移动应用分析平台大数据系统实践

以友盟移动应用分析平台为例

DTCC

2015中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2015

大数据技术探索和价值发现

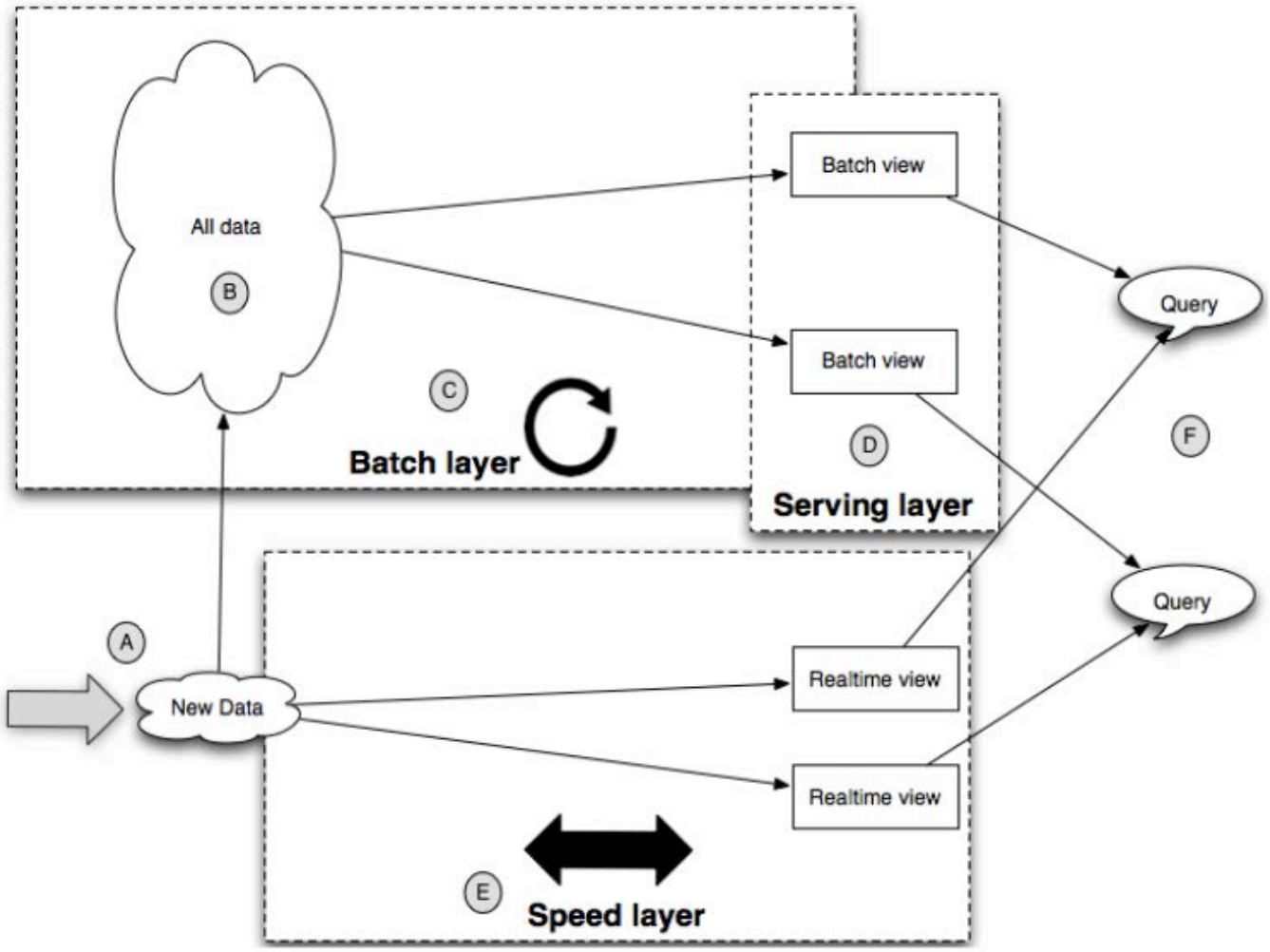
吴磊 wl@umeng.com
友盟数据平台架构师
2015.04.18

数据是移动互联网的主旋律

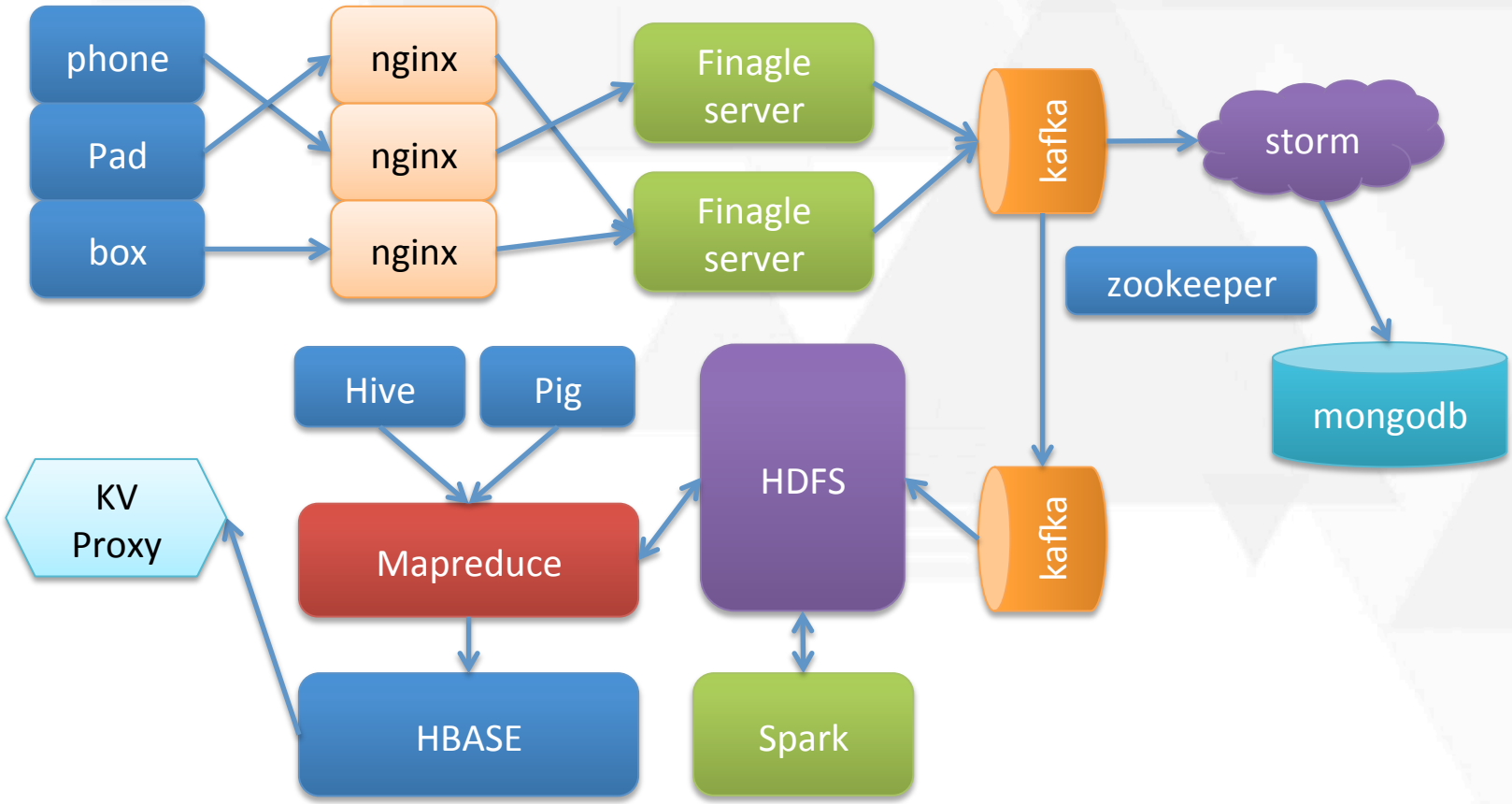
关于友盟移动分析平台

- 成立于2010年4月
- 目前涵盖52万app
- 目前处理的数据接近2PB
- 每天处理： 6.7 Billion Sessions
- 实时处理： 100k QPS
- 离线处理： 800+常规任务

基本架构思路

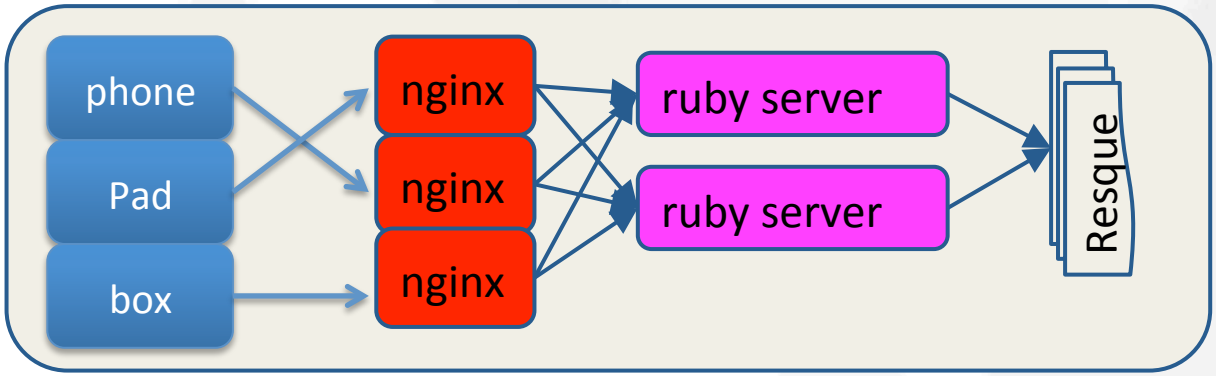

All data
B
C
Batch layer
Batch view
Batch view
D
Serving layer
Query
Query
F
A
New Data
Realtime view
Realtime view
E
Speed layer

平台基本架构


phone
Pad
box
nginx
nginx
nginx
Finagle server
Finagle server
kafka
storm
zookeeper
mongodb
Hive
Pig
KV Proxy
Mapreduce
HDFS
kafka
HBASE
Spark

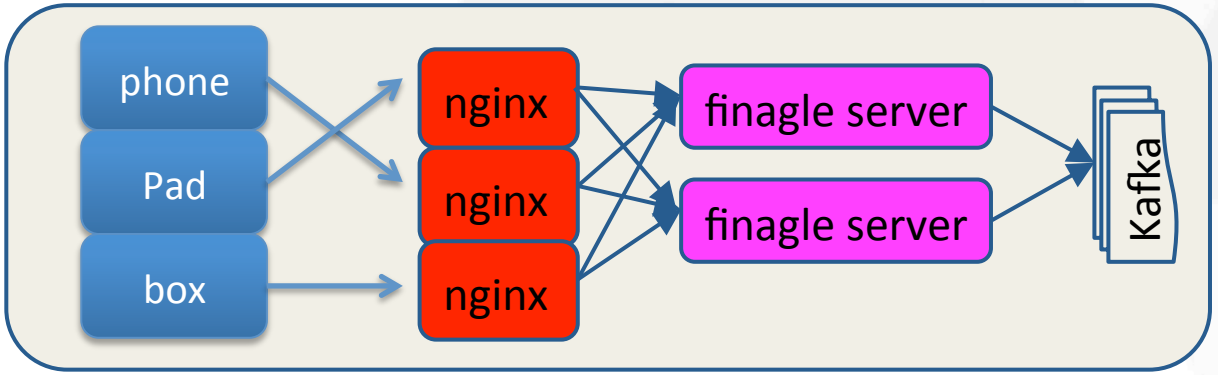
数据处理流程

数据采集


phone
Pad
box
nginx
nginx
nginx
ruby server
ruby server
Resque

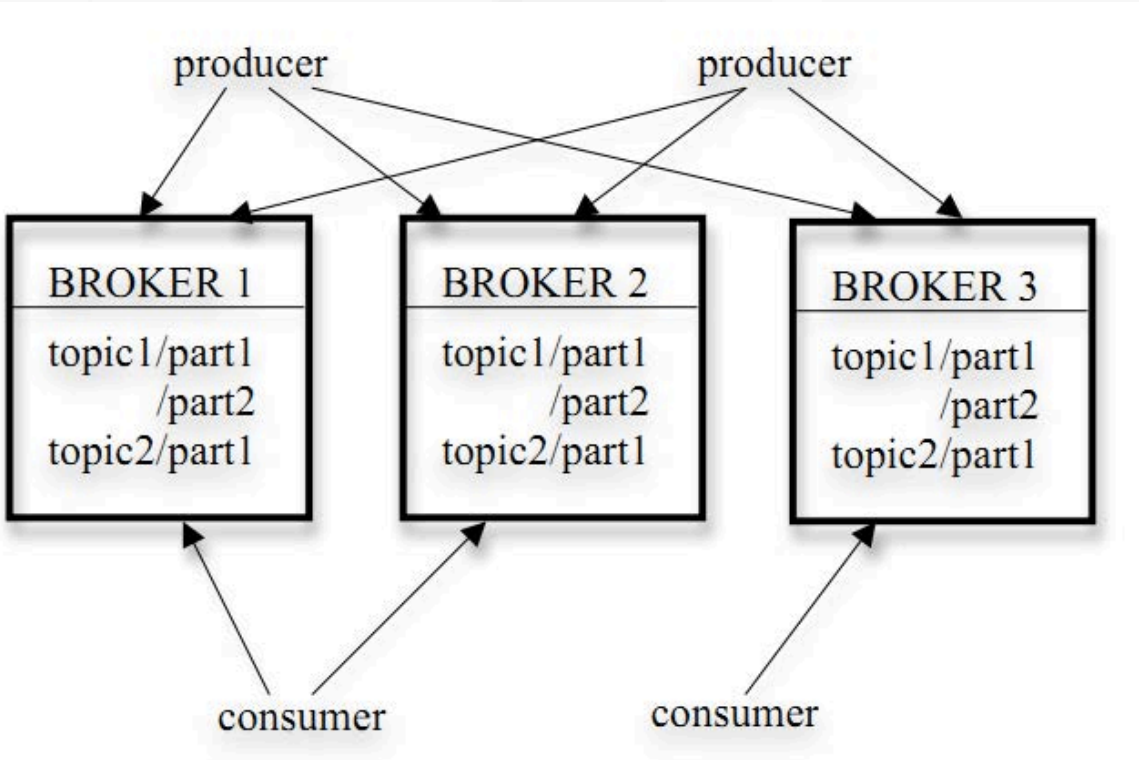
之前

现在


phone
Pad
box
nginx
nginx
nginx
finagle server
finagle server
Kafka

数据传输

- 数据总线(kafka)
  - 分布式
  - 吞吐量
  - 持久化
  - Mirror
  - Storm
  - MapReduce



数据处理

- 实时
  - Storm
- 离线
  - Map-Reduce 计算模型
  - 二进制数据格式
    - ElepantBrid
    - ProtoBuf
    - Compress

数据存储

- 在线数据存储
  - MongoDB
    - 数据规模(TB)
    - **数据特性**
- 离线数据存储
  - HDFS 存储
    - 数据规模(PB)
    - 二进制数据模型 (ElephantBird)
    - 数据压缩 (LZO, LZMA)
  - Hbase 存储
    - 数据规模(10TB)
    - 数据skew的影响
    - **数据聚合**
    - 随机读的需求
- 数据缓存
  - Redis

数据存储

- Hbase使用的经验
  - Data skew
  - 随机读的优化
  - 变随机读为顺序读
  - 异步访问
  - **数据聚合**
  - RowKey设计
  - 参数调优
  - 表预先切分
  - 客户端使用参数调优
  - **大批写使用**Bulk Load
  - 中间数据和小表使用文件更优
- Hbase使用的教训
  - 重视运维
  - 关注官方动态
  - 谨慎使用新特性

数据分析

- 数据统计分析
  - Storm实时计算 + MR离线计算
- 深度挖掘
  - Pig + Hive
- 分类聚类
  - MR + Spark
- 预测建模
  - MR + Spark

数据分析

- Hadoop 使用经验
  - Map Reduce 槽位混用
  - 磁盘空间优先的调度策略
  - 任务运行内存限制及调整，内存动态计算
  - 高比例的压缩算法
- Hadoop 使用的教训
  - 对于迭代式计算的支持
  - MR算法的僵化(二次排序，各种Join)

数据分析

- Pig 篇
  - Pig 应用场景
    - QA 测试
    - 数据深度挖掘
    - 聚类分类
  - Pig 优势
  - Pig 局限

数据分析

- Hive 篇
  - Hive 应用场景
    - 报表生产
  - Hive 优势
    - SQL
  - Hive 局限
    - UDF

数据分析

- SPARK篇
  - Spark 应用场景
    - 深度学习
  - Spark 优势
    - 更多的抽象模型
    - Mllib库
    - Spark SQL
    - 强大的语言支持
  - Spark 局限
    - 吞吐量

数据分析

- 任务调度
- 开源任务调度器
  - Azkaban
  - Oozie
- 友盟任务调度器

数据分析

友盟任务调度器

数据展示

- 数据服务
  - Rest API
  - ProtoBuf
  - 异步请求
  - 缓存 Redis
  - 预先聚合计算

监控报警

- Zabbix
- Ganglia
- BaconTower
- 报警内容
  - 任务失败
  - 任务延迟
  - JT, NN, DN下线
  - HM, RS 下线

总结

- RealTime + Batch
- Kafka
- MongoDB + Hbase
- Storm, MR, Pig, Hive , Spark
- ElephantBird, ProtoBuf
- LZO, LZMA
- Scheduler
- Zabbix, ganglia


ChinaUnix
ITPUB
THANKS