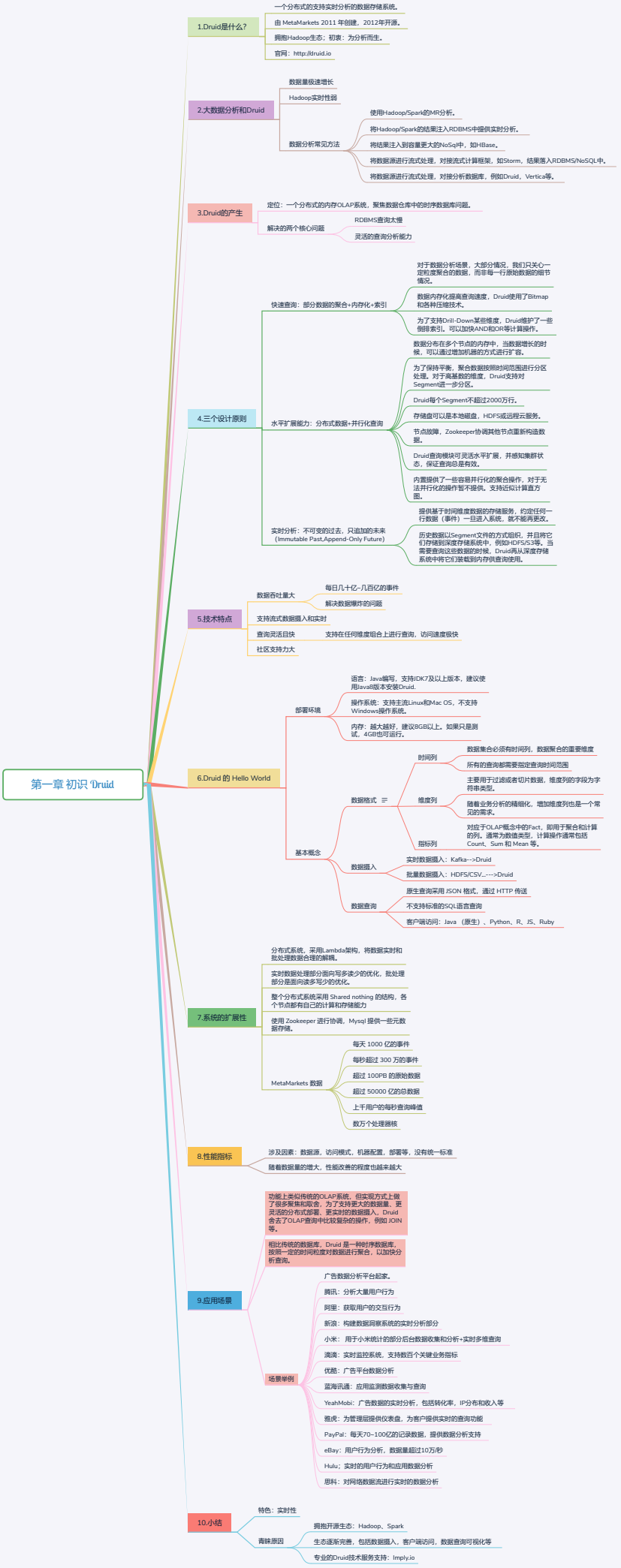
第一章 初识 Druid
1.Druid是什么?
一个分布式的支持实时分析的数据存储系统。
由 MetaMarkets 2011 年创建，2012年开源。
拥抱Hadoop生态；初衷：为分析而生。
官网：http://druid.io
2.大数据分析和Druid
数据量极速增长
Hadoop实时性弱
数据分析常见方法
使用Hadoop/Spark的MR分析。
将Hadoop/Spark的结果注入RDBMS中提供实时分析。
将结果注入到容量更大的NoSql中，如HBase。
将数据源进行流式处理，对接流式计算框架，如Storm，结果落入RDBMS/NoSQL中。
将数据源进行流式处理，对接分析数据库，例如Druid，Vertica等。
3.Druid的产生
定位：一个分布式的内存OLAP系统，聚焦数据仓库中的时序数据库问题。
解决的两个核心问题
RDBMS查询太慢
灵活的查询分析能力
4.三个设计原则
快速查询：部分数据的聚合+内存化+索引
对于数据分析场景，大部分情况，我们只关心一定粒度聚合的数据，而非每一行原始数据的细节情况。
数据内存化提高查询速度，Druid使用了Bitmap和各种压缩技术。
为了支持Drill-Down某些维度，Druid维护了一些倒排索引。可以加快AND和OR等计算操作。
水平扩展能力：分布式数据+并行化查询
数据分布在多个节点的内存中，当数据增长的时候，可以通过增加机器的方式进行扩容。
为了保持平衡，聚合数据按照时间范围进行分区处理。对于高基数的维度，Druid支持对Segment进一步分区。
Druid每个Segment不超过2000万行。
存储盘可以是本地磁盘，HDFS或远程云服务。
节点故障，Zookeeper协调其他节点重新构造数据。
Druid查询模块可灵活水平扩展，并感知集群状态，保证查询总是有效。
内置提供了一些容易并行化的聚合操作，对于无法并行化的操作暂不提供。支持近似计算直方图。
实时分析：不可变的过去，只追加的未来（Immutable Past,Append-Only Future）
提供基于时间维度数据的存储服务，约定任何一行数据（事件）一旦进入系统，就不能再更改。
历史数据以Segment文件的方式组织，并且将它们存储到深度存储系统中，例如HDFS/S3等。当需要查询这些数据的时候，Druid再从深度存储系统中将它们装载到内存供查询使用。
5.技术特点
数据吞吐量大
每日几十亿-几百亿的事件
解决数据爆炸的问题
支持流式数据摄入和实时
查询灵活且快
支持在任何维度组合上进行查询，访问速度极快
社区支持力大
6.Druid 的 Hello World
部署环境
语言：Java编写，支持JDK7及以上版本，建议使用Java8版本安装Druid.
操作系统：支持主流Linux和Mac OS，不支持Windows操作系统。
内存：越大越好，建议8GB以上。如果只是测试，4GB也可运行。
基本概念
数据格式 =
时间列
数据集合必须有时间列，数据聚合的重要维度
所有的查询都需要指定查询时间范围
维度列
主要用于过滤或者切片数据，维度列的字段为字符串类型。
随着业务分析的精细化，增加维度列也是一个常见的需求。
指标列
对应于OLAP概念中的Fact，即用于聚合和计算的列。通常为数值类型，计算操作通常包括 Count、Sum 和 Mean 等。
数据摄入
实时数据摄入：Kafka-->Druid
批量数据摄入：HDFS/CSV...--->Druid
数据查询
原生查询采用 JSON 格式，通过 HTTP 传送
不支持标准的SQL语言查询
客户端访问：Java（原生）、Python、R、JS、Ruby
7.系统的扩展性
分布式系统，采用Lambda架构，将数据实时和批处理数据合理的解耦。
实时数据处理部分面向写多读少的优化，批处理部分是面向读多写少的优化。
整个分布式系统采用 Shared nothing 的结构，各个节点都有自己的计算和存储能力
使用 Zookeeper 进行协调，Mysql 提供一些元数据存储。
MetaMarkets 数据
每天 1000 亿的事件
每秒超过 300 万的事件
超过 100PB 的原始数据
超过 50000 亿的总数据
上千用户的每秒查询峰值
数万个处理器核
8.性能指标
涉及因素：数据源，访问模式，机器配置，部署等，没有统一标准
随着数据量的增大，性能改善的程度也越来越大
9.应用场景
功能上类似传统的OLAP系统，但实现方式上做了很多聚集和取舍，为了支持更大的数据量、更灵活的分布式部署、更实时的数据摄入，Druid舍去了OLAP查询中比较复杂的操作，例如 JOIN 等。
相比传统的数据库，Druid 是一种时序数据库，按照一定的时间粒度对数据进行聚合，以加快分析查询。
场景举例
广告数据分析平台起家。
腾讯：分析大量用户行为
阿里：获取用户的交互行为
新浪：构建数据洞察系统的实时分析部分
小米：用于小米统计的部分后台数据收集和分析+实时多维查询
滴滴：实时监控系统，支持数百个关键业务指标
优酷：广告平台数据分析
蓝海讯通：应用监测数据收集与查询
YeahMobi：广告数据的实时分析，包括转化率，IP分布和收入等
雅虎：为管理层提供仪表盘，为客户提供实时的查询功能
PayPal：每天70~100亿的记录数据，提供数据分析支持
eBay：用户行为分析，数据量超过10万/秒
Hulu：实时的用户行为和应用数据分析
思科：对网络数据流进行实时的数据分析
10.小结
特色：实时性
青睐原因
拥抱开源生态：Hadoop、Spark
生态逐渐完善，包括数据摄入，客户端访问，数据查询可视化等
专业的Druid技术服务支持：Imply.io