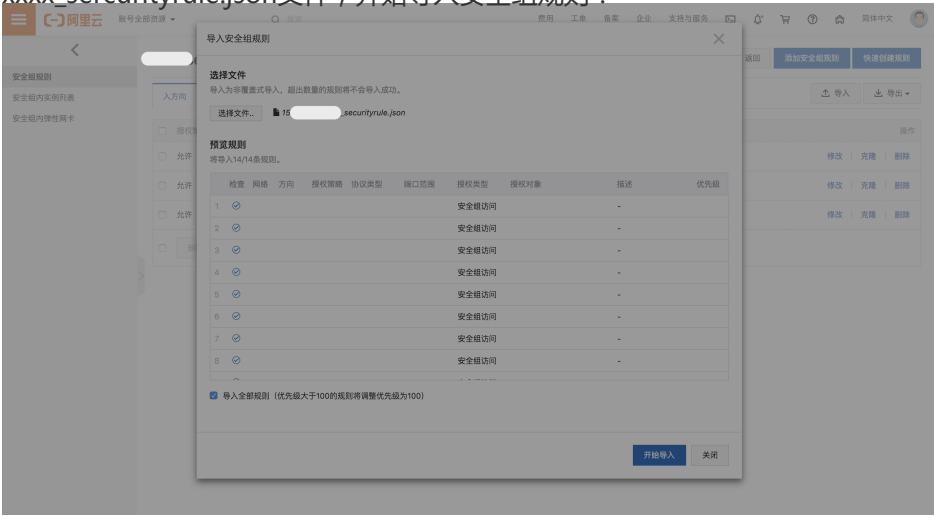
找到安全组(如果是完全自建)

在安全组中选择导入：



选择刚才下载的

xxxx_securityrule.json文件，开始导入安全组规则：



导入完成后BDS有权限可以

访问自建HBase的相关端口；自建HBase实例需要在BDS中绑定实例节点的host信息。进入bds页面，选择配置管理，配置bds.conf.hosts后，点击修改：

配置信息修改

配置项

生成安全组规则

常用配置项修改

配置名	默认值	现有值	配置项说明
bds.full.reader.bandwidth	10485760	10485760	设置单个HFile迁移任务的最大带宽
bds.full.task.pool.size	5	3	设置每个Worker节点迁移Hfile的线程数
bds.wal.retry.num	50	50	设置WAL迁移任务失败重试的次数
bds.task.pool.size	5	5	设置每个Worker处理WAL迁移的线程数
bds.wal.task.asyncwriter.size	1	1	单个WAL的写线程数
bds.job.monitor	true	true	监控开关
bds.conf.hosts		192.1 emr-worker-1.cluster 192.1 emr-worker-2.cluster 192.1 emr-header-1.cluster	配置迁移节点的/etc/hosts

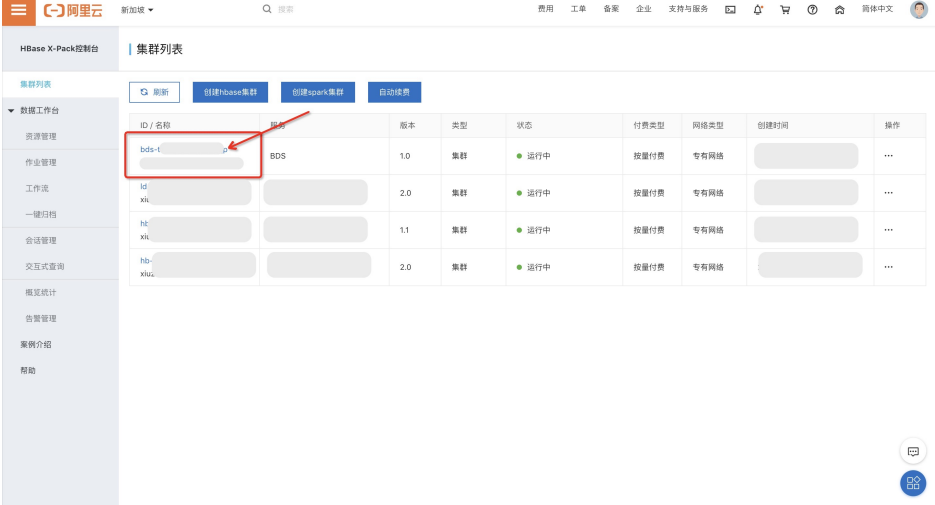
修改

EMR的hosts信息可以登入到

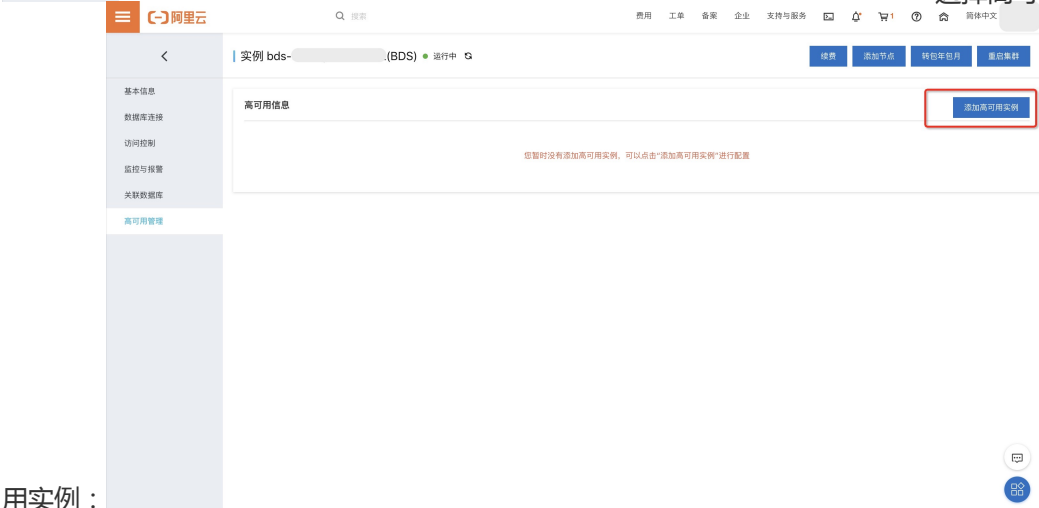
EMR的集群的ecs上查看/etc/hosts查看IP和域名，域名规则一般是emr-header-x.cluster-xxx/emr-worker-x.cluster-xxx

控制台新增高可用实例

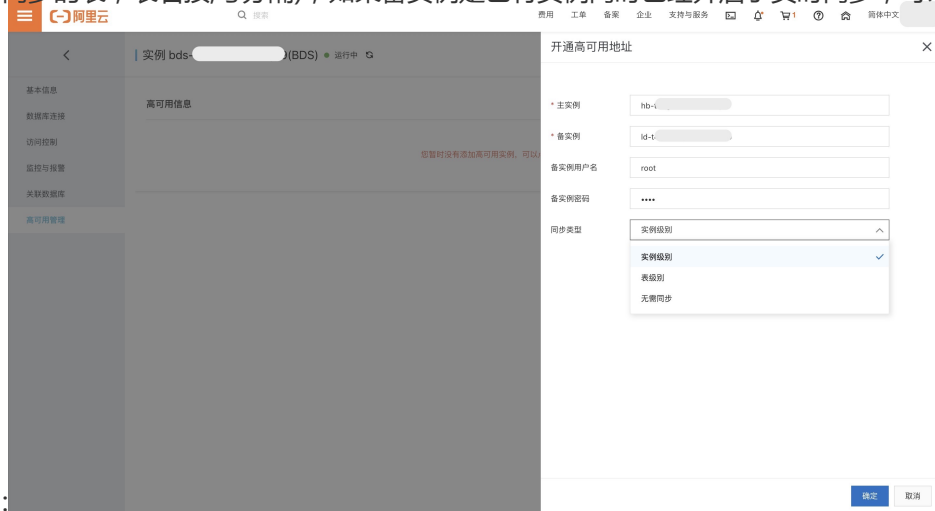
进入集群列表页面，点击新购买的bds实例：



选择高可用管理，添加高可

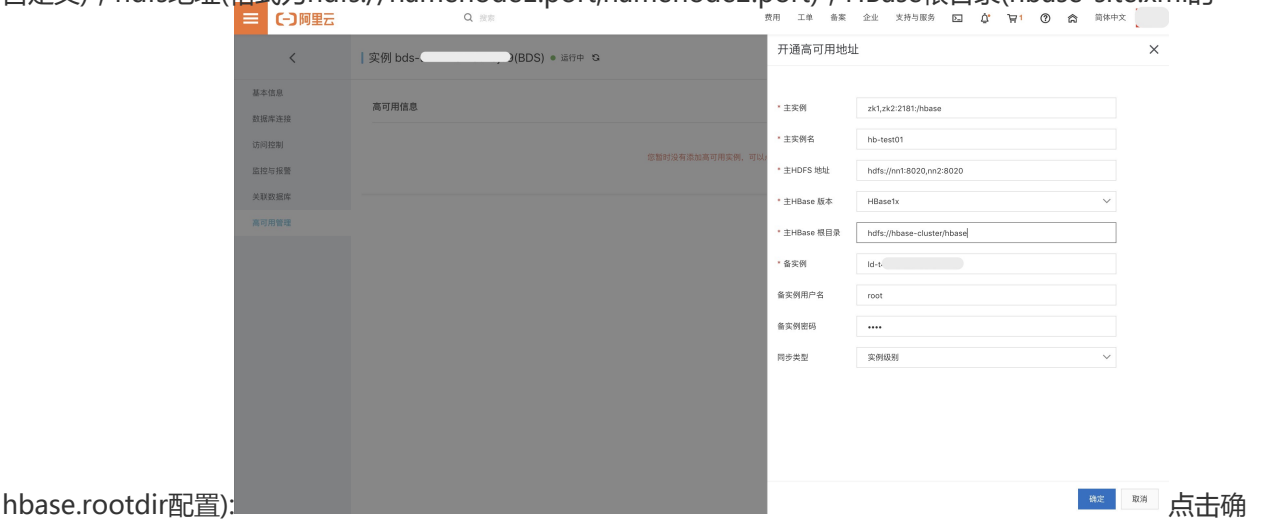


用实例：如果主备实例都是云HBase实例，则可以直接选择相应的实例名(对于增强版实例需要填写用户名和密码详见);如果备实例为新实例，高可用服务可以支持一键迁移功能，可以选择同步类型实例级别(同步主实例中全部表)和表级别(指定需要同步的表，表名按,号分隔)，如果备实例是已有实例同时已经开启了实时同步，可以选择无需同步跳过数据同步

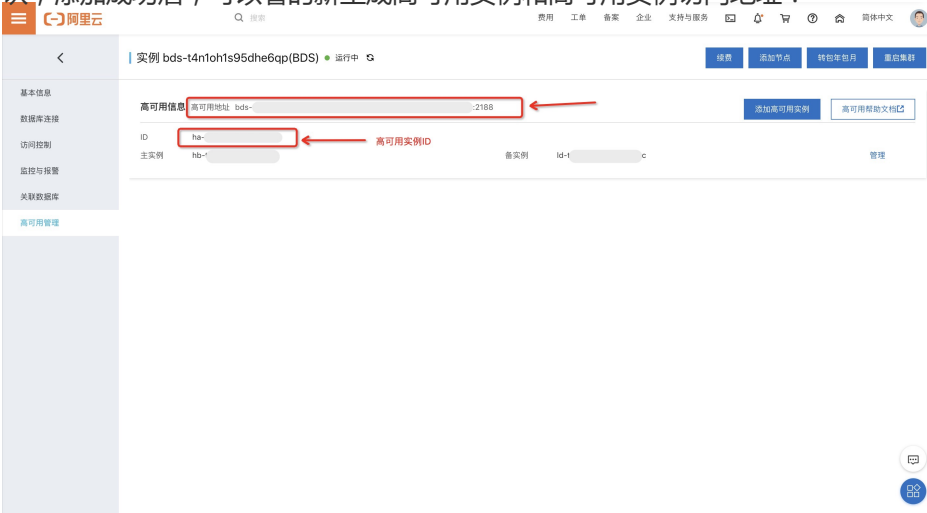


如果主备实例中存在自建实

例，则需要填写自建实例的链接信息，包括自建实例的链接地址(格式为zk1,zk2:port:/hbase),自建实例名(用户自定义)，hdfs地址(格式为hdfs://namenode1:port,namenode2:port)，HBase根目录(hbase-site.xml的



认，添加成功后，可以看的新生成高可用实例和高可用实例访问地址：



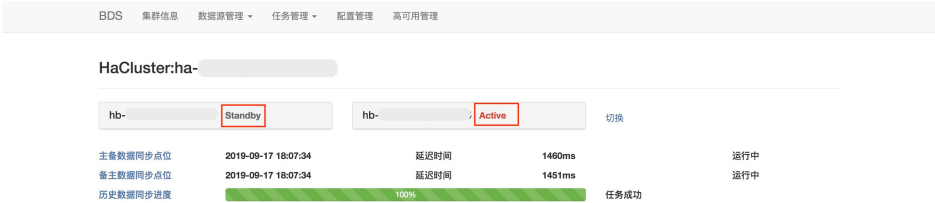
切换管理

点击管理(用户名和密码可以在数据库链接-》重置UI访问密码里面进行设置, 访问说明)可以跳到管理页面，如果选择了数据同步，可以在管理页面上看的主备实例实时数据双向同步的点位，同时也可以看的历史数据的同步进度，历史数据同步完成且实时点位正常更新就说明数据迁移成功：



可以点击切换，进行实例之

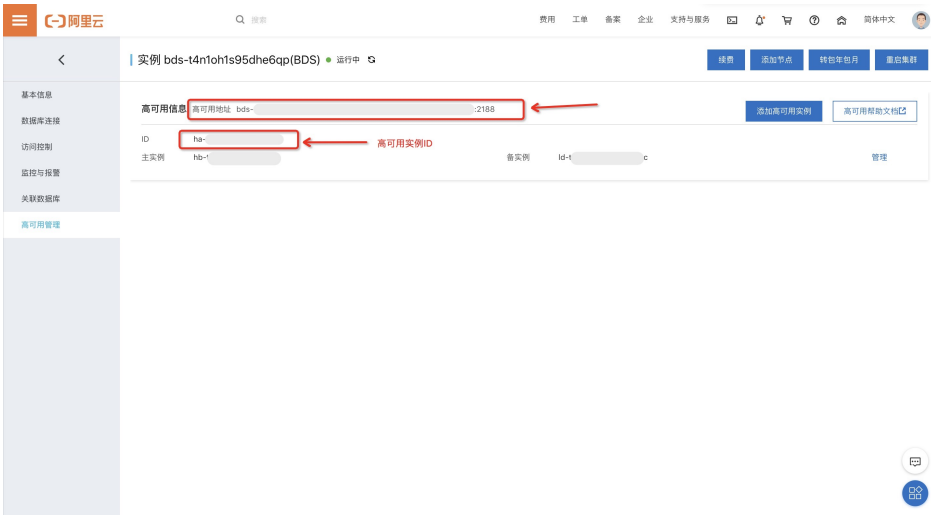
间的切换操作，操作成功后，可以看的Active状态的变化：



Java API访问

准备

1. 在高可用管理页面获取高可用链接地址和高可用实例ID：



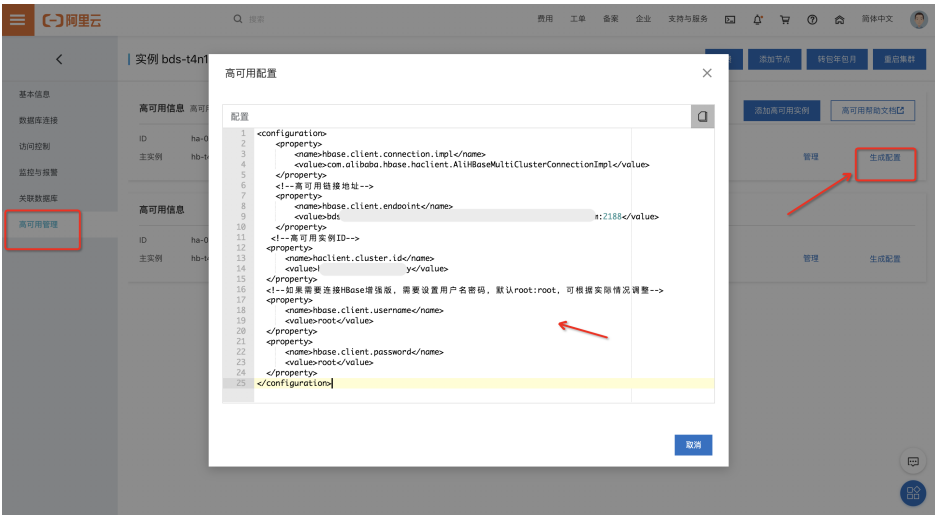
2. 完成JAVA SDK的安装

配置客户端参数

您可以通过以下两种方式，来配置客户端参数：
方式一：配置文件

```
<configuration>
<!-- 高可用链接地址 -->
<property>
<name>hbase.zookeeper.quorum</name>
<value>HOST:PORT</value>
</property>
<!--
高可用实例ID
-->
<property>
<name>haclient.cluster.id</name>
<value>ha-xxxxxx</value>
</property>
<!--
如果需要连接HBase增强版，需要设置用户名密码，默认root:root，可根据实际情况调整
-->
<property>
<name>hbase.client.username</name>
<value>root</value>
</property>
<property>
<name>hbase.client.password</name>
<value>root</value>
</property>
</configuration>
```

配置信息可以在高可用管理页面中获取：



方式二：代码

通过代码Create Configuration，然后增加相关配置

```
// 新建一个Configuration
Configuration conf = HBaseConfiguration.create();
// 高可用链接地址
conf.set("hbase.zookeeper.quorum", "HOST:PORT");
// 高可用实例ID
conf.set("haclient.cluster.id", "ha-xxxxxx");
// 设置用户名密码，默认root:root，可根据实际情况调整
conf.set("hbase.client.username", "root")
conf.set("hbase.client.password", "root")
```

创建连接

通过配置conf创建Connection，然后访问HBase

```
// 创建 HBase连接，在程序生命周期内只需创建一次，该连接线程安全，可以共享给所有线程使用。
// 在程序结束后，需要将Connection对象关闭，否则会造成连接泄露。
// 也可以采用try finally方式防止泄露
Connection connection = ConnectionFactory.createConnection(conf);
```

新建表

使用haclient，支持使用API，在主备实例上新建表，新建表会实时在Active实例上进行创建，同时会生成异步任务在standby实例上进行创建，并且会对更新实时同步任务更新，可以在高可用实例管理页面进行查看：

HA-HBase:

Active 链接数: 3

Standby 链接数: 0

切换

新建表

Active->Standby数据同步点位

延迟时间

369ms

运行中

Standby->Active数据同步点位

延迟时间

391ms

运行中

历史数据同步进度

100%

任务成功

切换日志

时间	切换记录
2019-09-23 22:19:10	FROM:
2019-09-23 22:19:58	FROM:
2019-09-23 22:23:14	FROM:
2019-09-23 22:28:30	FROM:
2019-09-26 11:26:59	FROM:
2019-10-09 10:41:25	FROM: