

HBase at JD

HBaseCon Asia 2017
xingbo peng & bang wen



主要内容

- 规模现状
- 应用的业务场景
- 高可用改进
- 运维实践
- 未来展望

规模现状

- 4个数据中心
- 3000+台服务器
- 30+个集群，单集群最大规模400+
- 业务实例600+
- 用户1000+

基础版本

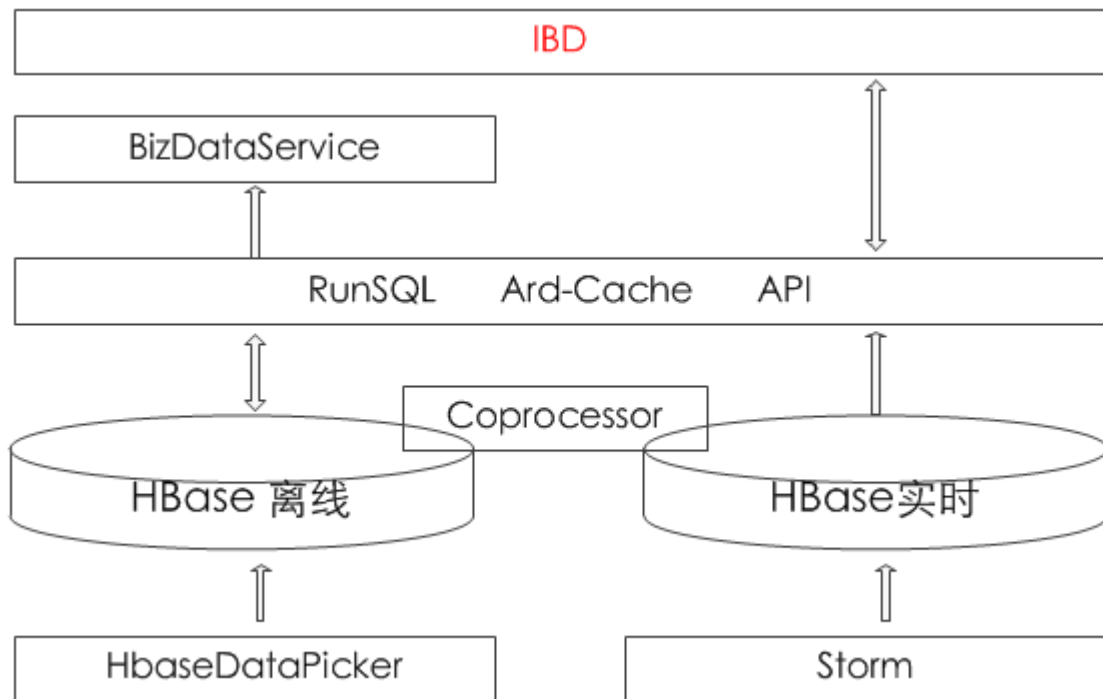
- 0.92->0.94 -> 0.96 -> 0.98 -> 1.0 ->1.1

应用业务场景

- 罗盘（商家、供应商、品牌）
- 个性化推荐
- 商品评价/会员PLUS
- 风控/白条/支付/资管（金融）
- 订单/青龙
- MDC/UMP/CAP/JDH/监控大屏

应用业务场景-商智

- 数据罗盘
- 存储600T
- 表620+
- 类sql查询
- 1100w/tps



应用业务场景-推荐

- 首页、商详、购物车
- 存储130T
- 表700+
- TP999<200ms
- 300w/tps

The screenshot displays the JD.com shopping cart. At the top, the JD.COM logo and '购物车' (Shopping Cart) are visible. A search bar contains the text '自营'. A yellow notification bar states: '您还没有登录！登录后购物车的商品将保存到您账号中' (You are not logged in! After logging in, the items in your shopping cart will be saved to your account) with a '立即登录' (Log In Now) button.

The cart header shows '全部商品 1 京东大药房' (All items 1 JD Pharmacy) and a delivery location of '北京朝阳区四环到五环之间' (Beijing Chaoyang District, between the 4th and 5th Ring Roads). The cart table has columns for '全选' (Select All), '商品' (Item), '单价' (Unit Price), '数量' (Quantity), '小计' (Subtotal), and '操作' (Action).

The cart contains one item: a Gree (GREE) 2匹定频 (2HP Fixed Frequency) Inverter Air Conditioner (KFR-50LW/(50551)NhAa-3). The unit price is ¥5999.00, and the quantity is 1. The subtotal is ¥5999.00. A discount of ¥200.00 is applied, resulting in a total of ¥5799.00. The item is marked as '京东自营' (JD Self-Operated) and '已免运费' (Free Shipping). A '满减' (Spend and Save) banner indicates that items over ¥5000.00 receive a ¥200.00 discount.

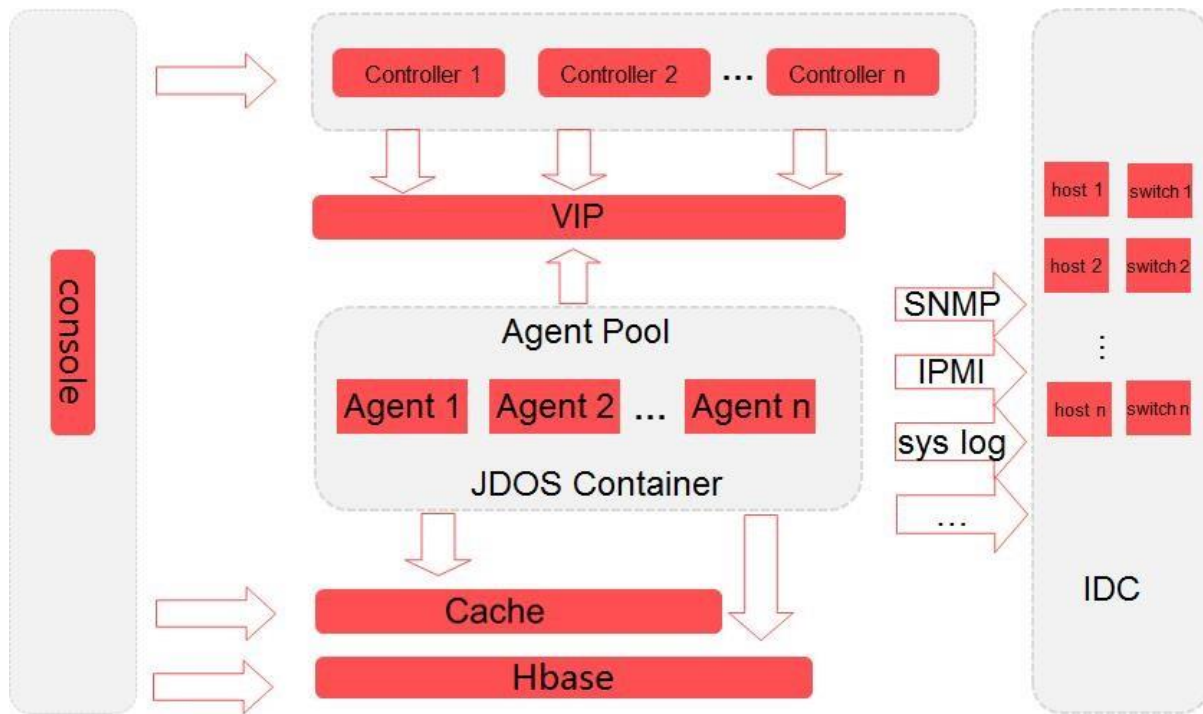
Below the cart, there are options to '去结账' (Check Out) and '去领券' (Get Coupons). A summary bar shows '已选择 1 件商品' (1 item selected), '总价: ¥5799.00' (Total Price: ¥5799.00), and '已节省: -¥200.00' (Savings: -¥200.00).

The '猜你喜欢' (You May Also Like) section features four recommended items:

- Pampers Baby Fresh XL Diapers:** 12-pack, ¥239.90.
- Gree (GREE) 1.5HP Inverter Air Conditioner:** ¥3049.00.
- Books: '分水岭' (The Great Divide) series:** ¥182.00.
- Gree (GREE) 1.5HP Inverter Air Conditioner:** ¥2999.00.

应用业务场景-MDC

- 统一监控平台
- 物理机5.8w, 容器近10w
- 用户2700+
- Metric 5大类, 31小类
- 日增量1.6 TB



应用业务场景-监控大屏

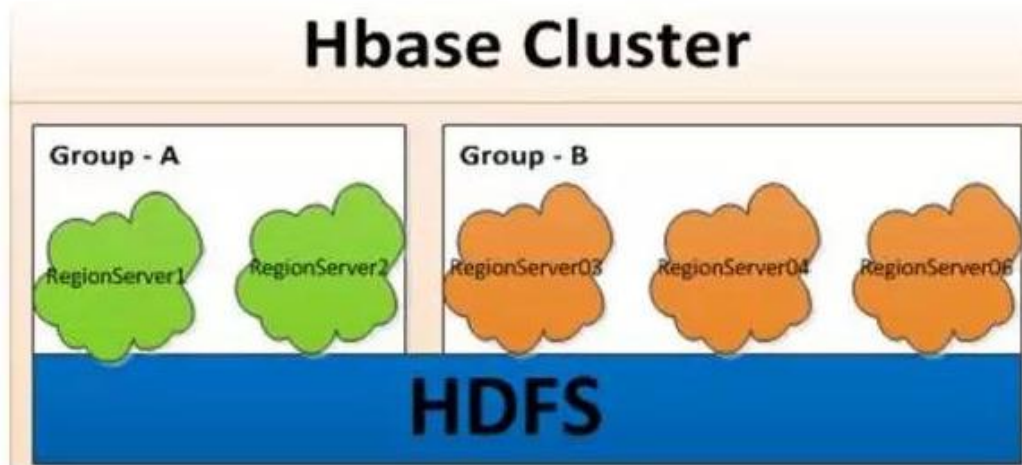


高可用改进

- RSGroup
- Security
- Replication
- Client
- Tuning
- Other

RS Group

- HBASE-6721 RSGROUP
- 多租户业务分级（独立、共享）
- 差异化配置管理
- 移动分组滚动重启
- 资源弹性管理
- 成本核算



RS Group

[Home](#)[Table Details](#)[Local Logs](#)[Log Level](#)[Debug Dump](#)[Metrics Dump](#)[HBase Configuration](#)

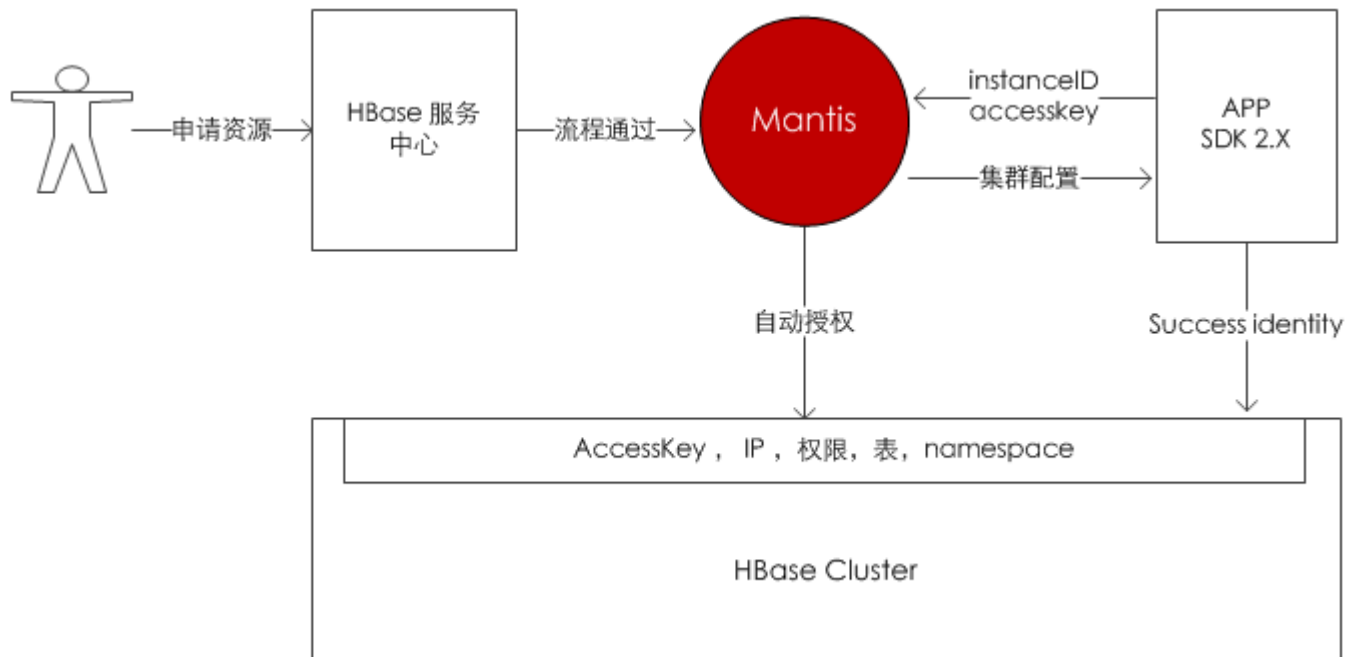
RS Group

Group	RegionServer Num	Table Num	Region Num	Load average	RequestPerSecond	StoreFile Size
pop_oplog	6	3	5584	930.67	18756	36971404
monitor	3	73	177	59.00	31859	730262
mdc	50	13	18484	369.68	29941	259529530
sd	30	35	9069	302.30	45132	78942185
friendsgo	3	29	663	221.00	866	4096335
ai_image	15	5	25	1.67	41	297
ptcp_jiaoyi_usertracer	30	11	25175	839.17	98560	172758503
tmp_qianyi	113	6	2881	25.50	12124	19880602
default	30	248	4782	159.40	40252	30802246
sr_mon	3	10	334	111.33	28093	1932907
fdm	20	105	5542	277.10	23239	26606976

Security

- Hadoop集群安全
 - JDHadoop 版本
 - Whitelist,Blacklist,Otherlist
 - SecurityPathLists
- HBase权限
 - 服务端：认证、授权， SASL+IP 实现
 - 客户端： InstanceID+AccessKey

HBase鉴权流程图



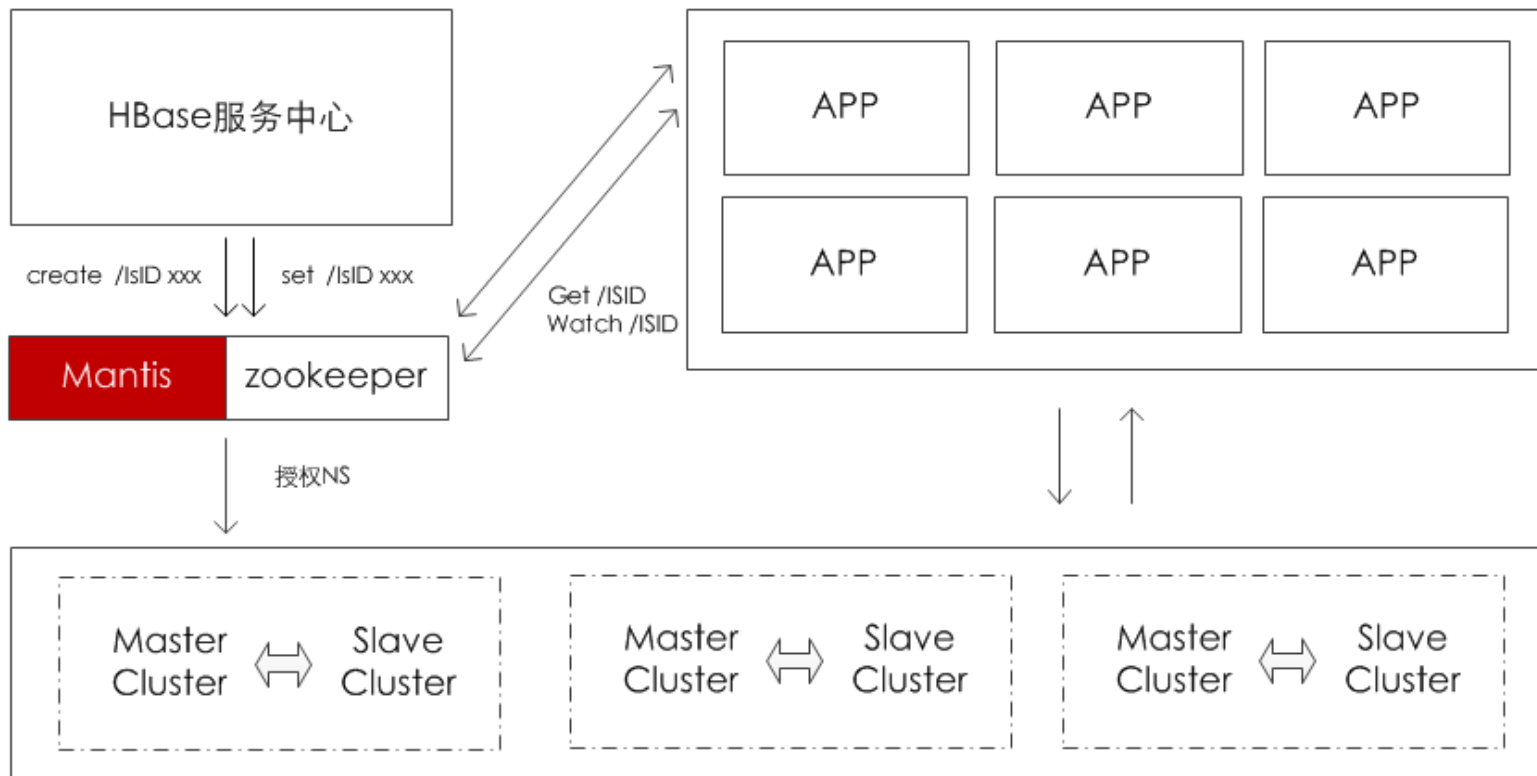
Replication

- HBASE-13153 Bulk Loaded HFile Replication
- HBASE-15809 Basic Replication WebUI
- HBASE-16447 Replication by namespaces config in peer
- HBASE-11393 Replication TableCfs should be a PB object rather than a string
- HBASE-17698 ReplicationEndpoint choosing sinks

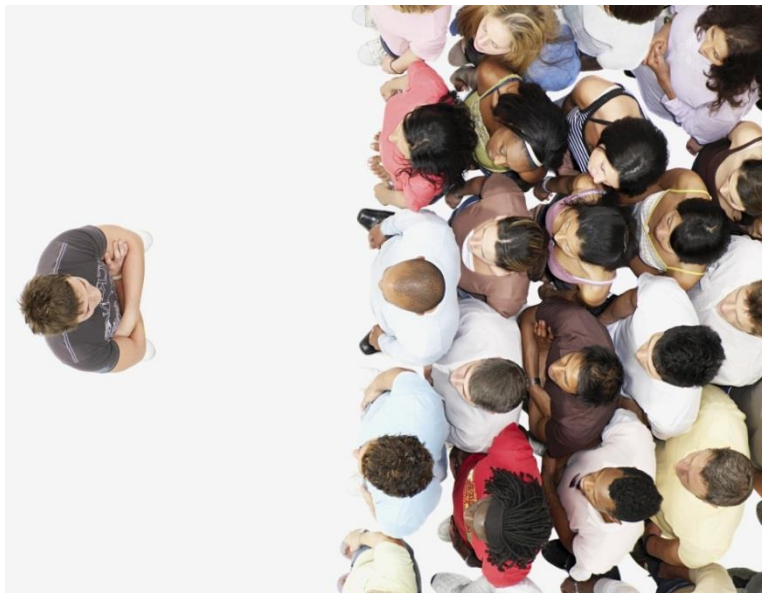
Client

- 一键切换
 - 主库异常，服务故障转移
 - 数据迁移
 - 秒级切换
 - 集群、分组、实例

一键切换-架构



一键切换-迁移



before



after

Tuning

- Compaction 优化
- Hlogs参数优化
- 堆外内存
- 读写队列分离
- GC参数优化
- Quota
- 分布式Log Replay

other

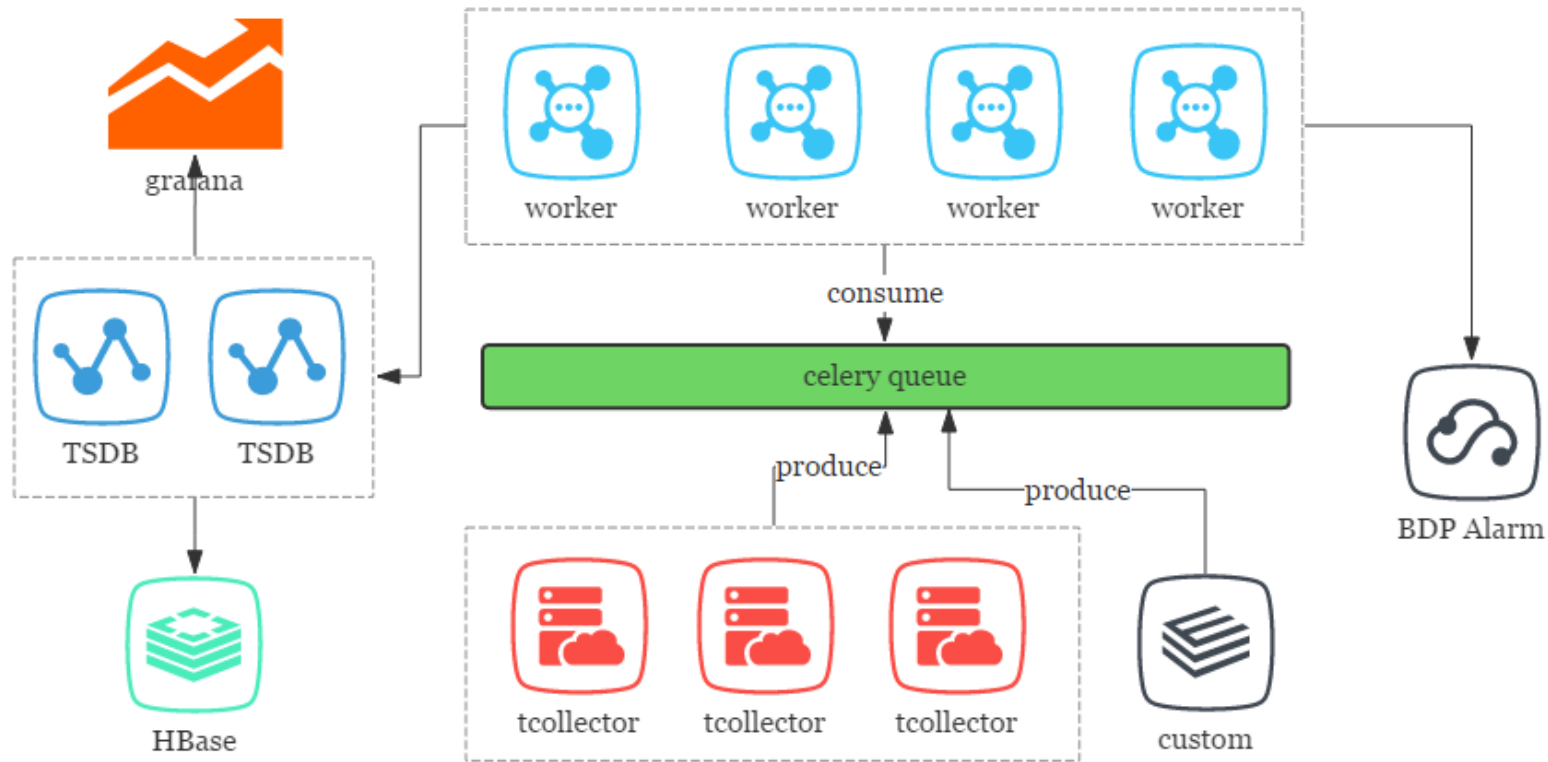
- HBASE-14644 Region in transition metric is broken
- HBASE-14963 Remove use of Guava Stopwatch from HBase client code
- HBASE-15319 clearJmxCache does not take effect actually

运维实践

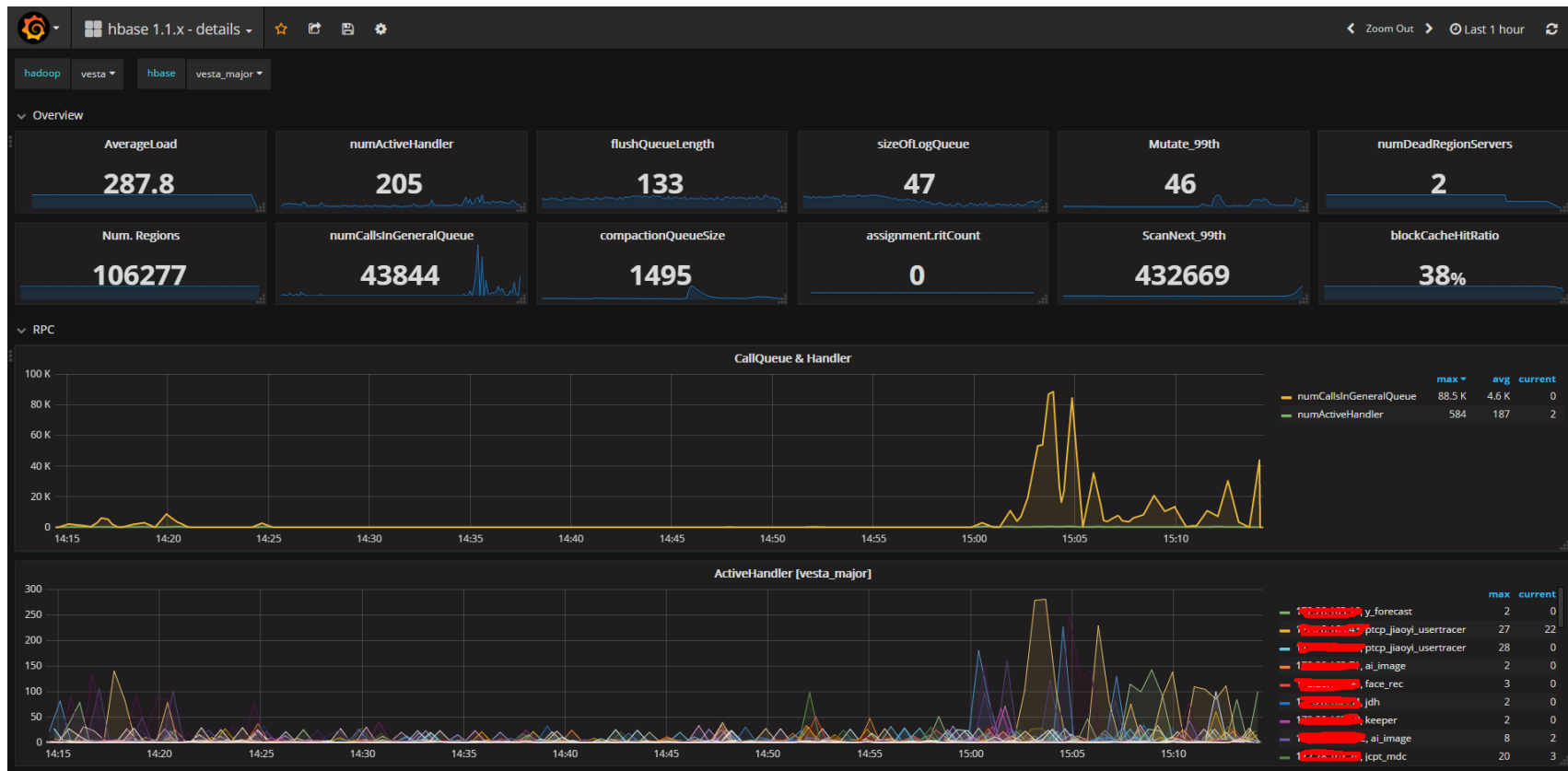
监控

- Mummut监控
 - 目标：集群jmx、日志、服务器运行指标等信息
 - 处理能力：秒级8Kw
 - 架构实现：tcollector + celery + redis + tsdb + grafana + bdp_phoenix
- 其它监控
 - ganglia、mdc、ops、ump
- 巡检
 - 基础服务、磁盘健康、swap、连接数、僵尸进程、集群配置

监控-Mammut



监控-Mammut

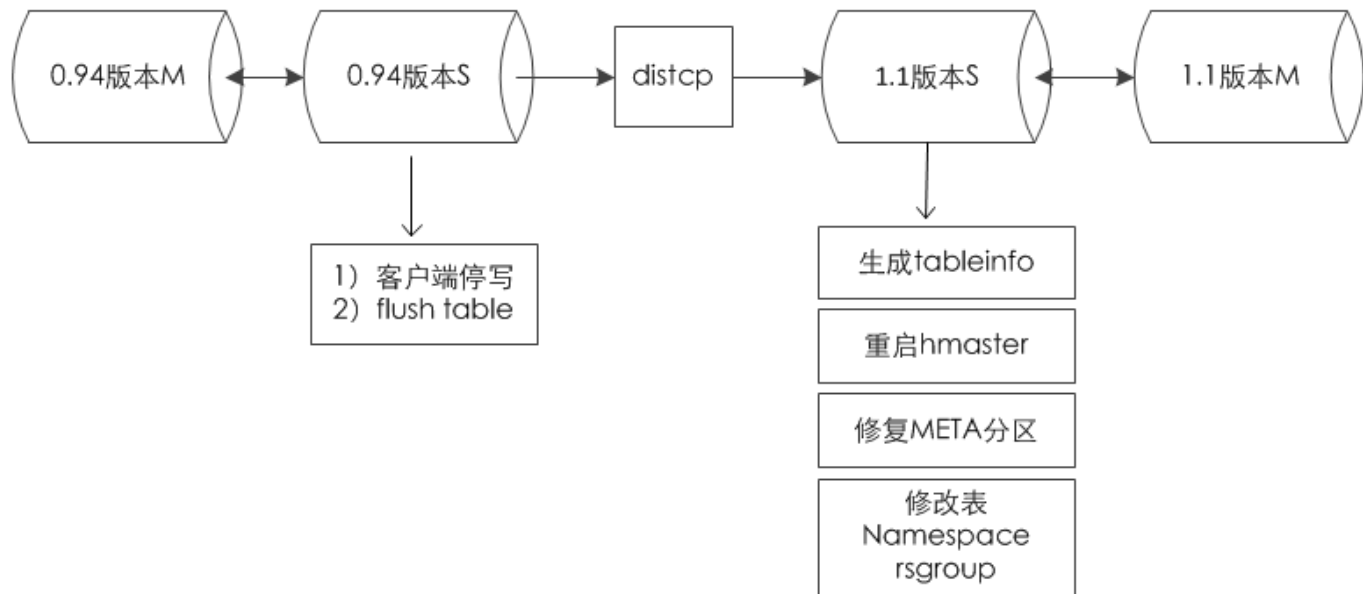


告警

- 报警方式：
 - 短信、邮件、咚咚
 - 监控中心7x24小时，方法性能调用次数、可用率
- Metric：
 - AverageLoad, sizeOfLogQueue, ritCount, DeadRegionServer, numActiveHandler, numCallGeneralQueue

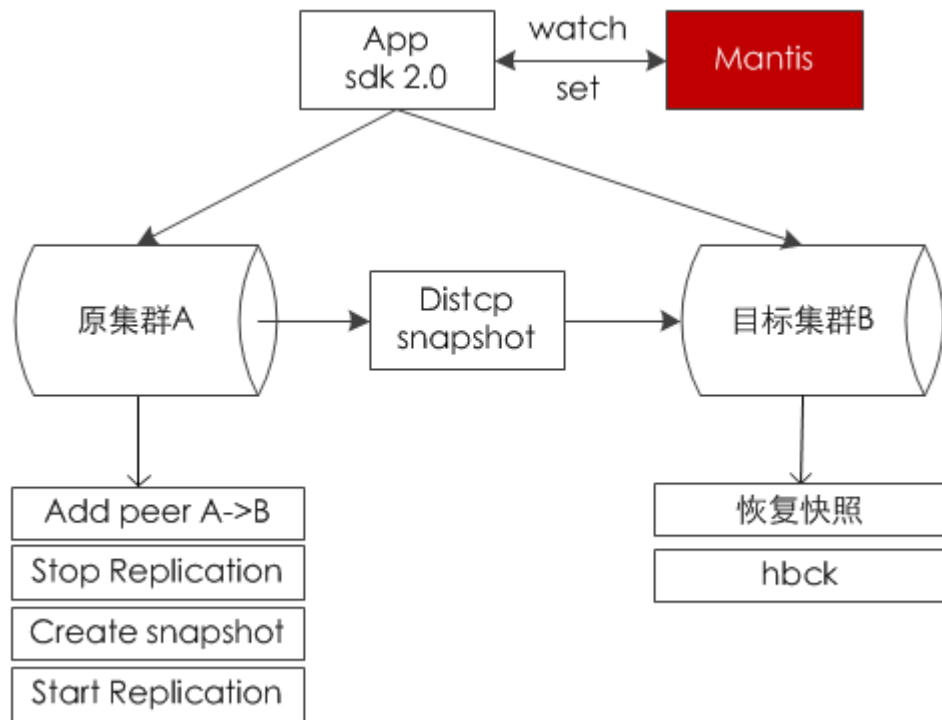
集群迁移

■ 0.94->1.1



集群迁移

- 1.1->1.1



管理平台

公告: [12.28] JDHBase使用手册, 点击查看

实例管理

实例管理

表管理

Phoenix管理

实例管理

表管理

关键字: 实例编号/项目名称

环境: 请选择

查询

高级搜索

系统级别: 请选择

集群: 请选择

版本: 请选择

新建实例

修改

实例下线

迁移/升级

推数授权

使用手册

问题反馈

<input type="checkbox"/>	实例编号	项目名称	集群	环境	版本	系统级别	状态	负责人	创建时间
<input checked="" type="checkbox"/>	SL10000000030...	██████████	adam	线上	1.0以上	0	已配置	██████████	2017-07-25 18:...
<input type="checkbox"/>	SL10000000030...	██████████	adam	线上	1.0以上	1	已配置	██████████	2017-07-25 14:...
<input type="checkbox"/>	SL10000000030...	██████████	lions	测试	1.0以上	0	已配置	██████████	2017-07-24 17:...
<input type="checkbox"/>	SL10000000030...	██████████	adam	线上	1.0以上	3	已配置	██████████	2017-07-24 15:...
<input type="checkbox"/>	SL10000000030...	██████████	adam	线上	1.0以上	0	已配置	██████████	2017-07-24 14:...
<input type="checkbox"/>	SL10000000030...	██████████	adam	线上	1.0以上	0	已配置	██████████	2017-07-20 18:...
<input type="checkbox"/>	SL10000000030...	██████████	K8S	预发布	1.0以上	0	已配置	██████████	2017-07-20 18:...

JDHBase生态

大数据平台BDP

JDHBase
服务中心

Plumber

流程中心

调度中心

UMP

推荐统

罗盘

JA

智能云

人脸识别

大屏

more
+

白条

京东到家

风控

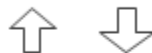
青龙

交易系统

资金流水

领航

jdhbase-client(SDK 2.x)



高可用容灾

自动分区

安全

一键主从切换

主从互备

sql/phoenix

kylin

多租户隔离

二级索引

协处理器

JDHBase

JDHadoop

运维平台

监控中心7x24小时

MDC/UMP/MAM

验收测试

自动部署

故障自动恢复

集群管理/分组

配置管理

白名单、绿色通道

基础设施

普通物理机

混合ssd+sas

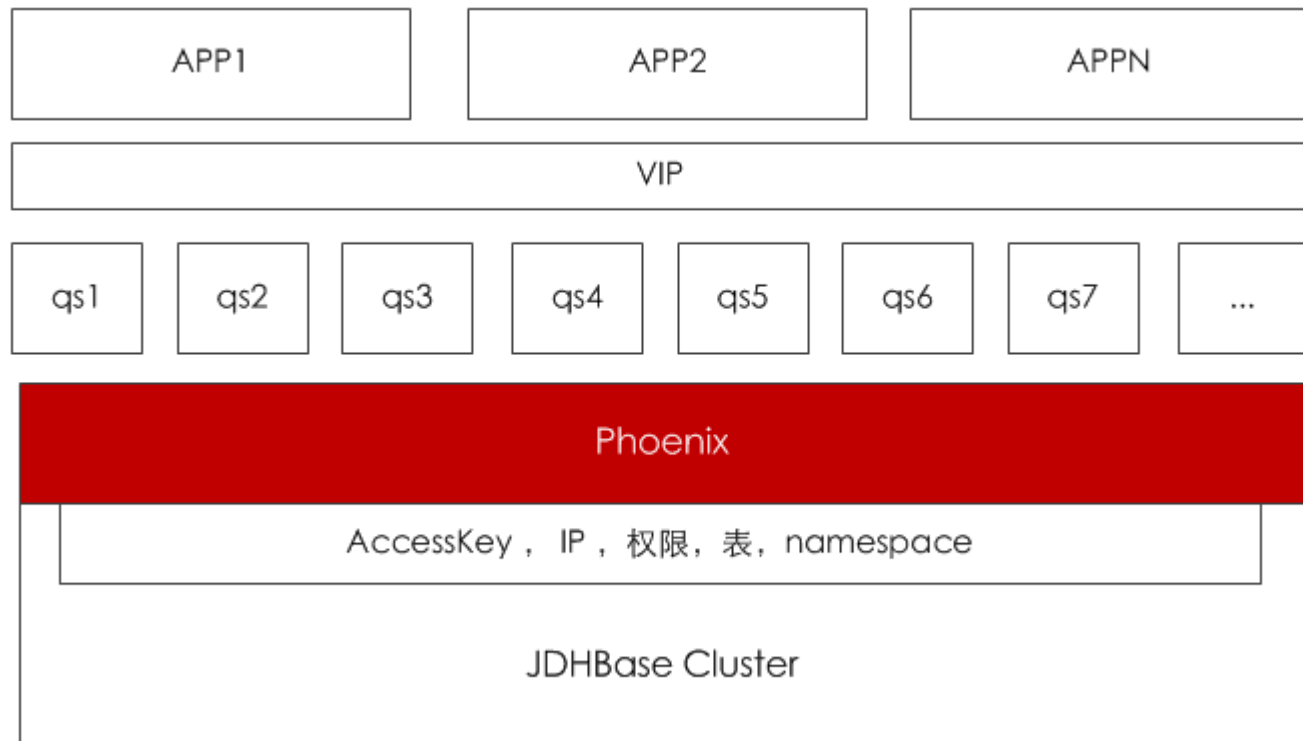
K8s

未来展望

Phoenix

- 版本：4.10
- Patch
 - PHOENIX-3757 System mutex table not being created in SYSTEM namespace when namespace mapping is enabled
 - PHOENIX-3756 Users lacking ADMIN on 'SYSTEM' HBase namespace can't connect to Phoenix
- 改进
 - 适配jdhbase认证机制
 - 支持多用户同时登陆，改造cache用户逻辑

Phoenix



容器化部署

- k8s, 预发布环境

感谢！

我们联系方式， hbase@jd.com

