

# 大数据组核心技术介绍

——蔡岩

# 培训目的及对象

- 目的
  - 介绍Hadoop家族主要成员
  - 介绍大数据组将要使用的主要技术
- 对象
  - 需要快速了解Hadoop家族的开发者、测试者
  - 有一定数据库使用经验的人员

# 资源推荐

- 书籍：  
《Hadoop权威指南 第三版》  
快速浏览，快速上手
- 玩转Hadoop——起手部署  
[https://github.com/yancai/mydoc/blob/master/hadoop/hadoop\\_ha\\_deploy.md](https://github.com/yancai/mydoc/blob/master/hadoop/hadoop_ha_deploy.md)
- awesome big data  
<https://github.com/onurakpolat/awesome-bigdata>

# 目录

- Hadoop简介
- Hadoop家族成员介绍

# Hadoop

## 简介



# Hadoop直观感受——大象的直观感受

- 起身慢
- 力量大
- 步子迈得大
- 成群结队
- 象群势不可挡
- 谁说大象不能跳



# Hadoop介绍

- 历史：
  - 2003-2004：Google提出GFS、MapReduce、BigTable
  - 2004-2005，Doug Cutting，源于Lucene的子项目Nutch（搜索系统）
  - 2006年Apache Hadoop独立发展
  - 2008年成为Apache顶级项目
- “Hadoop” ——Doug Cutting孩子所起名
- Yahoo——Hadoop的诞生和发展之地

# 性能一窥

- 2008年，900节点，1T数据排序，209秒
- 2009年，1460节点，1TB数据排序，173分钟



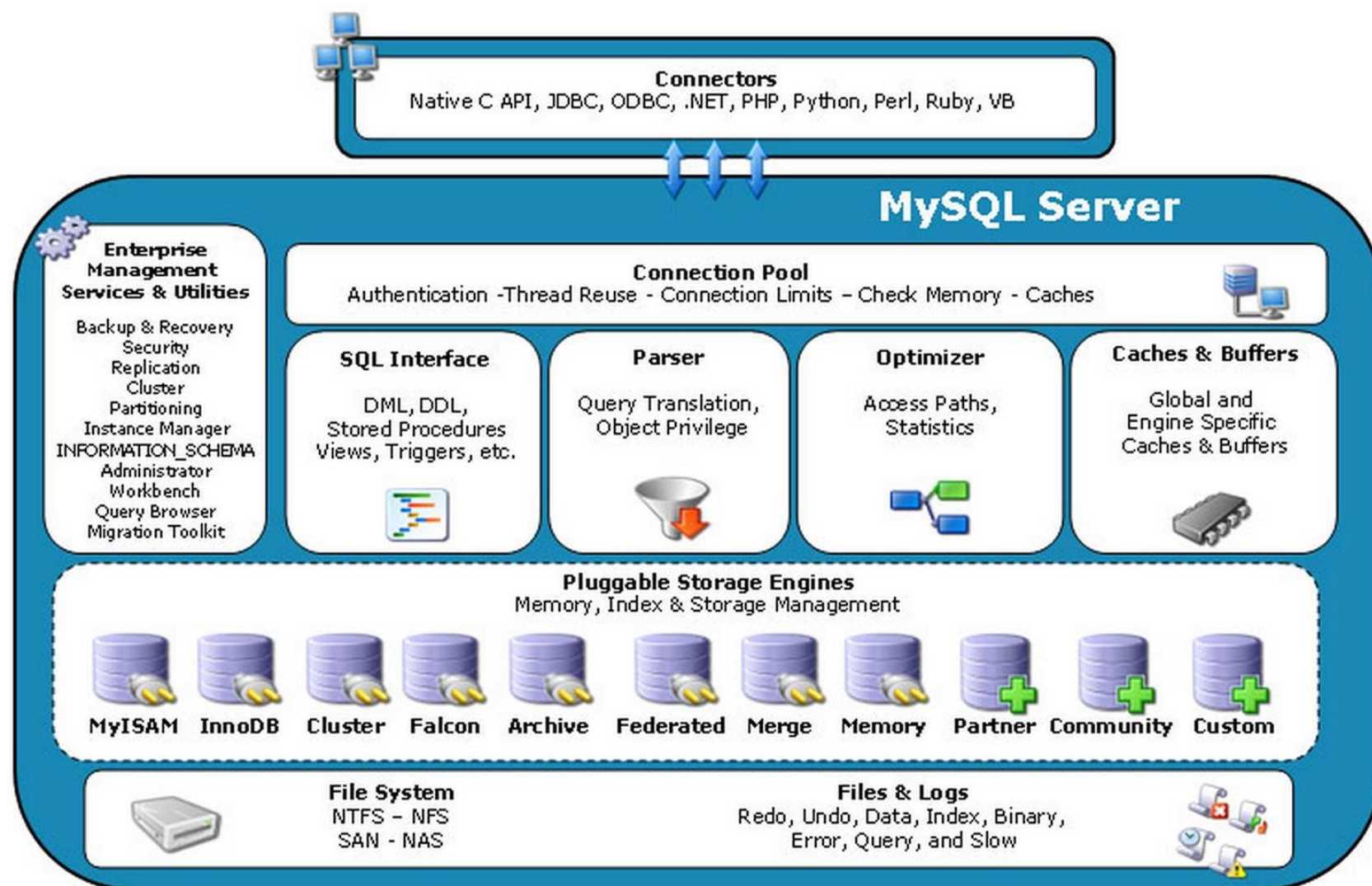
# Hadoop 家族成员



# 大数据组核心技术列表

- Zookeeper
- HDFS
- Parquet
- Map-Reduce、YARN
- Spark
- HBase
- Hive、Impala
- Sqoop
- Oozie
- Hue
- Kylin

# MySQL架构图





Cloudera Manager

Web Interface



Apache Ambari

**Ambari**



**Oozie** workflow



**Flume**  
Collector



**Pig**  
Scripting



**Hive**  
SQL Query



**Spark**  
process engine



**Impala**  
MPP  
SQL Query



**Kylin**  
OLAP Engine



**Zookeeper**  
Coordination



**Kafka**  
Message Queue



**YARN**  
Distributed Processing Framework



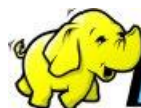
**Parquet** Columnar Storage Format



**HBase**  
Columnar Store



**Sqoop**  
Data Exchange



**hadoop** **HDFS** Hadoop Distributed File System

## 其它产品



Apache Thrift™



# Zookeeper



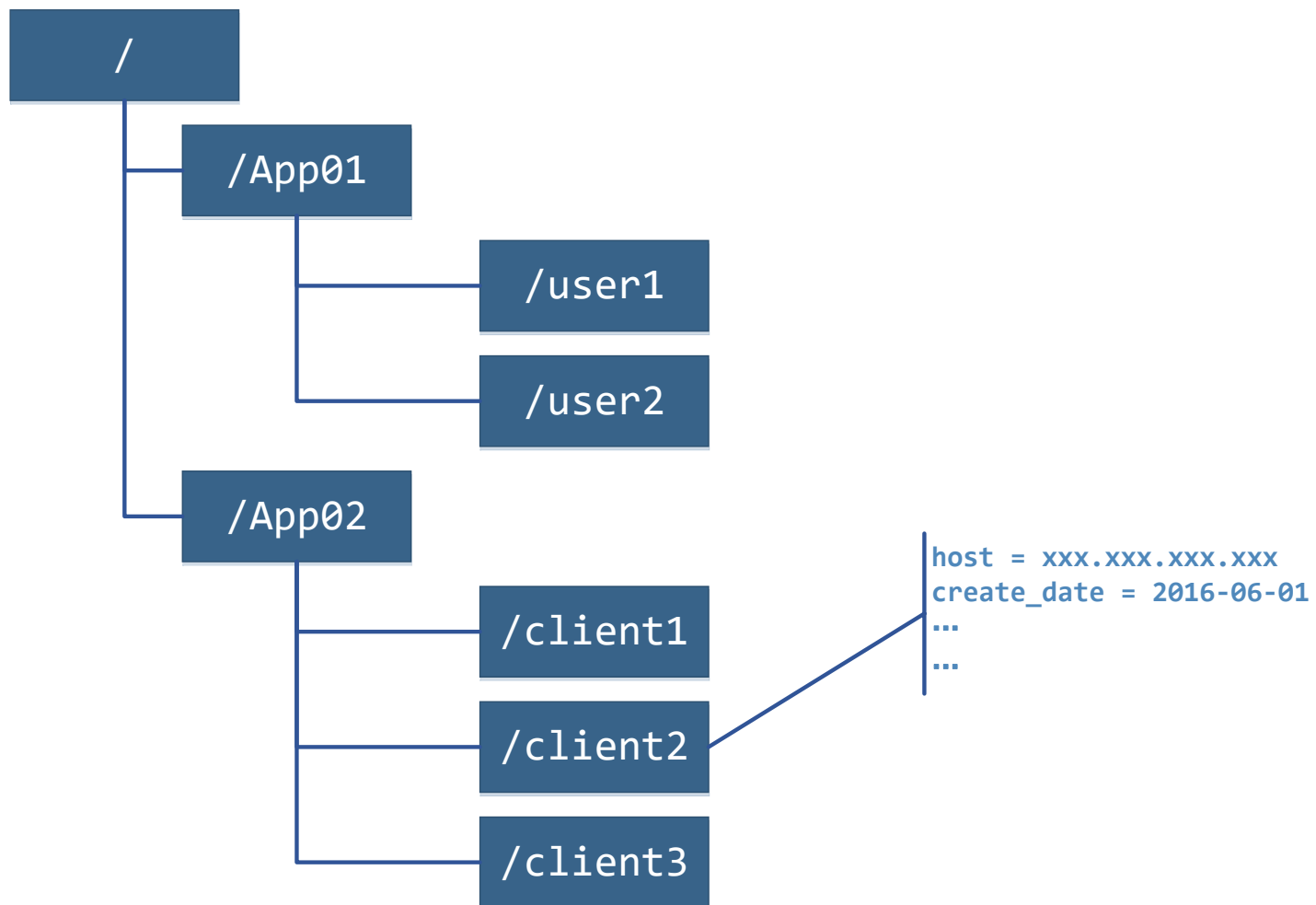
ZooKeeper is a centralized service for maintaining configuration information, naming, providing distributed synchronization, and providing group services.

- 文件系统 & 通知机制
- 功能
  - 命名服务
  - 配置管理
  - 集群管理
  - 分布式锁
  - 分布式队列

CAP原则：

1. C : Consistency  
一致性，数据一致更新，所有数据变动都是同步的  
**单调一致性**
2. A : Availability  
可用性，系统具有好的响应性能
3. P : Partition tolerance  
分区容错性

# Zookeeper目录结构





# Hadoop



The Apache™ Hadoop® project develops open-source software for reliable, scalable, distributed computing.

# HDFS



A distributed file system that provides high-throughput access to application data.

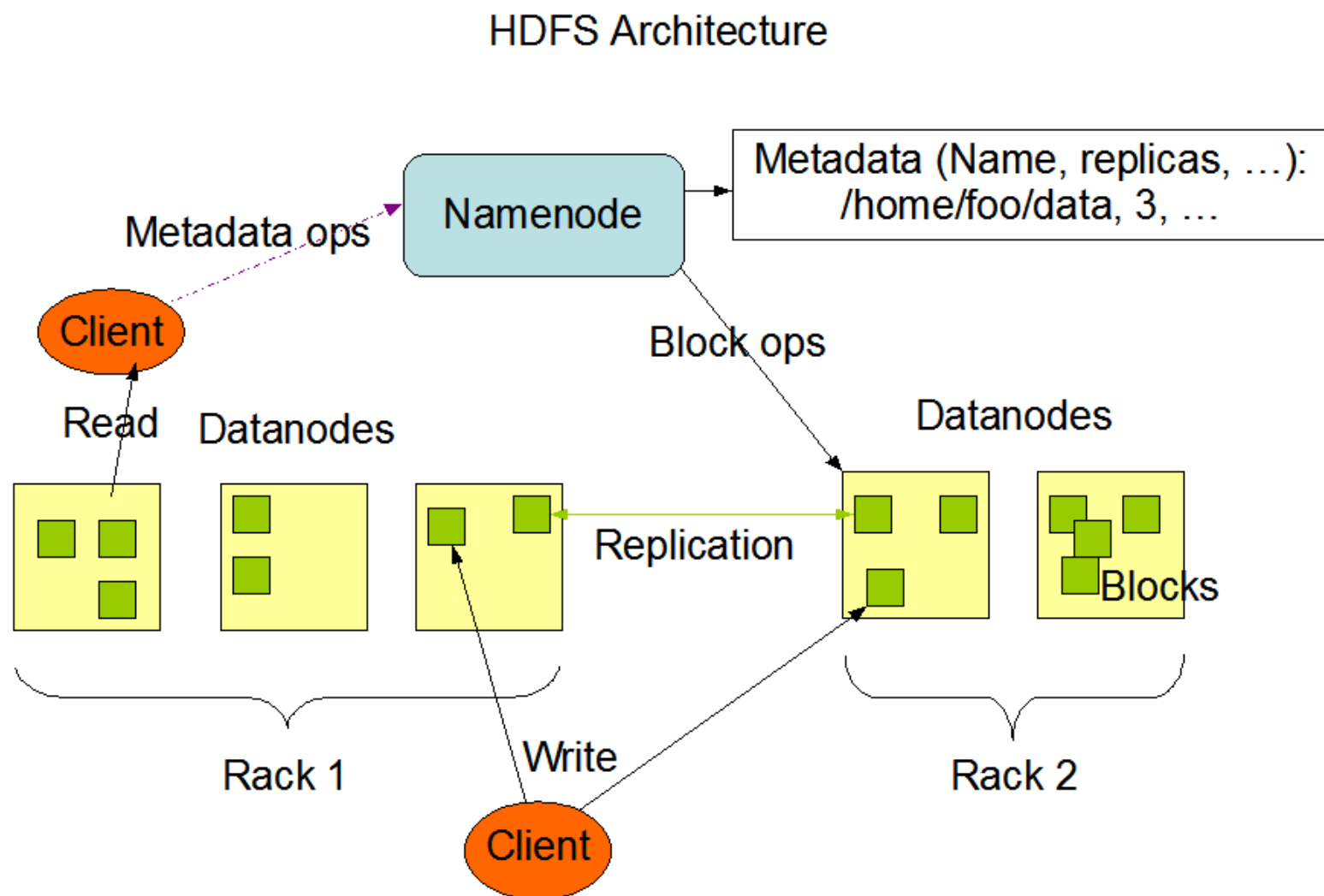
# HDFS简介

- HDFS ( Hadoop Distributed File System ) , Hadoop分布式文件系统
- 分布式 & 文件
- 特点：
  - 超大文件：GB、TB、PB
  - 流式数据访问：一次写入，多次读取
  - 普通硬件：低廉硬件
  - 延时：不适合低延时需求，高吞吐牺牲延时
  - 避免大量小文件：会占用namenode内存
  - 不支持多用户写入和修改文件：为了高吞吐

# HDFS主要概念

- 数据块：HDFS读/写的最小单位  
`dfs.blocksize = 128MB` (默认64MB, 主流128MB)  
目的：最小化寻址开销
- namenode：管理者，管理文件系统命名空间
- datanode：工作者，存储并检索数据，定期向namenode发送块信息
- HA：namenode高可用，active&standby

# HSFS架构

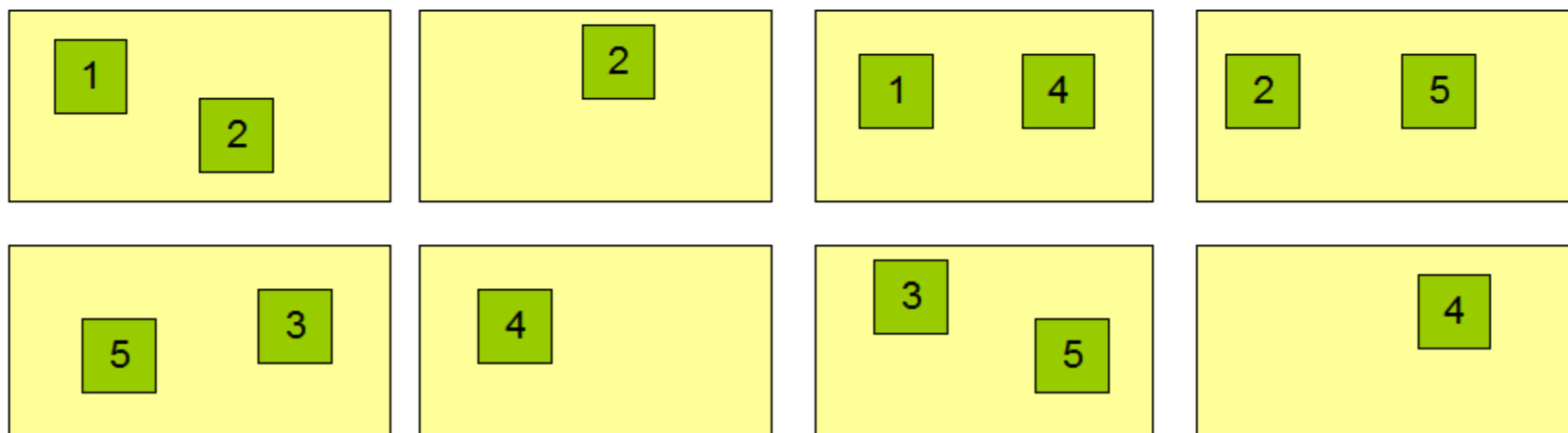


# HDFS块副本

## Block Replication

Namenode (Filename, numReplicas, block-ids, ...)  
/users/sameerp/data/part-0, r:2, {1,3}, ...  
/users/sameerp/data/part-1, r:3, {2,4,5}, ...

## Datanodes



# parquet



Apache Parquet is a columnar storage format available to any project in the Hadoop ecosystem, regardless of the choice of data processing framework, data model or programming language.

# Parquet简介

- Cloudera和Twitter共同开发
- 2015年成为Apache顶级项目
- 列式存储
- 支持嵌套格式
- 包含有schema
- Hadoop家族大部分产品都支持Parquet格式读写

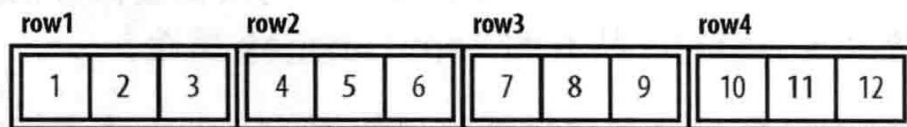


# 列式存储示意图

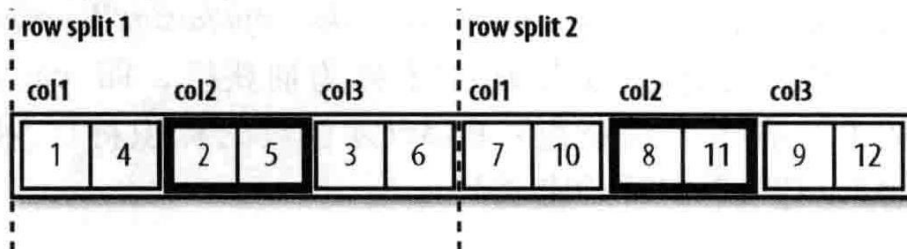
Logical table

	col1	col2	col3
row1	1	2	3
row2	4	5	6
row3	7	8	9
row4	10	11	12

Row-oriented layout (SequenceFile)



Column-oriented layout (RCFile)



# 行列存储对比

	行式存储	列式存储
优点	<ol style="list-style-type: none"><li>1. 数据保存在一起</li><li>2. INSERT/UPDATE方便</li><li>3. 数据压缩效率更高</li></ol>	<ol style="list-style-type: none"><li>1. 查询只有涉及的列会被读取</li><li>2. 投影操作高效</li><li>3. 任何列都可以作为索引</li></ol>
缺点	<ol style="list-style-type: none"><li>1. 选择操作只涉及一个列也会读取所有数据</li><li>2. 数据完整性要求高</li></ol>	<ol style="list-style-type: none"><li>1. 选择完成时，需要重新组装列</li><li>2. INSERT/UPDATE比较麻烦</li></ol>

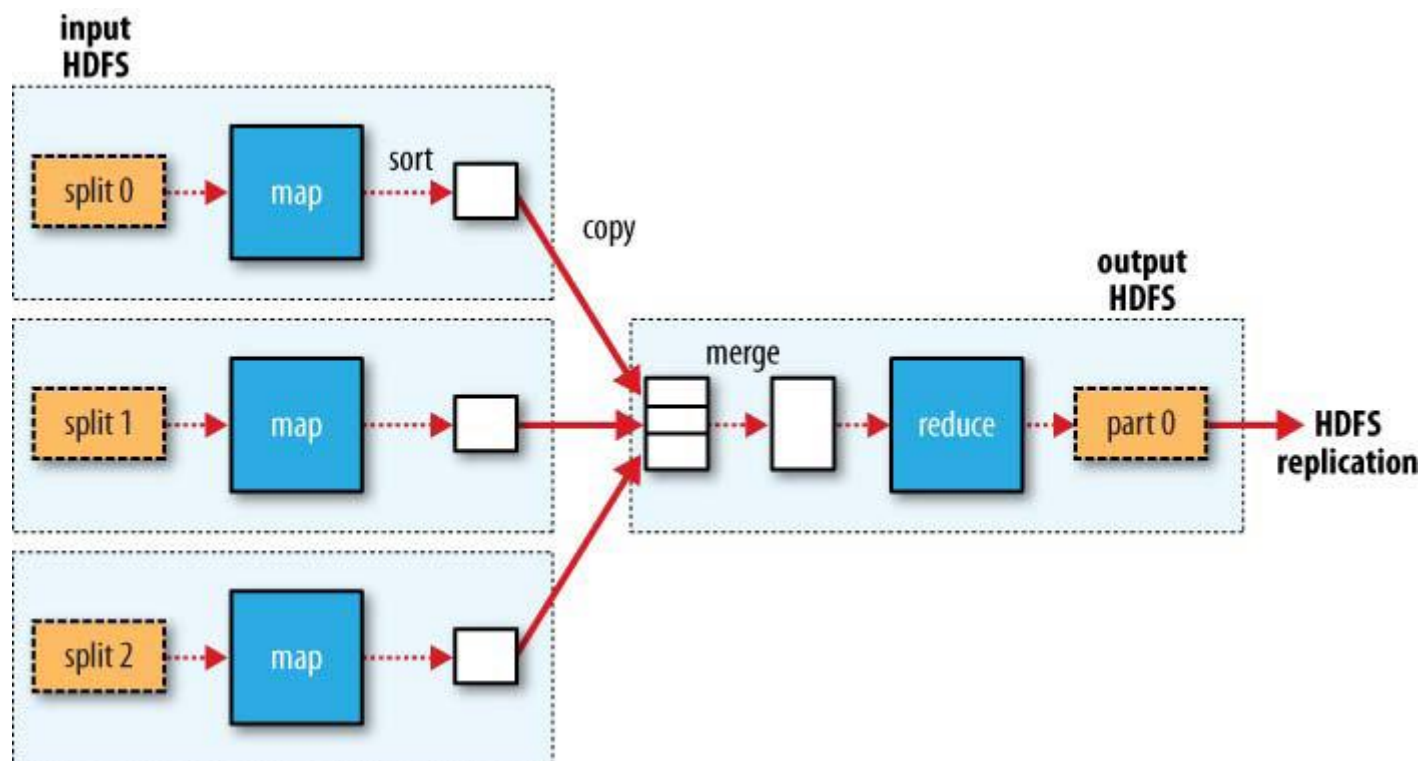
Parquet甜头：  
数据具有自解释能力——自身包含schema

# Map-Reduce



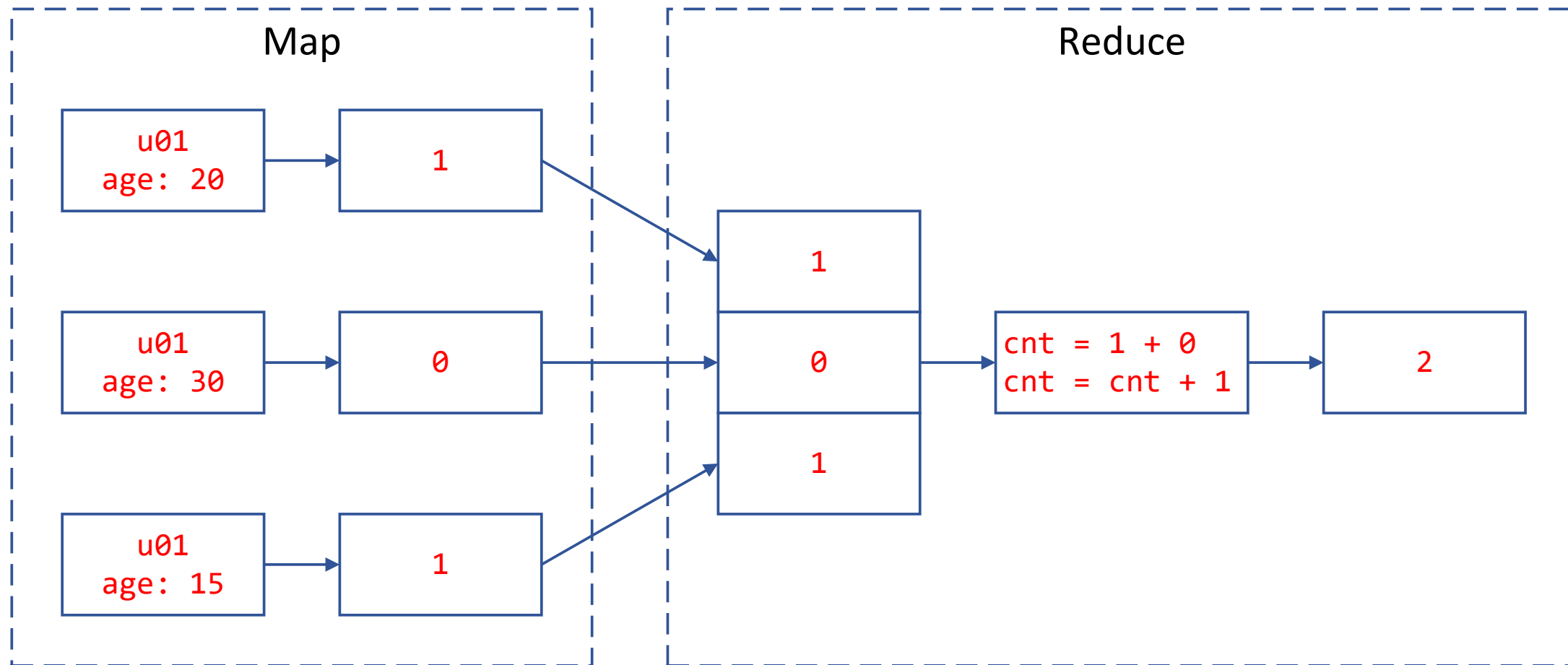
- Map-Reduce编程模型
- Map-Reduce计算引擎

# Map-Reduce编程模型简介



# Map-Reduce计算示例

示例：计算年龄小于25的用户数量



# Hadoop Map-Reduce计算引擎

- 别名：MR1，经典MapReduce，...
- 作用：并行化、任务调度、资源管理、机器间通信

- 主要角色：

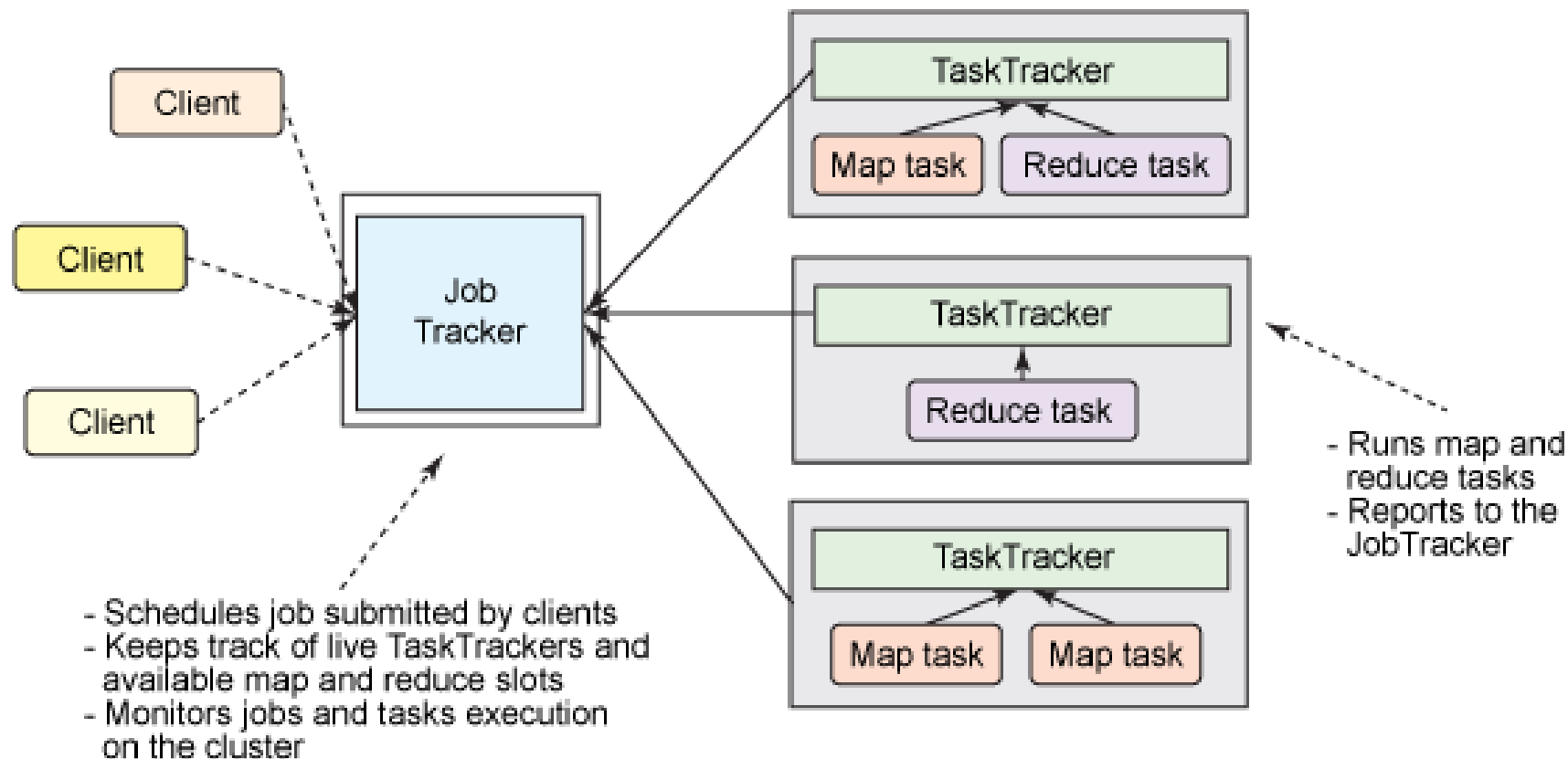
Client

JobTracker——协调集群上的所有作业

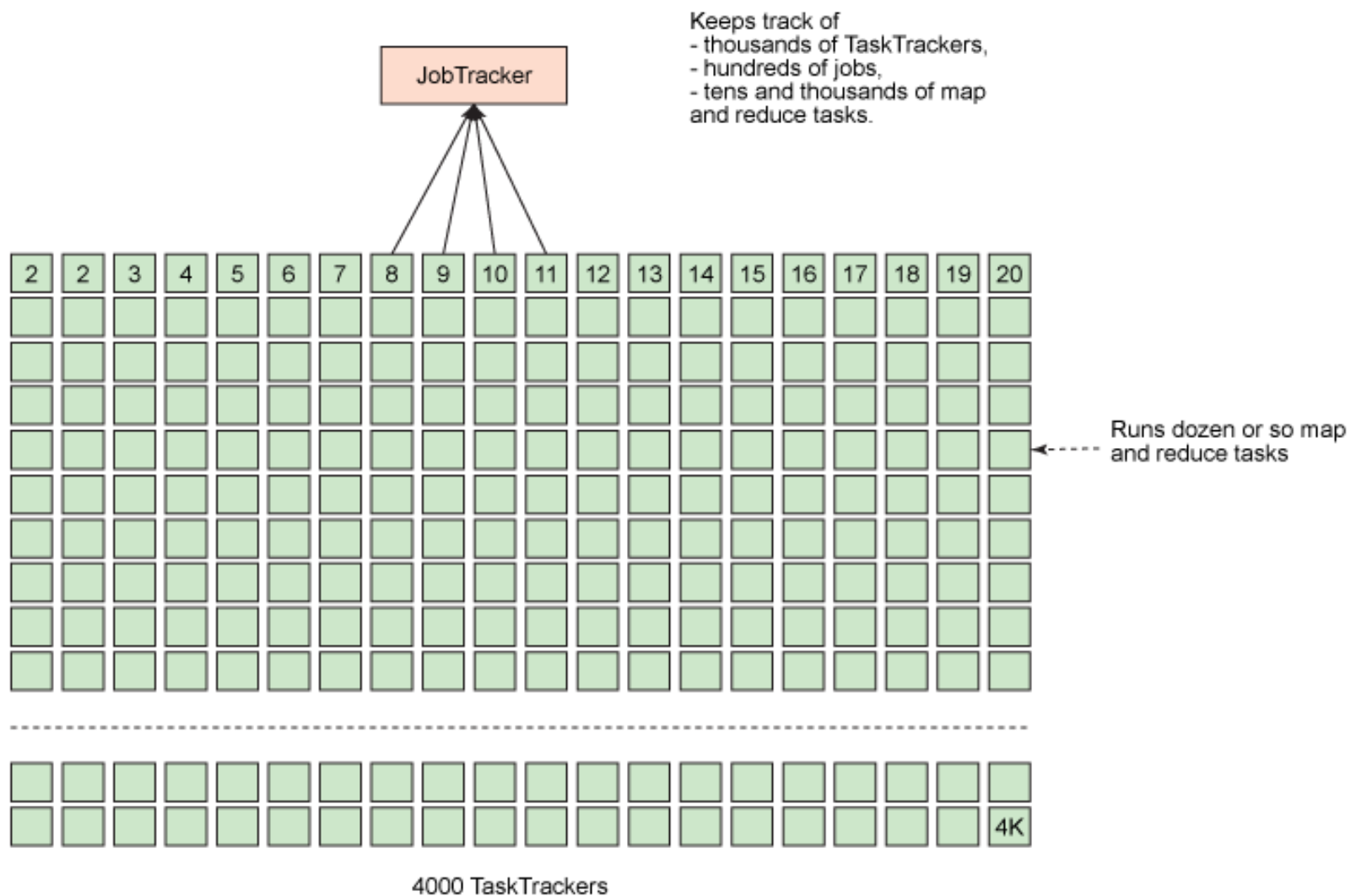
TaskTracker——运行job并反馈给JobTracker

MapReduce作业

# MapReduce1 架构







存在的问题：

1. 单个JobTracker成为瓶颈
2. slot固定，资源利用率低
3. 节点数达到极限

如何解决？

——引入YARN

# YARN

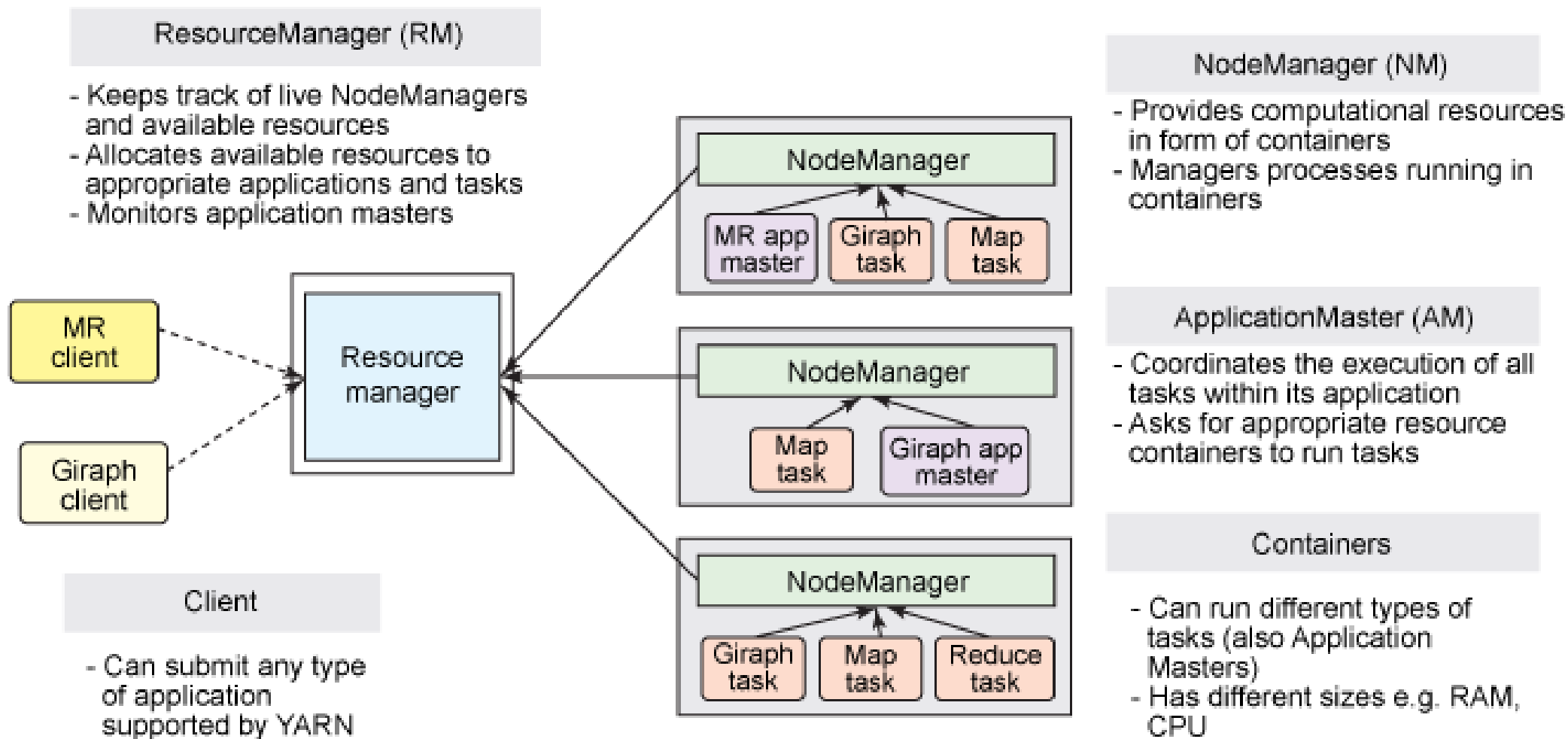


A framework for job  
scheduling and cluster  
resource management.

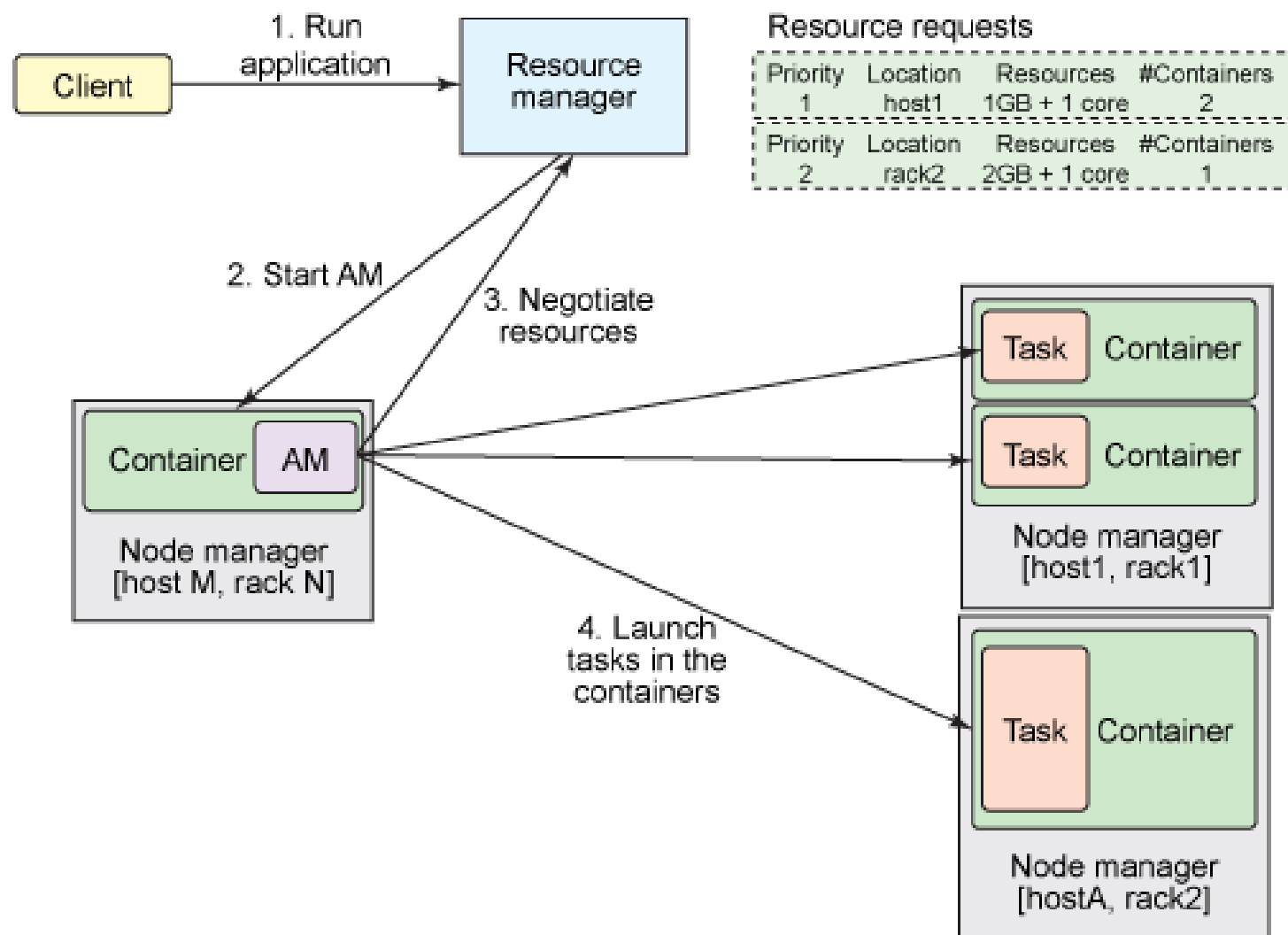
# YARN简介

- 2010 , Yahoo开始设计新的Hadoop架构
- 独立的资源管理功能
- 主要角色和名词：
  - Application——替代MR作业
  - Container——资源的抽象，资源划分的单位
  - ResourceManager——应用仲裁及资源管理
  - ApplicationMater——管理Application实例的所有任务
  - NodeManager——各节点的资源和管理者

# YARN架构



# YARN中应用提交执行示例





Lightning-fast cluster  
computing

Apache Spark™ is a fast and  
general engine for large-scale  
data processing.

# Spark简介

- 主页：<http://spark.apache.org/>
- 诞生于2009 UC Berkeley AMP lab，2010开源
- 是一个快速、通用的大规模数据处理**引擎**
- 类似于Hadoop的MapReduce，可伸缩、基于内存、计算中间结果无需写入磁盘。
- 快速：传统MR的100倍+
- 易用：多种语言支持：Java、Scala、Python、R等，提供REPL
- 通用：提供高级工具：SparkSQL、MLib、GraphX、Spark Streaming
- 集成Hadoop：易于和Hadoop结合，支持YARN、HBase、Hive等组件



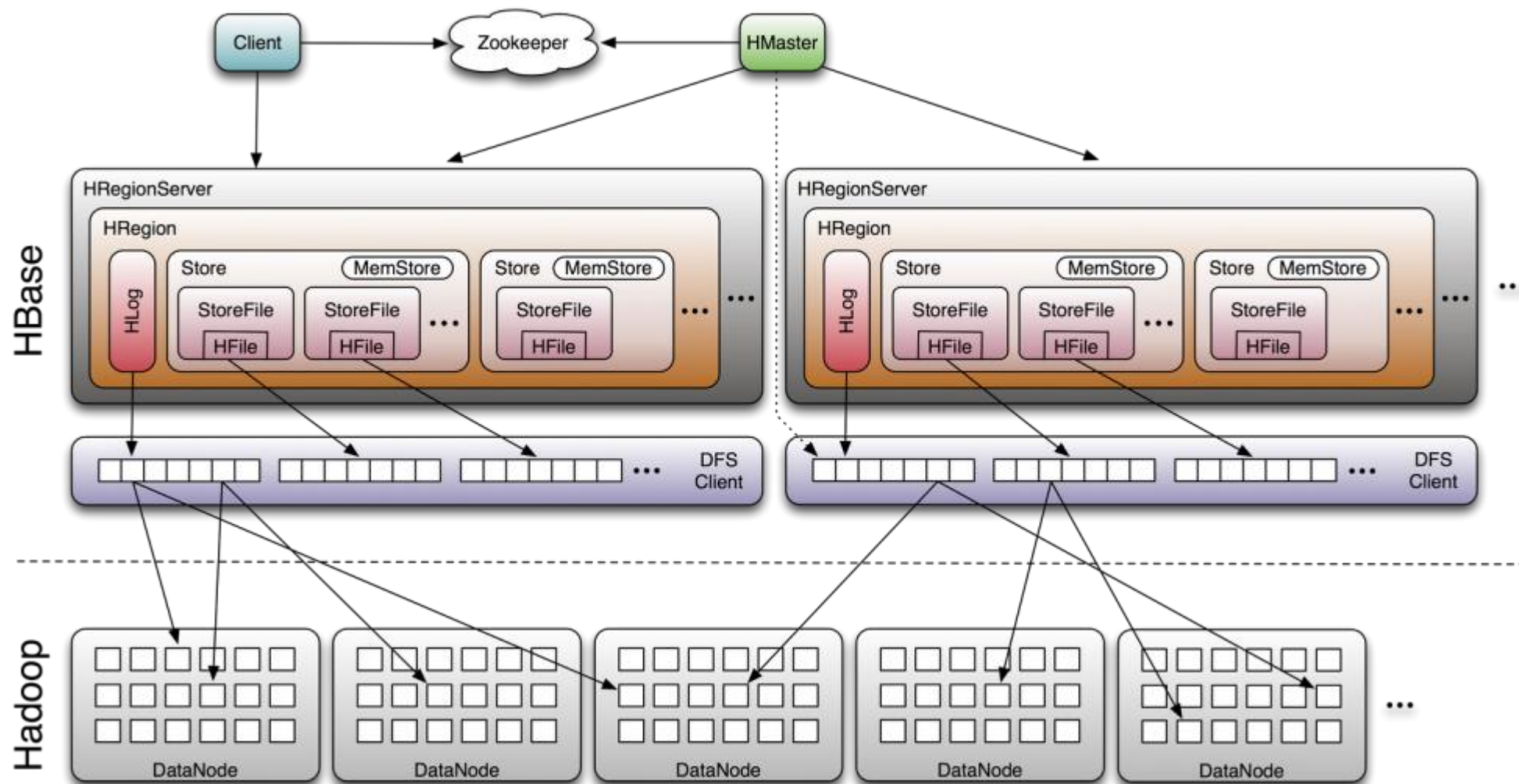
Apache HBase™ is the Hadoop database, a distributed, scalable, big data store.



# HBase简介

- Google BigTable的开源实现
- HBase——Hadoop Database
- 分布式、可伸缩、大数据存储
- 特点：
  - 列式存储
  - 存储量大
  - 数据多版本
  - 数据类型为字符串或二进制
  - 列可动态增加

# HBase架构



# HBase数据示例

Row						
RowKey	Column Family					
	Column			Column		
	user			school		
	username	age	...	schoolname	address	...
10102743	李华	21		第一中学	某街10号	
10102745	韩梅梅	22		第四中学	某市不知道	
				第三中学		

查询数据所需：table, rowkey, column family, column

# Hive



The Apache Hive™ data warehouse software facilitates reading, writing, and managing large datasets residing in distributed storage using SQL.

# Hive简介

- 2009年诞生于Facebook
- 构建数据仓库
- 提供库、表以及类SQL语句操作数据
- 主要角色：
  - cli：本地命令行客户端
  - hiveserver2：对外提供thrift接口远程调用
  - metastore server：元数据信息管理服务
  - warehouse：数据的实际存储位置，可以在本地磁盘，也可以在hdfs

# Impala

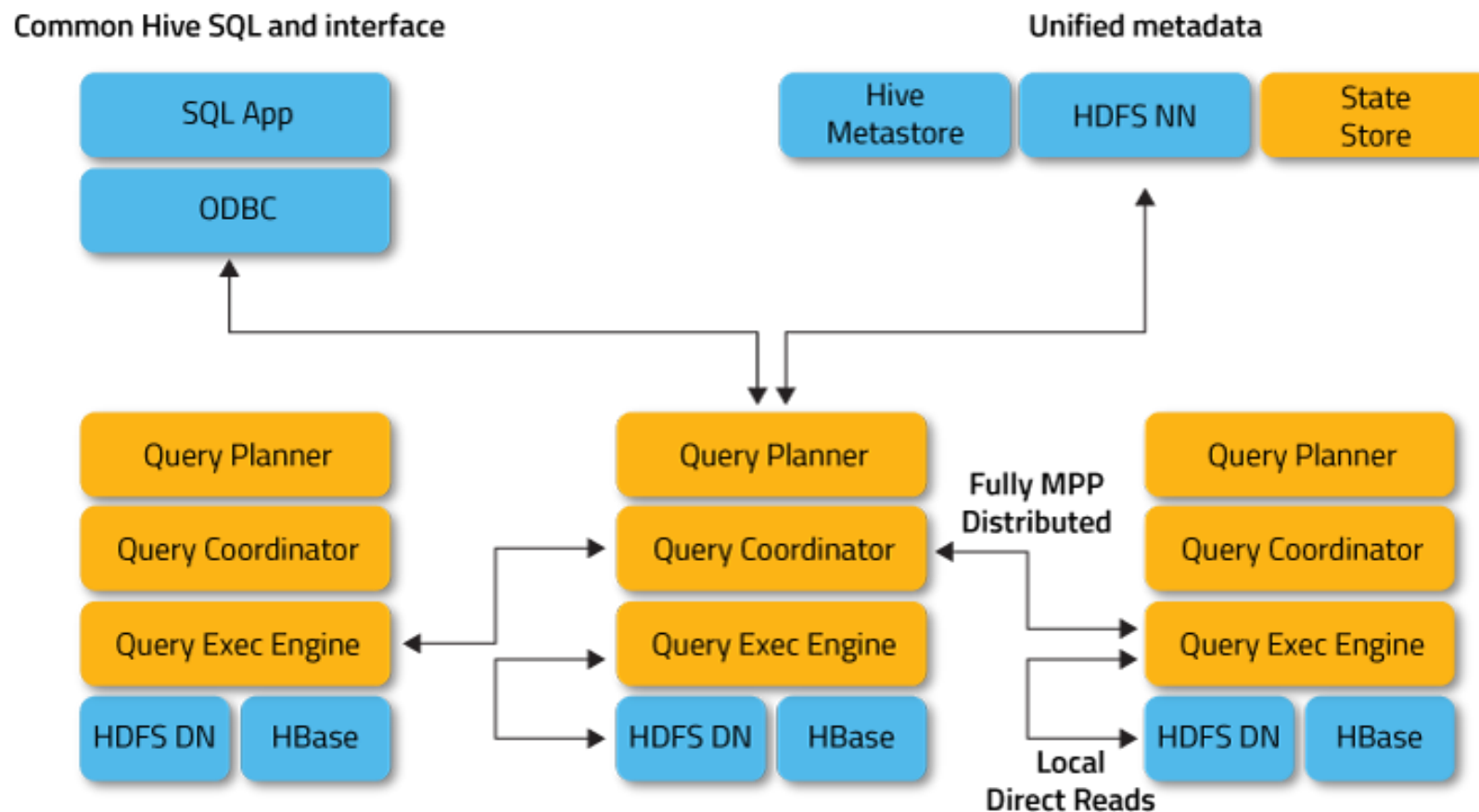


Apache Impala (incubating)  
is the open source, native  
analytic database  
for Apache Hadoop.

# Impala简介

- Apache孵化中
- Cloudera主导开发，源于Google Dremel
- 提供库、表以及SQL操作
- 速度**超级**快

# Hive架构





# Impala VS Hive

## 一亿数据测试结果

查询方式	hive+parquet ( 单位 : 秒 )	impala ( 单位 : 秒 )
单表 count(1)	1.367	2.29
单表 count(id)	85.691	16.88
单表 count(id) + where	75.071	8.2
单表 group by	264.317	6.31
单表 order by	80.867	7.56
join count	376.173	13.43
join group by	561.565	机器内存不足无法测试

# sqoop



Apache Sqoop(TM) is a tool designed for efficiently transferring bulk data between Apache Hadoop and structured datastores such as relational databases.

# sqoop介绍

- 2009年诞生
- 用于Hadoop和关系型数据库之间数据传输
- 支持全量、增量抽取



oozie



Oozie is a workflow scheduler system to manage Apache Hadoop jobs.

# oozie简介

- 工作流调度引擎
- workflow -> coordinator -> bundle
- 稍候上演示



Hue is a Web interface for analyzing data with Apache Hadoop.

# Hue简介

- 直接上演示

# Kylin



Extreme OLAP Engine for Big Data

Apache Kylin™ is an open source Distributed Analytics Engine designed to provide SQL interface and multi-dimensional analysis (OLAP) on Hadoop supporting extremely large datasets.



# Kylin简介

- eBay开发
- 提供SQL查询及多维数据分析的OLAP ( Online Analytical Processing ) 引擎
- 提供亚秒级查询海量数据的能力
- 空间换时间

# 建立cube

- cube建立相关：  
维度数量  
数据量

