



GOPS 2021
Shenzhen

GOPS

全球运维大会

2021
-XOPS 风向标



深圳站

中国·深圳

指导单位：



主办单位：



时间：2021年5月21日-22日

腾讯Elasticsearch数据迁移与容灾实践

高斌龙 腾讯云大数据工程师



高斌龙

腾讯云大数据工程师

高斌龙(bellen), 腾讯云大数据开发工程师, 目前专注于Elasticsearch云产品的研发工作。

目录

CONTENTS

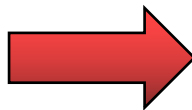
- ① 异构数据与ES同步
- ② ES集群间数据迁移
- ③ ES集群容灾实践

异构数据与ES同步

不同数据源与Elasticsearch之间数据导入导出

01

异构数据与ES同步

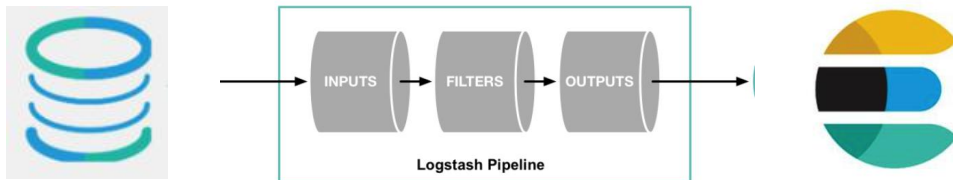


对象存储 COS

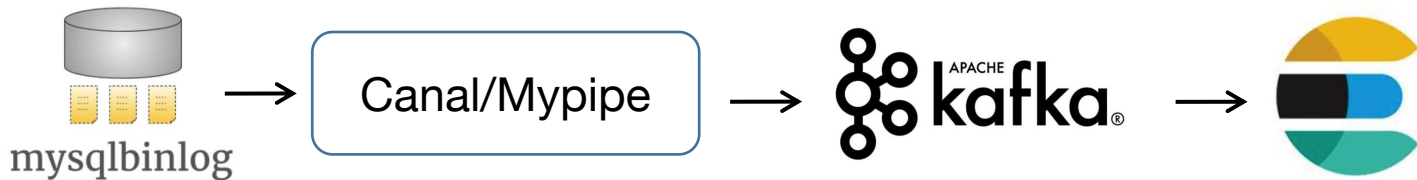
Amazon S3

1.1 mysql -> es

1. 离线同步(批量读+批量写)

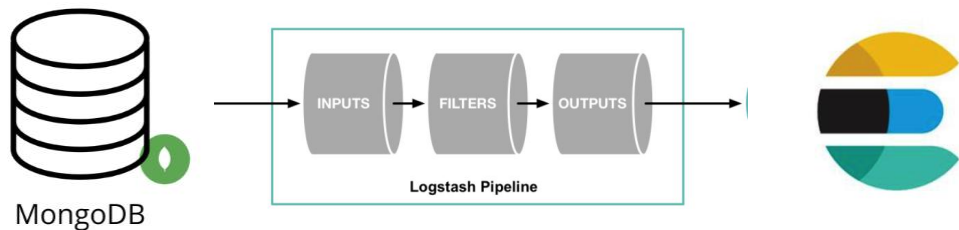


2. 实时同步(基于Binlog)

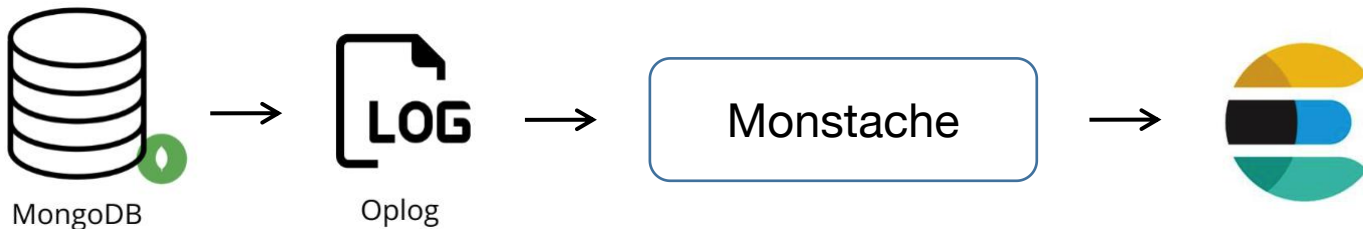


1.2 mongodb -> es

1. 离线同步(批量读+批量写)

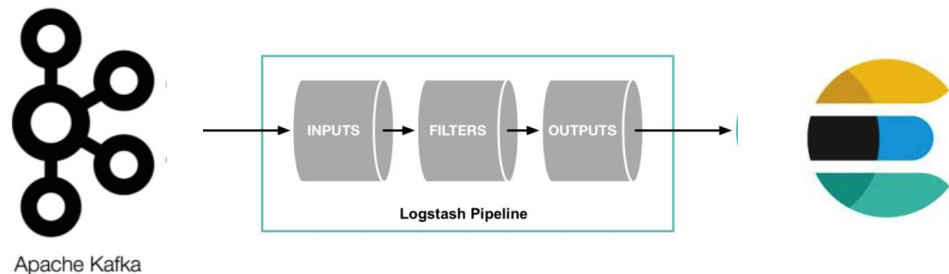


2. 实时同步(基于oplog)



1.3 kafka -> es

1. Logstash

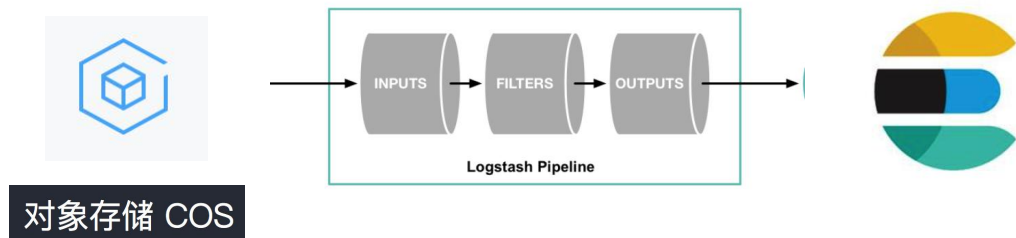


2. Serverless



1.4 cos -> es

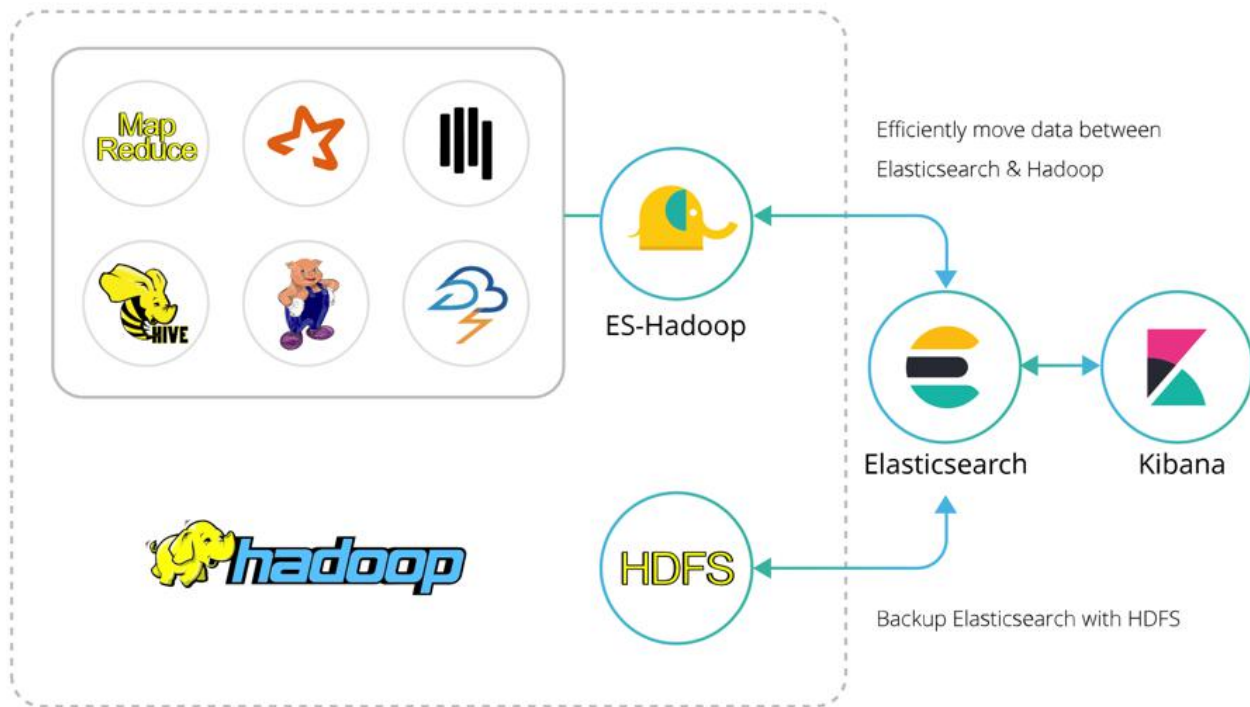
1. 全量同步



2. 增量同步



1.5 hadoop-> es



ES集群间数据迁移

满足跨机房、跨云数据同步需求

02

ES集群间数据迁移

1. 离线迁移

迁移过程中旧的集群**可以停服或者暂停写入**，迁移完成后业务切换新的集群进行读写

2. 在线迁移

迁移过程中旧的集群**不能停服，不能暂停写入**

2.1 离线迁移

1. elasticsearch-dump

简单易用、适用于数据量较小的场景(10GB以下)

2. Logstash

适用于需要对数据进行过滤或者预处理，或者源集群与目的集群版本跨度较大的场景

适用于100GB以下数据量

3. Reindex

直接调用ES的API进行迁移，目的集群的节点需要能够访问源集群

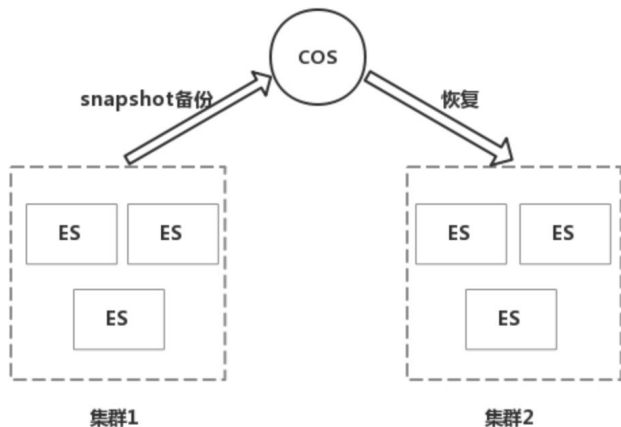
适用于100GB以下数据量

2.1 离线迁移

4. Snapshot

适用于数据量较大、对迁移速度要求较高的场景

实现原理：直接备份源集群节点磁盘上的文件，在目的端集群中进行恢复



2.2 在线迁移

1. 只有数据追加或者更新操作

数据量较小的情况下，可以使用Logstash，先全量、再增量迁移

2. 有删除操作

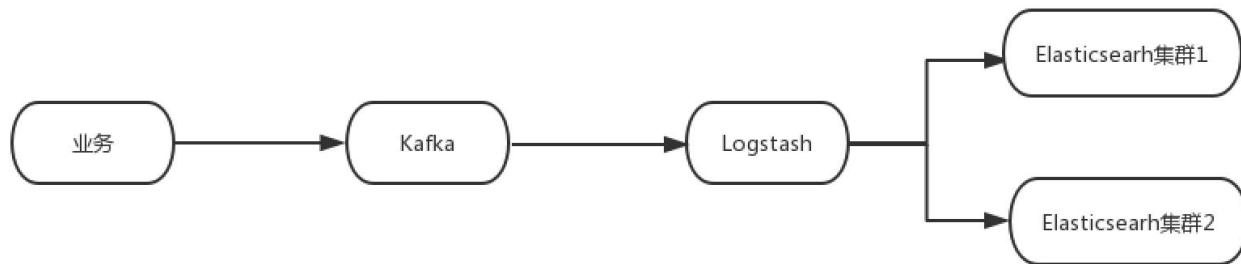
- snapshot+双写：存量索引通过snapshot迁移，增量索引进行双写
- 业务双写：业务数据都存储在kafka等消息中间件中
- CCR(跨集群复制)：开启集群间数据同步，数据一致性可以得到保证
- 集群融合：基于腾讯云ES集群的节点双网卡方案，实现集群融合

2.2 在线迁移之业务双写

- 业务通过kafka实现双写

写入过程中丢失数据风险较低

可以保证新旧集群的数据一致性

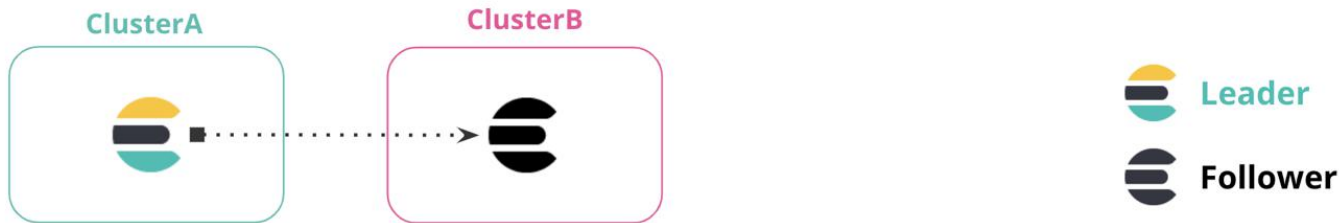


2.2 在线迁移之CCR

- 基于原生CCR特性

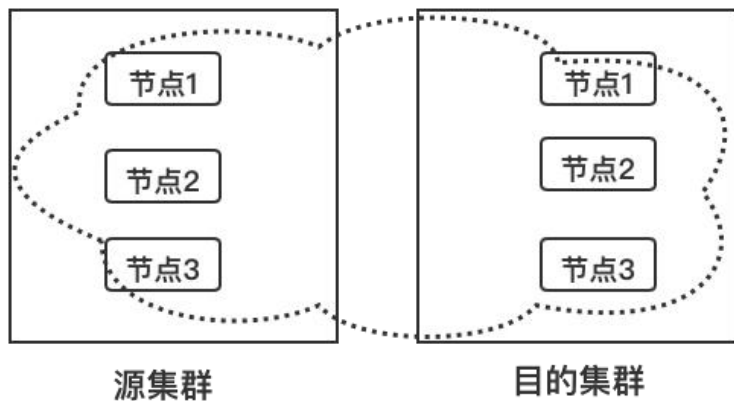
旧集群和新集群的版本 ≥ 6.5 ，且均为白金版集群

Leader Index必须开启index.soft_deletes.enabled(7.x版本默认开启)



2.2 在线迁移之集群融合

- 将两个集群融合为一个



ES容灾实践

跨机房、跨地域高可用

03

ES集群容灾

1. 同城容灾

主备集群（一主多备，部署在多个机房）

跨机房部署集群（专线打通，一个集群的节点分布在多个机房）

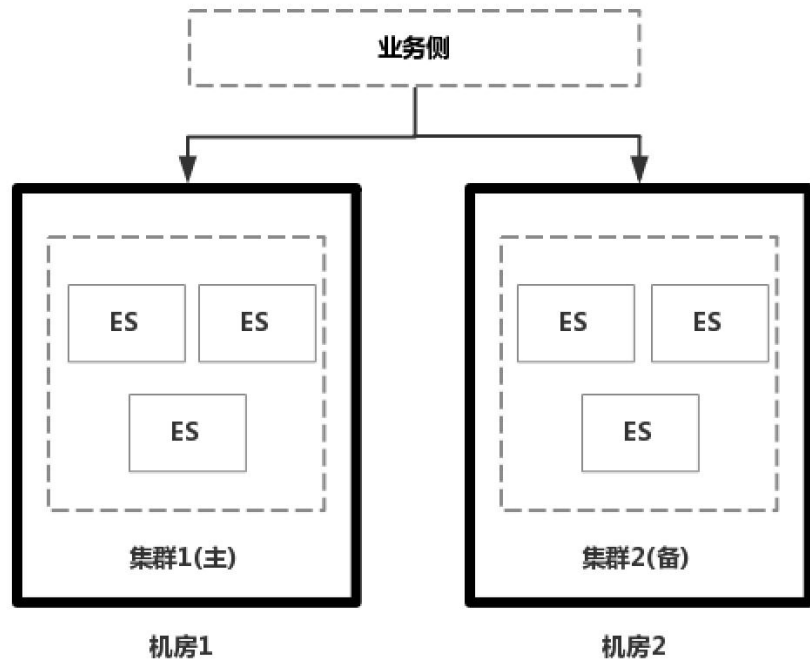
2. 异地容灾

主备集群（通过CCR实现主备同步）

3.1 同城容灾之主备集群

• 主备集群

主备集群分别部署在两个机房中
解决数据同步的问题（同步双写、
异步复制）



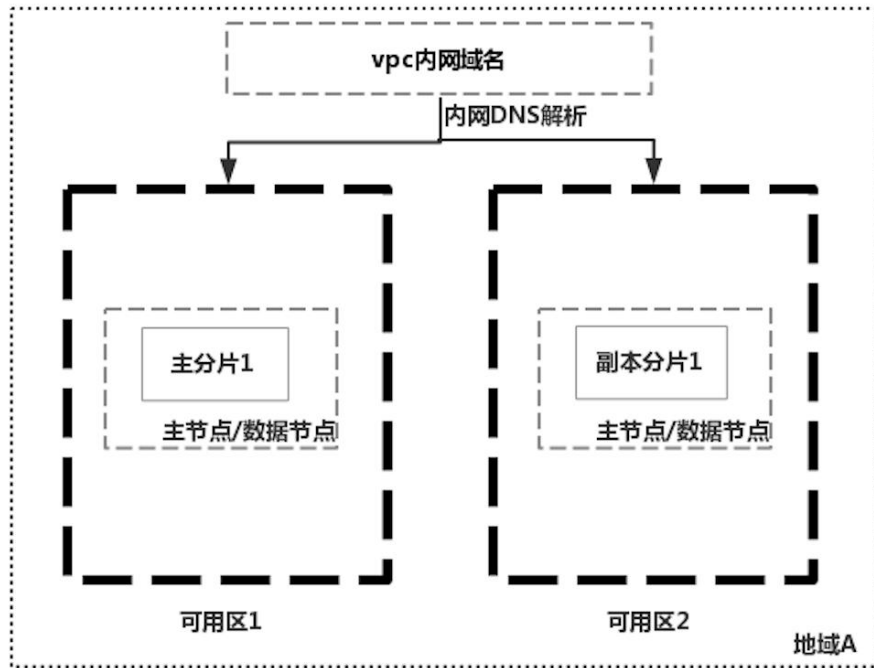
3.1 同城容灾之跨机房部署

• 跨机房部署集群

在两个机房中部署一个集群

主分片和副本分片分布在两个机房中

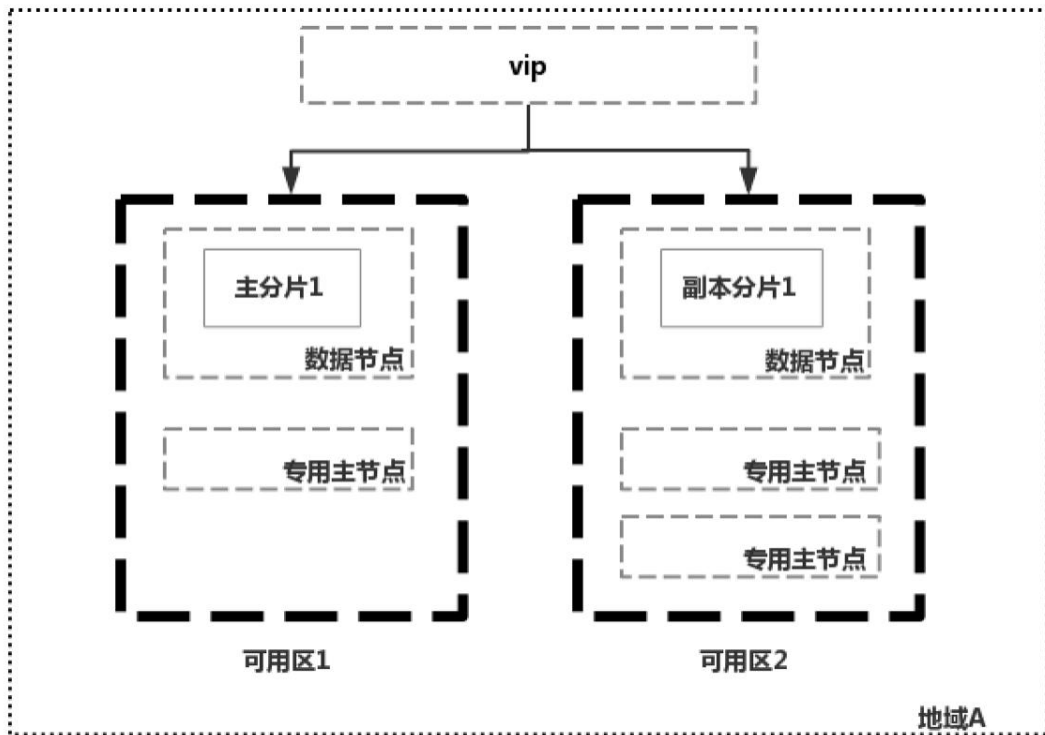
脑裂问题?



3.1 同城容灾之跨机房部署

- 跨机房部署集群

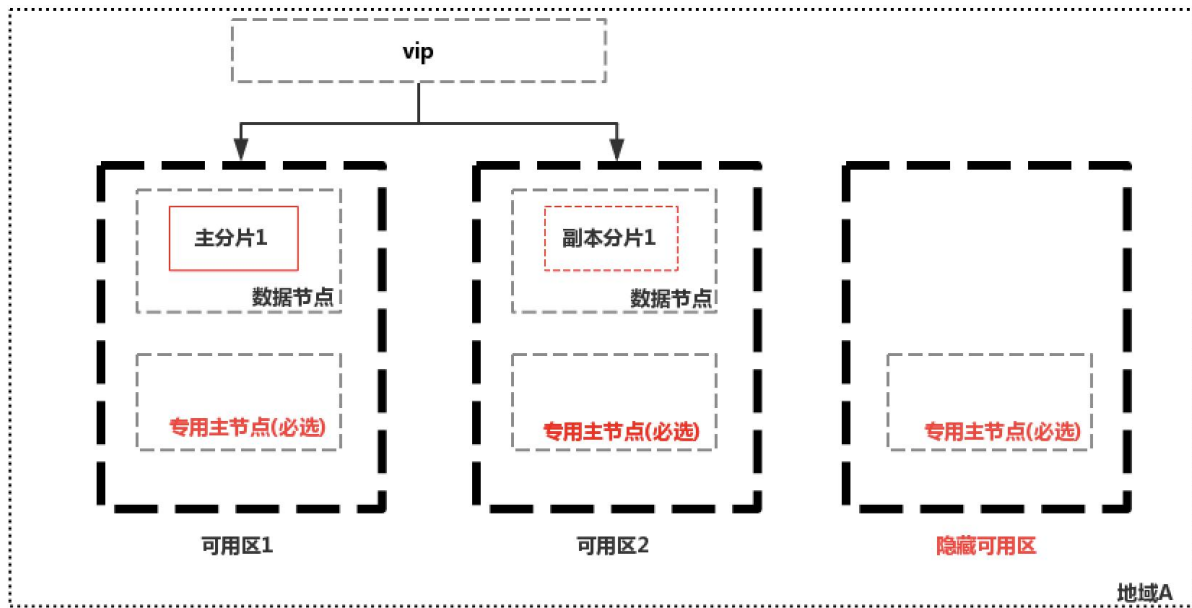
无法选主的问题？



3.1 同城容灾之跨机房部署

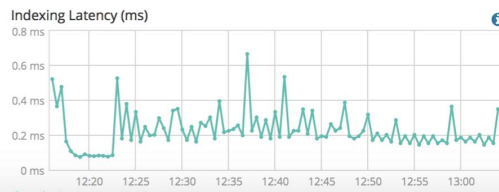
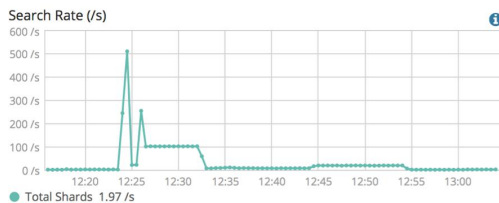
- 增加隐藏可用区

必选3个或5个专用主节点，分布在三个可用区

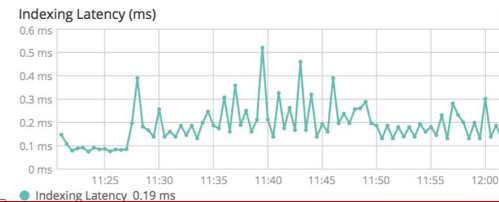
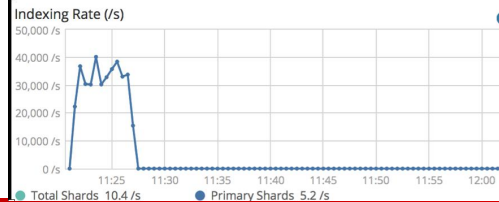
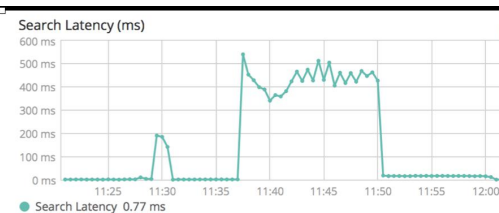
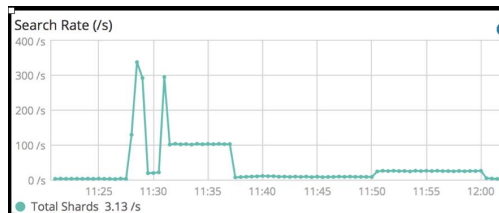


3.1 同城容灾之跨机房部署

• 性能损耗

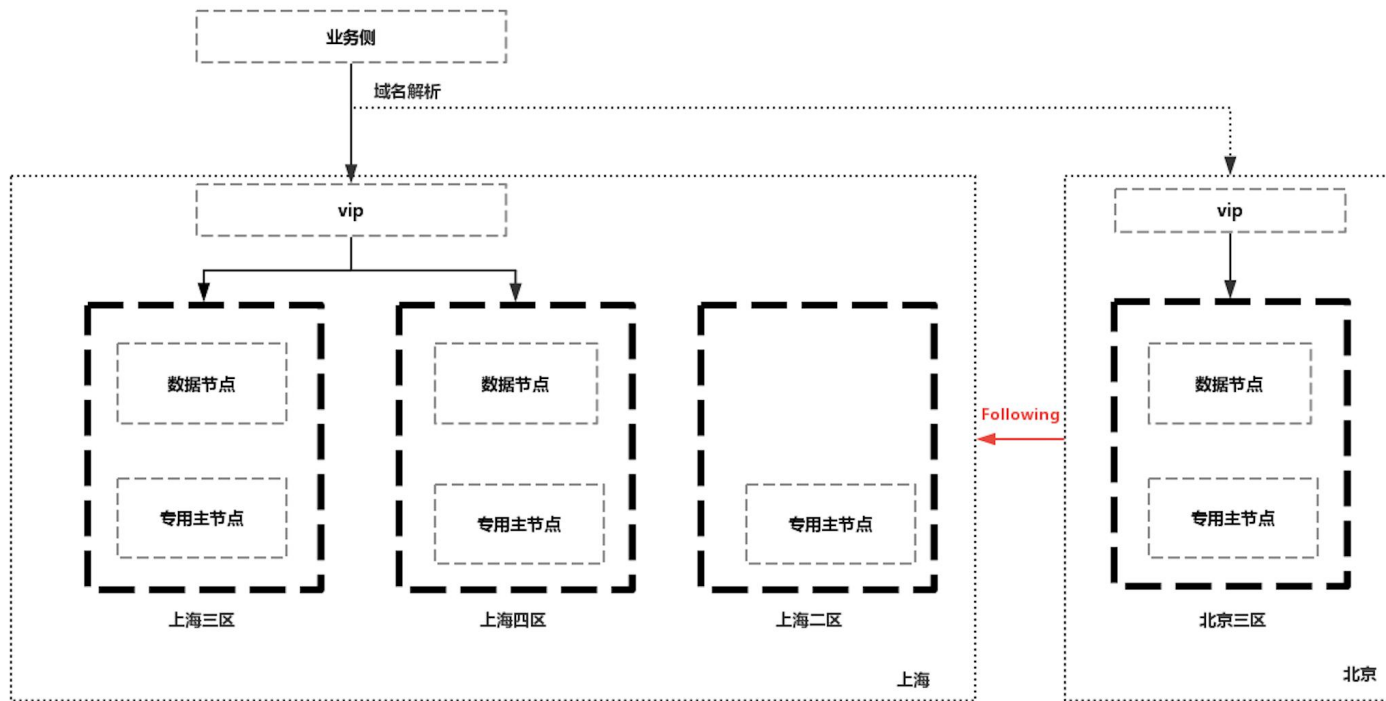


单可用区



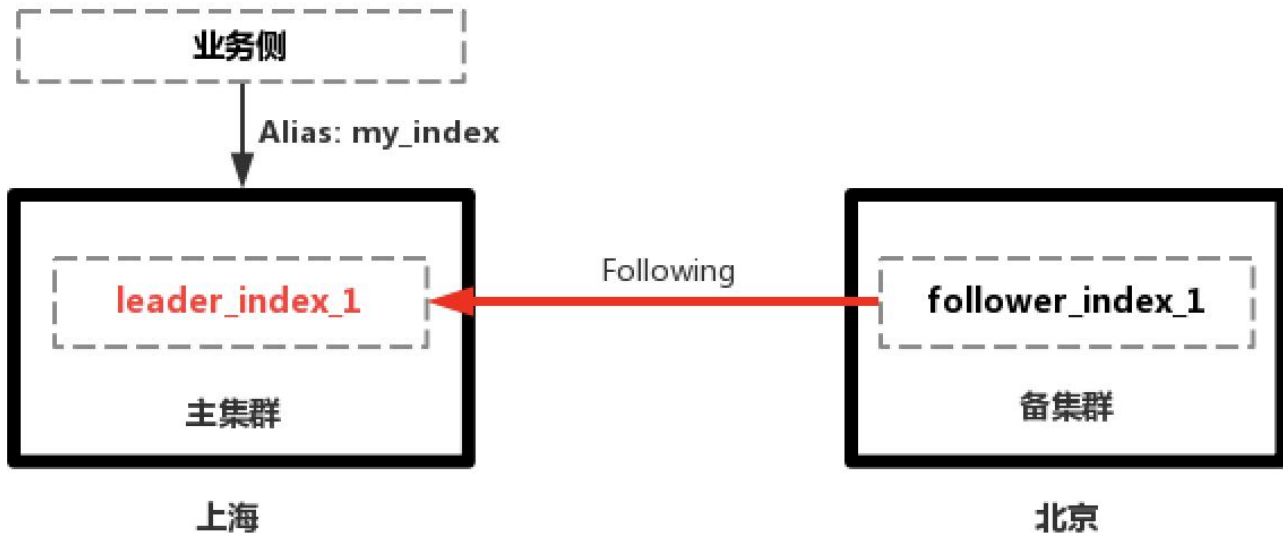
多可用区

3.2 异地容灾



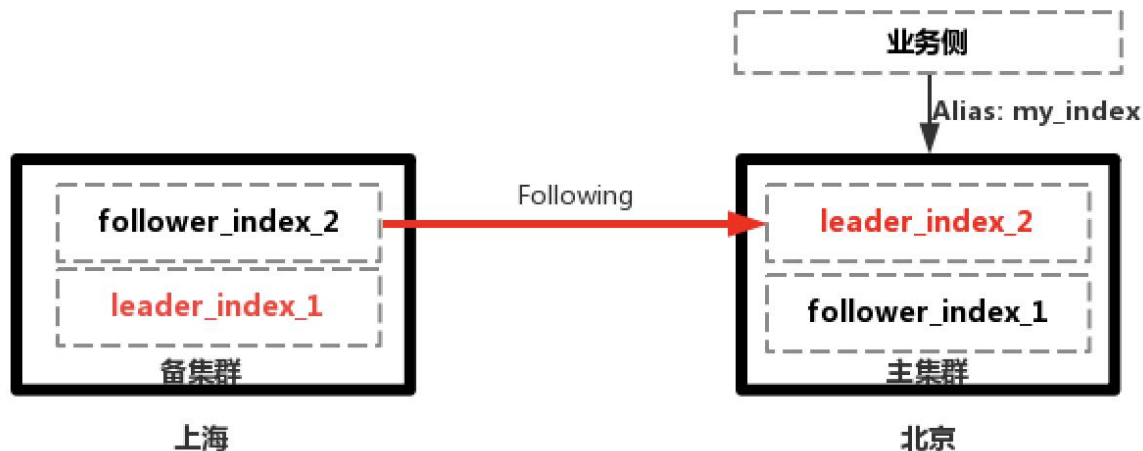
3.2 异地容灾

- 上海主集群正常提供服务
- 北京备集群从主集群同步数据



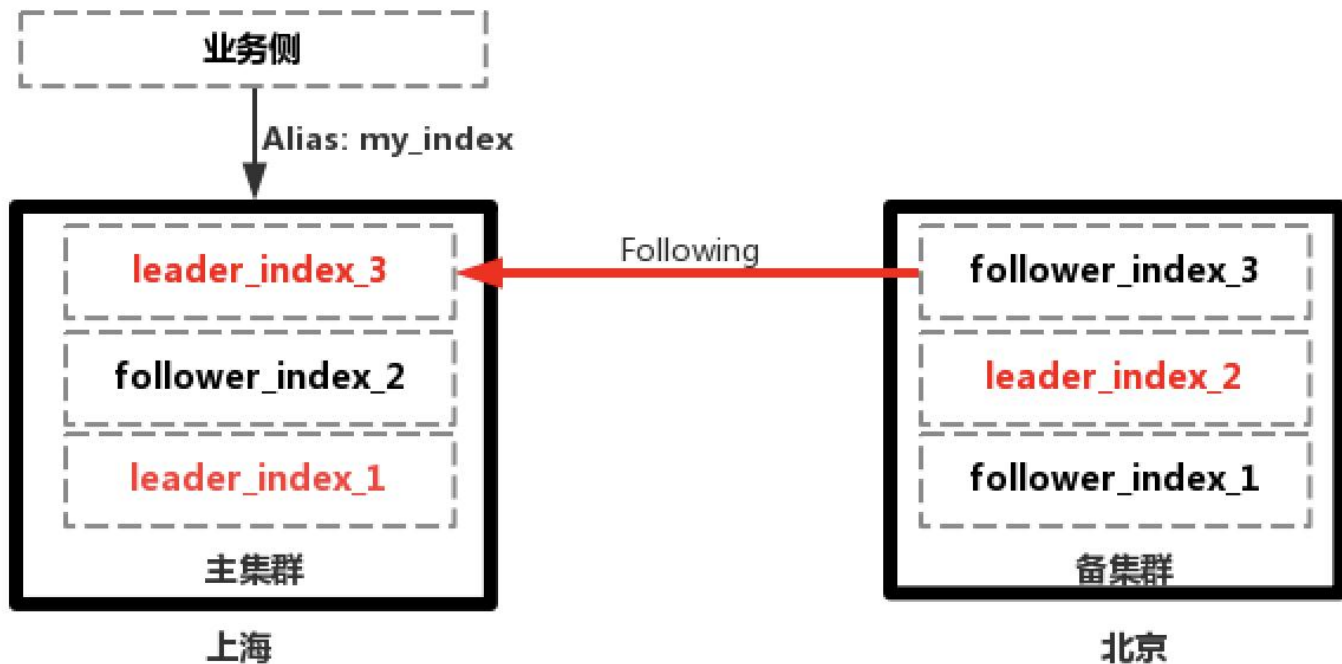
3.2 异地容灾

- 上海主集群故障
- 业务切换至北京集群进行读写
- 上海主集群恢复后从北京集群Follow数据



3.2 异地容灾

- 数据追平后，业务切换到上海集群进行读写





Thanks

高效运维社区
开放运维联盟

荣誉出品