

# 亿级大数据实时分析之旅

欧阳辰

# 我是谁？

欧阳辰



开发主管

ORACLE®

10<sup>g</sup>  
ORACLE  
DATABASE

企业软件

高级开发经理/工程师



互联网

10年

架构师主管



广告平台，大数据

- >15年的软件研发经验



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
世纪传媒

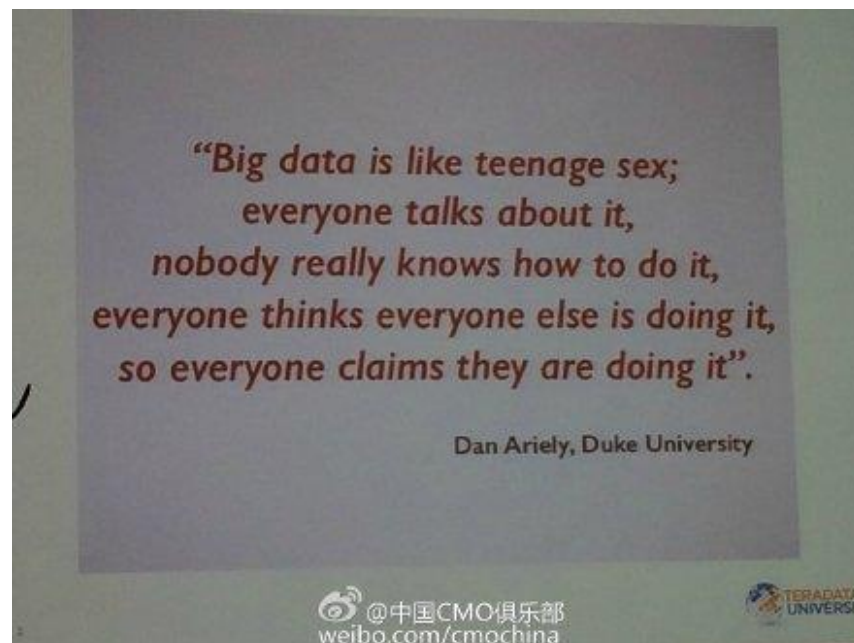
IT168

ChinaUnix

ITPUB

# 大数据的故事：价值为美

- 大量，快速，多样，变化等
- 采样的数据价值明显降低
- 实时数据的价值为最大
- 大数据外表光鲜亮丽，内则无可奈何
- 大数据的变现之殇



《大数据的8个小观点》



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

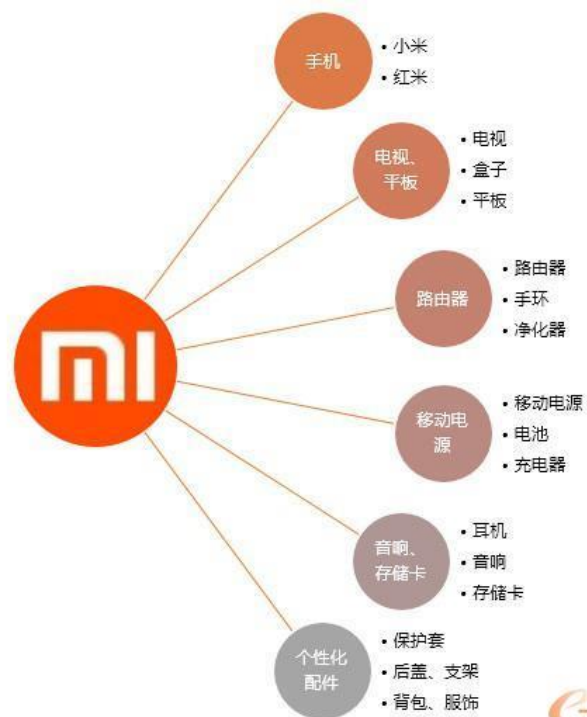
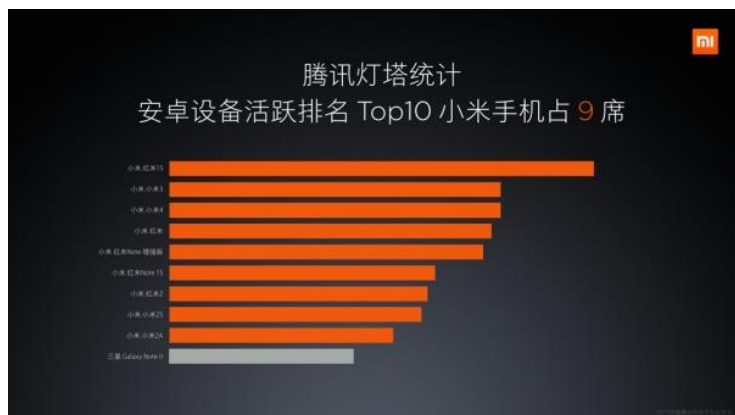
SequeMedia  
数据传媒

IT168

ChinaUnix

ITPUB

# 小米的大数据



e-works

“MIUI平台上日活跃量过千万的小米应用达21个”



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia

IT168

ChinaUnix

ITPUB

# 小米的大数据技术框架



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
媒体联盟

IT168

ChinaUnix

ITPUB

# 小米大数据应用

## 广告营销

- 点击预估
- 人群画像
- 营销DMP
- 精准营销

## 搜索和推荐

## 互联网金融

## 精细化运营

## 防黄牛

## 图片分析和处理



# 小米大数据实时分析场景案例



<http://dev.xiaomi.com>



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

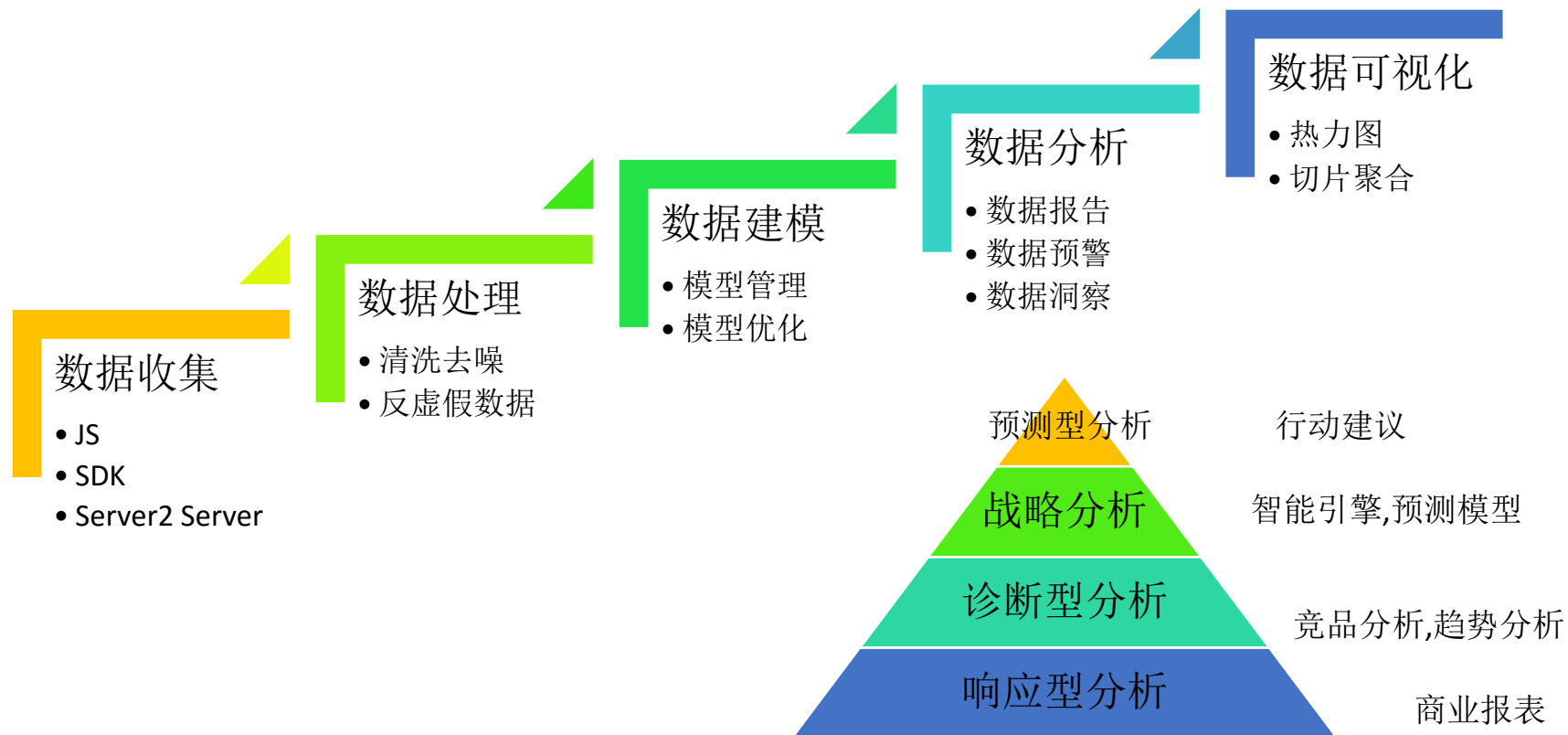
SequeMedia

IT168

ChinaUnix

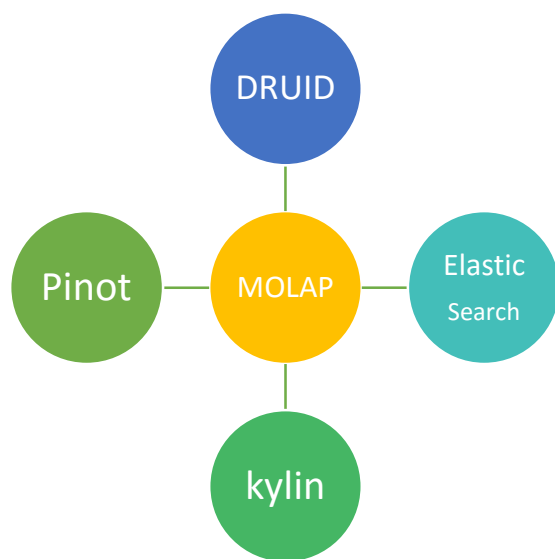
ITPUB

# 数据分析的几个步骤

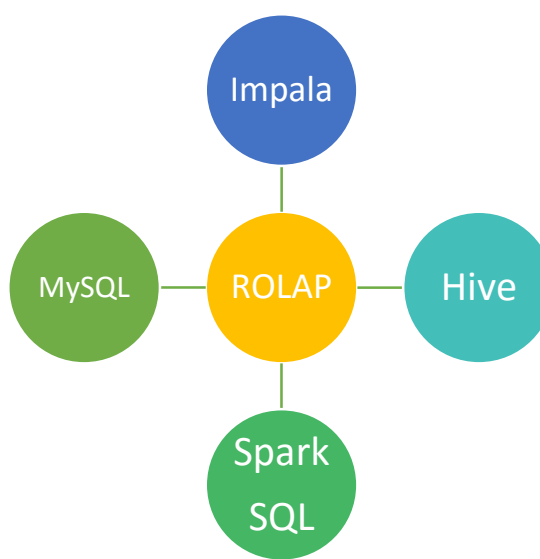




# 大数据分析工具



开源方案



商业方案



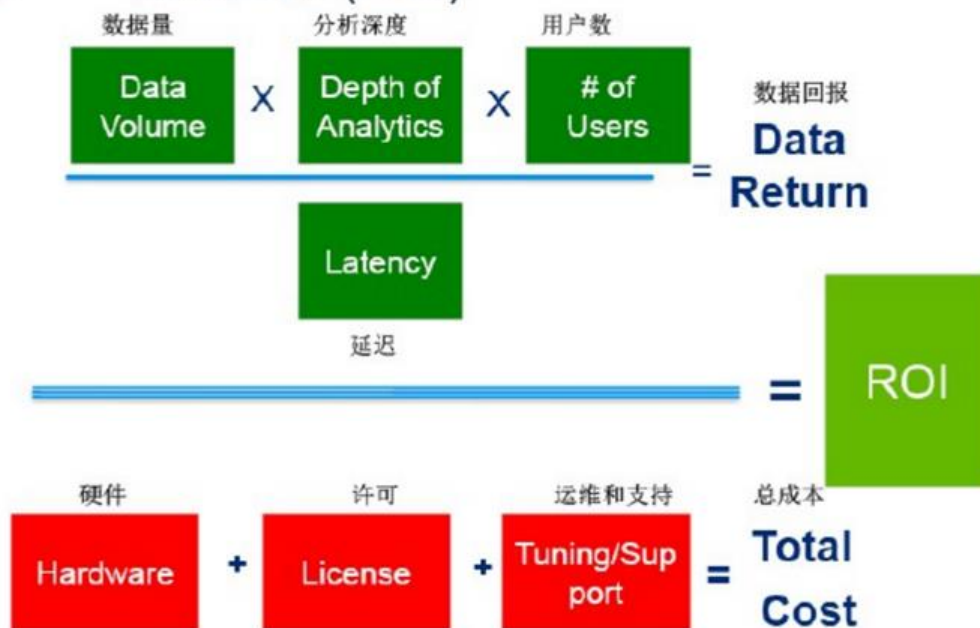
CAP理论



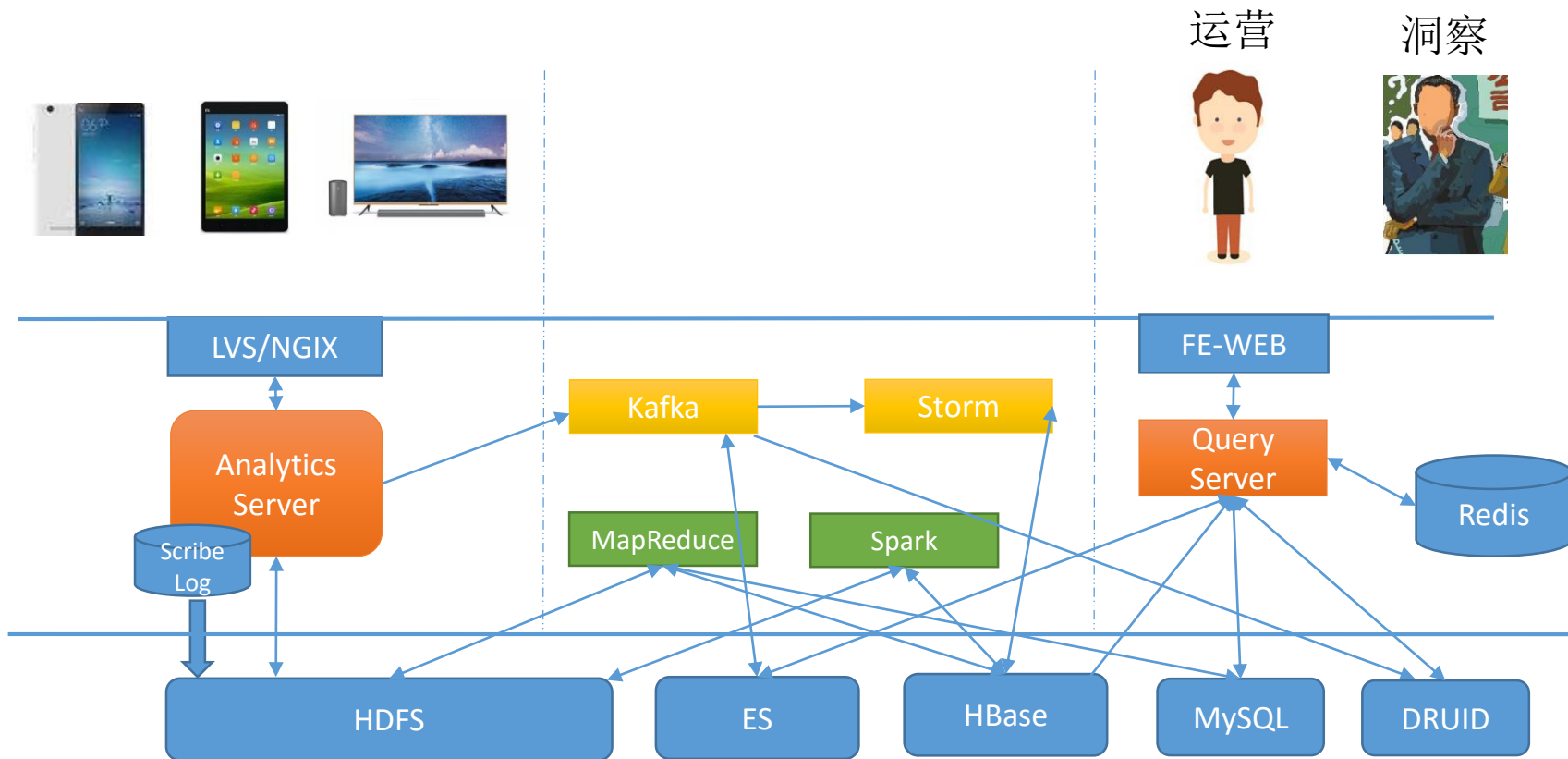
# 如何选择数据分析工具

- 度量数据分析系统效率的指标——信息回报 (ROI)

- 分析数据量越大，利用率越高
- 深度越深
- 实时性越强
- 总体成本越低
- 分析系统效率越高



# 小米数据统计分析平台-架构



# 一些分析工具的技术和应用



Pinot



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
媒体联盟

IT168

ChinaUnix

ITPUB

# 为什么青睐HBase?

- 天生为了大数据
- 改变Schema的时候平滑
- 扩容方便
- 成本考虑
- Facebook等公司做出了很好的示范
- 一些重要的Feature
  - CheckAndPut
  - Increment原子性
  - RowKey TTL
  - 单RegionServer内的强一致
  - 高I/O

我们有三位HBase Committer!



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
媒体联盟

IT168

ChinaUnix

ITPUB

# Hbase 在小米的改进



# HBase 在小米的改进

## 优化和改进

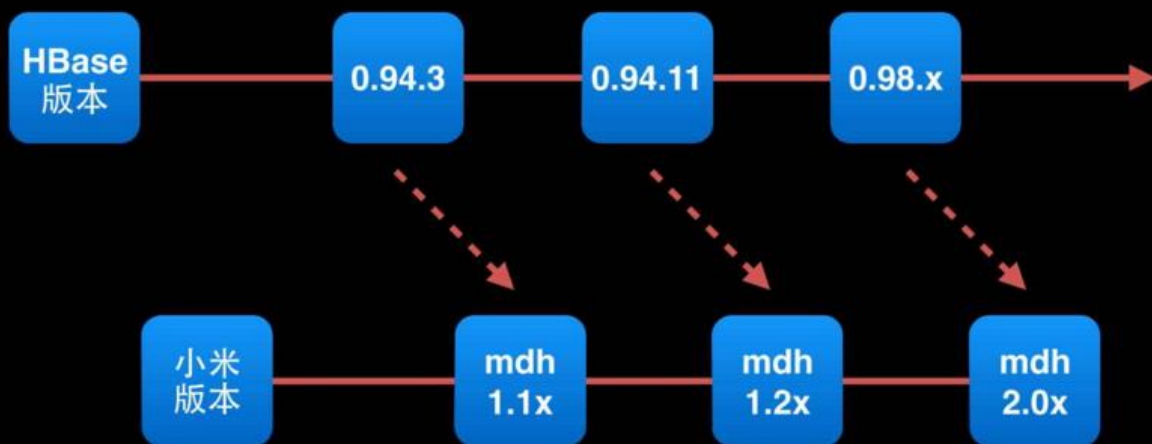
- 单机多实例，减少Heap大小
- BucketCache(Heap + Offheap)
- Compaction限速
- Read/Write Quota限制
- table/CF粒度的Replication限速
- 在线更新集群配置
- 新的HLog写模型
- 根据业务类型选择存储介质（磁盘/SSD）



# HBase 在小米

## 服务器端的演进

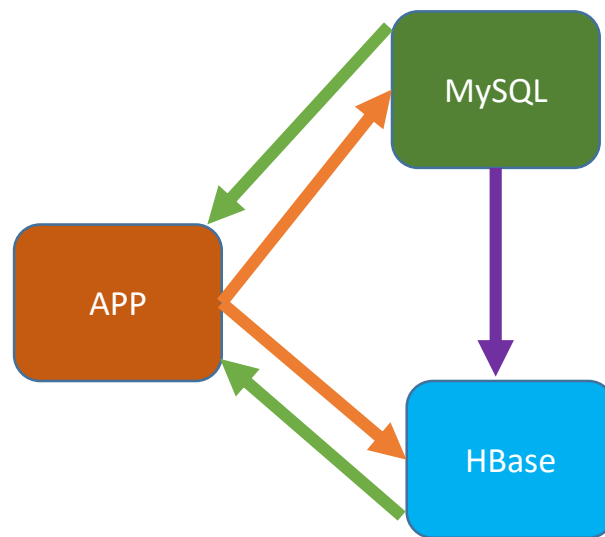
- 基于社区的Release版本+小米修改，不断测试和“趟坑”
- 修改积极反馈社区
- 定期跟进社区最新版本，与社区同步





# 如何从MySQL 平滑迁移到HBASE?

1. 双写HBase和mysql
2. 迁移历史数据（使用用老的时间戳）
3. 双读HBase和mysql，验证数据一一一致性
4. 灰度返回HBase结果



# 几种开源MOLAP分析工具的比较

	DRUID	Pinot	kylin
使用场景	实时处理分析	实时处理分析	OLAP分析引擎
开发语言	JAVA	JAVA	JAVA
接口协议	JSON	JSON	OLAP/JDBC
发布时间	2011	2015	2015
Sponsor	MetaMarkets /Yahoo	LinkedIn	eBay
技术	实时聚合	实时聚合	预处理, Cache

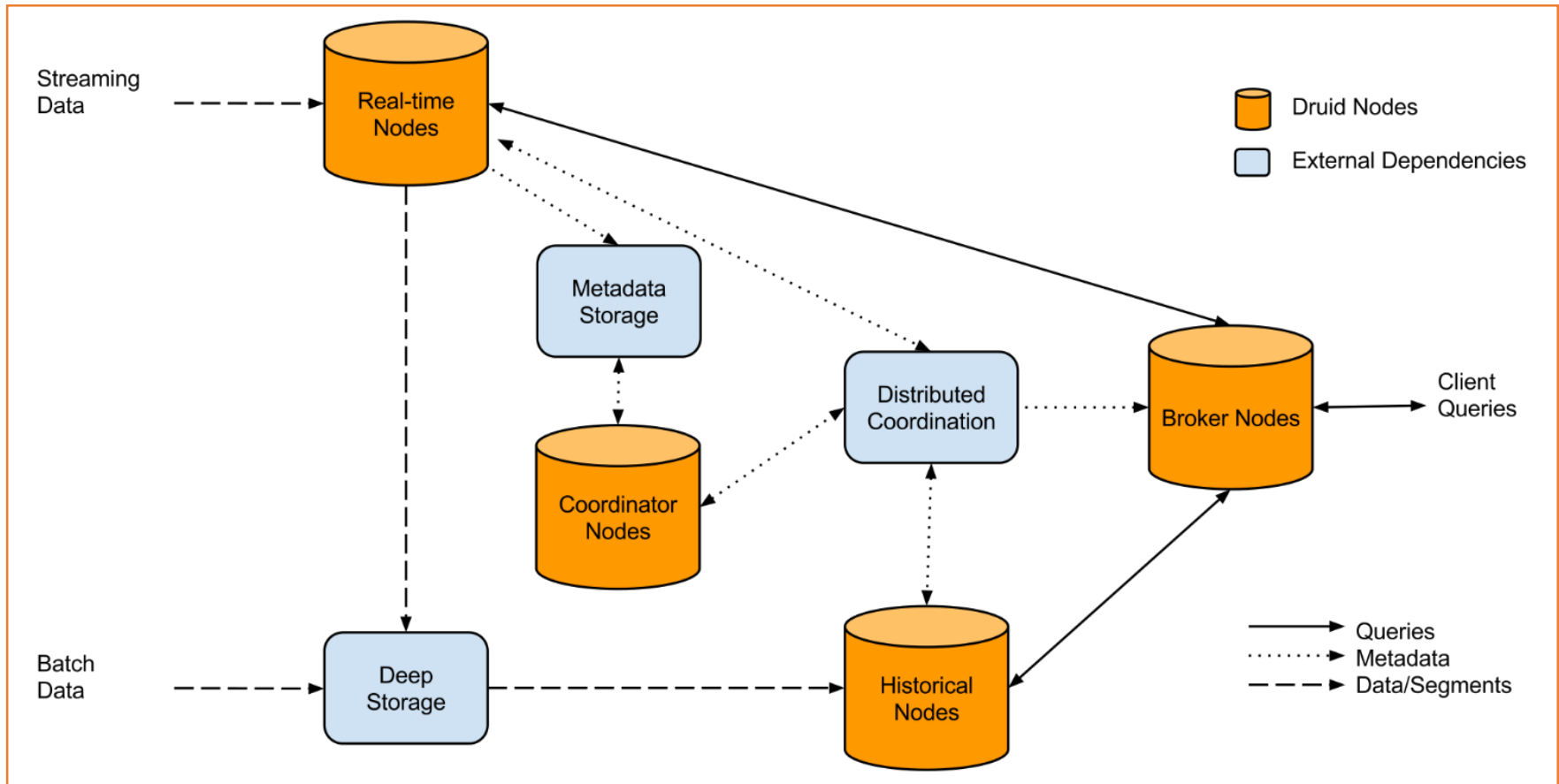


# DRUID 特点

- 来自METAMarkets，应用在Yahoo,OneAPM, 小米，阿里， Baidu
- 为分析而设计
  - 为OLAP而生，它支持各种filter、aggregator和查询类型
- 交互式查询
  - 低延迟数据，内部查询为毫秒级，
- 高可用性
  - 集群设计，去中性化规模的扩大和缩小不会造成数据丢失。
- 可伸缩
  - 现有的Druid部署每天处理数十亿事件和TB级数据。Druid被设计成PB级别。



# DRUID Architecture : <http://druid.io>



\*Pinot , PowerDrill , Dremel



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

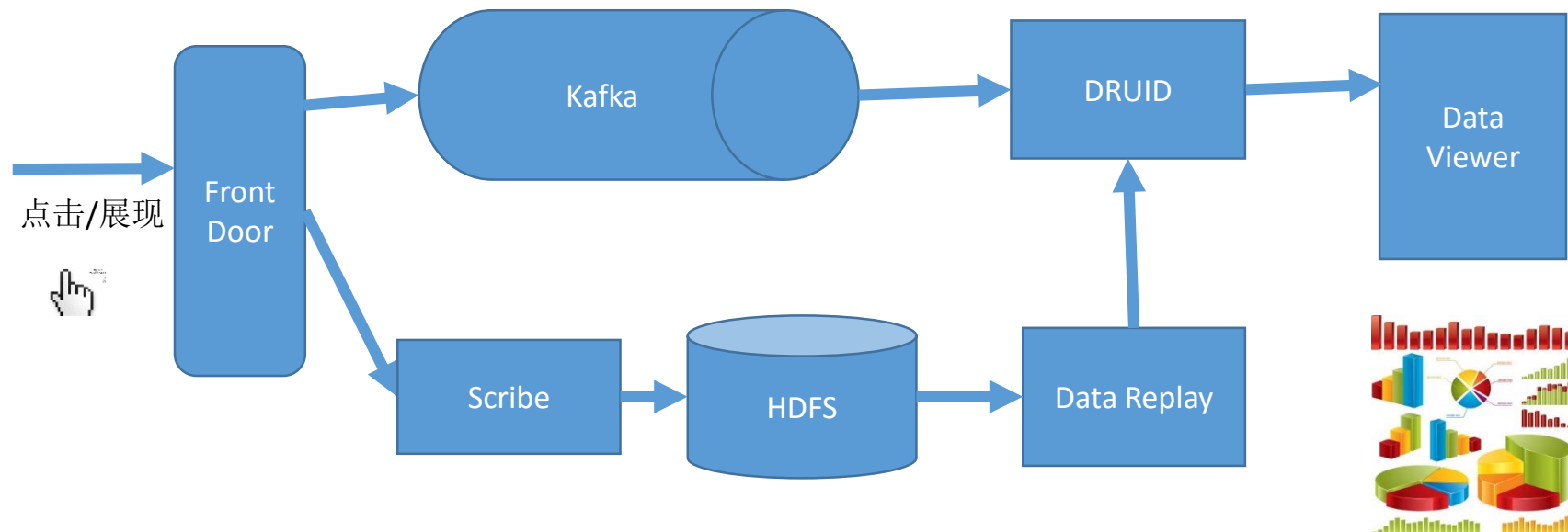
SequeMedia

IT168

ChinaUnix

ITPUB

# DRUID使用场景：广告实时统计分析架构图 (非计费部分)

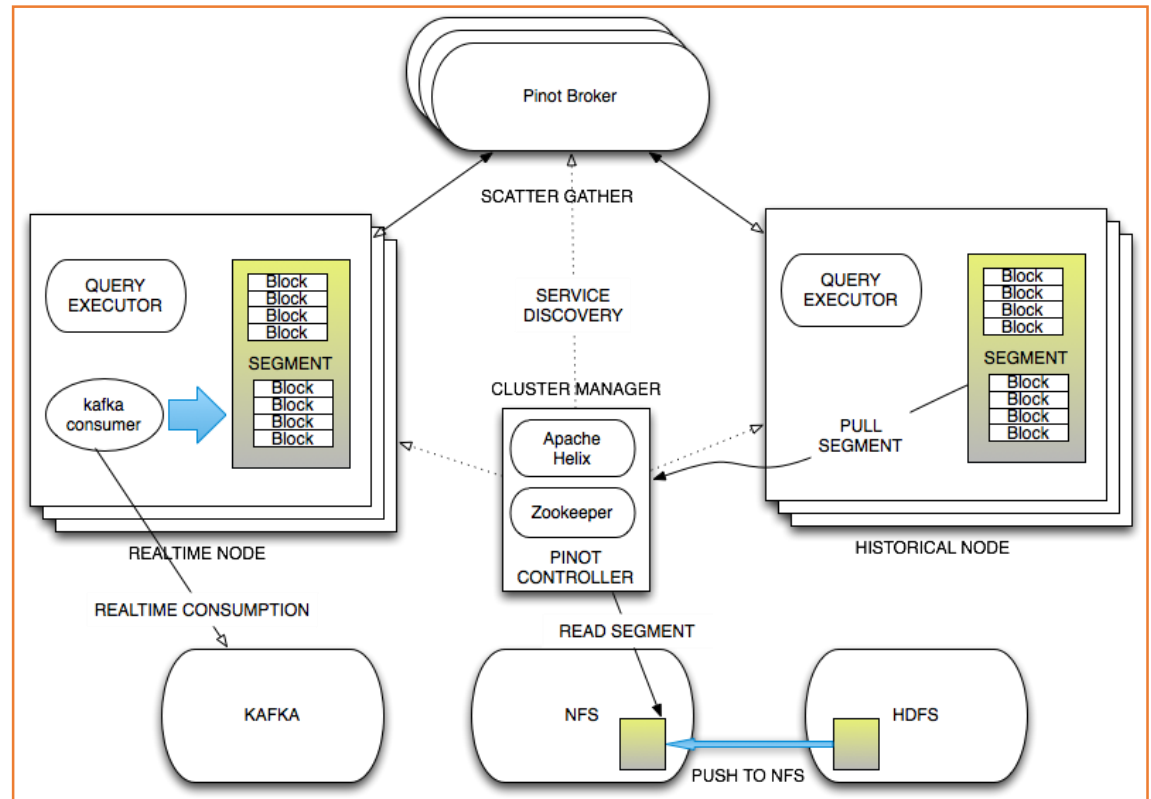
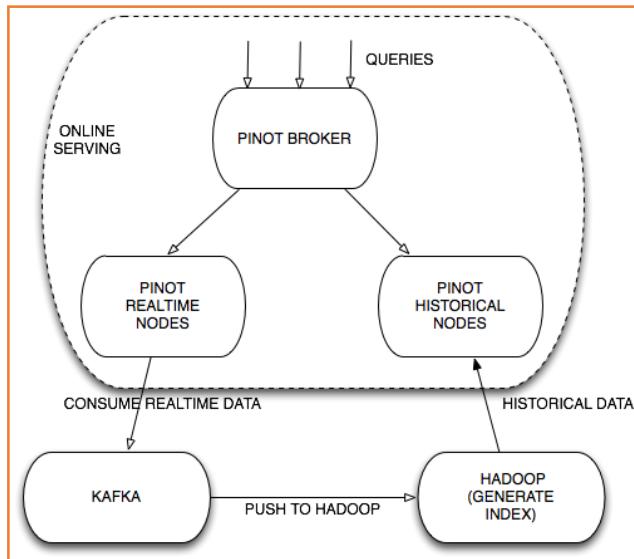


# 什么是Pinot

- 分布式的实时OLAP数据分析平台
- 用于LinkedIn内部，50个场景，250\*3 节点
  - “谁看了我的Profile”
  - “广告创建，跟踪”
  - “内部数据分析BI等”
- 功能
  - SQL -Like查询,不支持JOIN
  - 支持多种数据源Kafka, Hadoop
  - 自动数据过期等
  - 不支持UDF（开发中）



# Pinot Components Architecture



# Apache KYLIN <http://kylin.io>



麒麟：eBay开源的分析引擎

- 分布式分析引擎
- 提供大部分SQL查询
- 与BI工具集成，例如Tableau
- 完整的工具集合
  - 管理界面
  - 任务监控
  - 增量更新



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
数据传媒

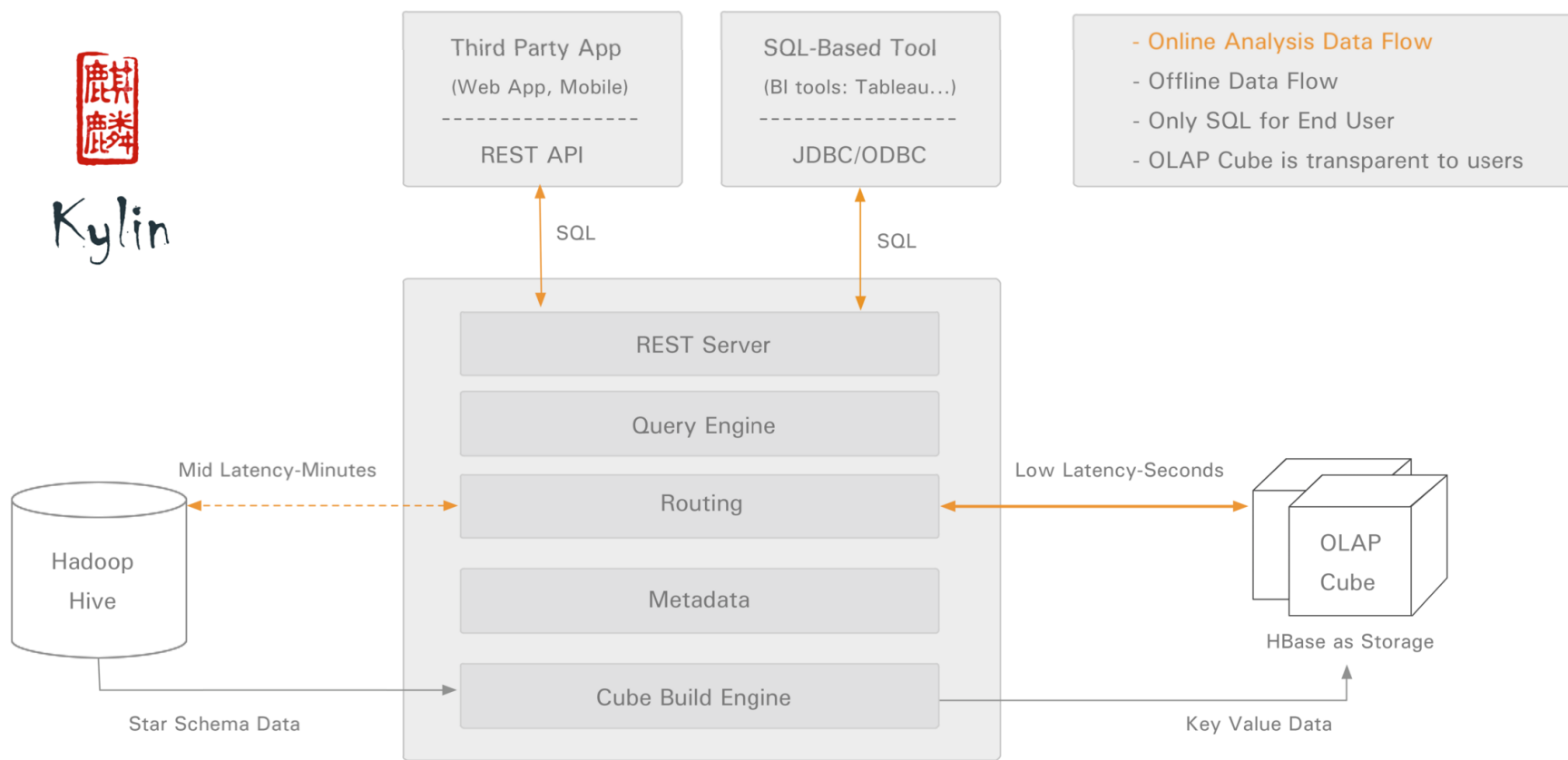
IT168

ChinaUnix

ITPUB



# 结构图



# Kylin使用的场景：Fast Cubing

## 场景一：

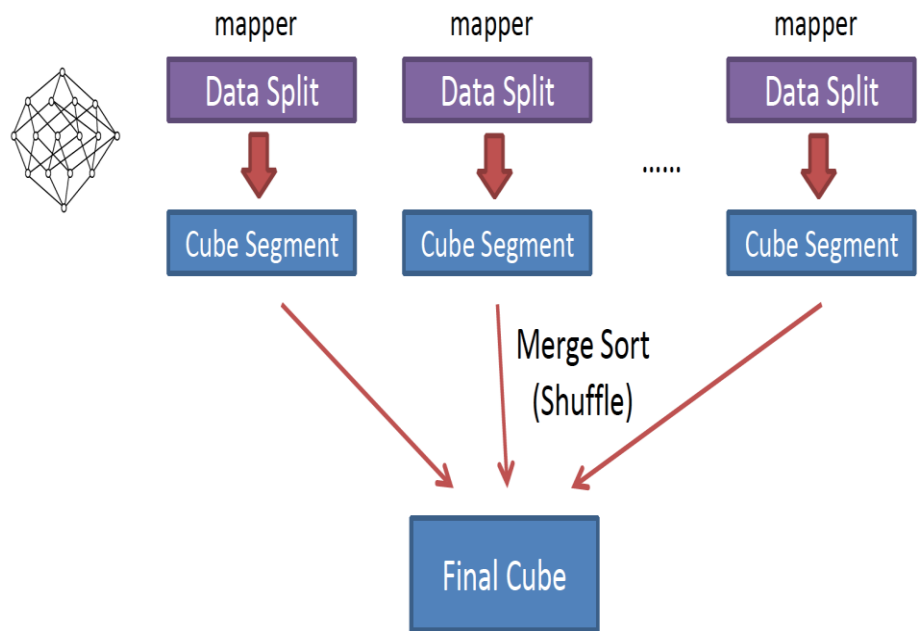
API请求的分析：响应时间，错误类型等

## 场景二：

一些广告返回类型的分析

查询延迟：秒级别

支持EXCEL客户端的多维表



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

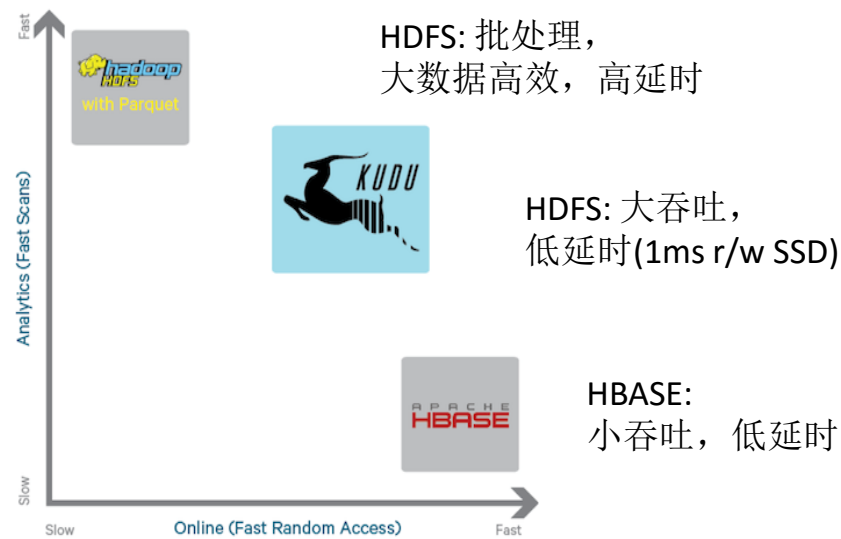
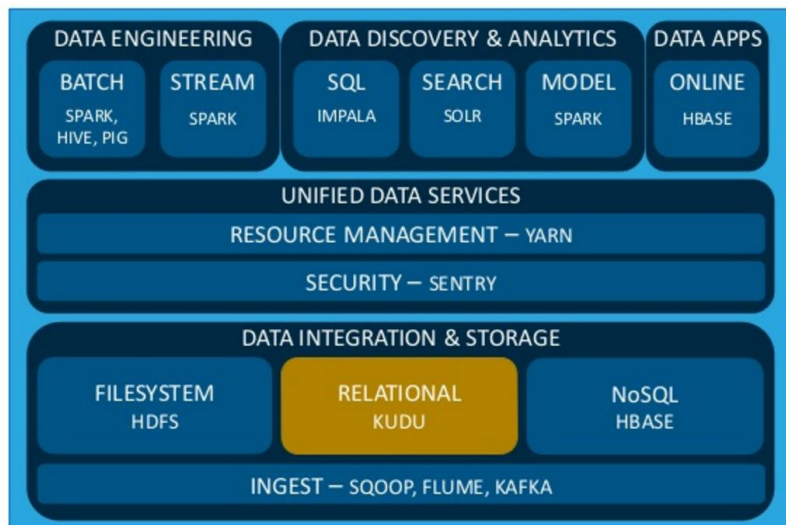
SequeMedia  
媒体传媒

IT168

ChinaUnix

ITPUB

# Apache KUDU



来源于Cloudera

场景：实时数据分析，数据报表，读写数据操作

小米应用：服务质量监控，问题排查工具



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

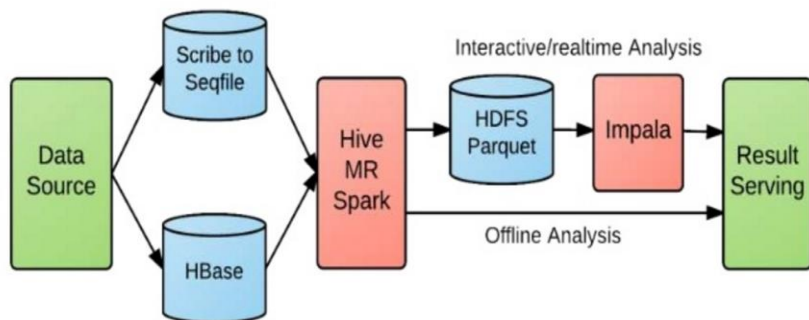
SequeMedia

IT168

ChinaUnix

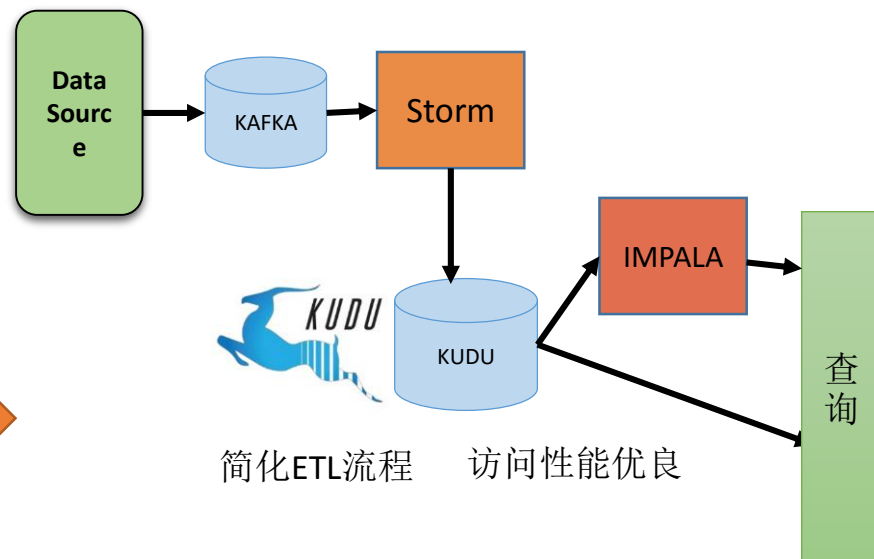
IT-PUB

# 小米KUDU的实践



KUDU之前的数据流

- ETL: 高延时
- LOG无序计算
- 需要等待完整数据, 才开始计算



## Xiaomi benchmark results

Query latency (seconds):



- HDFS parquet file replication = 3
- Kudu table replication = 3
- Each query run 5 times then averaged

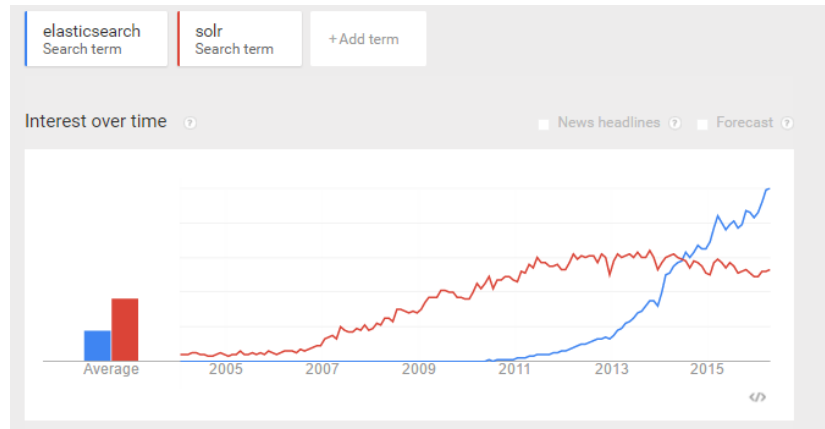


# Elastic Search

- 实时分布式搜索引擎和分析引擎
- 全文检索，结构化搜索和分析
- 核心引擎为Lucene，分布式产品

## 一些应用：

- Netflix: 2000 节点，多集群
- Wiki百科：全文搜索并且高亮
- Stack Overflow：全文索引
- GitHub:1300亿行代码
- 高盛：每天处理5TB数据，分析股票行情



小米的一些应用：搜索场景



- 将LOG进行索引，支持文本查询和计数
- 广告分析和查询

<https://www.elastic.co/use-cases>



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
媒体传媒

IT168

ChinaUnix

ITPUB

# 数据可视化

- **Meteorite:**

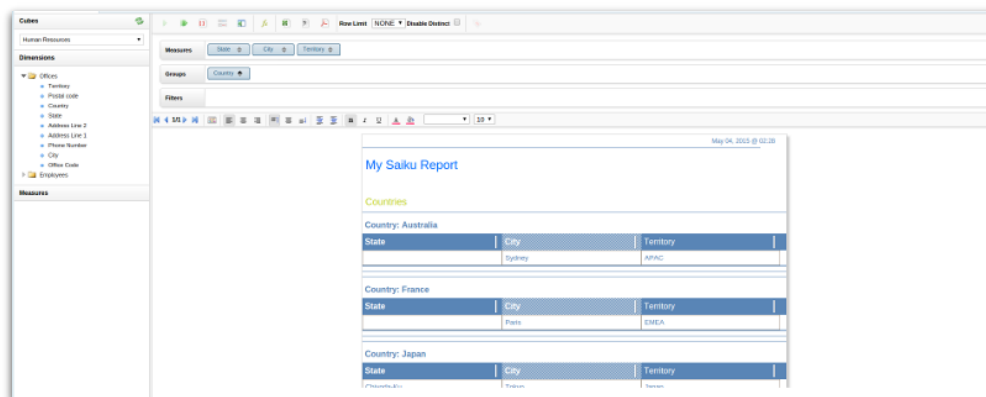
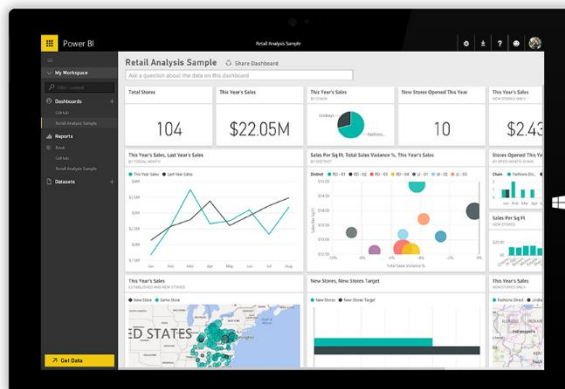
- Saiku

- **Microsoft:**

- Power BI
- Excel

- **Baidu**

- eChart



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia

IT168

ChinaUnix

ITPUB

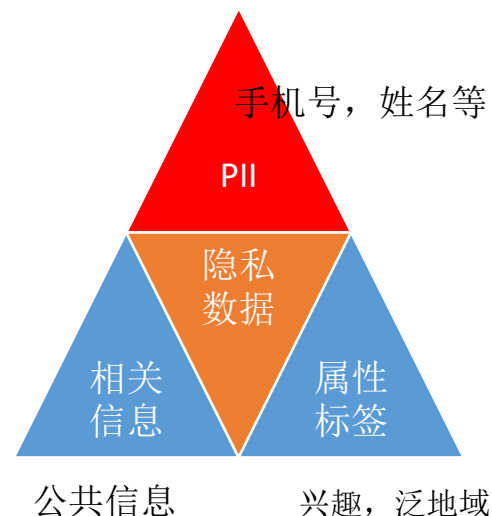
# Don't be evil : 用户隐私保护

**1890年**，沃伦和布兰代斯提出隐私权（The right to Privacy)

**2012年1月**，欧盟增加遗忘全力，(Right to be Forgotten)

**2006年4月14日**，欧盟《General Data Protection Regulation》

- 定义了数据控制者，处理者和数据主题 (Subject)
- 保护儿童数据
- 禁止收集个人特别数据（政治观点，性取向等）
- 数据可以转移到其他控制者，可遗忘
- 设立数据保护官(Data Protection Officer, DPO)



- PII信息: 个人标识信息
- Differential Privacy :差分隐私方法
- Quasi-identifier: 若干条件以定位某一用户
- K-Anonymity: 通过K个人信息推断某一用户
- L-Diversity: 数据值的多样性和精细粒度



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia

IT168

ChinaUnix

ITPUB

# 大数据实时分析的小结

- 没有业务应用的大数据都是耍流氓
- 技术选型没有想象中那么重要，实用和精通为妙
- 维度不够是一个永远的痛，无尽的伤
- 向保护你的眼睛一样保护用户的权利和隐私





# 道阻且长，行则将至



[www.ouyangchen.com](http://www.ouyangchen.com)

“互联居”

广告架构，大数据和观察

“致力于中国互联网广告技术的繁荣”！



不忘初心，方得始终



DTCC

2016年中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia  
媒体联盟

IT168

ChinaUnix

ITPUB