



HBase in Youzan

HBase 在有赞的平台实践和应用



目录 / Contents

01

关于我们

02

管控平台

03

改造与业务实践

04

未来展望



01

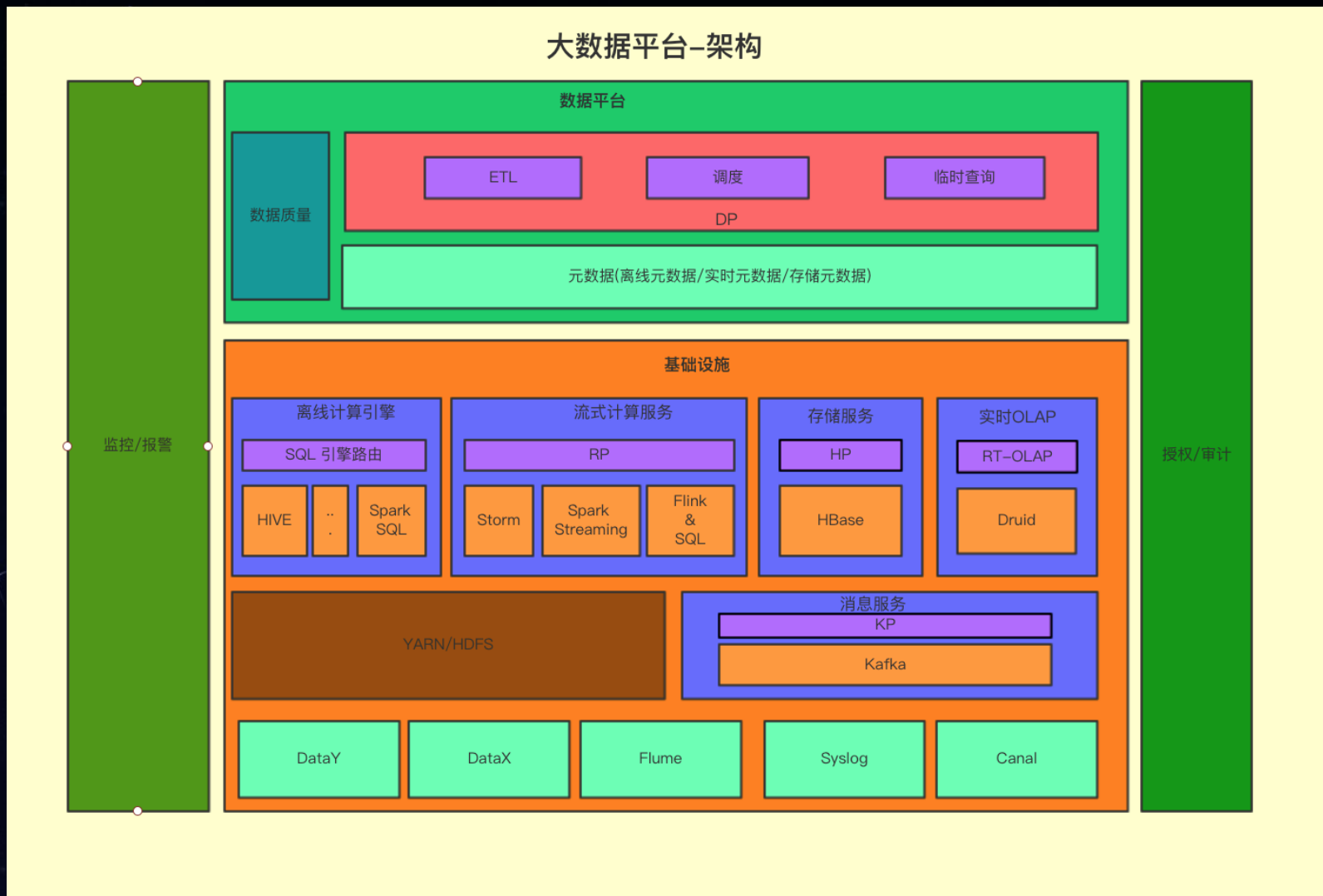
关于我们

有赞大数据平台



1

团队介绍





1 HBase 产品定位



在线存储

多集群容灾存储，在线实时访问



离线存储

离线批量导入导出，非敏感实时读取，
数据冷备



监控数据存储

OpenTSDB 组件后端存储

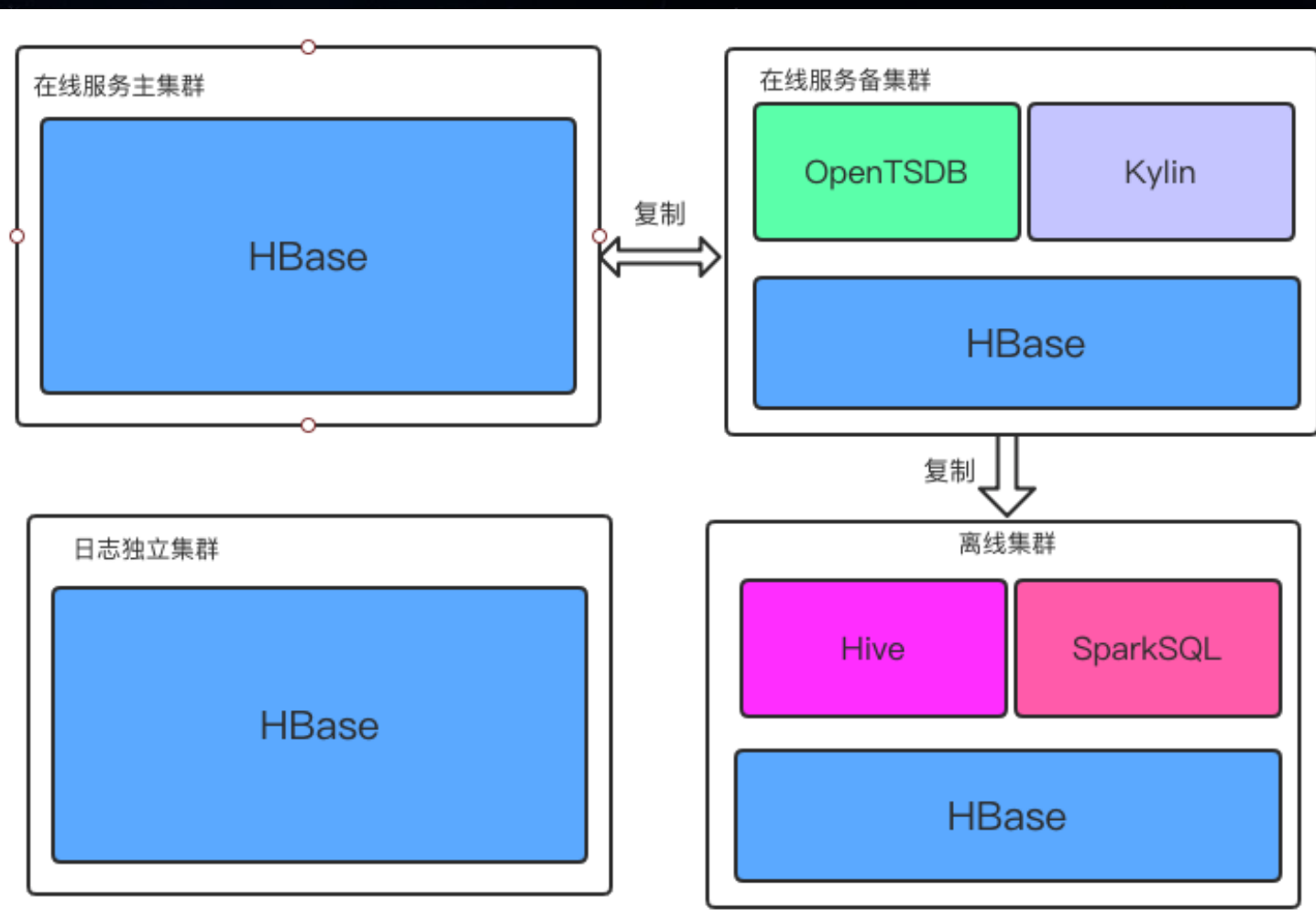


Kylin服务

Kylin的后端存储



1 集群规划



四大集群

在线服务主、备集群

离线集群

日志业务独立集群

在线服务集群互为主备

离线集群服务计算

日志业务独立集群专为日志服务



02

管控平台

监控系统与平台工具

痛点.....

混布问题

只有一个离线大集群，部署了很多组件

业务管控问题

不知道接入了多少业务，不知道哪些业务需要什么样的SLA，出了问题找不到对接人

监控系统问题

原有的监控不够灵活，告警配置单一，难扩展

人肉运维问题

没有成熟好用的迁移工具等运维工具，全靠人肉运维



2

管控平台

监控系统

全面掌握集群运转情况，
给告警系统提供基础数据



强大的迁移工具

离线到在线；在线到在线



业务管控系统

将项目-资源-人等重要信息整合



易用的查询工具

面向业务用户，易用而安全





2 管控平台/业务管控



工单

用户申请建表统一需
要走工单流程审批，
所有操作有迹可循

权限

业务用户可以根据需
求申请项目级别的权
限，查看其它表的元
信息与监控数据

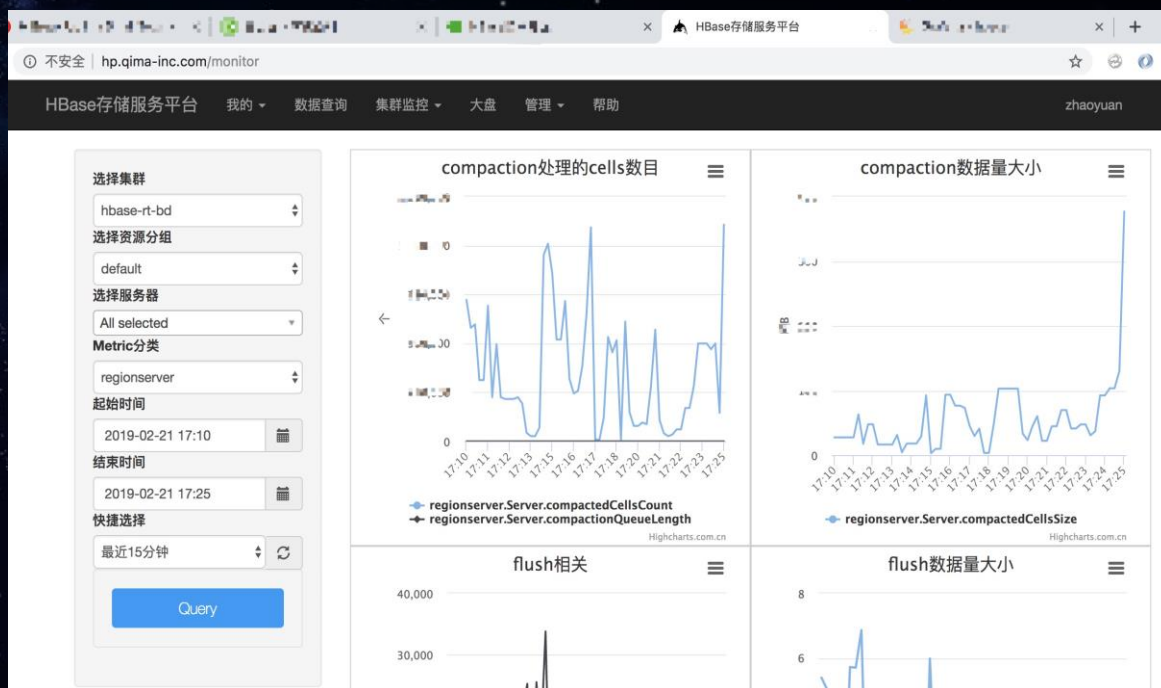
大盘

核心指标监控有数据
大盘，哪些表占用了
最多的资源一目了然



2

管控平台/监控系统



系统级别监控

关键指标监控

系统级别，RS级别

相比原有监控系统的优势

支持RegionServer逻辑分组

指标逻辑分组



2

管控平台/监控系统



复制延迟监控

最大复制延迟和平均复制延迟
多集群聚合，非聚合查询支持

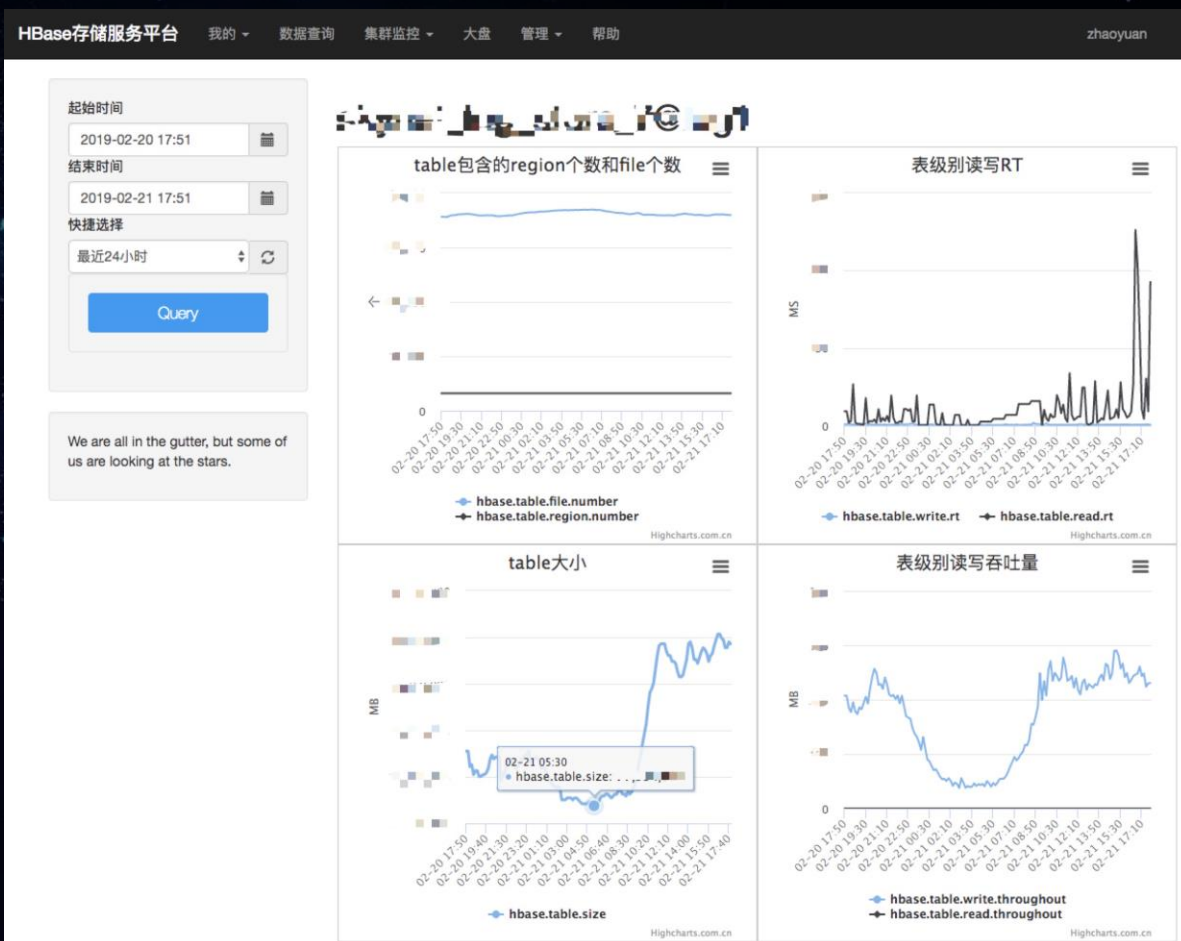
意义

最终一致性语义的下SLA的确定
提供延迟告警的监控数据



2

管控平台/监控系统



表级别监控

表级别监控数据采集，汇总，按表展示

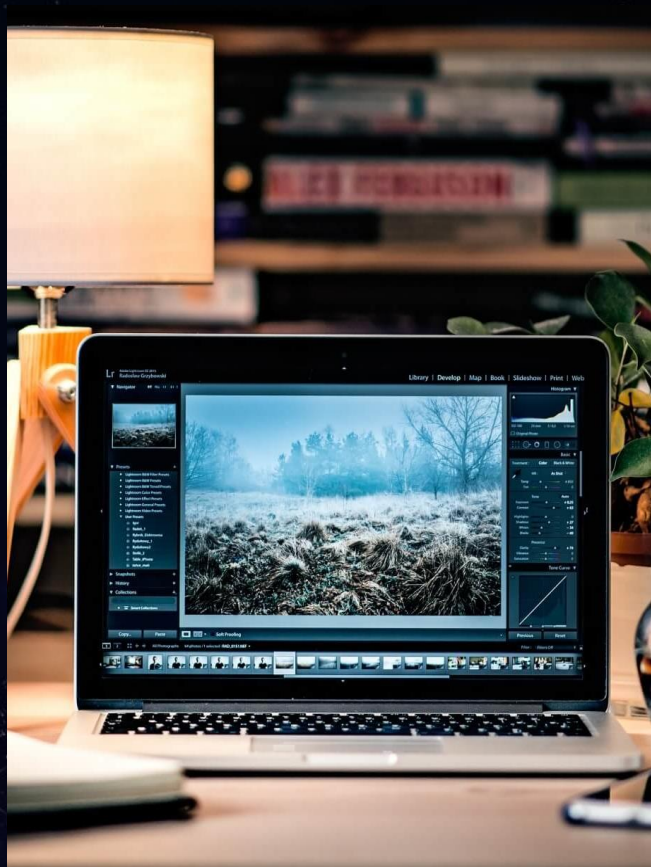
意义

补全原生监控数据的短板

关键指标一目了然



2 管控平台/告警



定制化告警

RIT , RS dead count , Replication lag

Phone , sms , wechat....

Group , Person



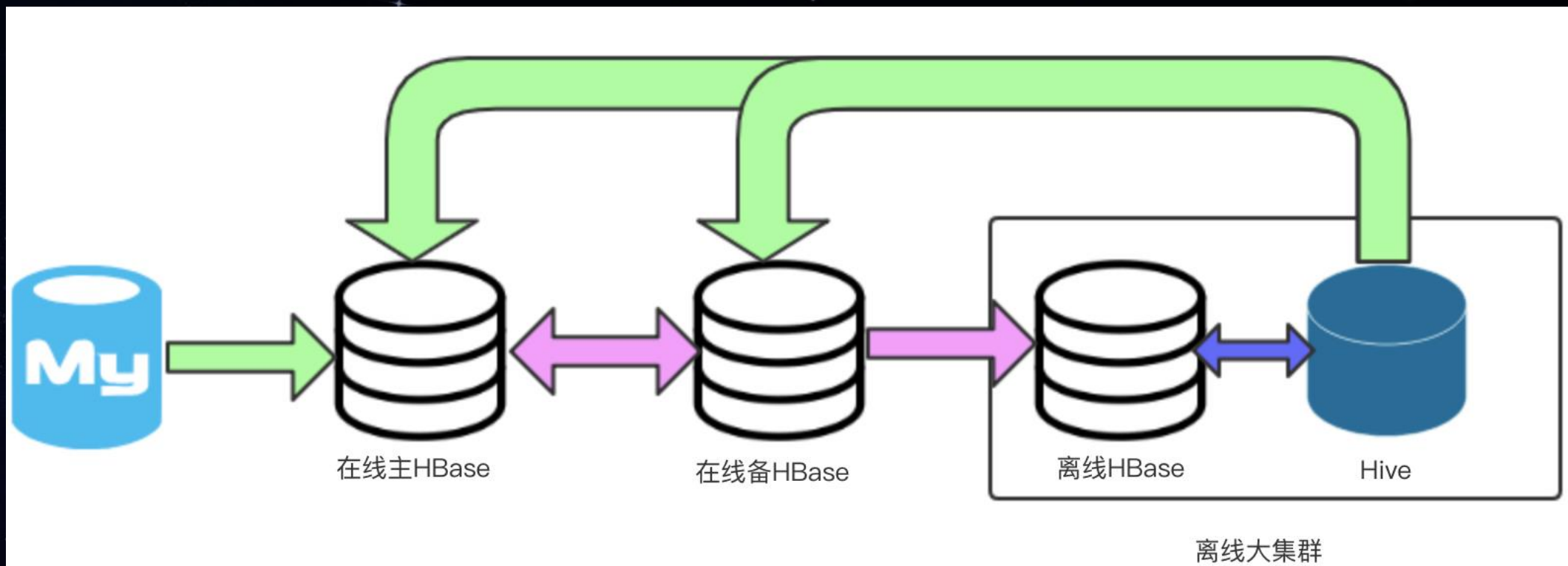
2

管控平台/效率工具



2

管控平台/迁移工具

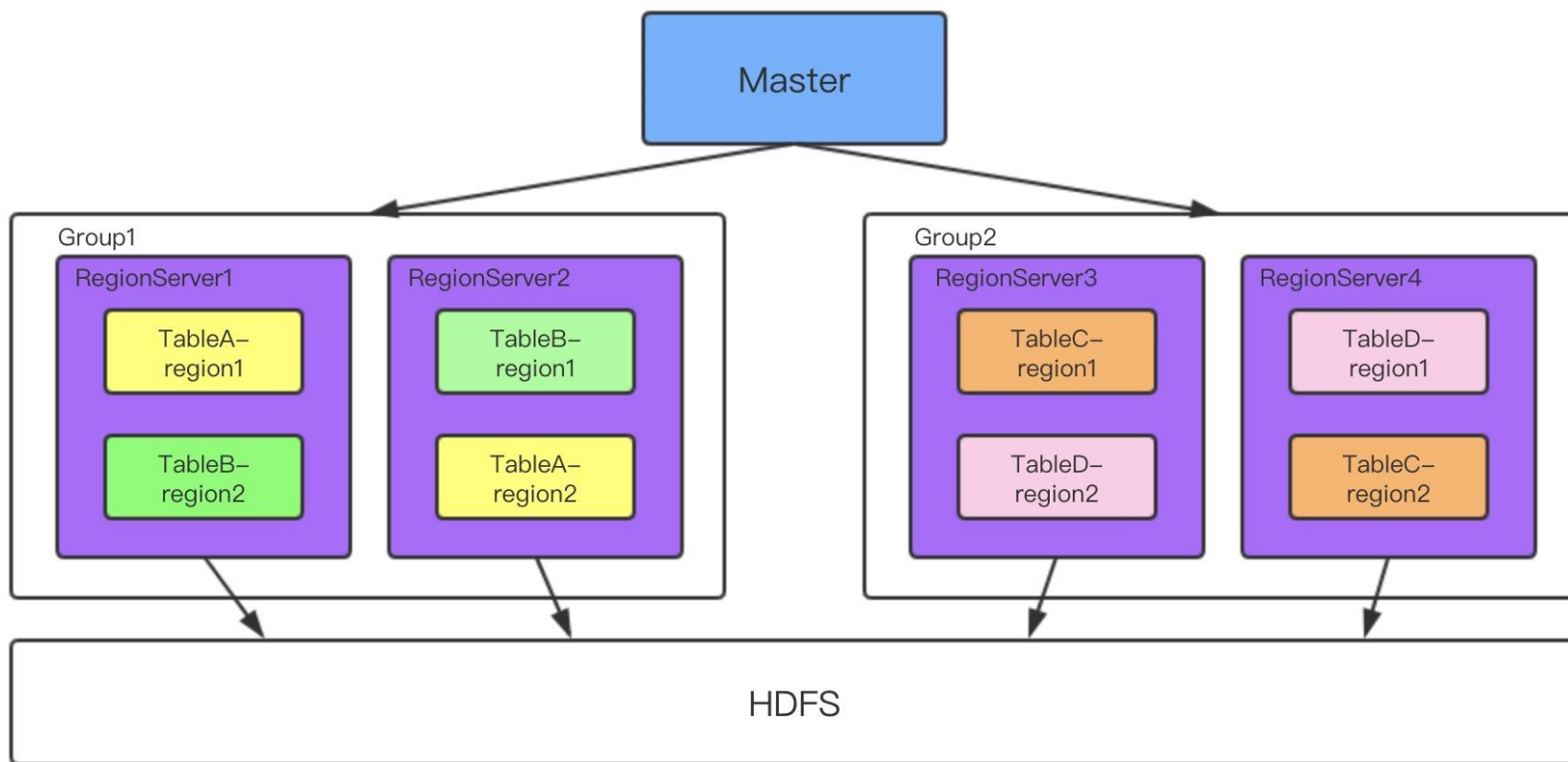




03

改造与业务实践

我们在1.2.6版本之上做了哪些改造，我们为何要做这些改造，这些改造给业务带来了哪些价值？



资源隔离

HBASE-6721

指定一些RegionServer成为一个group，表挂载在不同的分组里，只在该分组里balance。在HBase层面不用组之间的表互不影响

解决的问题

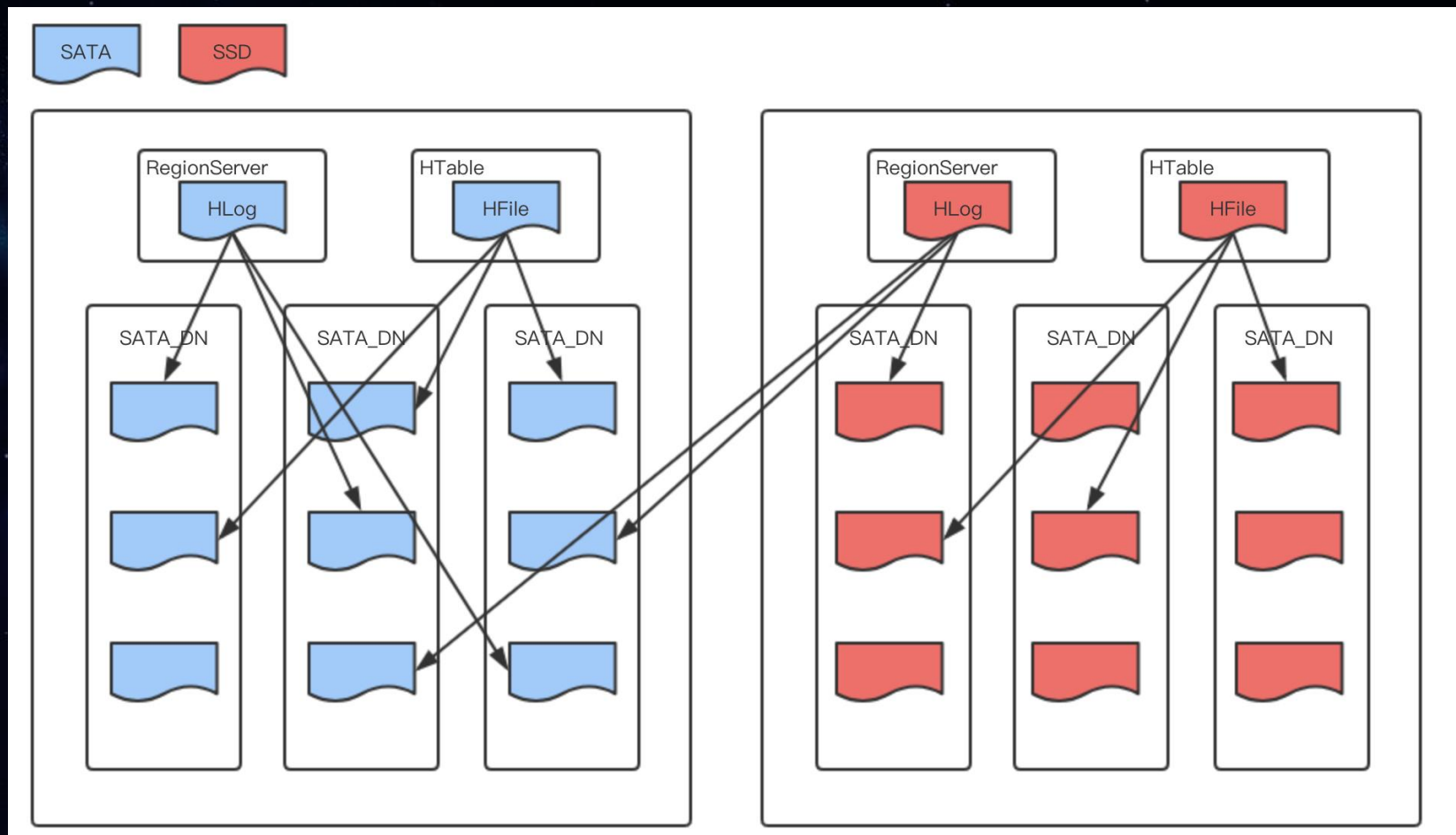
业务隔离，降低管理成本。

问题：HDFS层没有隔离，实际在IO与网卡等层面依然会互相影响



3

系统改造

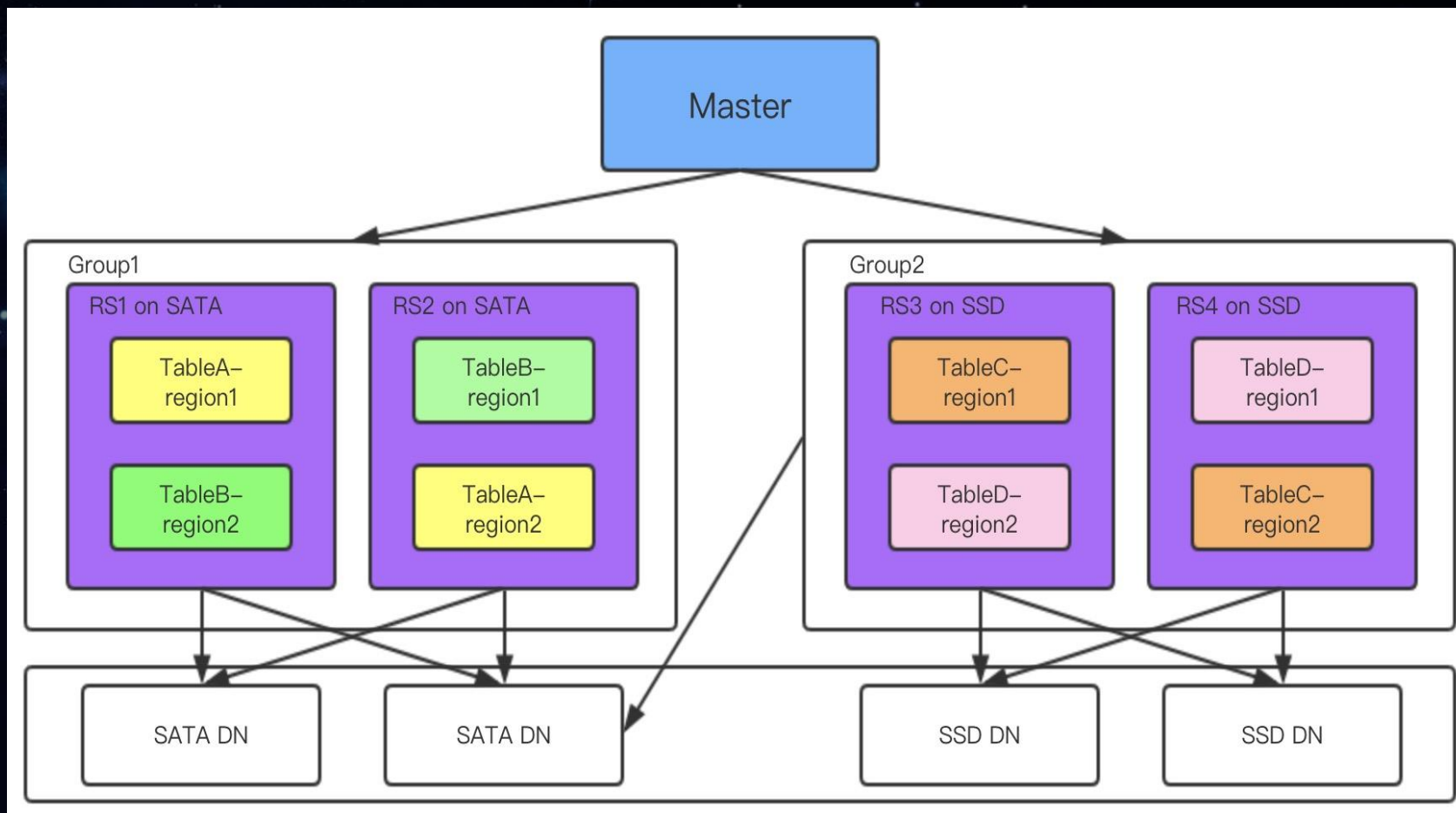


异构存储

HDFS 2.6版本支持SSD存储介质,HBASE-14061
可以指定RegionServer存储HLog的格式,以及table
cf级别的存储方式;
可以指定存储格式为HOT,ONE_SSD,ALL_SSD。

解决的问题

1. 支持同集群内对多种机型的资源利用
2. ALL_SSD, 提供低读写延迟
3. HOT, 提供大容量存储空间
4. ONE_SSD, 融合上面两种的优点



优化资源隔离

利用RSGroup 结合异构存储，完成HBase 层面与HDFS层面的双重隔离。

利用廉价的SATA存储较多的副本，SSD存储本地副本，或者所有的副本都存储在SSD DN上，达到最好的效果。

解决的问题

- 1.HDFS层面的隔离
- 2.提供满足不同SLA的分组



3 Bug Fixed

01

监控数据漏失

02

Split_policy无法设置

HBASE-19340

03

缓存不生效问题

HBASE-20697



3

业务实践

已接入的业务

粉丝详情

低延迟，高可用

调用链

重写低读，高吞吐

天网日志

消息

订单详情

风控数据

商品详情

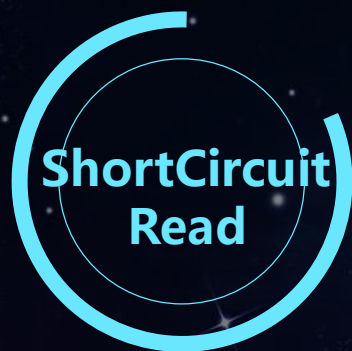
其他.....





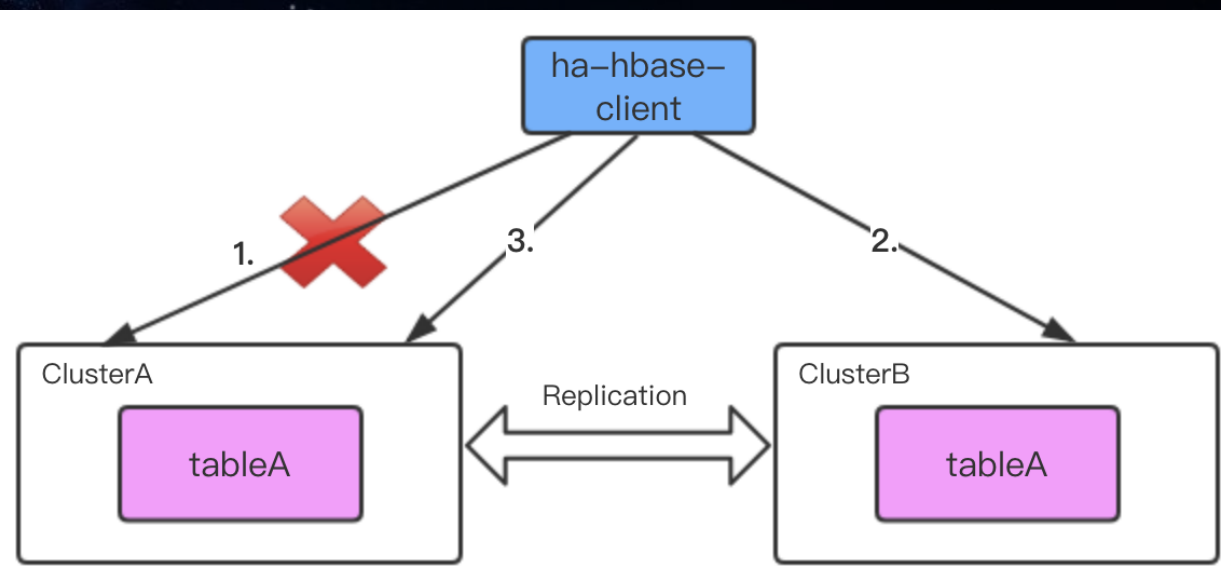
3

业务实践/粉丝详情





3 优化改造/ha-hbase-client



高可用客户端/ha-hbase-client

支持单次请求路由到备集群

支持熔断

支持集群健康检查

业务无感知

降低开发成本

高可用



3

业务实践 / 调用链

高吞吐

单条数据平均下来比较大
写TPS很高，读TPS较少



读写延迟不敏感

在写不发生阻塞的情况，对读写rt要求不高



改造前情况

多台cassandra ssd机型
负载高
运维成本高



痛点

运维成本偏高
集群负载高，服务不稳定
Cassandra优化做的不够，掌控能力差





3

业务实践/调用链

预分区

HDFS
副本数

避免MC



3

业务实践/调用链

更少的机器 更稳定的服务 更好的扩展性



我们用更少的机器，更低的负载完成了对调用链的支持，即使调用链流量翻倍，我们可以较快的横向扩展，且不会影响SLA



04

未来规划

实时数仓的重要组成部分



4

未来规划



跟进社区，反馈社区

获取社区红利，将发现的问题解决的bug，产出patch回馈社区



服务更多的业务

完善存储平台，开发更多更有用的工具，接入更多的业务



实时数仓的重要组成部分

与其他组件有机结合，而不单纯只做一个存储的组件，它将会是实时数仓的重要组成部分





CHTC
中国HBase技术社区

云栖社区

THANK YOU

有赞大数据平台团队·赵原



CHTC
中国HBase技术社区

云栖社区



社区管理员



HBase 技术社区公众号



HBase + Spark 钉钉社区大群