



大数据实时处理技术及其应用

李松林
2013年4月

Agenda



- 大数据的现状
- 业务场景思考
- 技术架构及其应用
- Q&A

1亿用户

京东拥有众多优质用户，如果他们都是姚明的话，连起来可以绕地球五圈，而且，这个数字每年都在翻倍增长！



100分钟售后满意服务

京东率先在行业推出“售后100分”服务，100分钟内为客户处理完售后问题。



1百万单

在京东，平均每天会产生100万个订单，平均每秒钟超过12个订单，他们来自全球各地，注意，是全球各地啊，我们在超过30个国家都开通了业务。



2亿人次

京东每日页面浏览量高达2亿人次，这个数字相当于2012春运铁路40天发送旅客的总和，注意，京东的是一天。



数百万种商品

京东拥有数百万种优质商品，是您一站式购物的理想商城，在这里，您可以选购到涉及衣食住行的各种商品和服务，甚至是车子、房子……



1百万m²仓储面积

京东在六地建有物流中心，在25个城市的66个仓库，拥有100万平方米的仓储面积，这个数字还在高速增长，未来有用户的地方就有我们的身影！



211限时达极速配送服务

在京东，鼠标一点，商品极速送达，全国125个区县的用户可以享受到上午下单、下午抵达，下午下单、次日抵达的极速配送服务，还有349个区县的用户可以实现次日达服务。



3万员工

在您的快乐购物的背后，有京东30000名员工的汗水和付出，这里面有采销人员、配送员、仓储人员、客服mm、程序猿……



6百亿元年交易额

京东平台去年的交易额超过600亿元，相当于每秒钟有超过2000元的交易额，京东已经连续数年超过150%的快速增长，这背后有广大用户的支持和厚爱。



关于京东 (续)

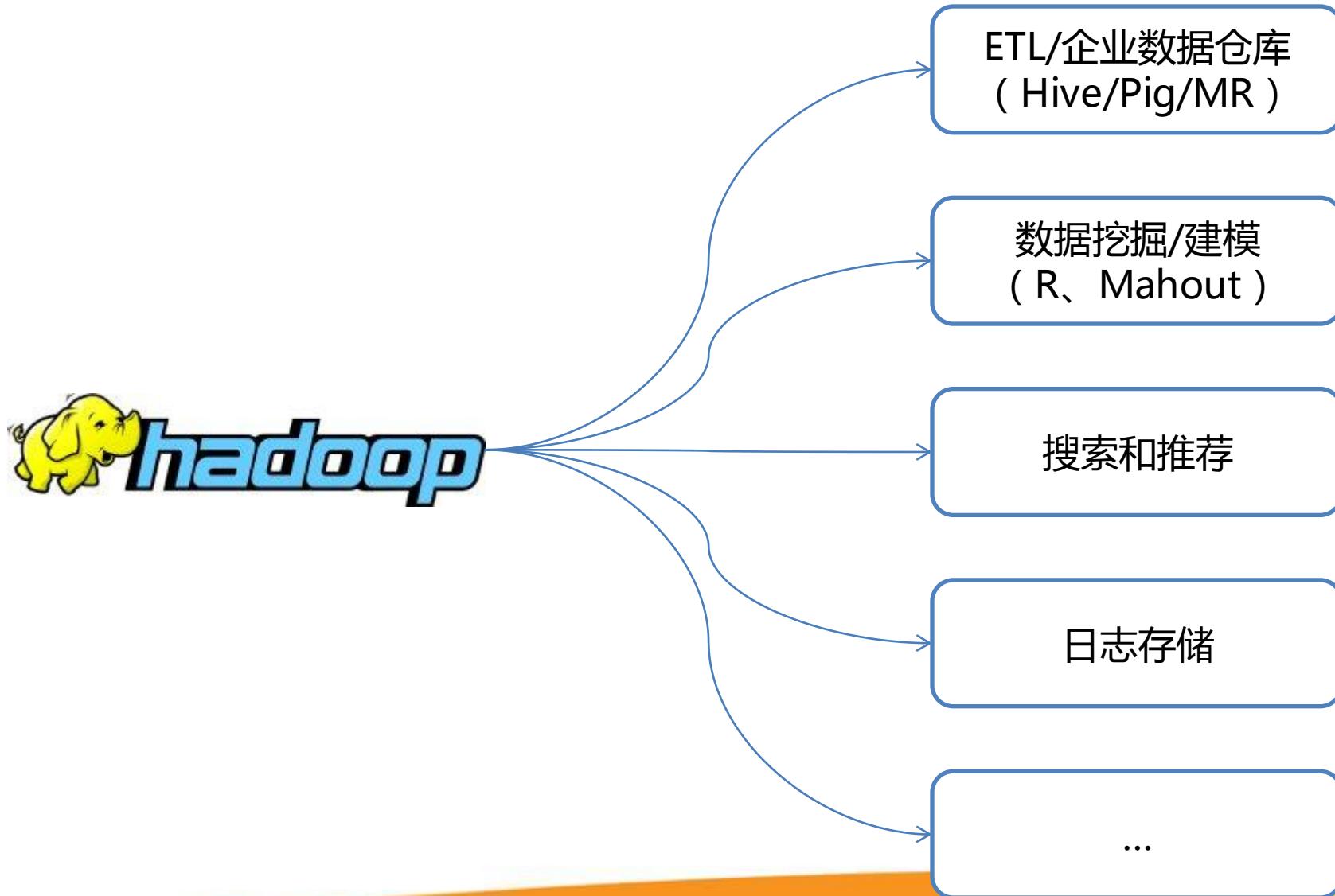


- 京东拥有覆盖企业全部价值链的稳定系统，通过持续优化打造开放平台，全面提升用户体验。



- 3 V : 速度、容量、类型
- 数据的来源更加丰富
 - 商品、订单、…
 - 社交 …
 - 营销 …
- 数据间的关联性更加复杂
 - 人与人
 - 商品与商品
 - …
- 数据的价值
 - 时效性
 - 新商业模式

大规模数据处理更加容易



- “Next Click”
- 运营智能
- 风险控制
- 互动分析
- ...

MapReduce批量处理
=

延迟较长
无法满足用户的实时需求
调度开销较大



大数据包括三部分



	批处理与分析	近实时分析	实时流处理
实时性	离线	准实时/实时	实时
处理时间	分钟到小时	毫秒到秒	持续不断
数据量	TB - PB	GB - TB	持续
编程模型	MapReduce	Queries	DAG
用户	分析师/开发者	分析师/开发者	开发者
成本	中	高	高
应用	ETL/数据挖掘/ 预处理…	数据决策分析 /…	…

大数据实时处理的思考

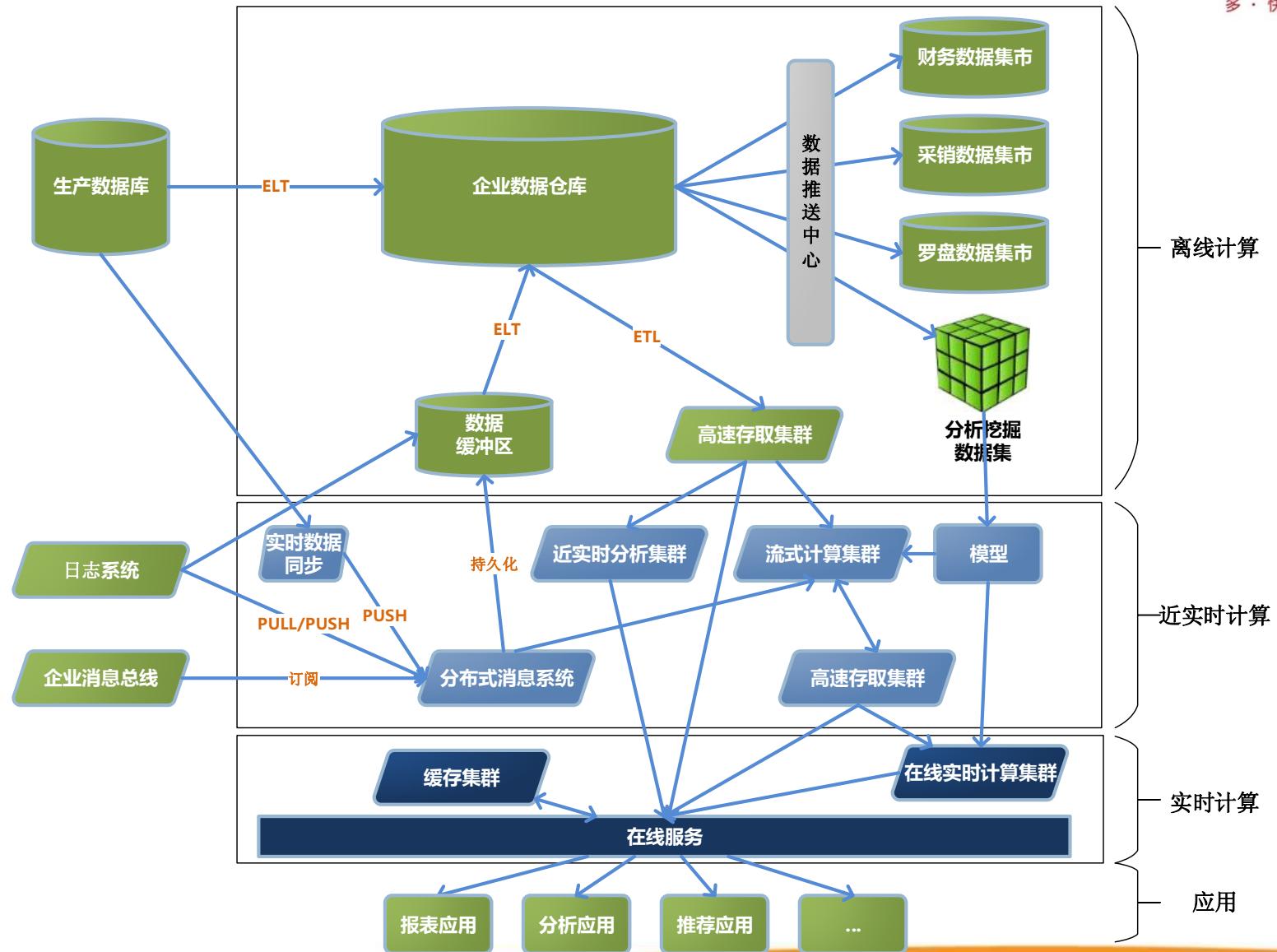


- 模型
 - 海量数据
 - 数据量大
 - 并发数高
 - 多个数据源整合
 - 预定义好的数据模型
 - 去规格化
 - 数据任务依赖关系简单
 - 推和拉的问题
 - 拉比推好

- 性能
 - 高并发需求
 - 大容量需求
 - GB - TB 级后台数据处理吞吐
 - 高速度需求
 - 从数据产生到处理完成结果延迟要求到秒级
 - 计算需要在短时间内完成
 - 批处理预算
 - 硬件支持
 - 内存、CPU、网络
 - 容错
 - 水平扩展

- 服务
 - 关联获取价值，维度按需定制
 - 互动分析、报表等完成价值交付
 - 与其他在线生产系统进行数据对接（数据反哺）
 - 计算即服务

大数据实时处理架构

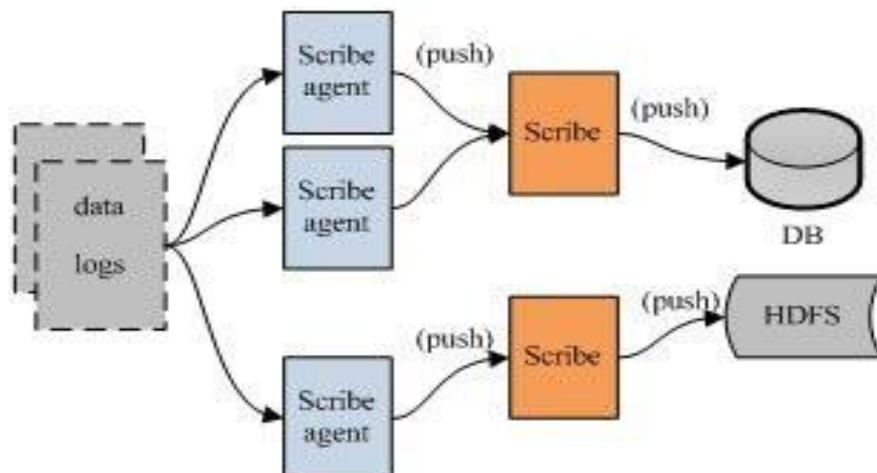
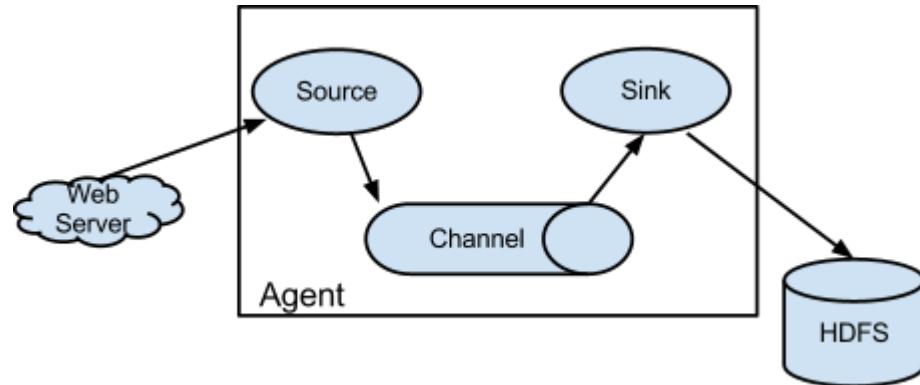


大数据实时处理技术

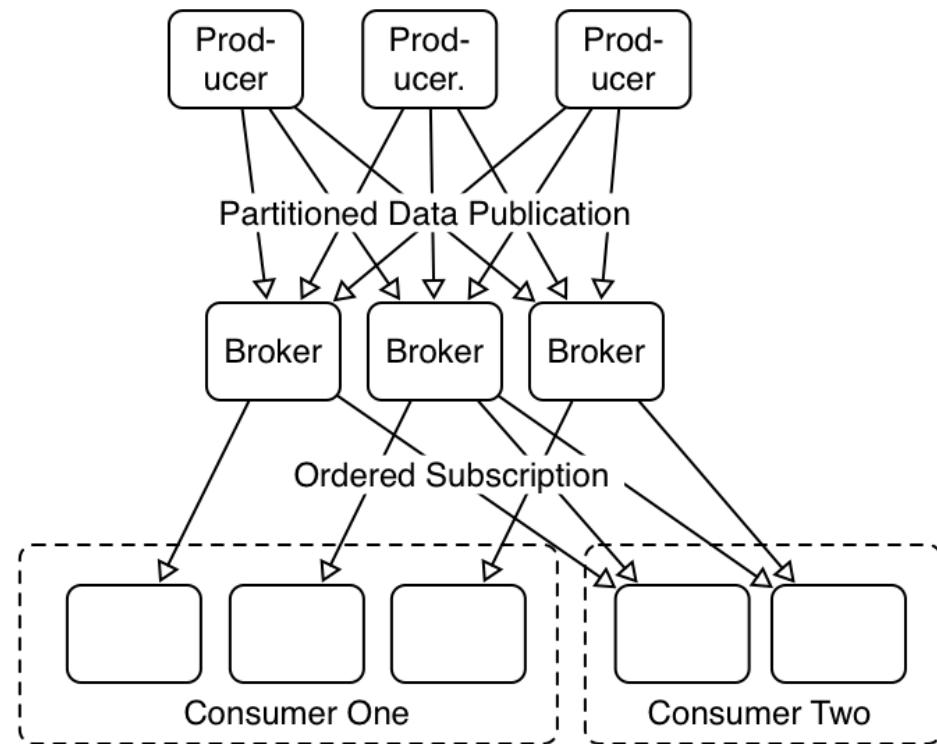


JD.COM 京东
多·快·好·省

- 数据传输
 - 日志（用户行为、...）
 - 批量同步
 - 消息队列
 - ...
- 开源技术
 - Flume
 - Scribe
 - Kafka
 - ...



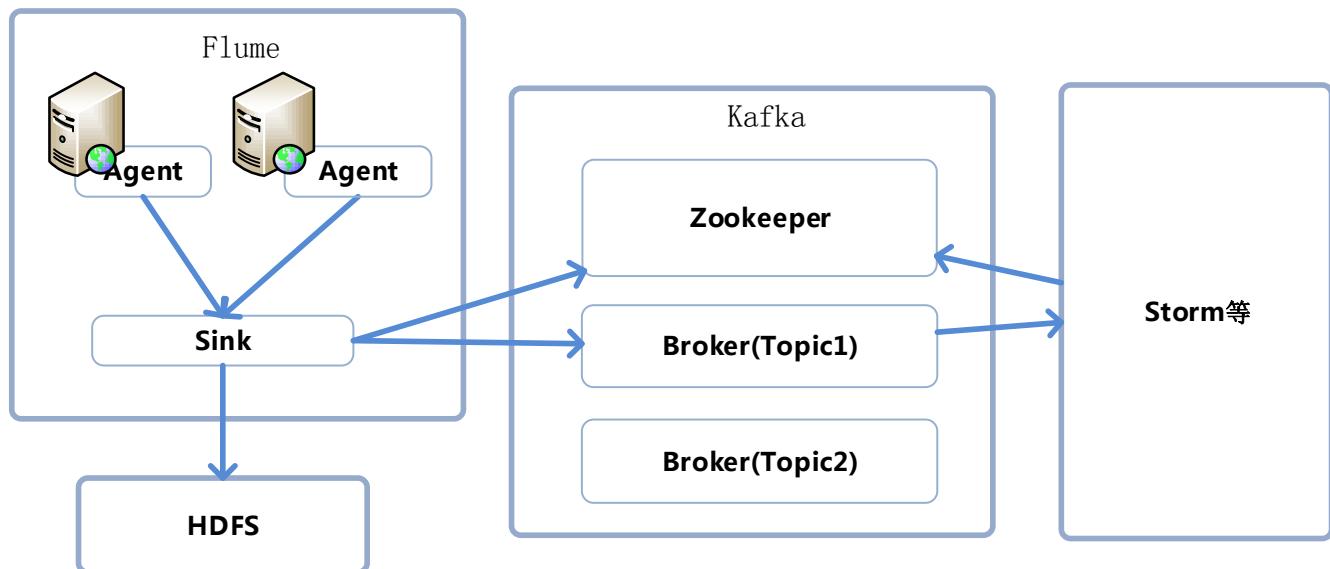
- Apache项目: <http://kafka.apache.org/>
- 一个分布式的发布/订阅消息系统
- 术语
 - Topics
 - 消息分组
 - Brokers
 - 消息存储
 - Producers
 - 消息生产者
 - Consumers
 - 消息消费者



大数据实时处理技术



- 几个点
 - 解耦
 - 缓冲
 - 容错
 - 透明
 - 跨数据中心数据分发



大数据实时处理技术



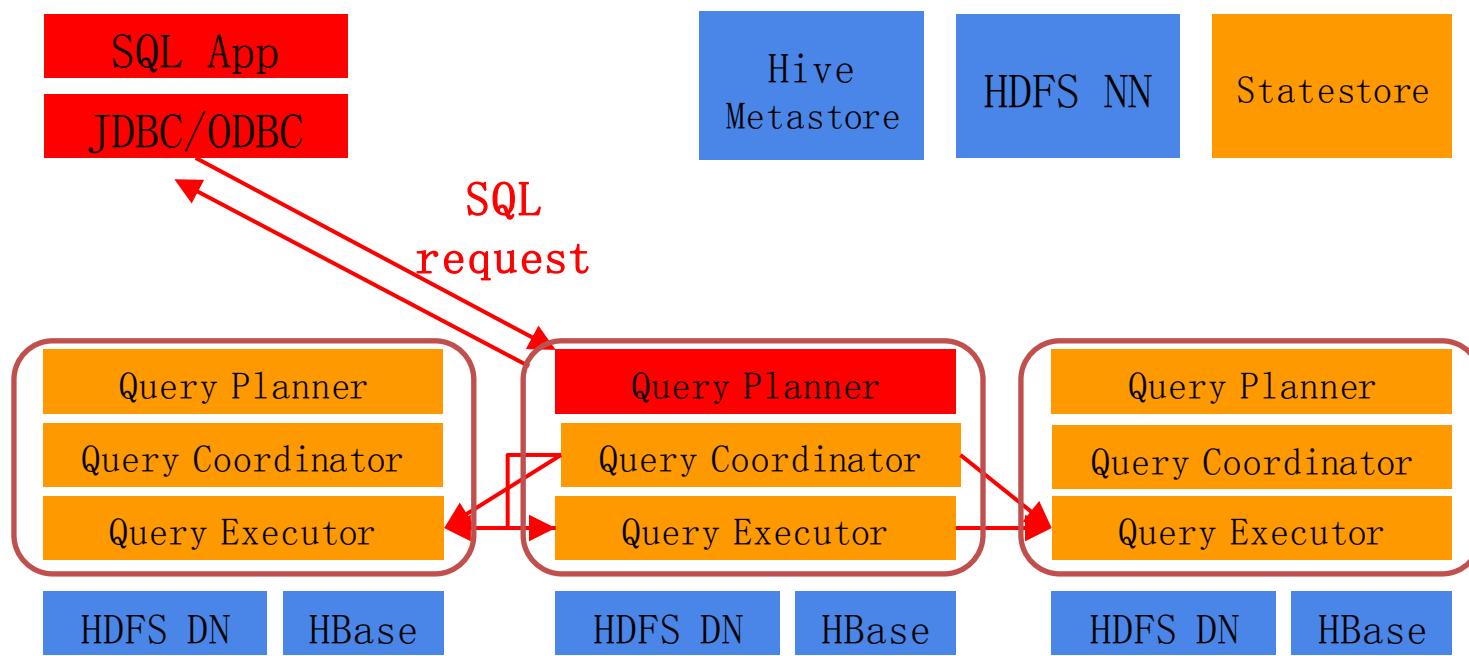
- 存储
 - 大容量低速存储
 - 高速存储
 - KV存储
- 开源NoSQL数据存储
 - Hadoop
 - HBase
 - Cassandra
 - MongoDB
 - Redis
 - ...
- 数据库Sharding
- 合适的就是最好的

大数据实时处理技术

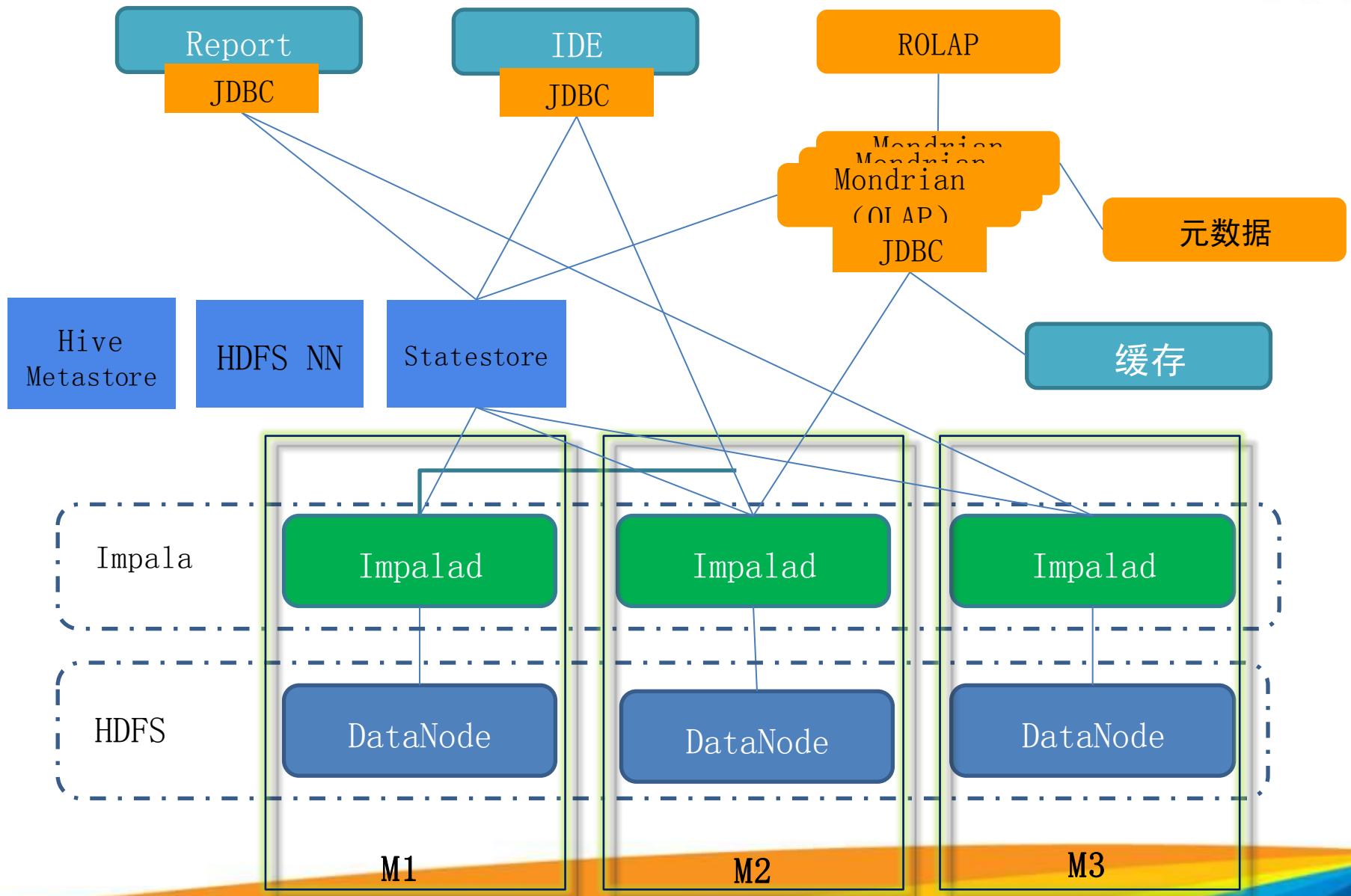


- 计算
 - 可加计算、不可加计算
 - 实时数据的实时计算
 - 实时数据的计算
 - 数据的实时计算
- 开源计算框架
 - Storm
 - Impala
 - ...

- Cloudera公司贡献
- 一种通用的SQL查询引擎（Hive语法）
- 与Hadoop整合在一起



大数据实时分析



大数据实时分析



Cubes

order_ana1

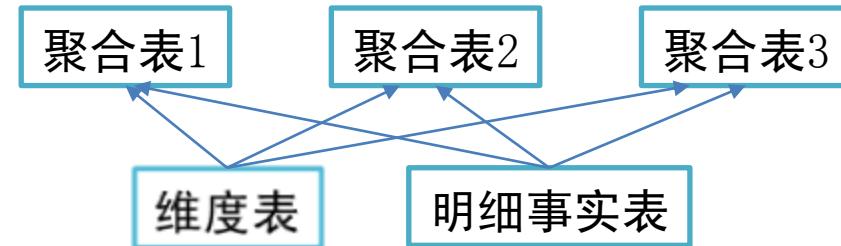
Dimensions

- 主供应商
- 分公司
- 商品
- 手机客户端
- 父销售订单
- 用户
- 配送方式

Measures

Measures

- (SKU运费金额)
- (优惠前单价)
- (优惠前金额)
- (优惠后单价)
- (优惠后金额)
- (其它优惠金额)
- (单品促销优惠金额)
- (团购优惠金额)
- (总优惠金额)
- (销售数量)



Columns (其它优惠金额) ◆ (总优惠金额) ◆ (销售数量) ◆

Rows Subd_Num ◆ Delv_Way_Cd ◆

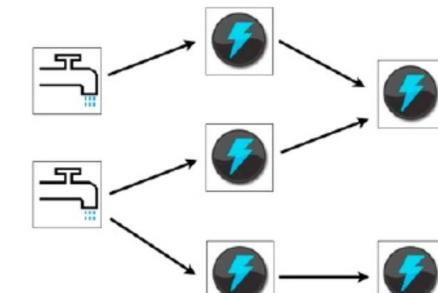
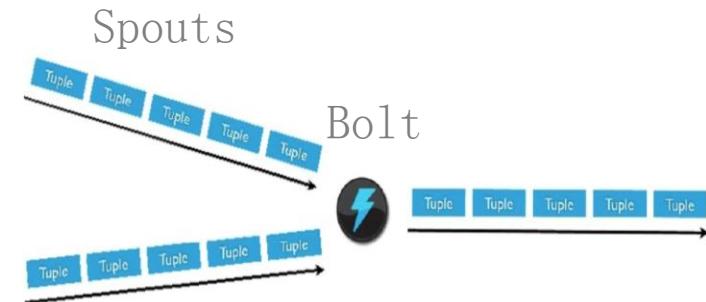
Filter

Subd_Num	Delv_Way_Cd	(销售数量)	(总优惠金额)	(其它优惠金额)
10	0	1	240	0
	1		0	0
2933		1	89	0
	4		156.1	0
64		9,509	383,907.5	0
70		1,234,903	59,896,441.61	0

流式计算Storm



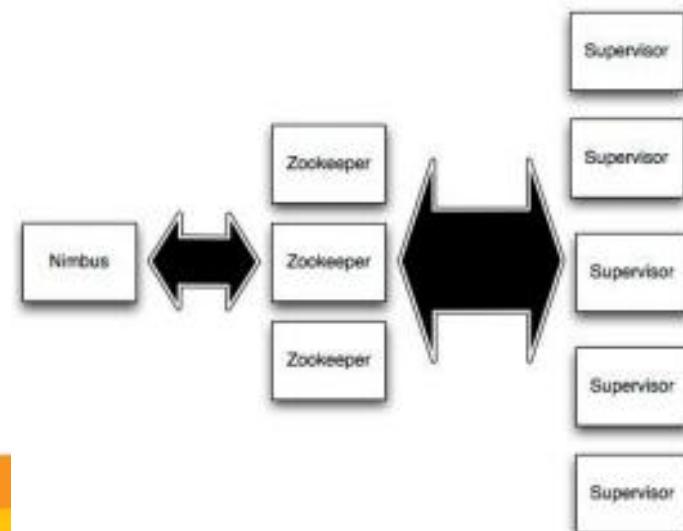
- Twitter开源的分布式处理框架
- 基本概念
 - Streams (流)
 - 元组序列
 - Spouts
 - 流的源头
 - Bolts
 - Functions, Filters, Joins, Agg
- 优点
 - 可扩展、容错、易用…
 - 在内存中执行



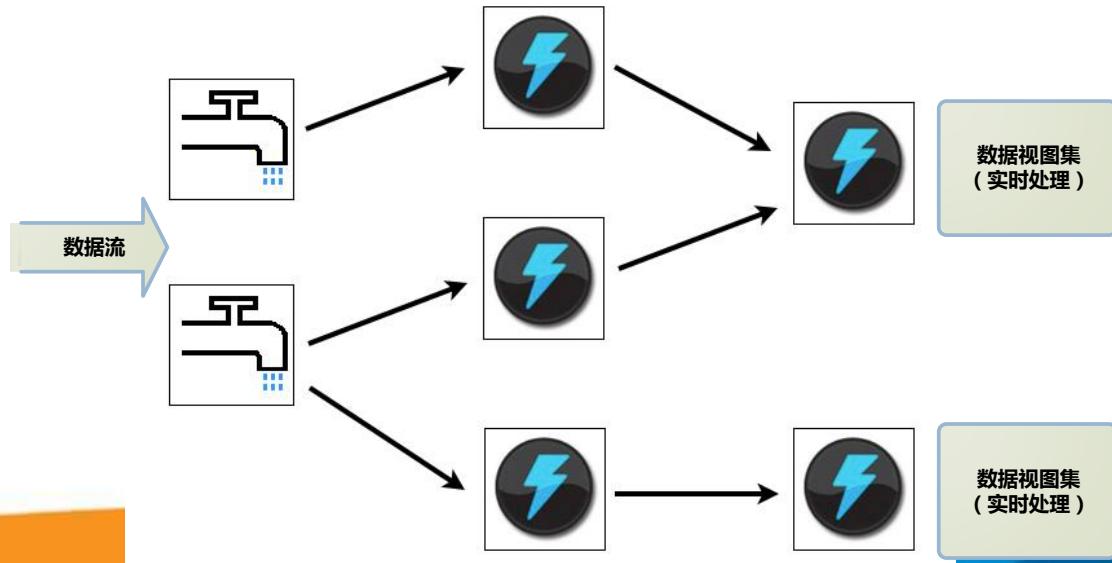
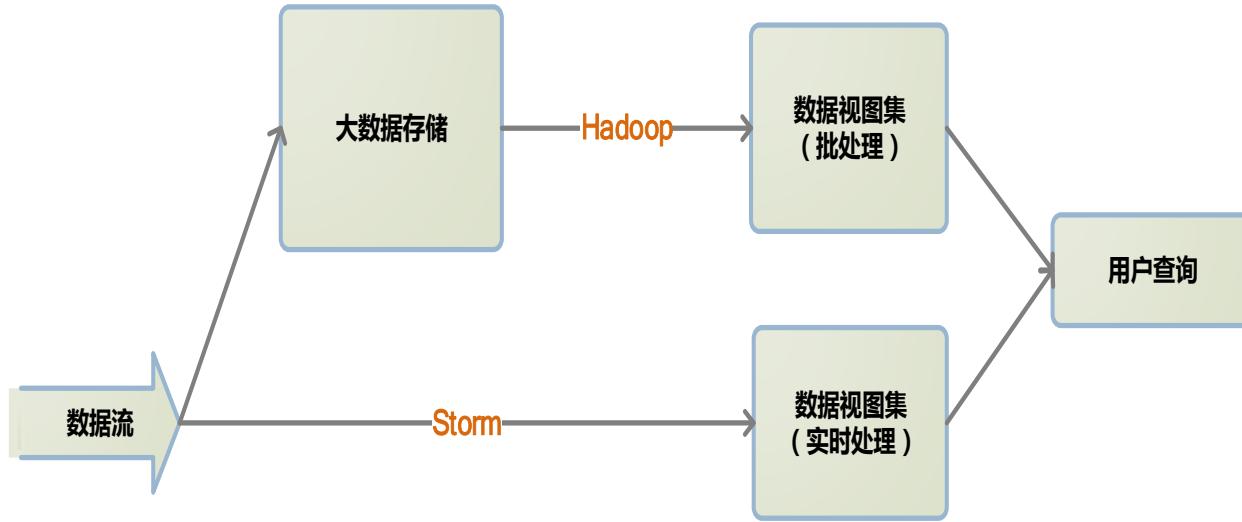
Topologies

Storm的部署

- Nimbus
 - 主控节点，用于任务分配，集群任务监控等
- Zookeeper
 - 集群中协调，共有数据的存放(如心跳信息)
- Supervisor
 - 对应一台物理机，用于启动worker
- Worker
 - 工作进程，负责启动task，以及通过zeromq进行tuple的分发，与接收。
- Task
 - 工作线程，任务的处理



Storm的应用模式



流式计算 (续)



流式计算 (续)



- 事件驱动实现



最后更新时间: ..

一天内搜索词排名

排名	搜索词(搜索次数)	排名	搜索词(搜索次数)
Top1	羽绒服 (38936)	Top11	电磁炉 (6482)
Top2	加湿器 (21608)	Top12	保温杯 (6467)
Top3	耐克 (11187)	Top13	热卖冬装 (5824)
Top4	洗衣机 (10935)	Top14	靴子 (5733)
Top5	手机 (9701)	Top15	羽绒服女款中长款 (5641)
Top6	耐克男鞋 (8378)	Top16	热水器 (5371)
Top7	ipad (8044)	Top17	吹风机 (5307)
Top8	耳机 (7553)	Top18	夹克 (5255)
Top9	冲锋衣 (7541)	Top19	短靴 (5233)
Top10	微波炉 (7018)	Top20	iphone4s (5178)

一分钟内搜索词排名

排名	搜索词(搜索次数)	排名	搜索词(搜索次数)
Top1	皮衣 (20)	Top11	平板电脑 (9)
Top2	自行车 (16)	Top12	皮草 (9)
Top3	鼠标 (16)	Top13	三星 (9)
Top4	u盘 (15)	Top14	保暖内衣 (9)
Top5	无线路由器 (14)	Top15	保温壶 (9)
Top6	羽绒服女 (12)	Top16	韩都衣舍 (9)
Top7	键盘 (10)	Top17	笔记本 (8)
Top8	羽毛球 (10)	Top18	笔记本电脑 (8)
Top9	夹克 (9)	Top19	包 (8)
Top10	雅诗兰黛 (9)	Top20	休闲高帮 (7)

流式计算 (续)



- 注意
 - 内存泄露
 - 消息堆积
 - 算法模块拆分



- 服务和应用
- 价值展现
 - 分析可视化
 - 数据可视化
 - 数据反哺
 - 计算即服务
- 仔细思考其价值
 - 实时的统计：最流行
 - 广告CTR预测
 - ETL：格式转换、重复值过滤、…
 - 运营需求：资源调派
 - ...

- 对系统的压力
 - 数据量
 - 数据展现
 - 数据读写和传输
- 解决方法
 - 前端和后端解耦
 - 缓存的应用
 - JS发挥前端的能力
 - 压缩
 - 排队
 - 异步、非阻塞I0模型
 - 线程池
 - 事件驱动
 - 后端更强劲
 - 数据库集群：分库、分表、分区
 - NoSQL数据库：Hbase、MongoDB等



谢谢

@随风逐梦12345

[HTTP://WEIBO.COM/TECHFASHION](http://WEIBO.COM/TECHFASHION)