

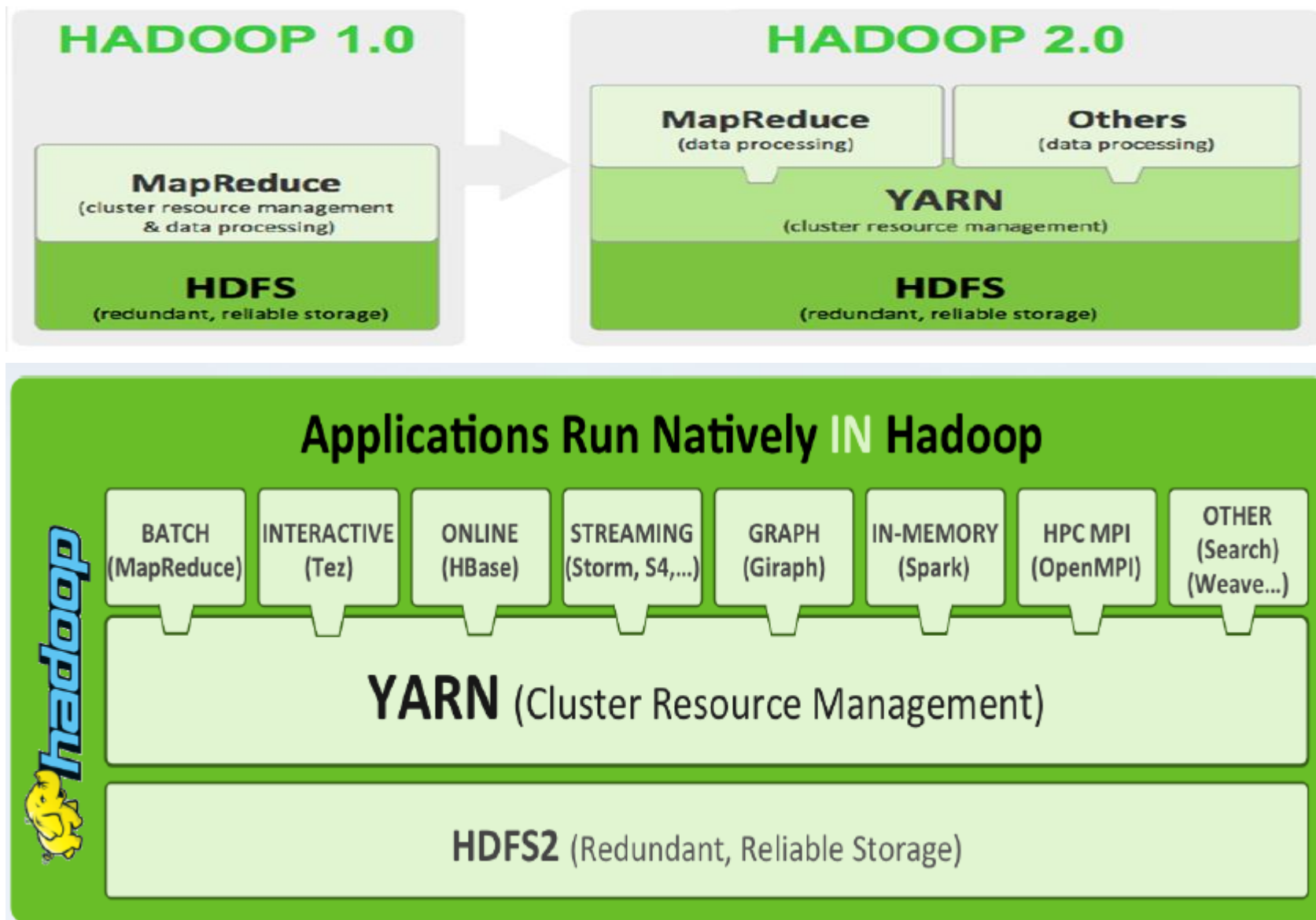
A decorative graphic on the left side of the slide showing a globe with several glowing blue and white orbits or data paths swirling around it.

Spark在中国移动省 公司试点应用

广州市浩微数据技术 邢刚

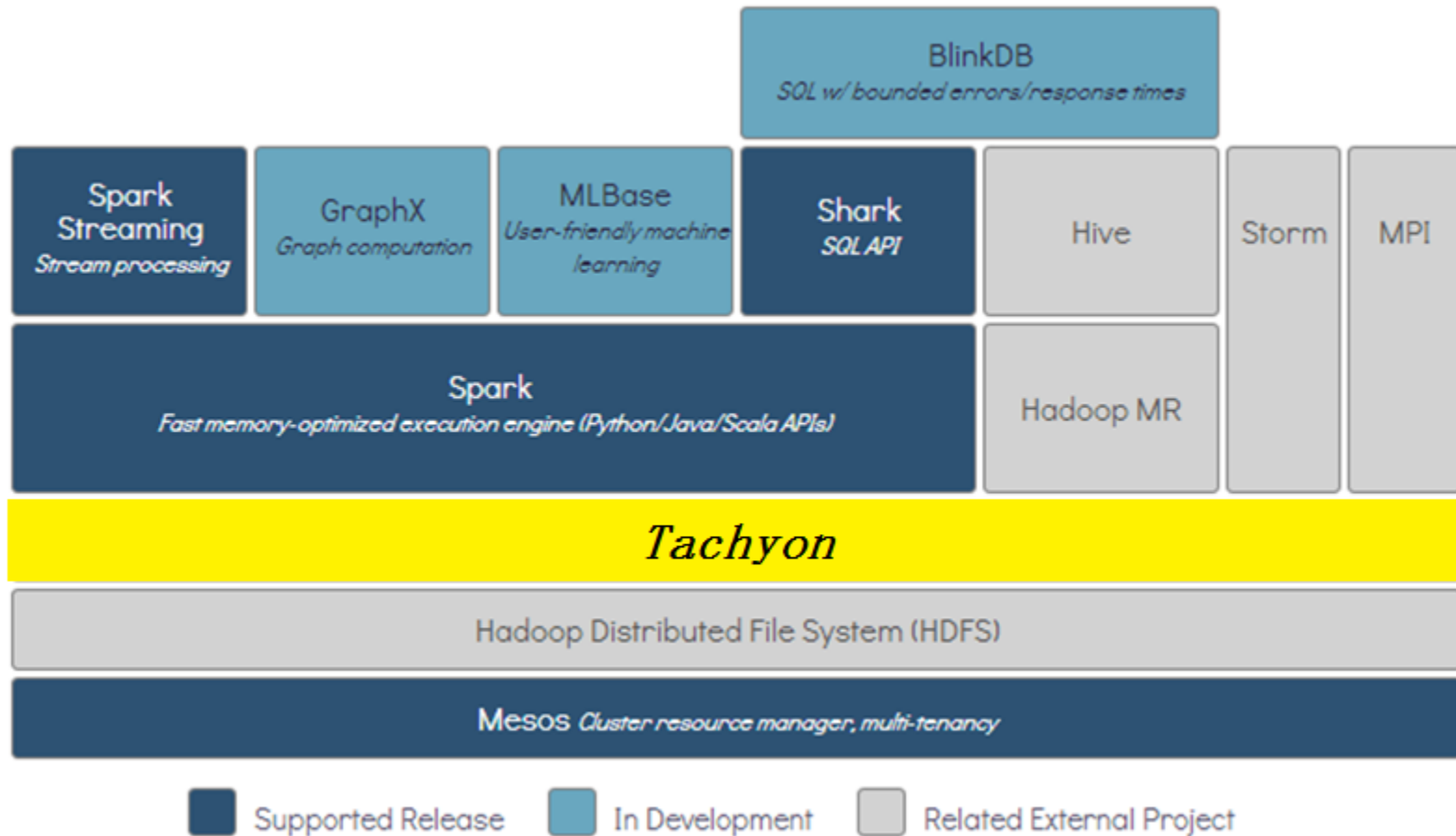
13538839570@139.com

Hadoop 2.0架构推出



选择BDAS架构

- BDAS的三个核心产品，分别是**Spark**实现批量MR算法，**Spark Streaming**实现实时计算和**Shark**实现交互式查询。这三个方面分别对应BI的**ETL/数据挖掘**、**实时经分/实时营销**和**OLAP/即席查询**三个方面应用。
- 比对测试的三个产品：MapReduce、Storm和Impala/Vertica。



- ❑ **企业**关注自己的业务，IT是实现**业务**运营的手段，IT大多通过外包或部分外包实现。企业喜欢通用的产品，希望能遵循标准进行实施，减少IT风险。
- ❑ **互联网运营商**喜欢个性化，IT是他们的核心，他们都自己干，不喜欢用通用产品，不喜欢被标准所约束。越是个性化的IT越能体现竞争力，越不易被复制。

GSM SNMP
eTom SQL
OSI IOE

尽量采购标准产品，由集成商集成并运维，形成IT能力支撑企业运营。关注方案的可替换性



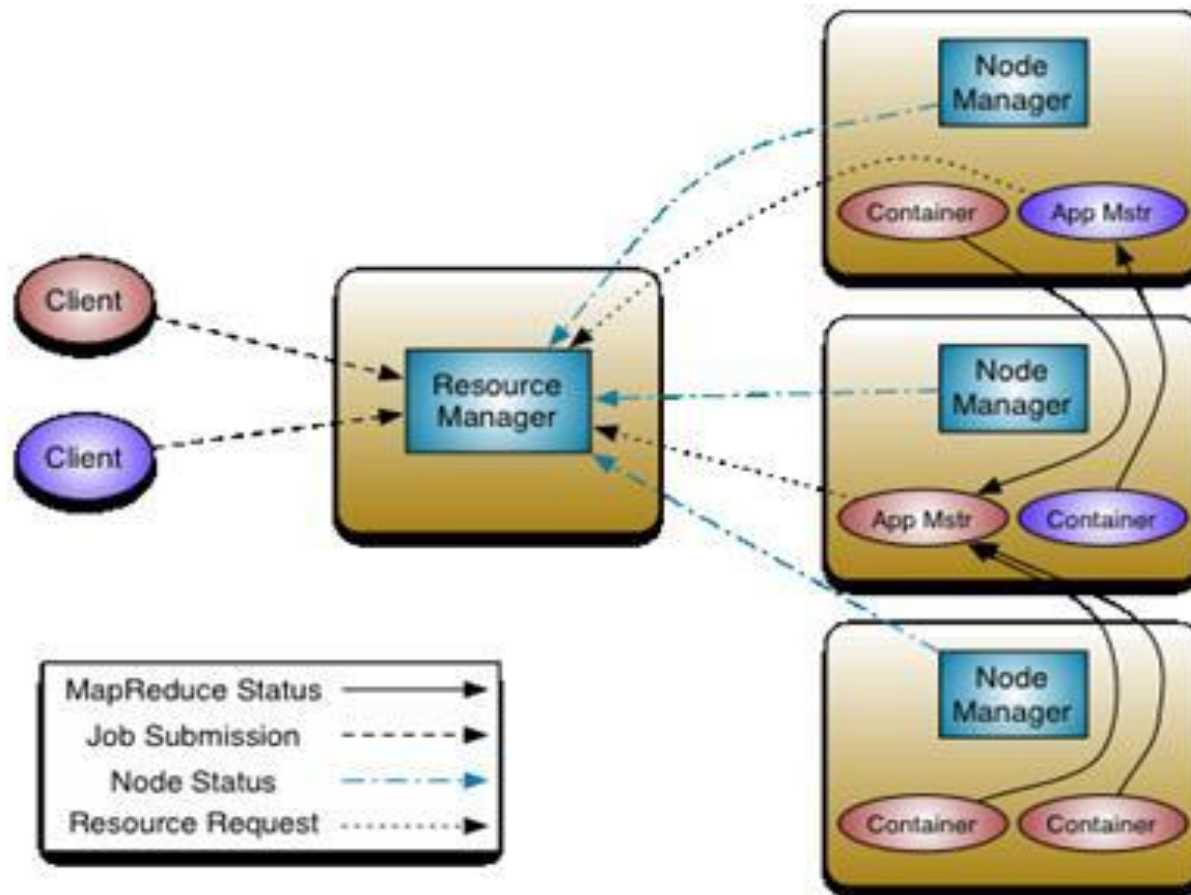
自己研发最适合自己的业务的产品，进行专业的优化（效率和成本），对外提供服务，成熟后通过开源引入更多开发者

Spark在中国移动省公司 试点应用

- Spark的可用性测试
- Spark的试点应用
- 问答

Hadoop 2.0架构

- 最主要改变：将 JobTracker 两个主要的功能分离成单独的组件，这两个功能是资源管理 YARN和任务调度 / 监控。
- 移动研究院的BigCloud在Hadoop1.0基础上，实现了ETL和常用数据挖掘算法。

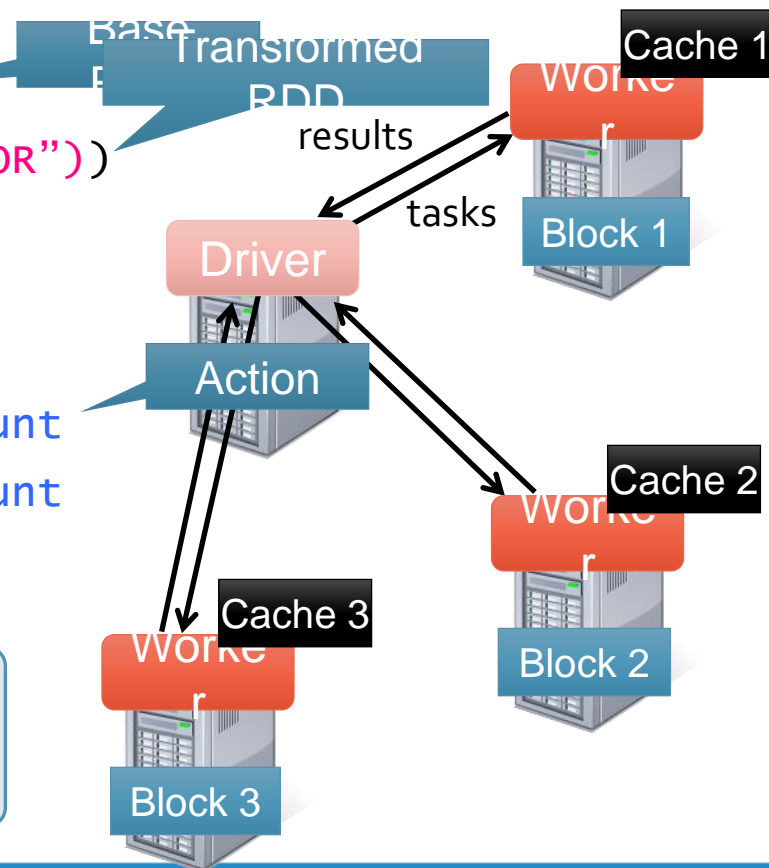


- 基于Hadoop的HDFS，Spark采用Driver、Worker的主从结构，由Driver节点调度，负责任务分配、资源管理。Worker节点主要是存放数据和进行计算。
- 第一次从外设读取数据，之后主要在内存计算，内存不足的情况下采用LRU算法，与磁盘交换数据。
- 案例中涉及到RDD、Tranform、Action等操作，处理方式是MR+类似SQL的集合语句，包括map、reduce、count、groupby、join、union等

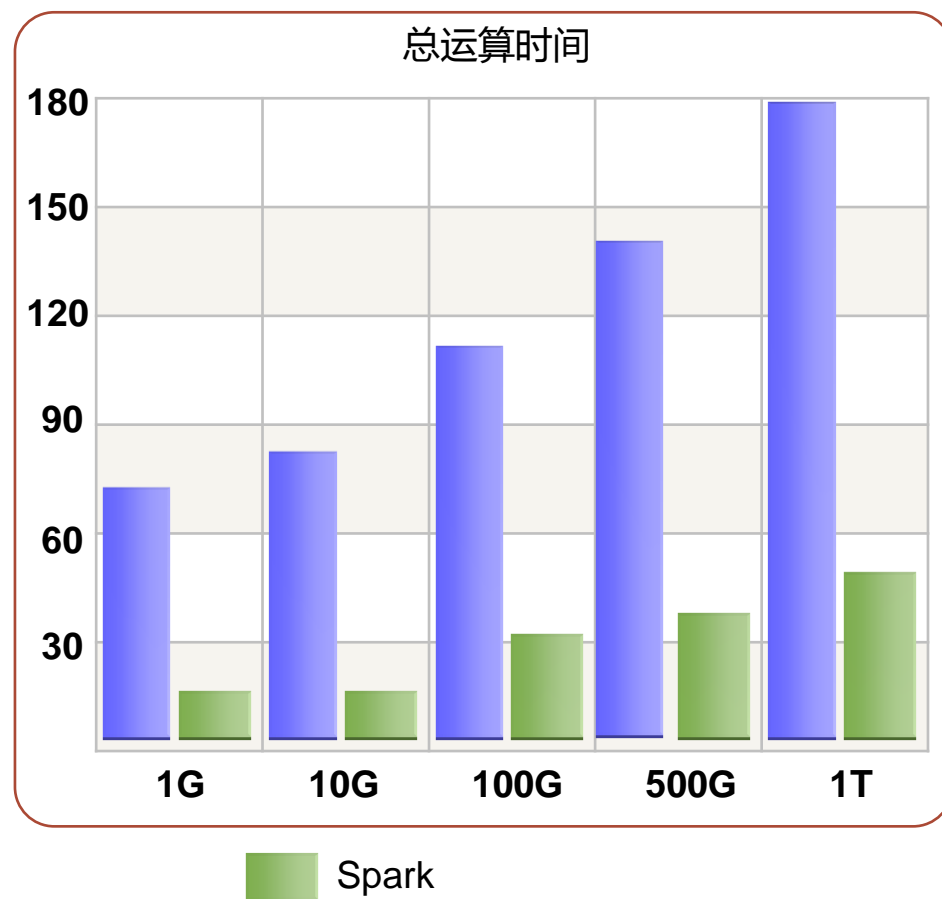
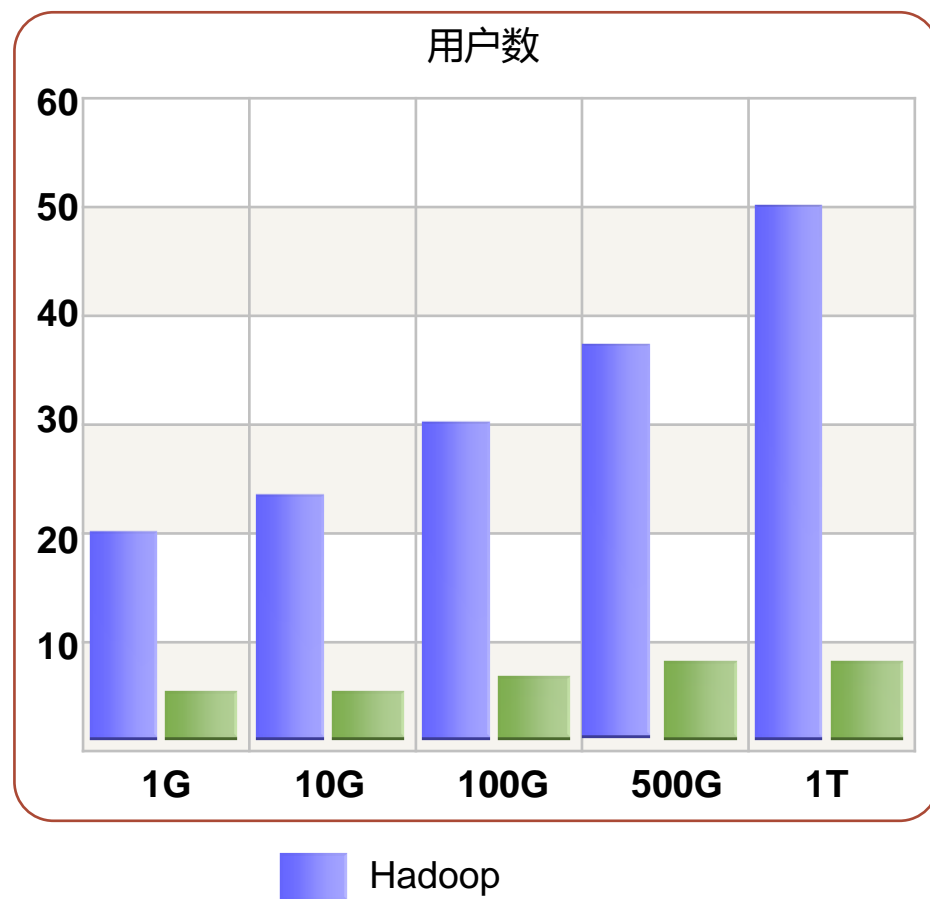
```
lines = spark.textFile("hdfs://...")
errors = lines.filter(_.startsWith("ERROR"))
messages = errors.map(_.split('\t')(2))
cachedMsgs = messages.cache()

cachedMsgs.filter(_.contains("foo")).count
cachedMsgs.filter(_.contains("bar")).count
. . .
```

Result: scaled to 1 TB data in 5-7 sec
(vs 170 sec for on-disk data)



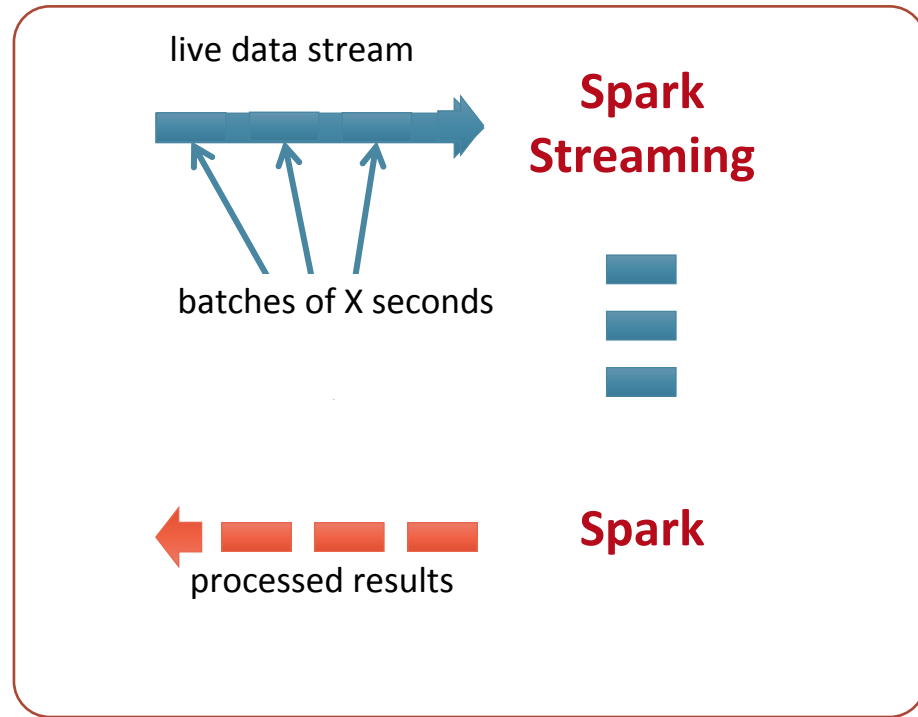
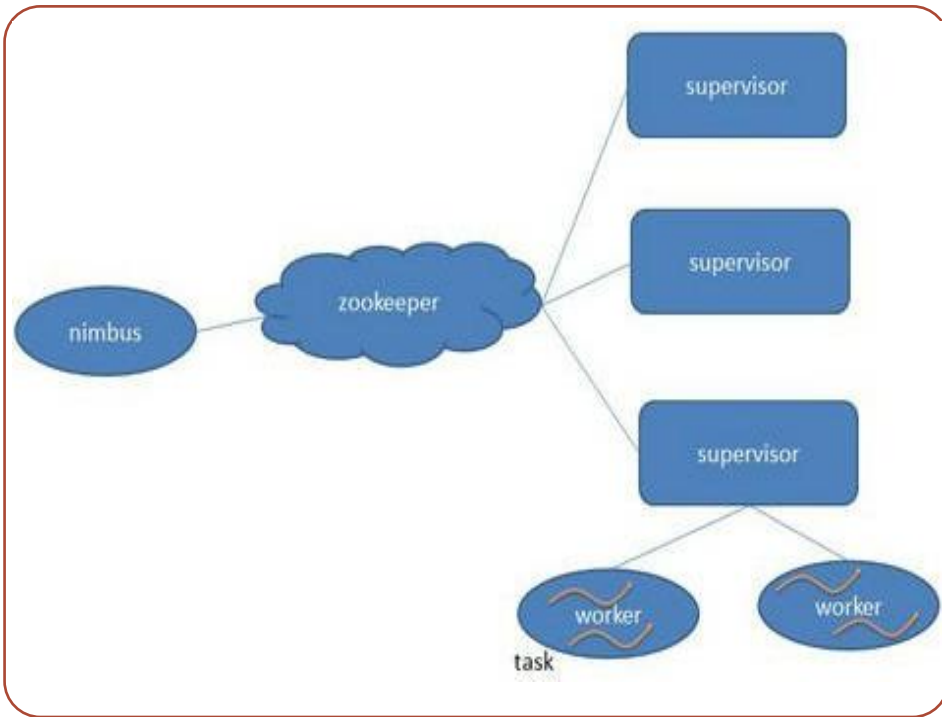
采用同样的算法，对1G、10G、100G、500G和1T的详单数据计算：通信用户数、每品牌用户数、总通话时间、每日通话时间、总运算时间的从时间、地域和品牌对比



从以上的计算看出：Spark计算所耗费的时间是MapReduce的1/6。

流处理的两种不同实现方式

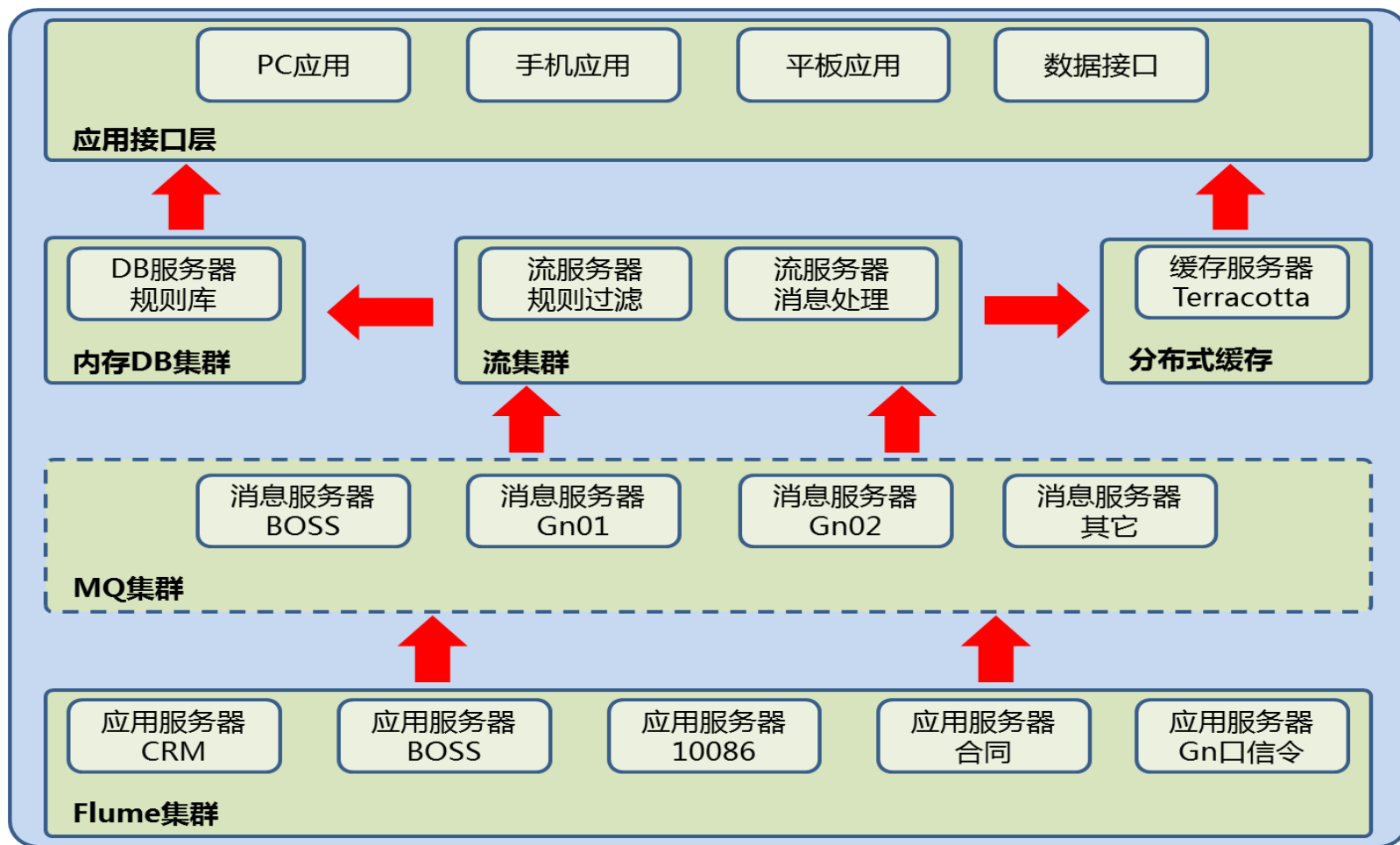
- ❑ 1、Streaming运行在Spark上，转化一系列毫秒级数据集的批量计算。
- ❑ 2、Storm采用消息机制，转化一系列元组处理。



- Batch sizes as low as ½ second, latency of about 1 second
- Storm每个Tuple的处理时间都是毫秒级别

流处理通用架构

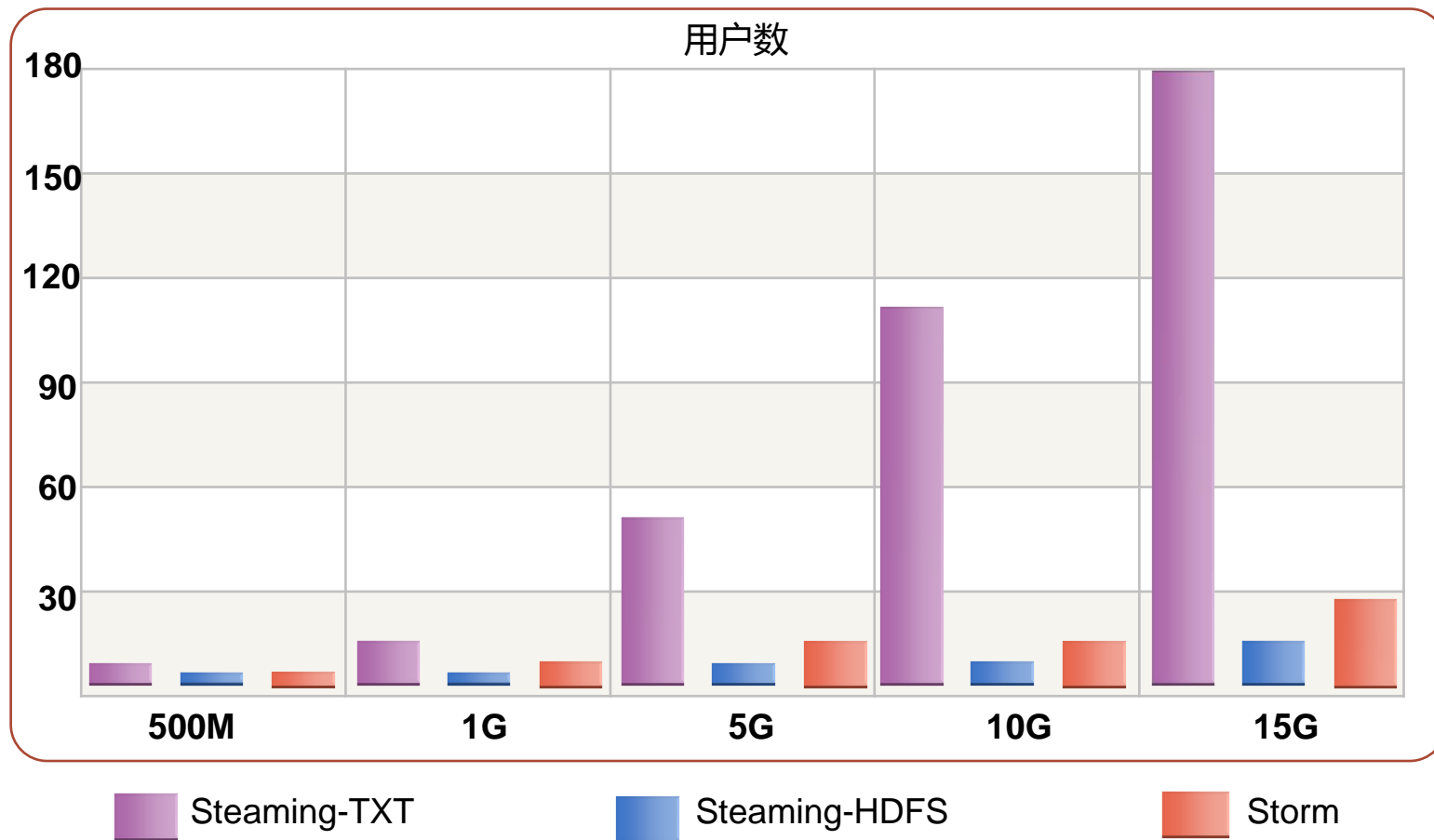
- 1、采用Flume采集数据，通过消息传递到Storm/Streaming，流处理后，放入内存DB和缓存，提供应用调用接口或服务。
- 2、根据接口的数据量和发送频率不同，来定义集群的大小。



- Streaming不需要MQ集群，可以直接连接到Flume集群。

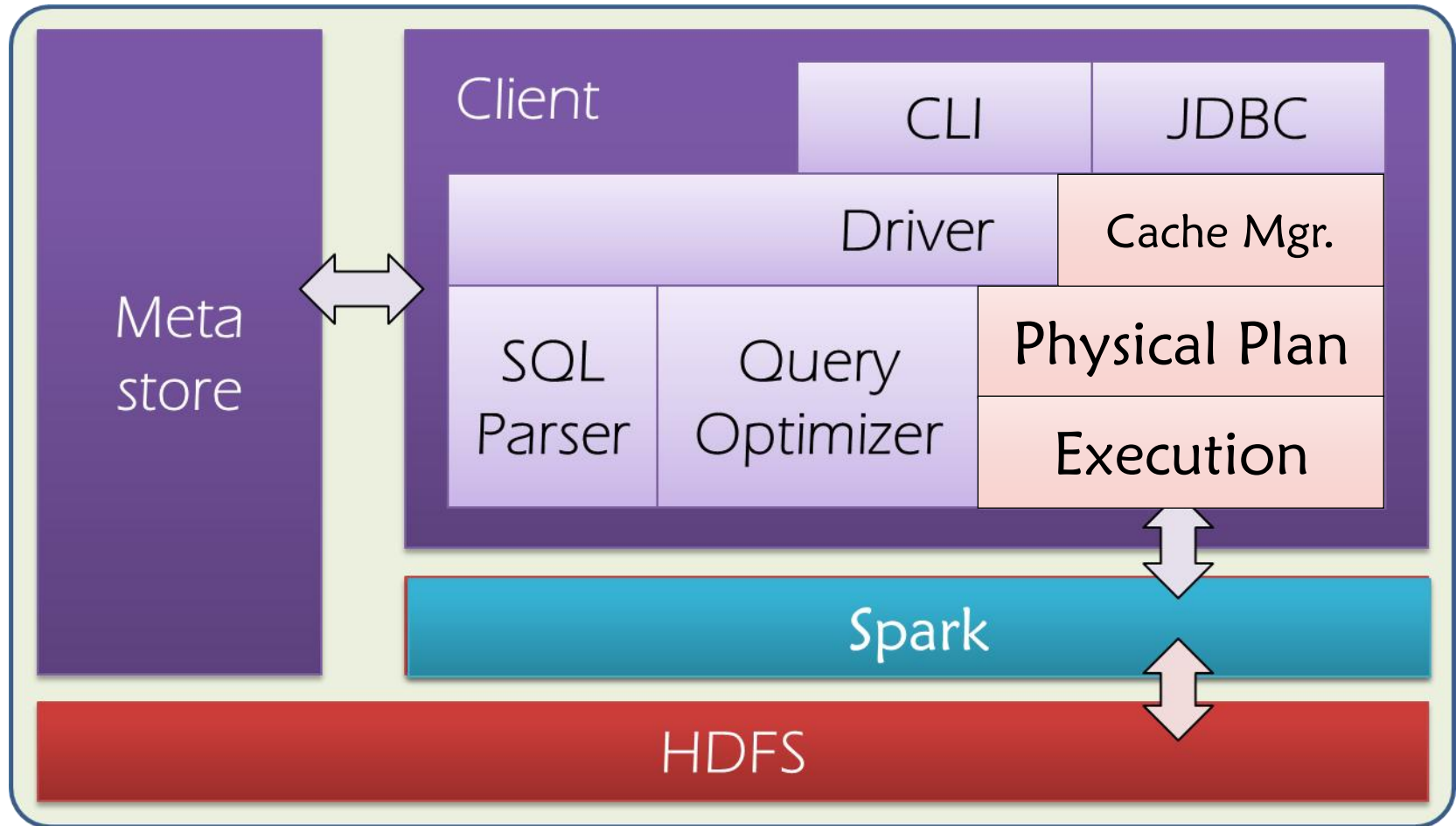
Streaming VS Storm

采用同样的算法，对500M、1G、5G、10G和15G的详单数据计算：通信用户数、每品牌用户数、总通话时间、每日通话时间、总运算时间的时间对比

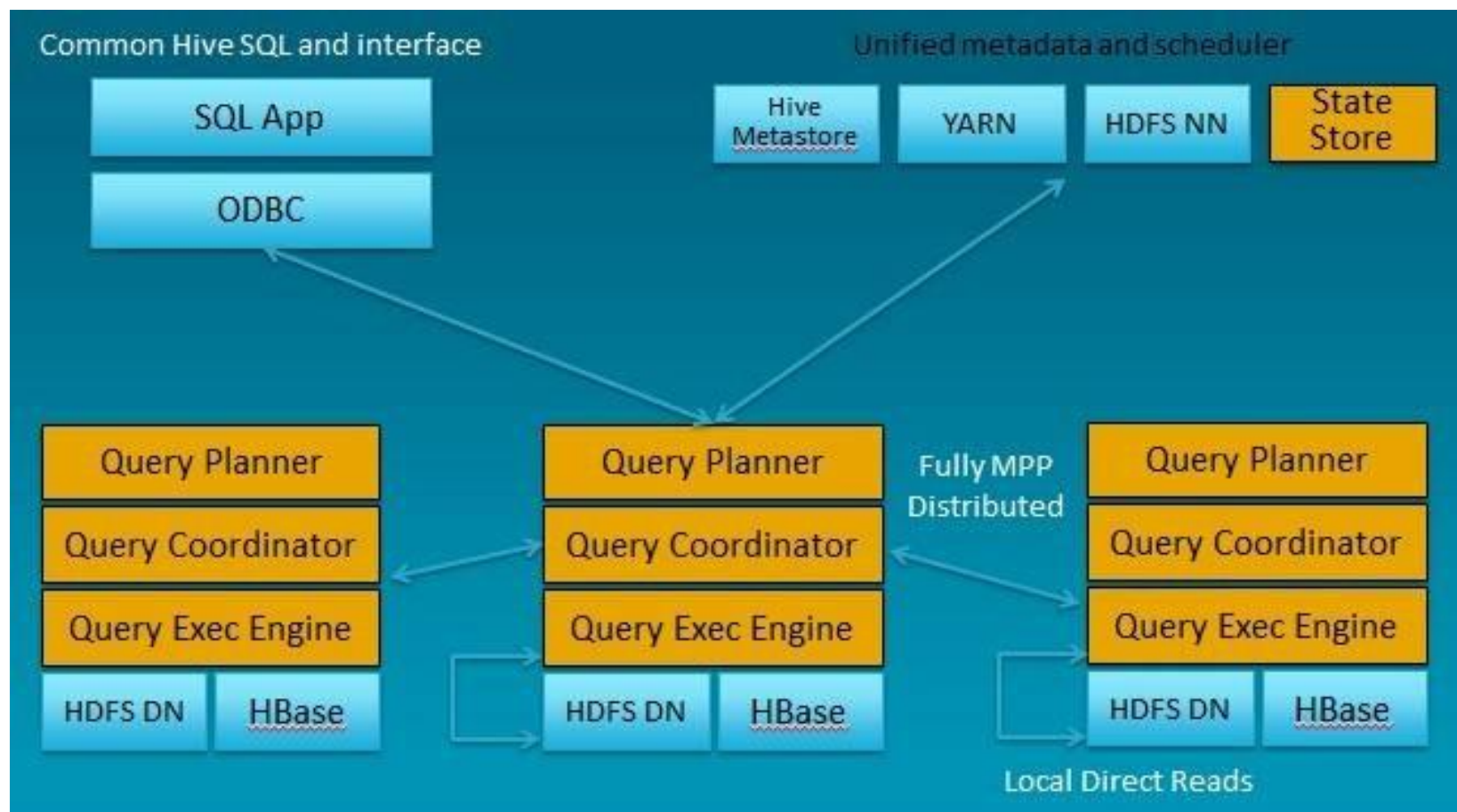


有三个测试，Storm是直接从文本文件读，结果写入磁盘；Streaming有两种方式，一个是从文本读入HDFS再计算，一个是从HDFS读取，第一种方式速度很慢，第二种方式则非常快。

- 在Hive的架构基础上，改写了“内存管理”、“执行计划”和“执行模块”三个模块，使HQL能够跑在Spark上。



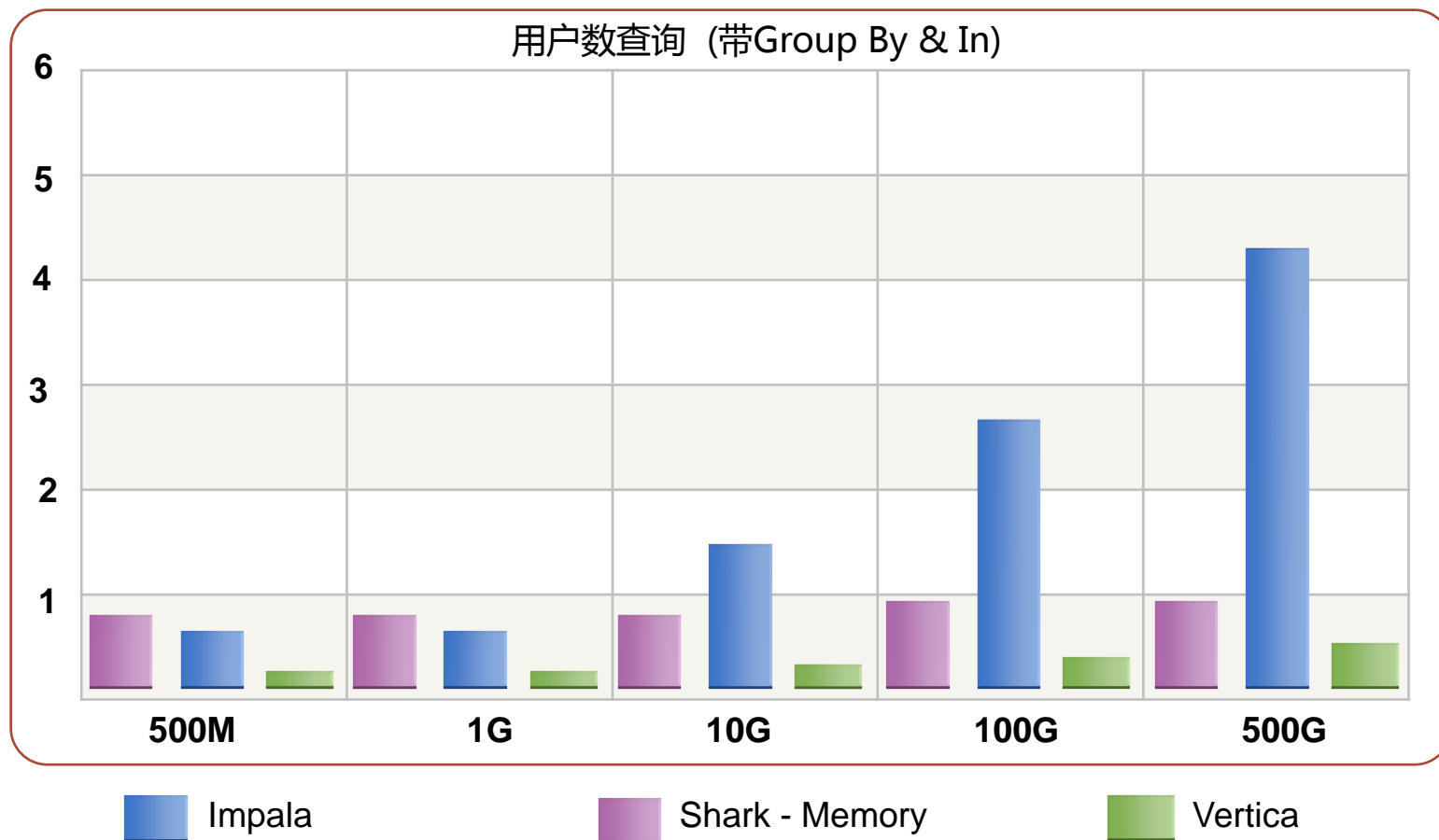
- 1、不同于Shark，Impala是直接在HDFS之上解释运行SQL。
- 2、采用C++编程，相对JVM，在内存管理性能有较大提升。



- 具有本地计算、分布式存储、共享Hive架构的特点。

Shark Vs Impala Vs Vertica

采用同样的算法，对500M、1G、10G、100G和500G的详单数据计算：通信用户数、每品牌用户数、总通话时间、每日通话时间、总运算时间的时间对比



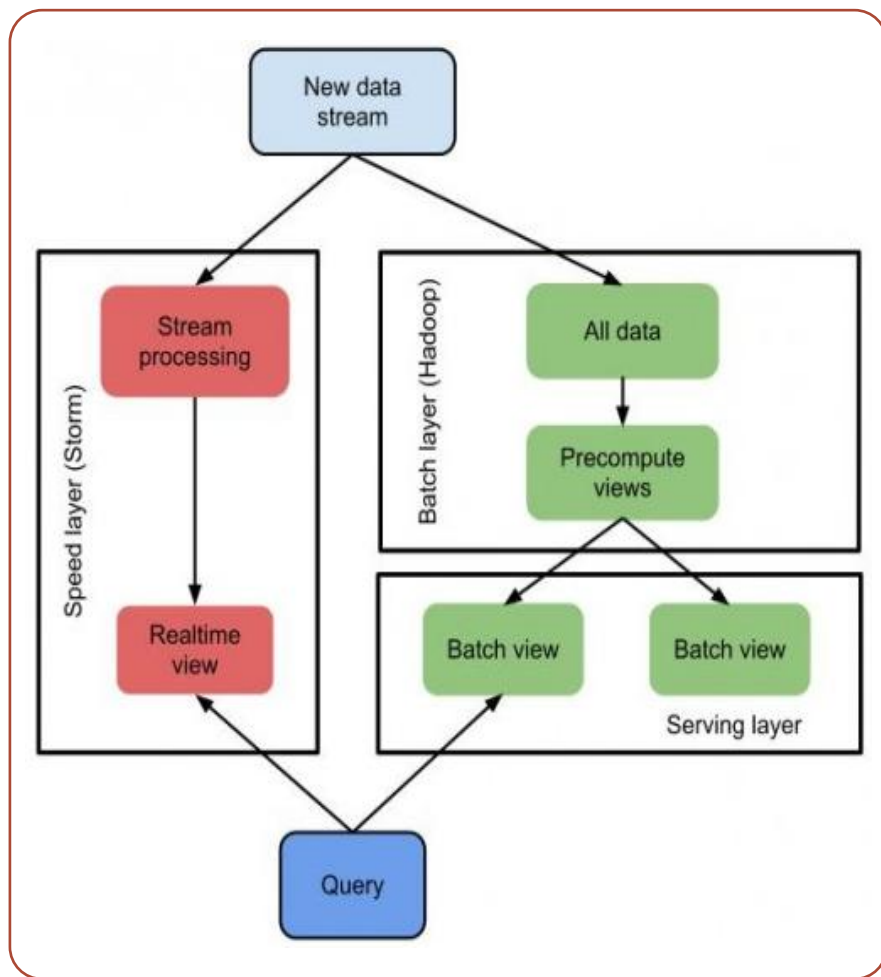
从上图看出：Impala的查询速度递增。Spark的速度保持稳定，因为大部分查询是在内存中进行。Vertica一直保持很快的速度，数据量增大，时延有少量提升，30%以内。

Spark在中国移动省公司 试点应用

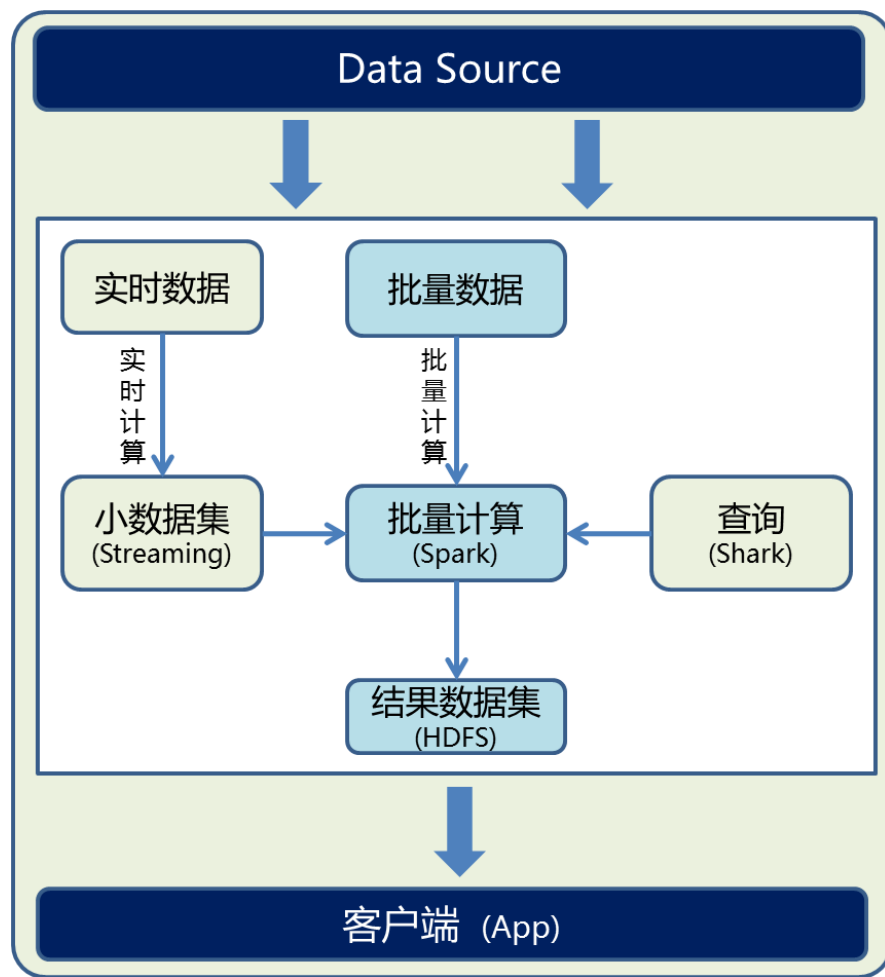
- Spark的可用性测试
- Spark的试点应用
- 问答

1、架构简化 – 混合式到单一架构

将Hadoop+Storm的架构，简化为Spark架构。实现一键式安装和配置，线程级别的任务监控和告警，减轻硬件集群、软件维护、任务监控和应用开发的难度。后续要做成统一的硬件、计算平台资源池，发展到云计算。

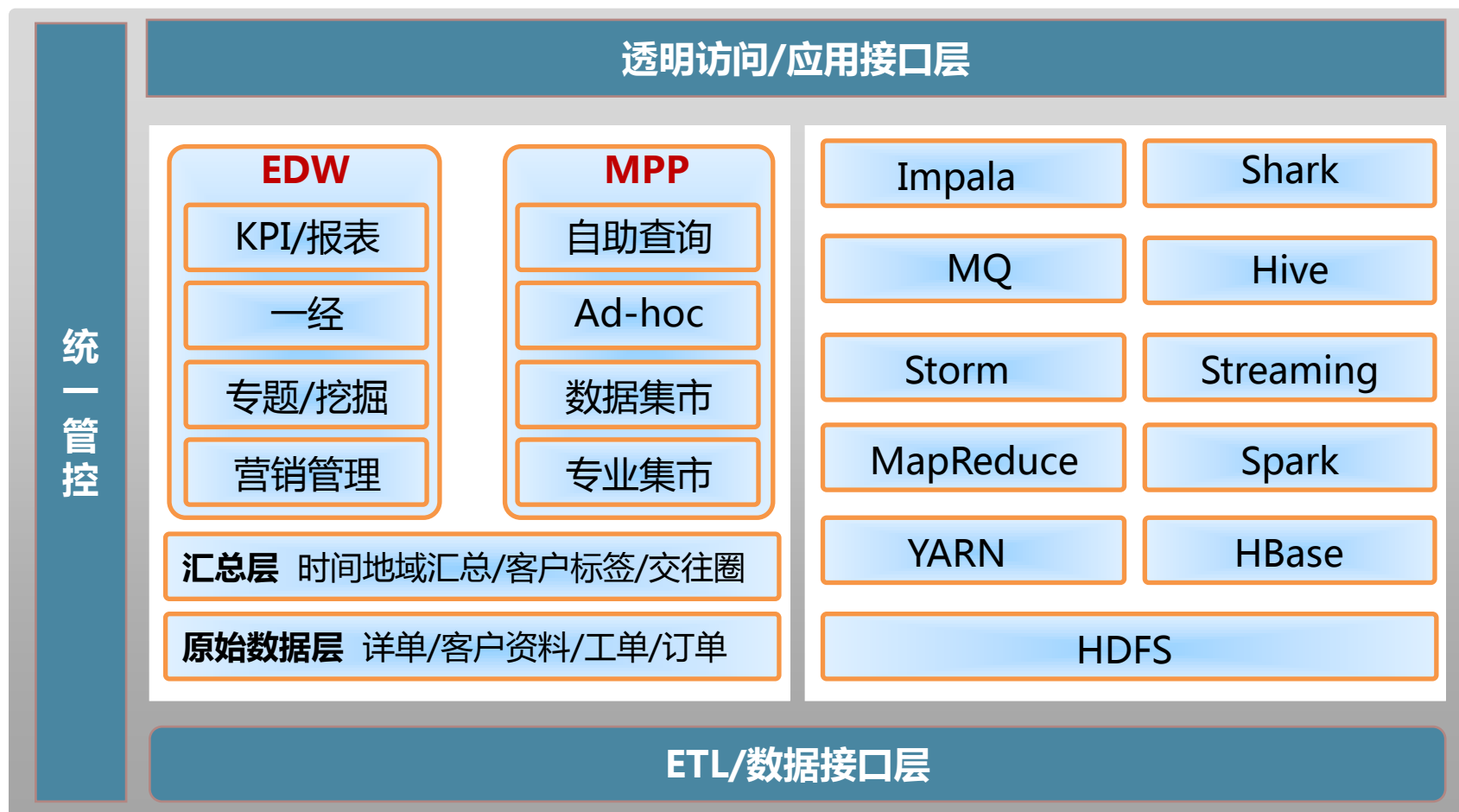


Lambda架构



Spark架构

混搭架构是当前运营商的自然选择，根据数据的热度和存储成本来分布。实现三者的有效融合，以提供最大的计算能力。



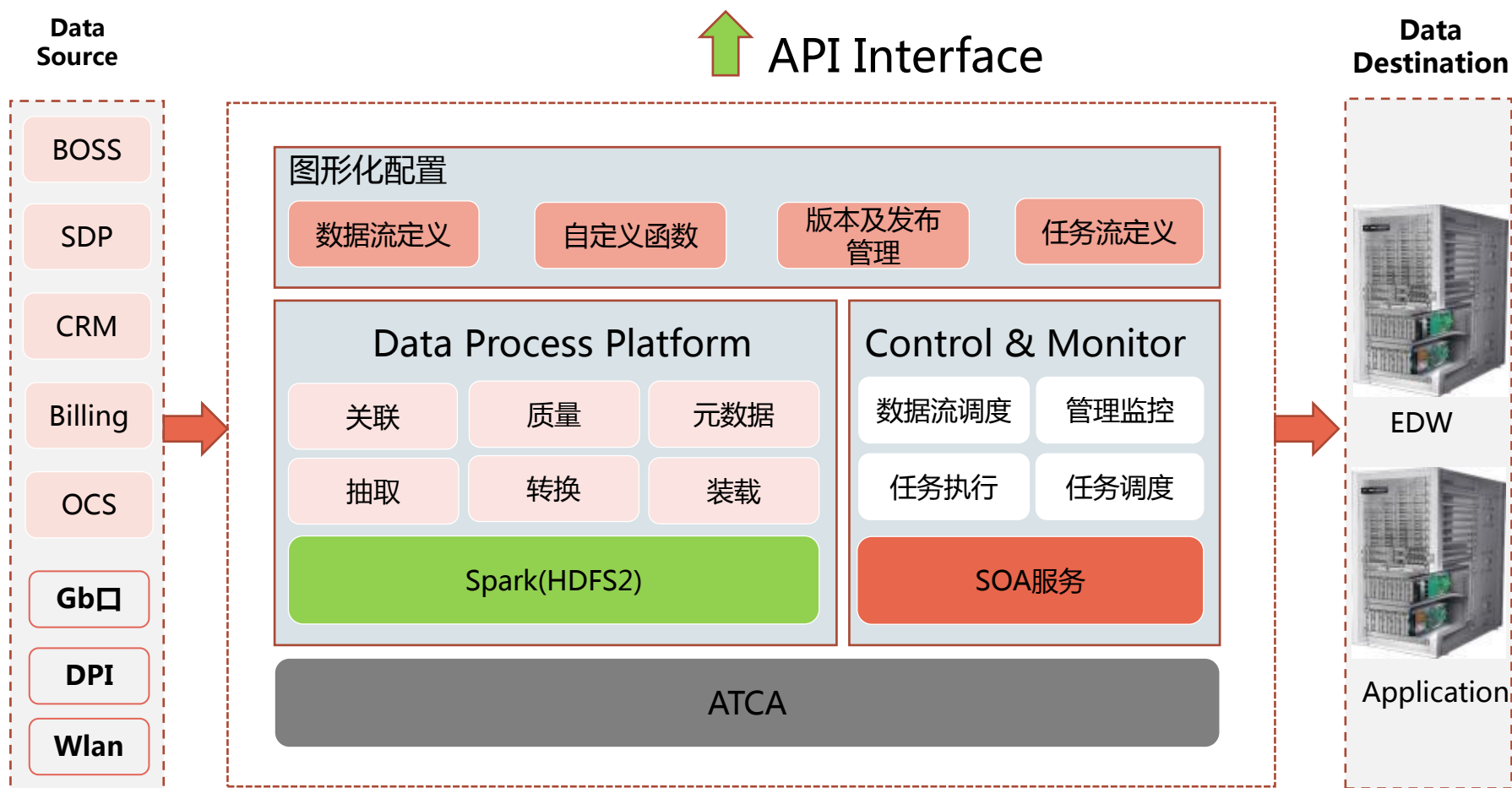
2、ETL – 选择开源产品或者自主研发

Kettle是流行的开源ETL产品，具有可视化设计、支持多种数据接口、具有集群功能、支持二次开发等特点，有着广泛的应用。

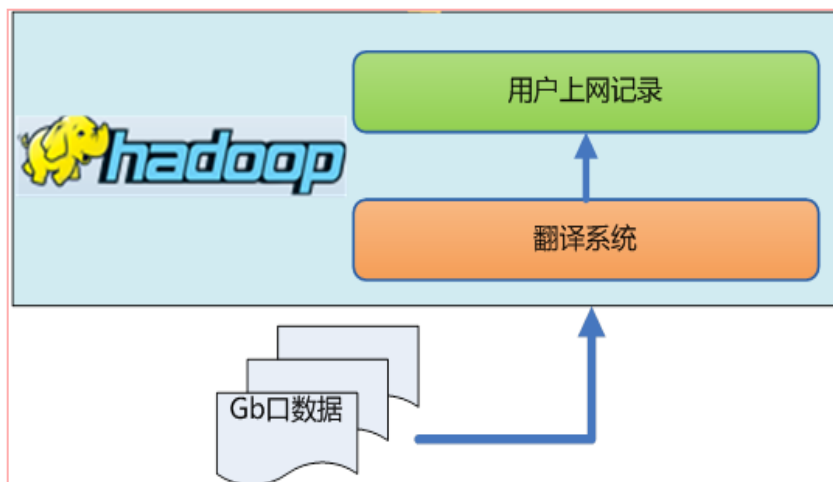
- Kettle有自己的集群，采用的Web Service的方式，基于Jetty实现。Web Service提供转换运行环境，并提供主服务器、从服务器和客户端之间的通信和控制，通信协议是httpClient，采用XML格式。
- Kettle通过插件(Plugins)来扩展，hadoop链接插件在“大数据”下。HDFS文件通过HDFS API读写。
- Kettle通过Hadoop's Distributed Cache来分发数据到HDFS，将job转换为MapReduce的job，从而实现在Hadoop上运行。
- Pandaho labs已经启动“kettle on spark”项目，完成时间未定。
- 对于具有大数据量的通信行业，大都采用专业的ETL工具，提供数据处理的吞吐量。
- 采用Kettle的可视化设计部分，将运行部分迁移到Spark上，是一种可能的方案。不采用系统提供插件式开发，对原架构进行大的修改，开发难度较大。

自主研发ETL – 支持批量和实时接口

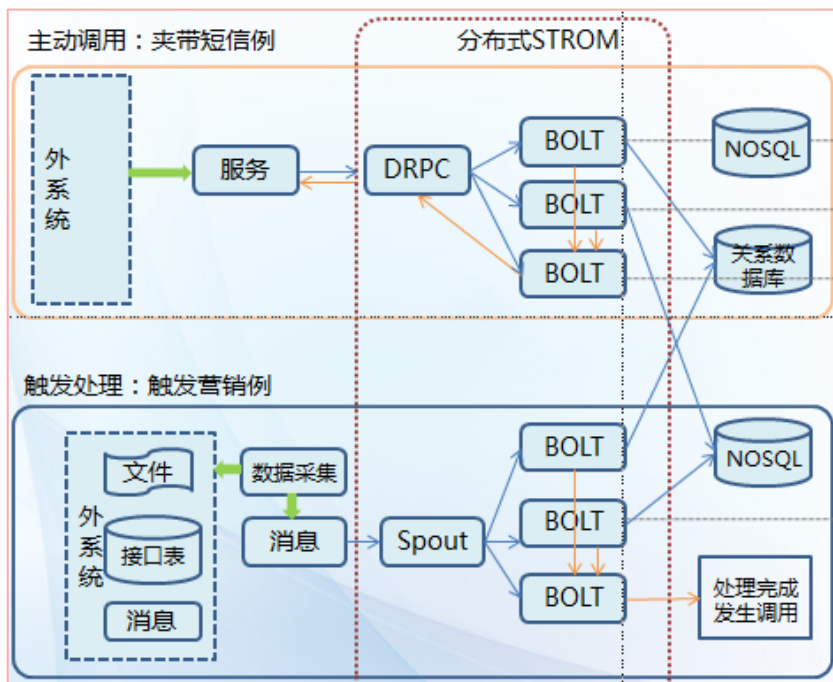
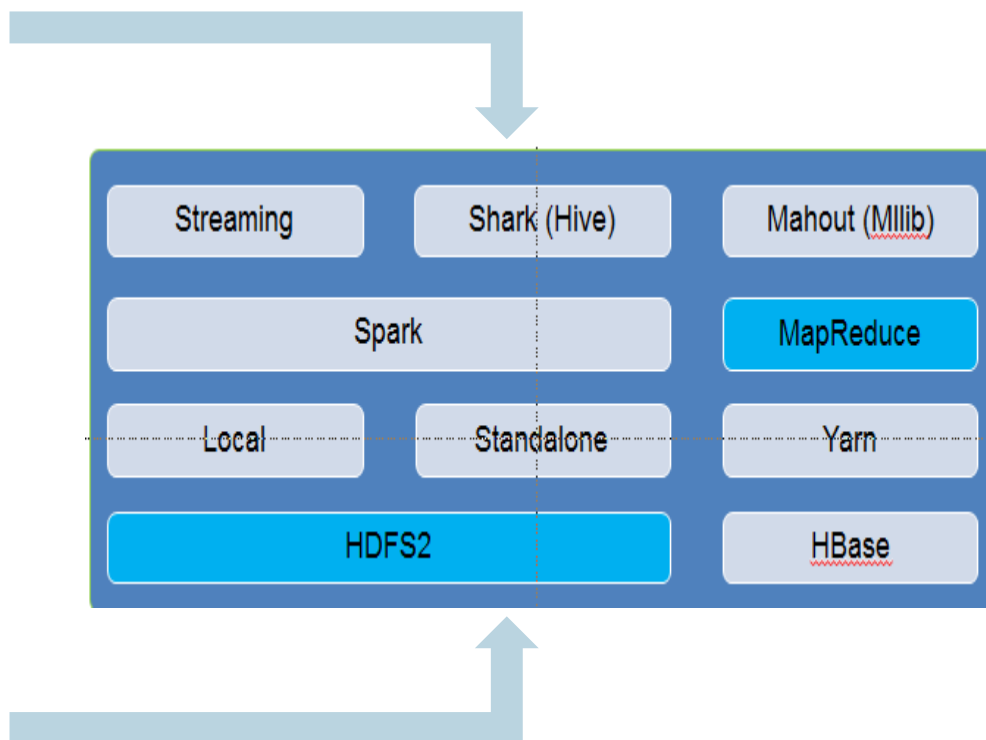
- 1、采用Tableau作为图形化配置和管理工具，将ETL过程、原子处理等转化为Spark的Task
- 2、离线批量接口和实时接口采用同样的配置，只有处理的时间间隔属性不同



3、业务系统 - 改造和对比



改造**流量经营**系统：用Spark替换MapReduce，迁移“翻译系统”的Java代码。目标：提升性能3-5倍。



对比**实时营销**系统：用Streaming替换storm，采用Scala重新开发。目标：原系统对比业务满足能力和性能。

Spark在中国移动省公司 试点应用

- Spark的可用性测试
- Spark的试点应用
- 问答

Q&A

欢迎加盟，一起推广**Spark**在中国的应用！

浩瀚数据 微观分析

广州市浩微数据技术有限公司

GUANGZHOU DATA Insights Tech CO., LTD.

地址：广州市天河区勤天大厦1916A室

邮编：510000

电话：+86 020 85553476

传真：+86 020 85544566

邮箱：trustds@163.com