



Alibaba Developer
Conference

淘宝网商品管理 技术实践之路

About Me

淘宝网-产品技术部
交易平台-商品平台-商品管理

姓名：刘抚狄
花名：府笛

@TB府笛

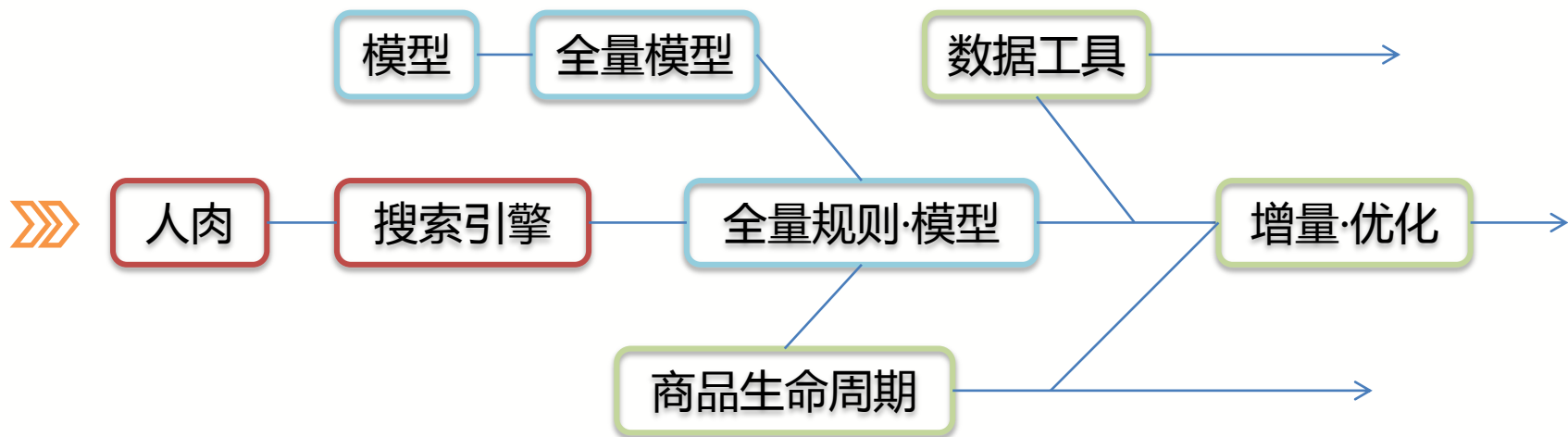
内容提要

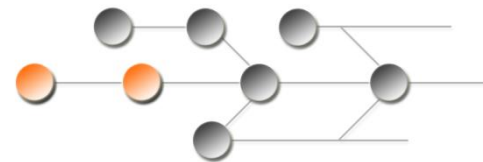
技术挑战
商品管理系统的演化过程
技术细节
展望

技术挑战@淘宝商品管理

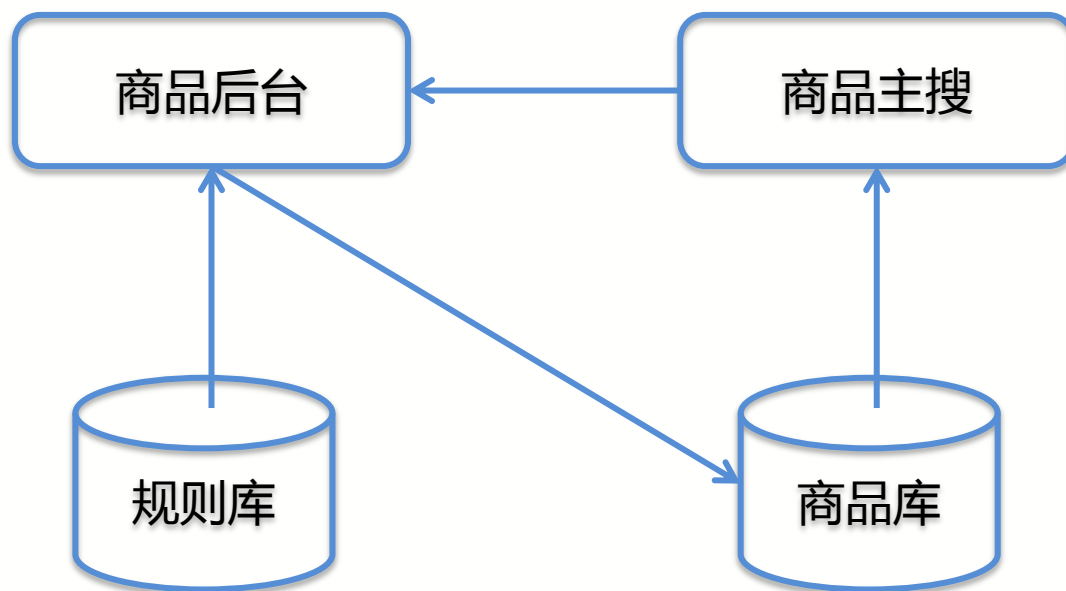
十亿级商品数
百万级用户数
每天处理TB级数据
数据沉淀
成本控制
业务多变
上千条业务规则

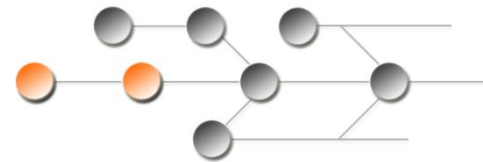
发展过程@淘宝商品管理





第一阶段：基于搜索





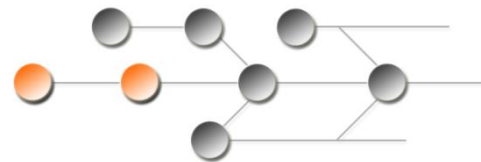
基于搜索的商品管理

› 实现

存储用户设置的规则，生成Query
定时任务触发，多进程/线程执行

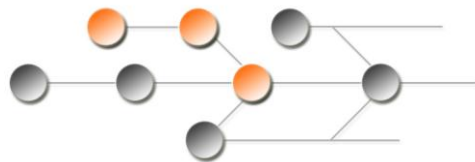
› 特点

作为搜索的子集，所见即所得
方案简单，快速满足需求

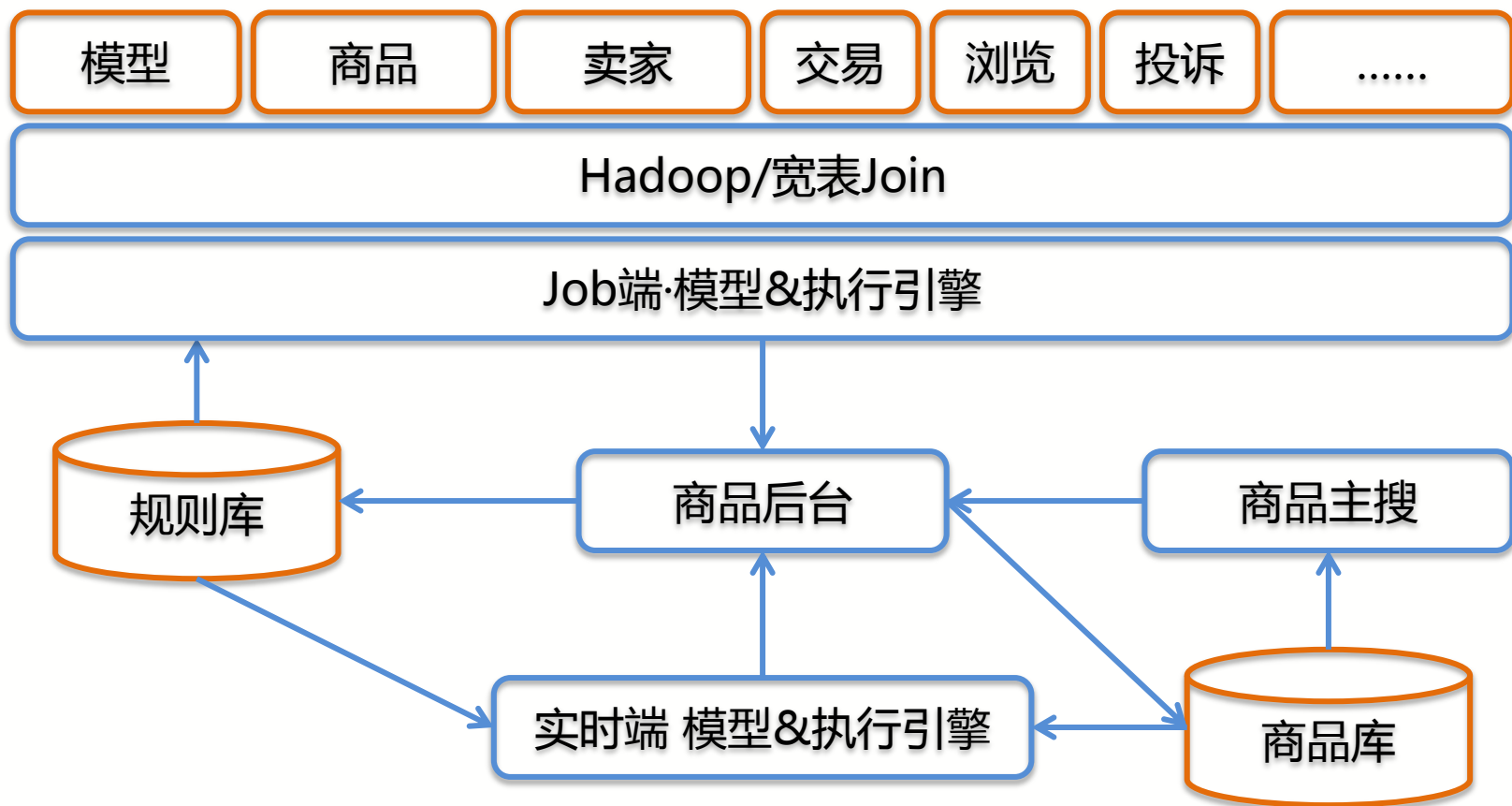


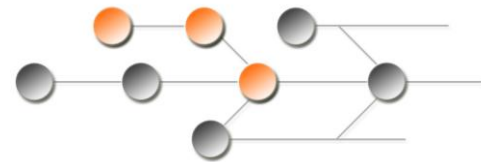
第一阶段：改进

- ▮ 翻页查询，时间长，成本高，引擎鸭梨山大
- ▮ 索引优化、预排序导致的截断，数据不全
- ▮ 指标有限，在不影响主流业务的情况下难以扩展
- ▮ 更换数据源，思考如何才能进行全量管理
- ▮ 开始考虑引入模型和算法资源，降低人肉成本



第二阶段：全量





基于Hadoop的全量商品管理

》实现

全量部分：Hadoop + 宽表 + DSL

实时部分：Notify(MQ) + DSL + 实时构造宽表

保留搜索引擎逻辑

》特点

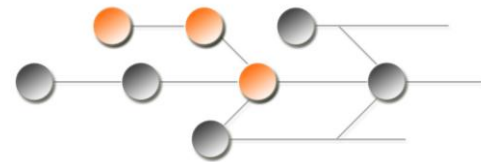
全网所有数据都在全量覆盖范围

术业专攻，发挥BI、网安、反作弊等团队技术优势

大量使用在搜索引擎中不提供的指标

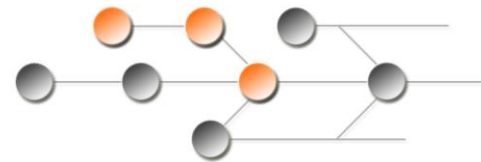
更好的发挥DW、BI资源作用

能够实现较为复杂的业务逻辑



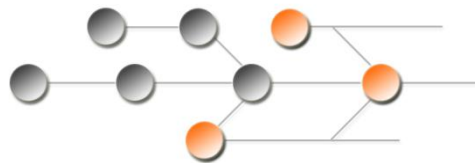
第二阶段：技术细节

- › 引入Hadoop进行全量数据处理
 - › 调度系统
 - › 通用Join任务
 - › SST的宽表数据结构实现
 - › DSL和执行引擎
-
- › why not HIVE?

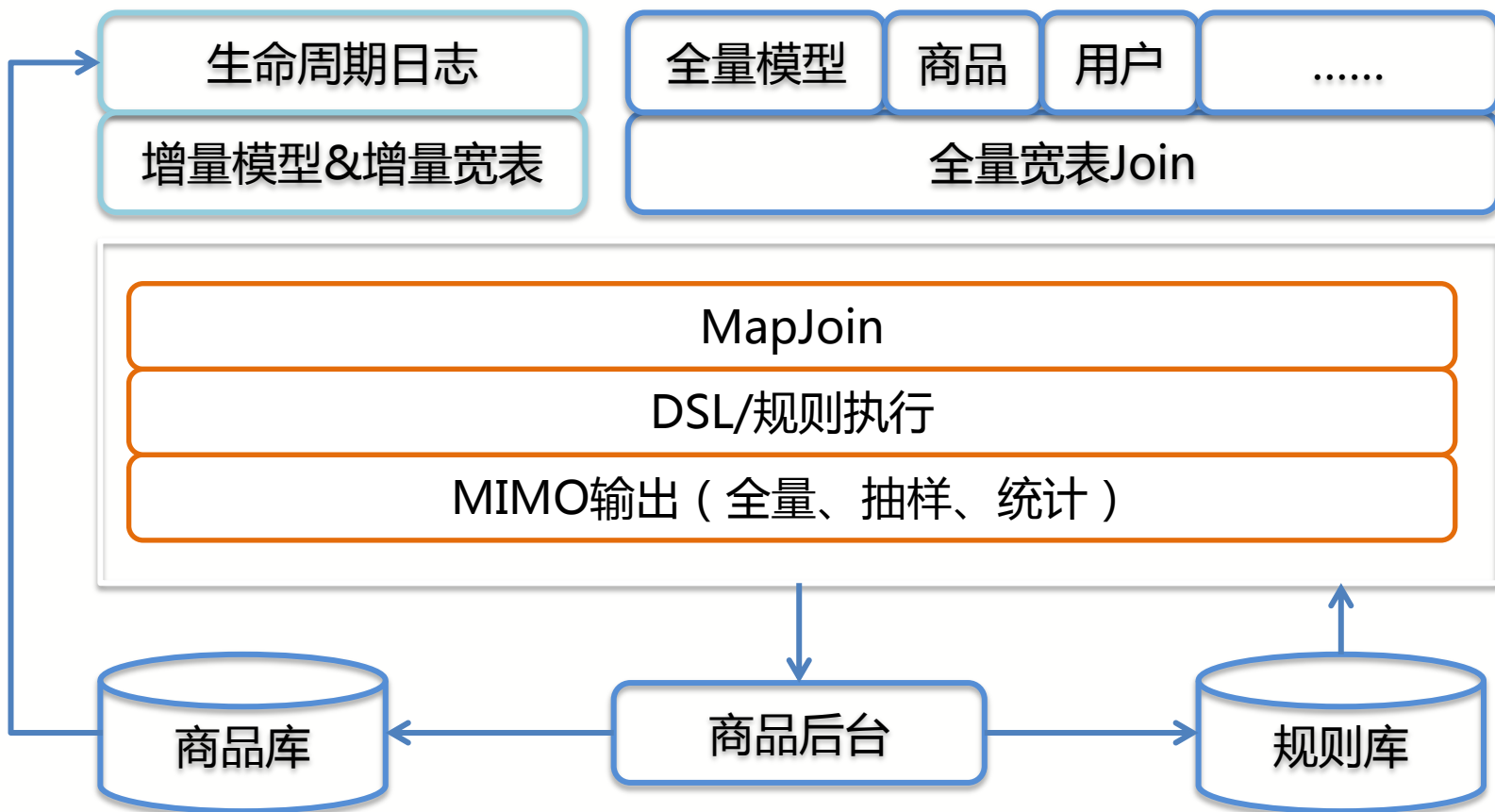


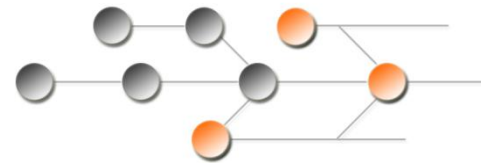
第二阶段：改进

- ▮ 全量数据源更新延时长
- ▮ 搜索查询优化的业务风险
- ▮ 实时端的业务描述能力有限
- ▮ 丰富的基础工具，但没有深度整合
- ▮ 应当引入帮助卖家的前端产品



第三阶段：增量





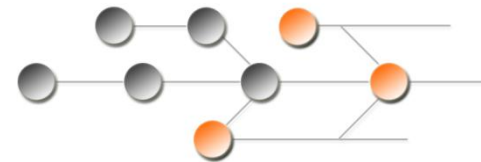
基于Hadoop的增量商品管理

》 实现

不再使用搜索引擎实现和原先的实时系统
通过消息队列接收变更，Dump到HDFS
经过整理的Dump文件以MapJoin的方式与全量
宽表Merge
对于需要三个Job输出的抽样、全量和统计数据，
借助MIMO在一个Job完成

》 特点

单一系统，覆盖增量和全量



第三阶段：技术细节

- › 商品生命周期系统
- › MapJoin
- › MIMO
- › 数据工具

技术@淘宝商品管理

- › 通用Join和宽表
- › 宽表和SST
- › DSL
- › 商品生命周期系统
- › MapJoin
- › MIMO
- › 数据工具

技术细节：通用Join和宽表

- › 处理多种不同格式的数据源
- › 层级Key
- › 性能优化
- › 基于SequenceFile或MapFile
- › MultipleInputs → MIMO

层级表逻辑结构

用户ID=1	用户数据	
用户ID=1	商品ID=11	商品数据
用户ID=1	商品ID=22	商品数据
用户ID=1	商品ID=33	商品数据
用户ID=2	用户数据	
用户ID=2	商品ID=12	商品数据
用户ID=2	商品ID=45	商品数据
.....		

技术细节：SST

- › Static Search Table
 - › 只读，全局有序
 - › 宽表+二分索引，索引数据分离
 - › 多级聚集存储，多级KeyValue查询
 - › 无需回流数据
 - › 实现FileFormat接口
-
- › 存储Join结果，或者规则输出
 - › 典型的SST应用，100GB压缩数据产生约20MB索引

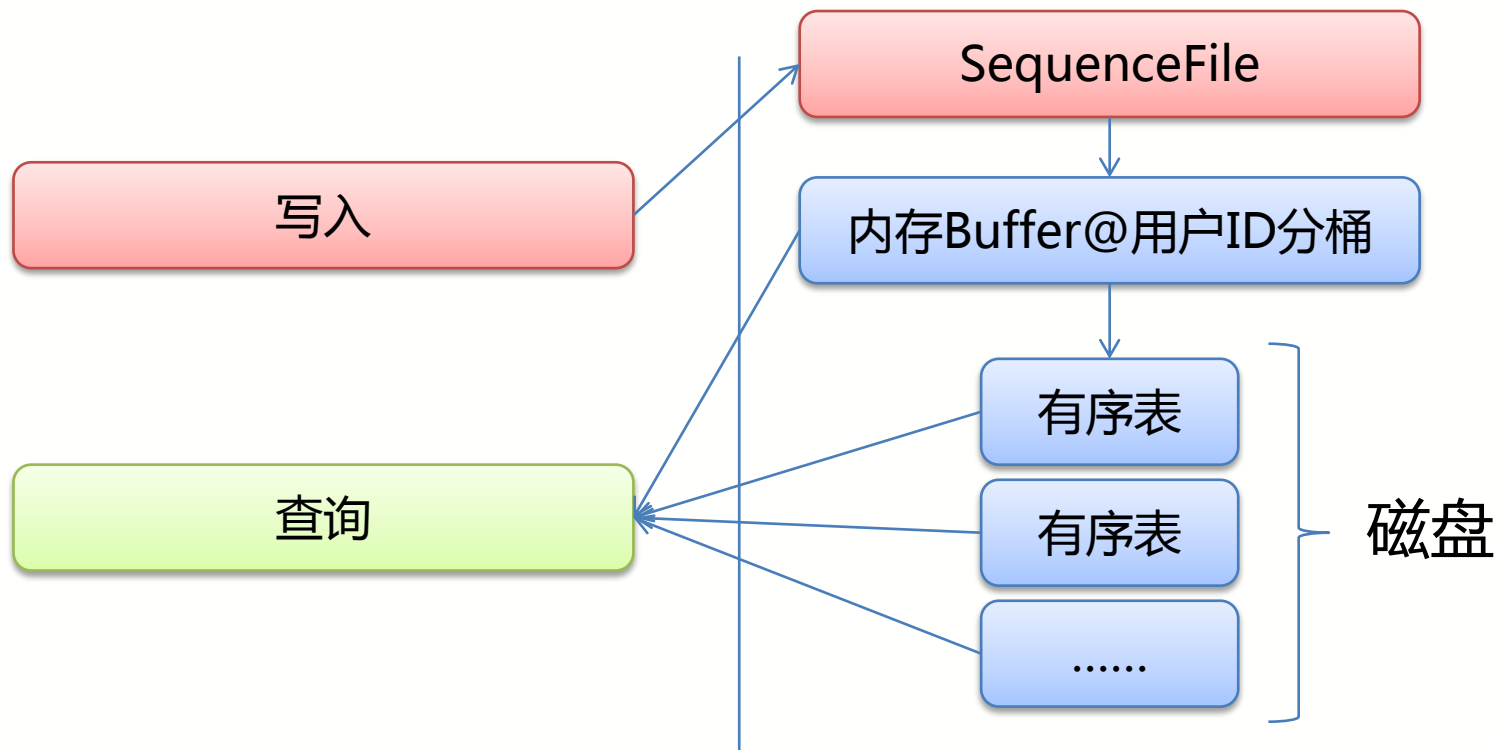
技术细节：DSL

- ▮ 语法糖
- ▮ 类似Java语法
- ▮ 强类型系统
- ▮ 共享变量
- ▮ 语法静态检测
- ▮ 在宽表上下文中帮助快速预判
- ▮ 用于简单的数据判断和处理逻辑，如筛选、报表

技术细节：商品生命周期系统

- ▮ 每天1E+条记录，20GB+日志，存储200天
- ▮ 每秒写峰值3k-5k，设计目标20k
- ▮ 基于SST
- ▮ 成本：两台虚拟机 + 4TB HDFS
- ▮ 扩展：商品快照系统
- ▮ 用于实时后台查询、商品增量数据DUMP

实时读写结构



全量只读结构：时间倒序+用户ID索引

用户ID	时间+商品ID	变更记录	时间+商品ID	变更记录	时间+商品ID	变更记录
用户ID	时间+商品ID	变更记录	变更记录			
.....						

全量按存档日期分段存储

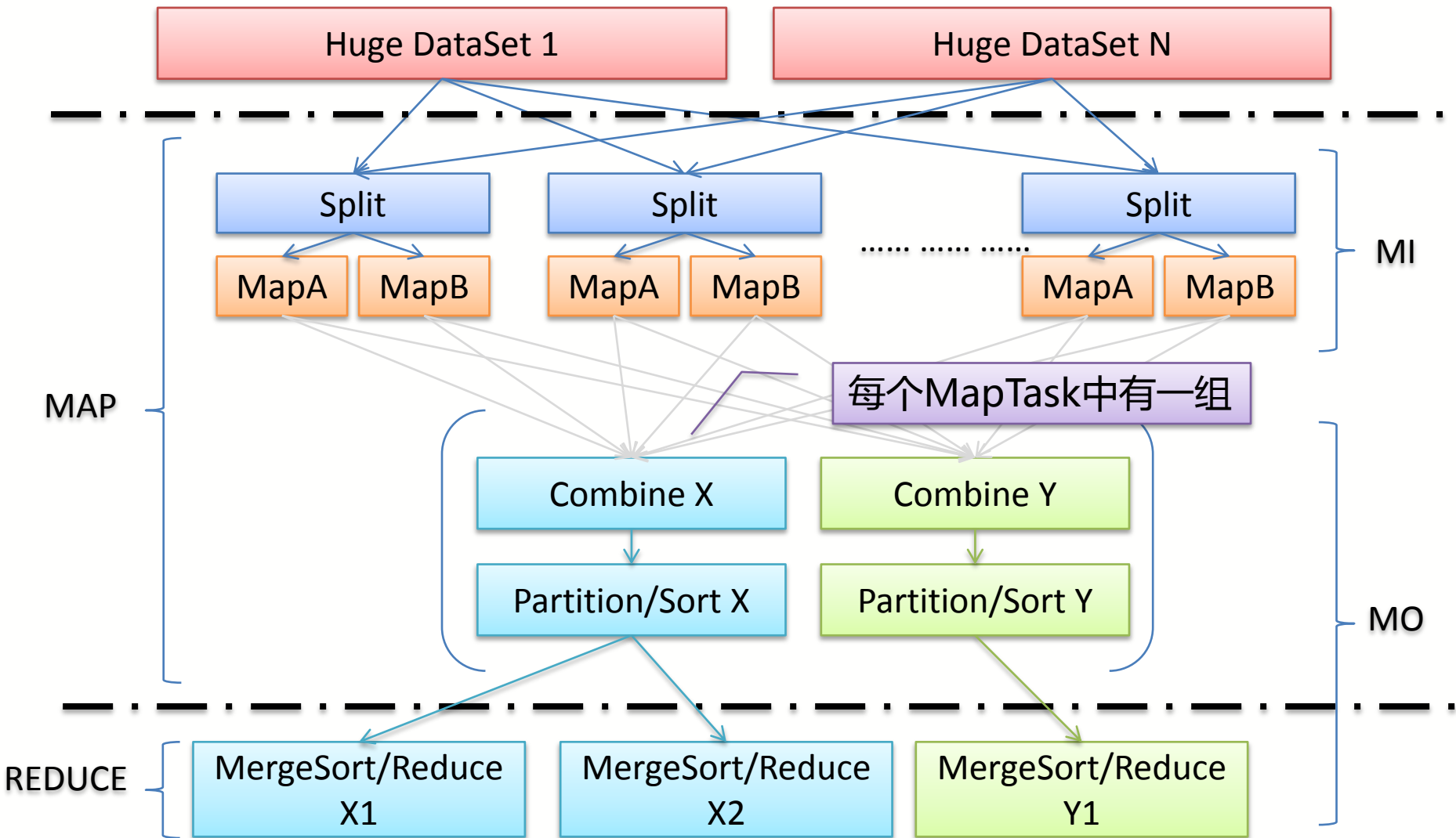
3	7	14	28	28	28	...
---	---	----	----	----	----	-----

技术细节：MapJoin

- ▮ 两张百GB级别表，在Map端完成归并和业务计算
 - ▮ 利用有序表特性，多路归并
 - ▮ 有序表支持KV-Seek
- ▮ 场景：全量数据和增量数据合并，运行规则，扫描过程在Map端完成，避免Hadoop通过Reduce过程归并数据；虽然无法免去网络传输，但极大的减少了增量任务的运行时间

技术细节：MIMO

- › MultipleInputs/MultipleOutputs的扩展
- › 处理不同格式、路径输入
- › 提供更好用的TaskContext
- › 独立的
 - PartitionerClass
 - CombinerClass
 - ReducerClass
 - OutputFormat
 - NumReducers
 -
- › 用于一个Job产出多组不同格式和数量的数据



技术细节：数据工具

- › 基于DSL、SST、MIMO等组件
- › Reporter 模式的实现
 - sum(“库存总数” , item_stock_quantity)
 - max(“最高价格” , item_price)
 - average(“平均价格” , item_price)
- › 通过MIMO定制Combiner、Reducer为筛选任务开辟一条数据统计通道
- › 场景：报表、影响评估、数据分析等

展望

- ▮ 实时
- ▮ 通用
- ▮ 开放
- ▮ 体验

谢谢大家！



Q&A

淘宝网