数领先机 智赢未来
第九届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2018

基于Spark、NoSQL的实时数据处理实践

张学敏

DTCC 2018
2018.05.10 – 12 北京国际会议中心
ChinaUnix ITPUB

概要

- 关于我（们）
- 数据、流程和架构
- 业务诉求
- 技术和架构
- 挑战和方案
- 未来展望

DTCC2018

概要

- 关于我（们）
- 数据、流程和架构
- 业务诉求
- 技术和架构
- 挑战和方案
- 未来展望

DTCC2018

关于我（们）

我

- 拥有7年技术实战经验，先后就职于锐安、新浪微博、TalkingData，曾任新浪研发中心大数据TeamLeader，2015年加入TD数据中心，负责公司的数据处理、数据服务等工作。

数据中心/治理

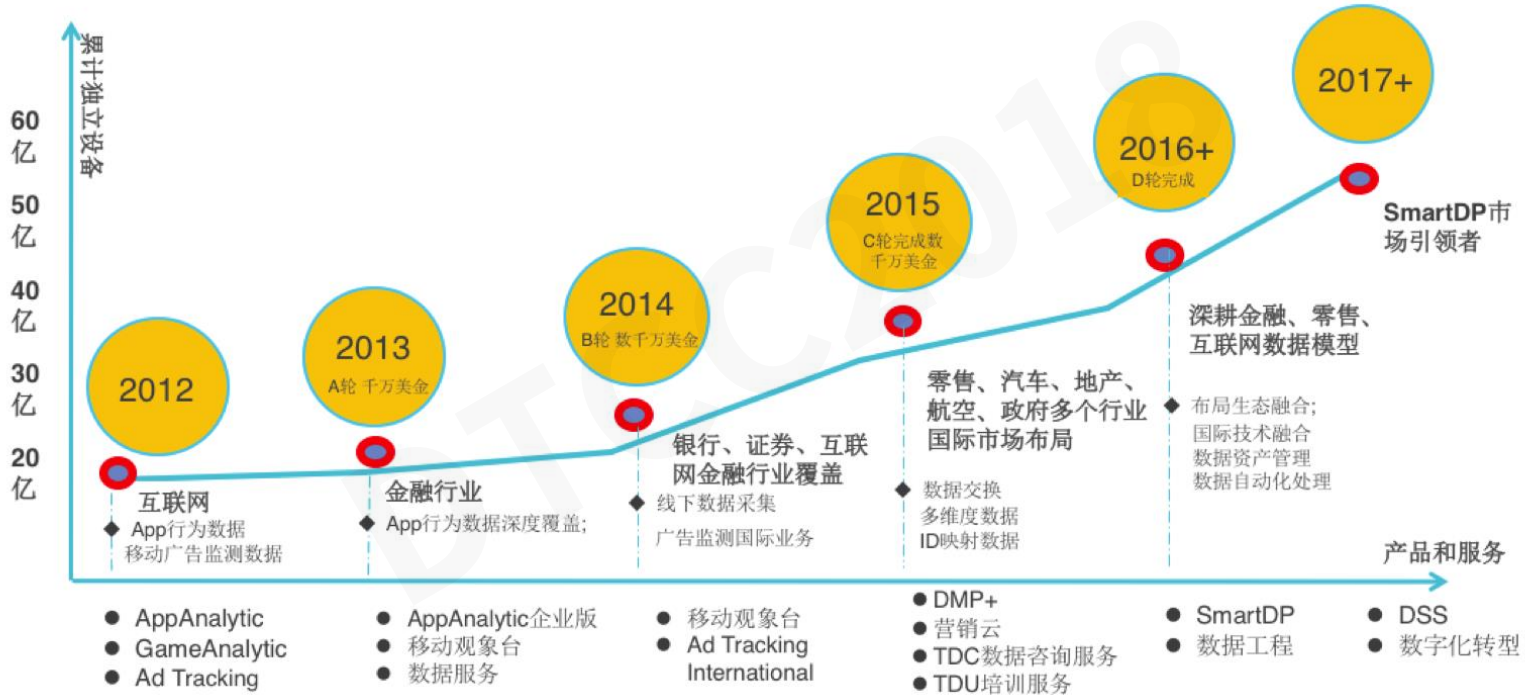
- 主要负责公司数据标准的制定、实施、监督，数据质量体系建设、质量审核及数据资产和数据服务的建设、管理等工作。

TalkingData

- 成立于2011年，是国内领先的独立第三方移动数据服务平台。TD一直致力于数据的深耕与数据价值的挖掘，从数据的采集、处理到数据的分析，再到数据的应用与咨询，TD已经形成了一套以“智能数据平台（SmartDP）”为主的完整数据应用体系；构筑了一套以数据商业化平台、数据服务平台，及数据合作平台为核心的数据生态。目前，TD的平均月活跃用户为7亿，为超过12万款移动应用，以及10万应用开发者提供服务。覆盖的客户主要为金融、地产、快消、零售、出行、政府等行业中的领军企业，连续三年实现业务的三倍快速增长。

关于我（们）

TalkingData发展历程



DTCC2018
DTCC 2018 数领先机 智赢未来
IT168 ChinaUnix ITPUB

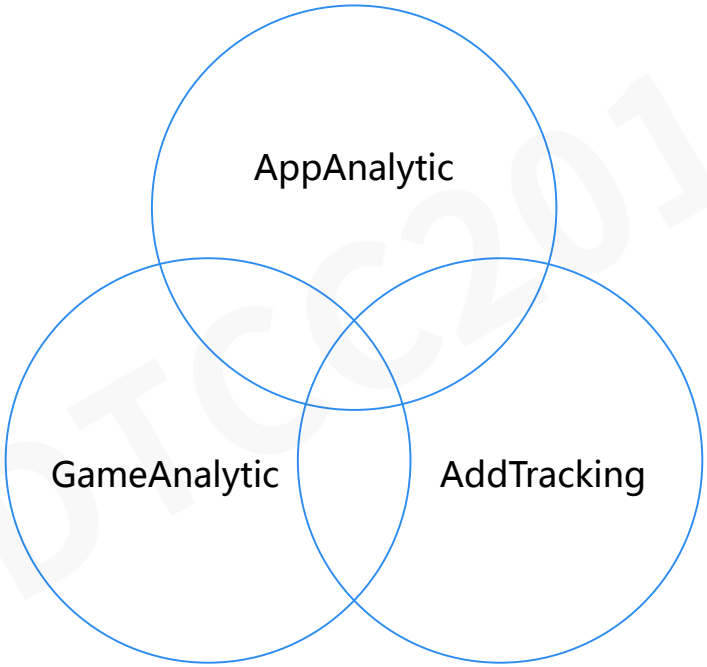
概要

- 关于我（们）
- ➢数据、流程和架构
- 业务诉求
- 技术和架构
- 挑战和方案
- 未来展望

DTCC2018

数据流程和架构

- 主要数据来源



DTCC2018

数据流程和架构

- 数据内容

设备信息	业务信息	上下文信息
• 设备ID • 设备软件信息 • 设备硬件信息	• 业务事件 • 会话信息 • 行为状态	• 网络 • 位置 • 传感器

DTCC2018

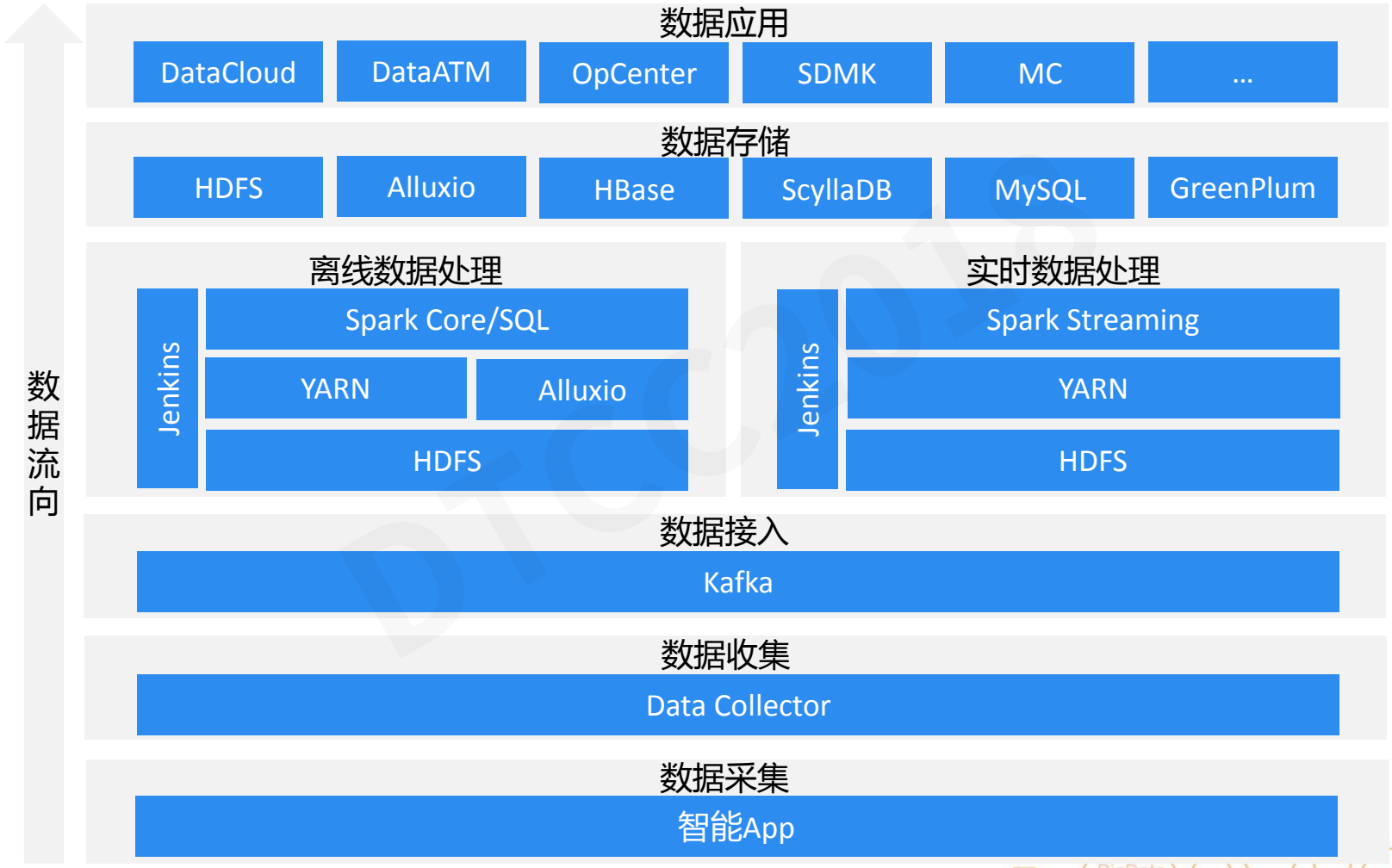
数据流程和架构

- 数据体量

活跃设备	事件(日志)数	存储大小
2.5/6.5亿+ 日/月活跃智能设备	370亿 每天处理事件	17T+ 每天新增日志量

DTCC2018

数据流程和架构


数据流向
数据应用
DataCloud
DataATM
OpCenter
SDMK
MC
...
数据存储
HDFS
Alluxio
HBase
ScyllaDB
MySQL
GreenPlum
离线数据处理
Jenkins
Spark Core/SQL
YARN
Alluxio
HDFS
实时数据处理
Jenkins
Spark Streaming
YARN
HDFS
数据接入
Kafka
数据收集
Data Collector
数据采集
智能App

概要

- 关于我（们）
- 数据、流程和架构
- ➢业务诉求
- 技术和架构
- 挑战和方案
- 未来展望

DTCC2018

业务诉求

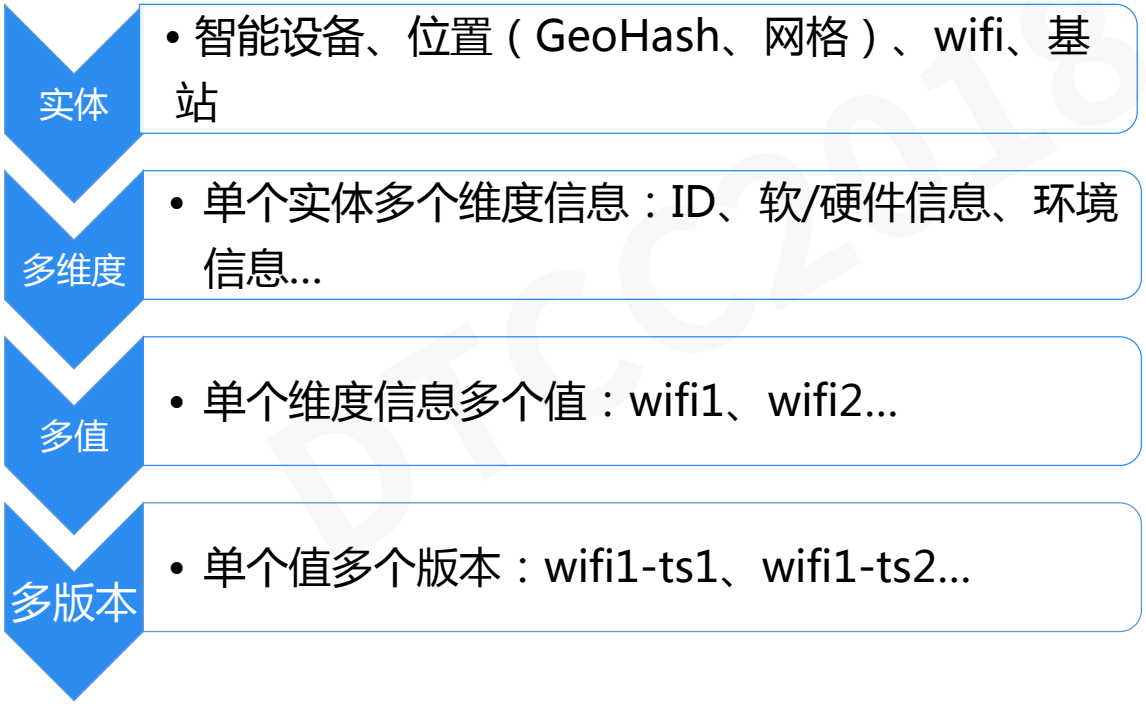
- 新的数据处理、服务诉求

数据修正	• 事件数据时间修正（离线计算：数据到达延迟）
时序数据	• 面向实体或者指标的时序数据需求（离线计算：时间断面）
实时处理	• 面向实体，实时数据处理 • 位置数据丰富
实时查询	• 面向实体，多维度、多值、多版本的查询

DTCC2018

业务诉求

- 面向实体，多维度、多值、多版本的查询



举个例子

tdid1，imei1，wifi1，timestamp1
tdid1，imei1，wifi1，timestamp2
tdid1，imei1，wifi1，timestamp3
tdid1，imei2，wifi2，timestamp4
tdid1，imei1，wifi2，timestamp5

DTCC2018
数领先机 智赢未来
IT168 ChinaUnix ITPUB

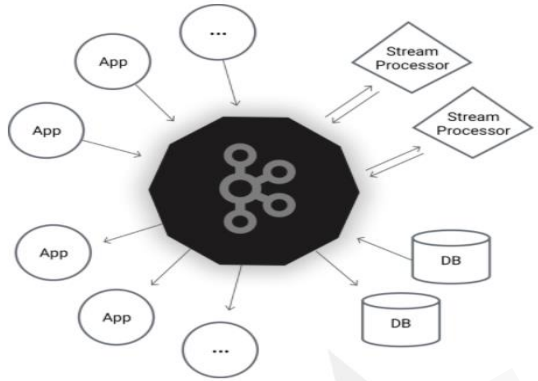
概要

- 关于我（们）
- 数据、流程和架构
- 业务诉求
- ➢技术和架构
- 挑战和方案
- 未来展望

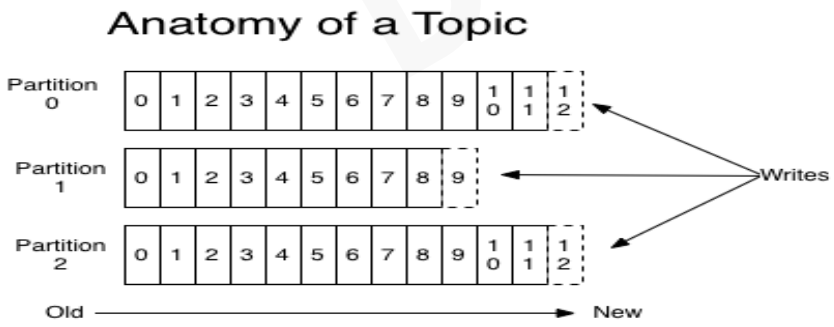
DTCC2018

技术和架构

- 数据接入--Kafka



Kafka® is used for building real-time data pipelines and streaming apps. It is horizontally scalable, fault-tolerant, wicked fast, and runs in production in thousands of companies.



选型原因

- 可扩展
- 高吞吐
- 高容错
- 较成熟
- 与其他组件好集成

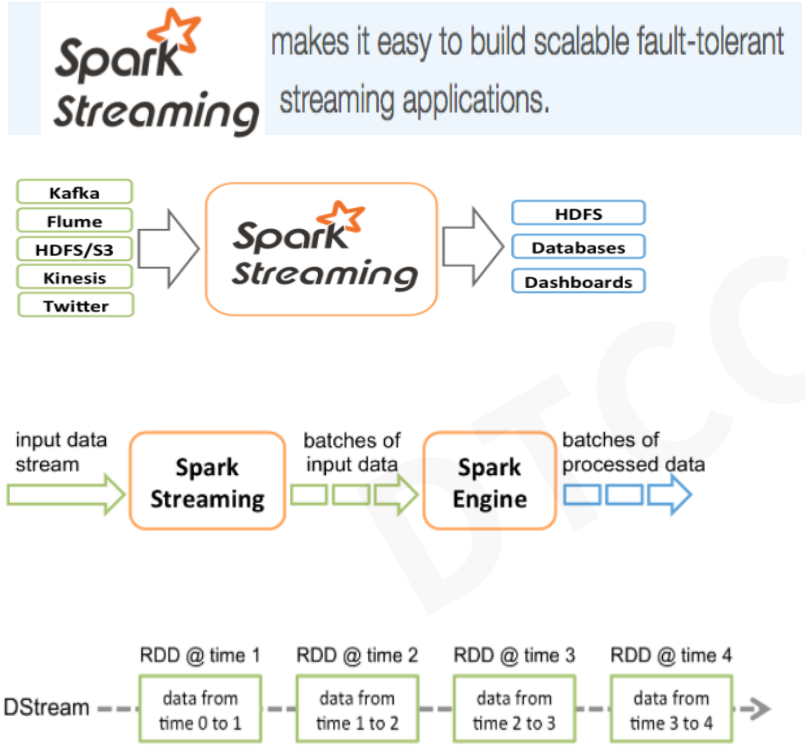
主要劣势

- 不保证数据有序
- 管理工具不够完善

DTCC2018
DTCC 2018 数领先机 智赢未来
IT168.com ChinaUnix ITPUB

技术和架构

- 实时数据处理--Spark Streaming



选型原因

- 可扩展
- 高吞吐
- 高容错
- 支持窗口函数
- 支持SQL
- 技术统一
- 调度、资源管理统一

主要劣势

- 微批、延迟高

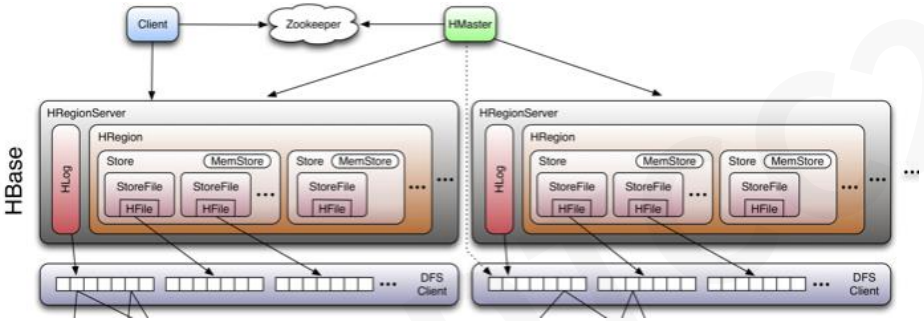
DTCC2018

技术和架构

- 数据存储--HBase

APACHE HBASE

Apache HBase™ is the Hadoop database, a distributed, scalable, big data store.



Row Key	Time Stamp	ColumnFamily contents	ColumnFamily anchor	ColumnFamily people
"com.cnn.www"	t9		anchor:cnnsi.com = "CNN"	
"com.cnn.www"	t8		anchor:my.look.ca = "CNN.com"	
"com.cnn.www"	t6	contents:html = "<html>…"		

选型原因

- 可扩展
- 高吞吐
- 高容错
- 较成熟
- 低延迟 (vs HDFS)
- Free Schema

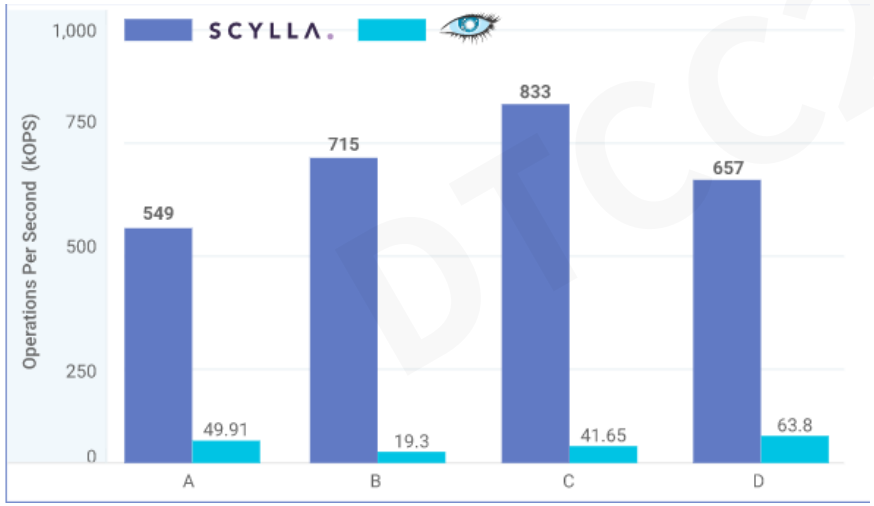
主要劣势

- 运维成本相对较高(compact、split、flush)
- 延迟不稳定(GC、缓存命中)

DTCC 2018 数领先机 智赢未来
IT168.com ChinaUnix ITPUB

技术和架构

- 数据存储--ScyllaDB



选型原因

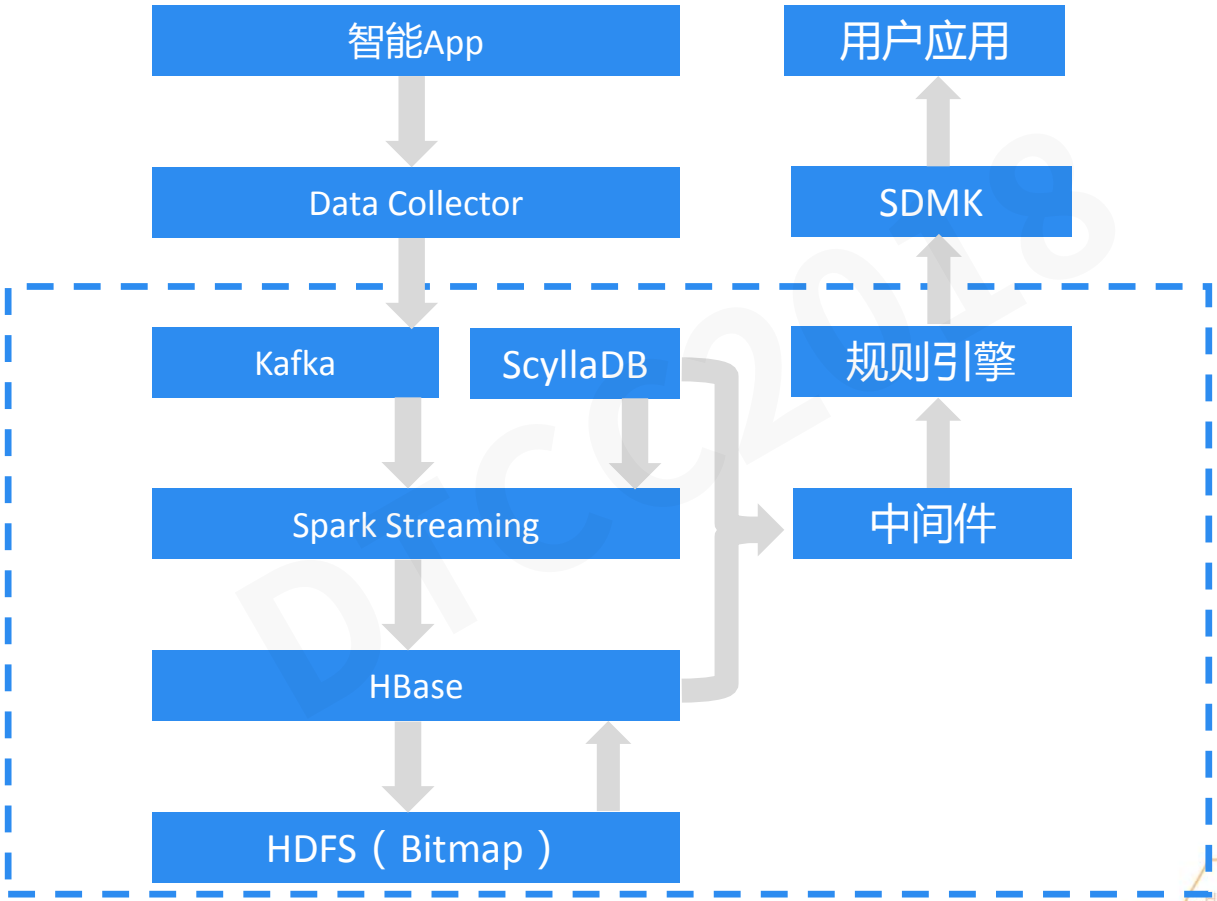
- 可扩展
- 高吞吐
- 高容错
- 低延迟
- 延迟稳定

主要劣势

- 项目较新
- Bug、使用坑多

DTCC2018

技术和架构


智能App
Data Collector
Kafka
ScyllaDB
Spark Streaming
HBase
HDFS（Bitmap）
用户应用
SDMK
规则引擎
中间件

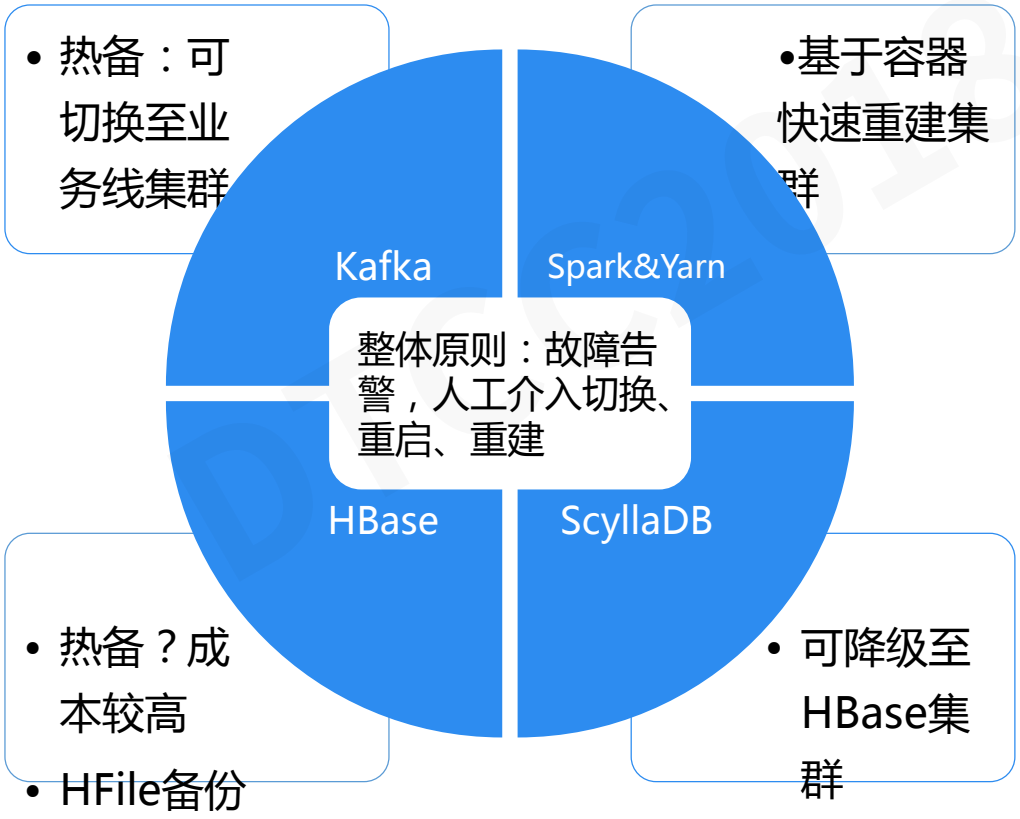
概要

- 关于我（们）
- 数据、流程和架构
- 业务诉求
- 技术和架构
- ➢挑战和方案
- 未来展望

DTCC2018

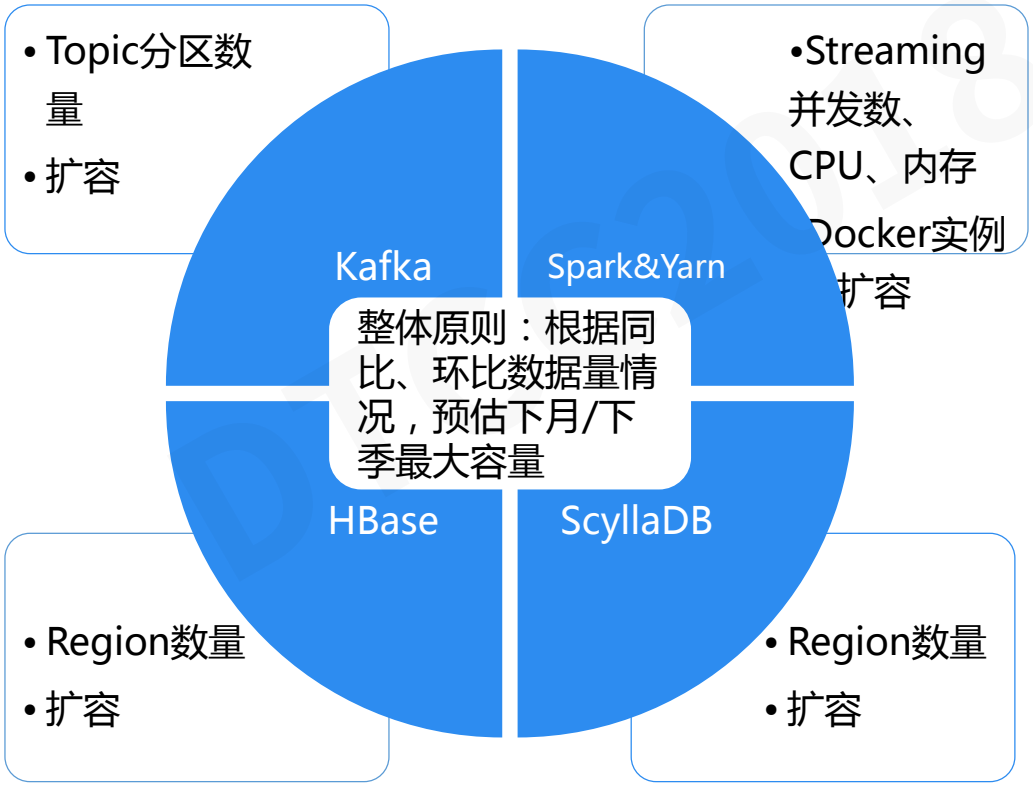
挑战和方案

- 服务稳定性


• 热备：可切换至业务线集群
•基于容器快速重建集群
Kafka
Spark&Yarn
整体原则：故障告警，人工介入切换、重启、重建
HBase
ScyllaDB
• 热备？成本较高
• HFile备份
• 可降级至HBase集群

挑战和方案

- 容量预估


• Topic分区数量
• 扩容
•Streaming并发数、CPU、内存
Docker实例
扩容
Kafka
Spark&Yarn
整体原则：根据同比、环比数据量情况，预估下月/下季最大容量
HBase
ScyllaDB
• Region数量
• 扩容
• Region数量
• 扩容

挑战和方案

- 数据正确性&一致性

没有银弹

根据业务定方案

- 使用HBase的version特性去重，保证数据没有重复
- Spark Streaming + Kafka保证“at last once”

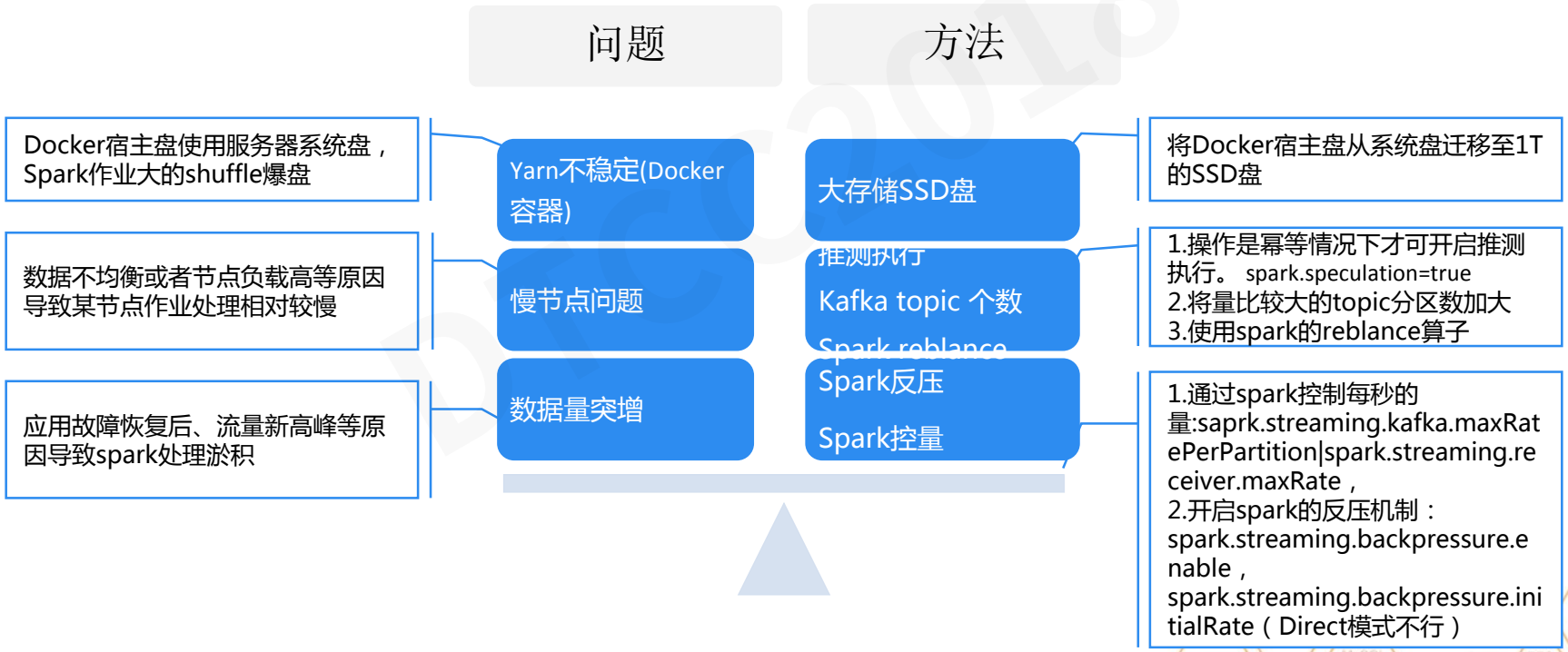
一些建议

- 让批次幂等：
  - 针对每个partition的数据产生一个uniqueID，只有这个uniqueID相关所有数据都被完全计算，才算成功，否则失败回滚。如果重复执行到uniqueID，会跳过。

DTCC2018

挑战和方案

- Spark Streaming--性能稳定性



挑战和方案

- Spark Streaming--优雅的停止

发送SIGTERM信号

- 发送SiGTERM信号给作业的Driver
  1. 设置spark.streaming.stopGracefullyOnShutdown为true；
  2. 在Spark UI上找到Drive所在节点；
  3. 登陆节点找到Driver进程ID；
  4. 执行*kill -SIGTERM <AM-PID>*。

其他事件触发

- 在独立线程事件触发调用ssc.stop(true, true)
  1. 定义事件，比如HDFS上标识文件、监听socket、启动RESTfull服务等，并调用ssc.stop(true, true)；
  2. 触发事件。

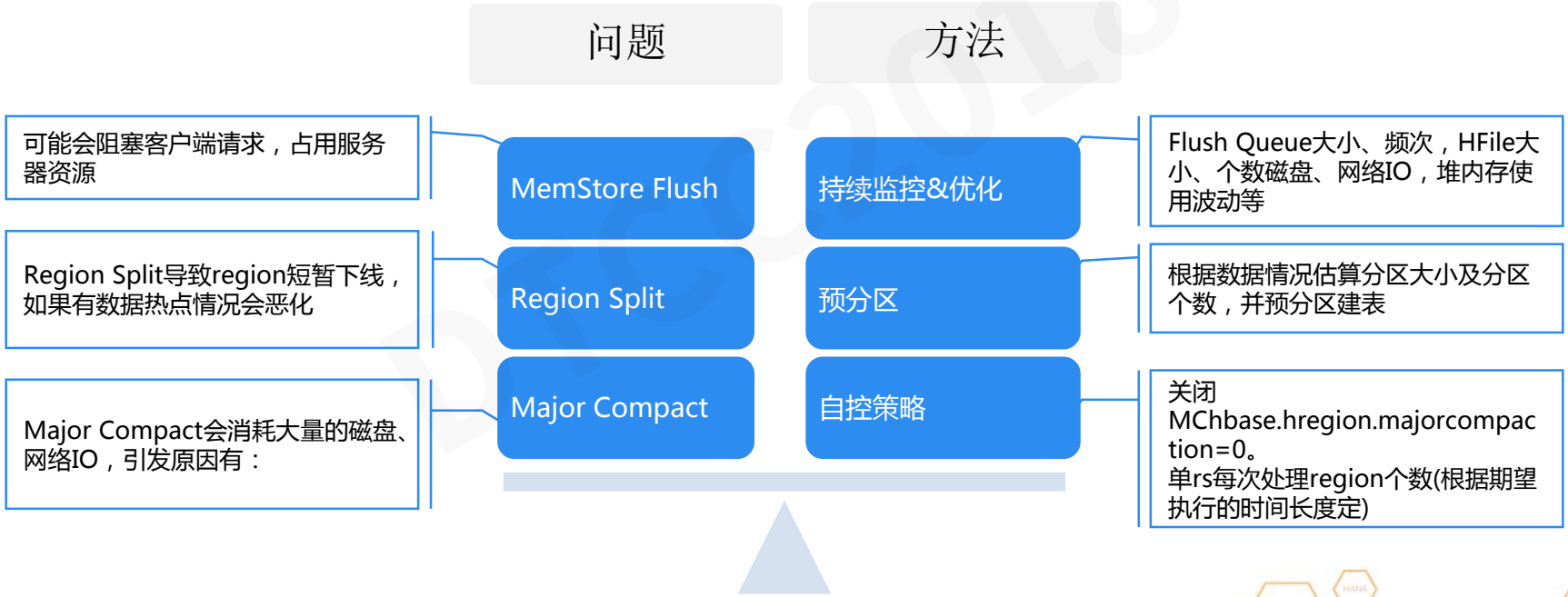
挑战和方案

- Spark Streaming--其他建议
  1. spark.streaming.kafka.maxRetries
  2. spark.yarn.maxAppAttempts
  3. spark.yarn.am.attemptFailuresValidityInterval
  4. spark.yarn.max.executor.failures
  5. spark.yarn.executor.failuresValidityInterval

DTCC2018

挑战和方案

- HBase--性能稳定性



挑战和方案

- HBase--其他建议

1. hbase.regionserver.thread.compaction.large/small
2. hbase.hstore.flusher.count
3. hbase.regionserver.optionalcacheflushinterva

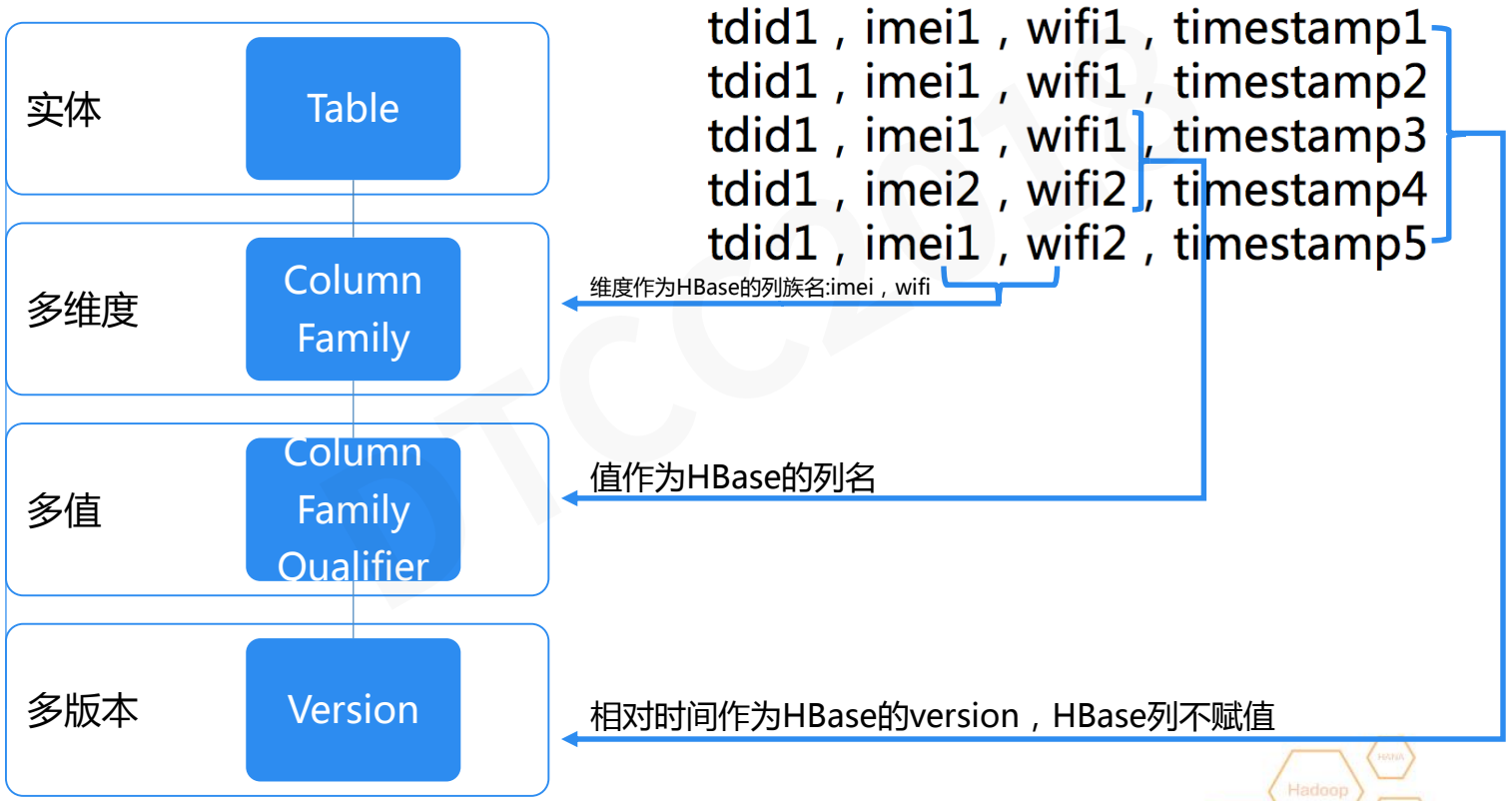
4. hbase.hregion.memstore.flush.size
5. hbase.hregion.memstore.block.multiplier
6. hbase.hregion.percolumnfamilyflush.size.lower.bound

7. hbase.regionserver.global.memstore.size
8. hfile.block.cache.size
9. hbase.regionserver.global.memstore.size.lower.limit
   （hbase.regionserver.global.memstore.lowerLimit）

DTCC2018

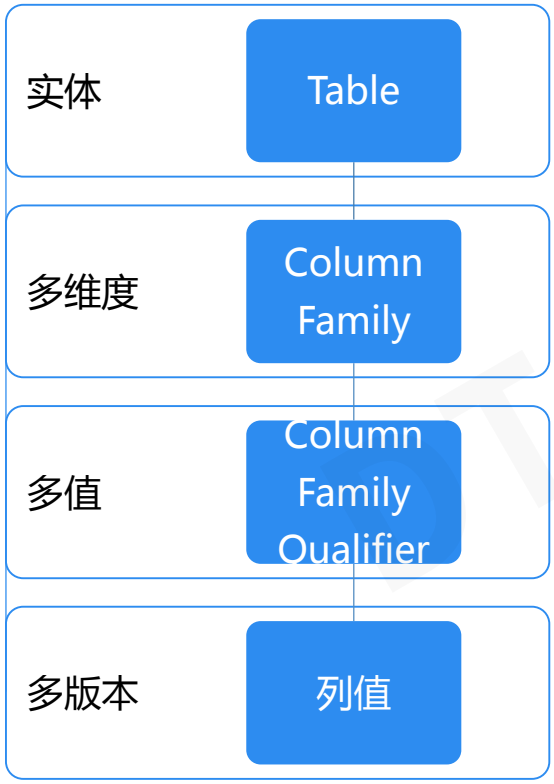
挑战和方案

- 面向实体 VS HBase数据模型



挑战和方案

- 大时间窗口查询及整体存储量



tdid1 , imei1 , wifi1 , timestamp1
tdid1 , imei1 , wifi1 , timestamp2
tdid1 , imei1 , wifi1 , timestamp3
tdid1 , imei2 , wifi2 , timestamp4
tdid1 , imei1 , wifi2 , timestamp5

值稀疏，出现频次高

tdid1 , imei1 , bitmap[1,2,3,5]
tdid1 , imei2 , bitmap[4]
tdid1 , wifi1 , bitmap[1,2,3]
tdid1 , wifi2 , bitmap[4,5]

bitmap字节作为HBase的列值

概要

- 关于我（们）
- 数据、流程和架构
- 业务诉求
- 技术和架构
- 挑战和方案
- 未来展望

DTCC2018

未来展望

- Docker挂载多盘/多版本Image
- Spark Streaming 日志收集
- Executor JVM监控
- HBase Region Replicas

DTCC2018

THANKS


HANA
Hadoop
SQL
MySQL
DB2
BigDate
Spark
SQL Server
NewSQL
Oracle
MPP
NoSQL
OLAP

讲师申请
联系电话（微信号）：18612470168

关注“ITPUB”更多
技术干货等你来拿~

与百度外卖、京东、魅族等先后合作系列分享活动

微学堂

让学习更简单

微学堂是以ChinaUnix、ITPUB所组建的微信群为载体，定期邀请嘉宾对热点话题、技术难题、新产品发布等进行移动端的在线直播活动。

截至目前，累计举办活动期数60+，参与人次40000+。

ITPUB学院

ITPUB学院是盛拓传媒IT168企业事业部（ITPUB）旗下
企业级在线学习咨询平台
历经18年技术社区平台发展
汇聚5000万技术用户
紧随企业一线IT技术需求
打造全方式技术培训与技术咨询服务
提供包括企业应用方案培训咨询（包括企业内训）
个人实战技能培训（包括认证培训）
在内的全方位IT技术培训咨询服务

ITPUB学院讲师均来自于企业
一些工程师、架构师、技术经理和CTO
大会演讲专家1800+
社区版主和博客专家500+

培训特色

无限次免费播放
随时随地在线观看
碎片化时间集中学习
聚焦知识点详细解读
讲师在线答疑
强大的技术人脉圈

八大课程体系

基础架构设计与建设
大数据平台
应用架构设计与开发
系统运维与数据库
传统企业数字化转型
人工智能
区块链
移动开发与SEO

联系我们

联系人：黄老师
电　话：010-59127187
邮　箱：edu@itpub.net
网　址：edu.itpub.net
培训微信号：18500940168