




Spark User Beijing Meetup

Spark在百度大数据生态上的 应用与实践

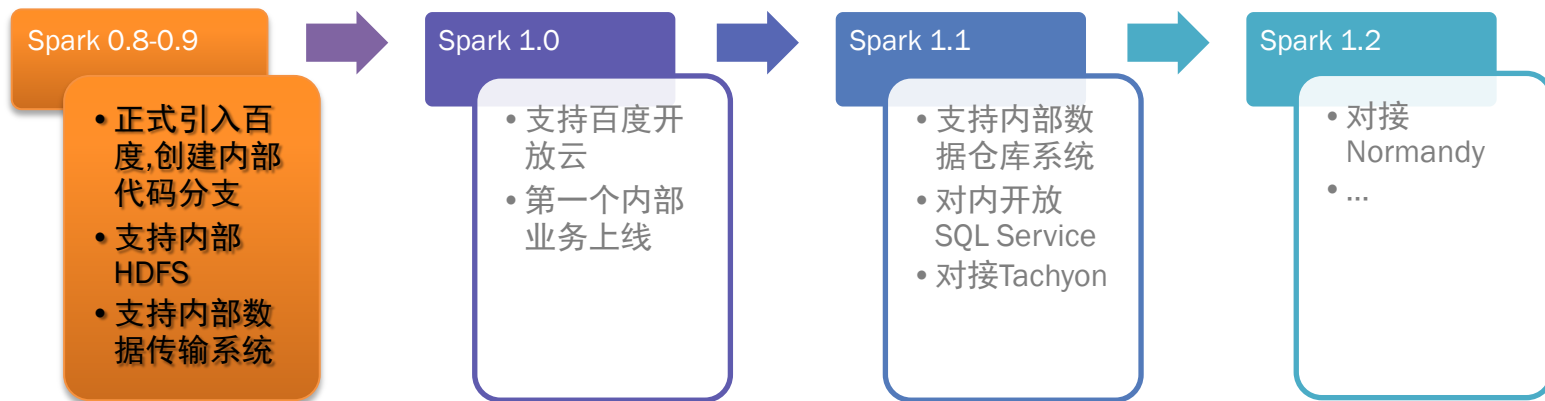
汪凯 百度基础架构部
2015年6月

- 
- 一、Spark 在 百度 的发展历程
 - 二、工程（产品）实践案例



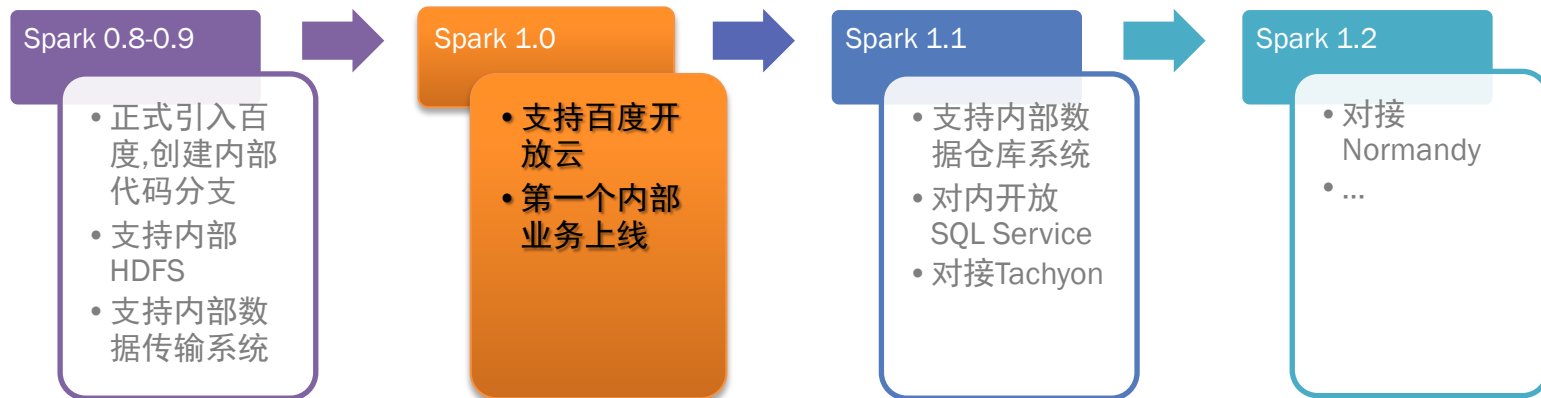
Spark 在 百度发展历程

Spark In Baidu



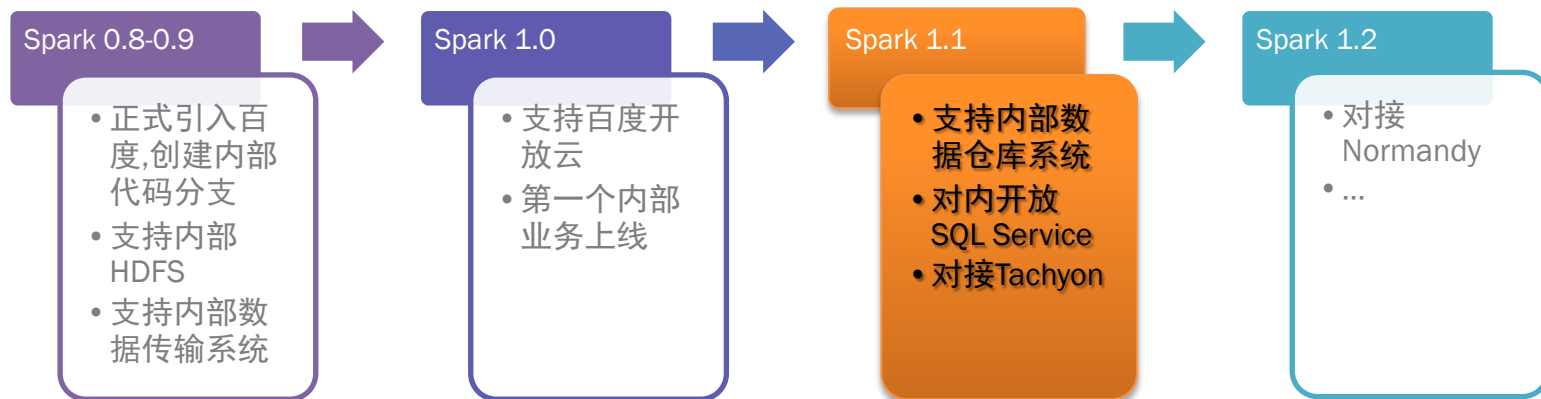
- ✓ 兼容百度内部的 HDFS(PETA)集群
- ✓ 支持BIGPIPE 数据源, 多一种流式计算选择

Spark In Baidu



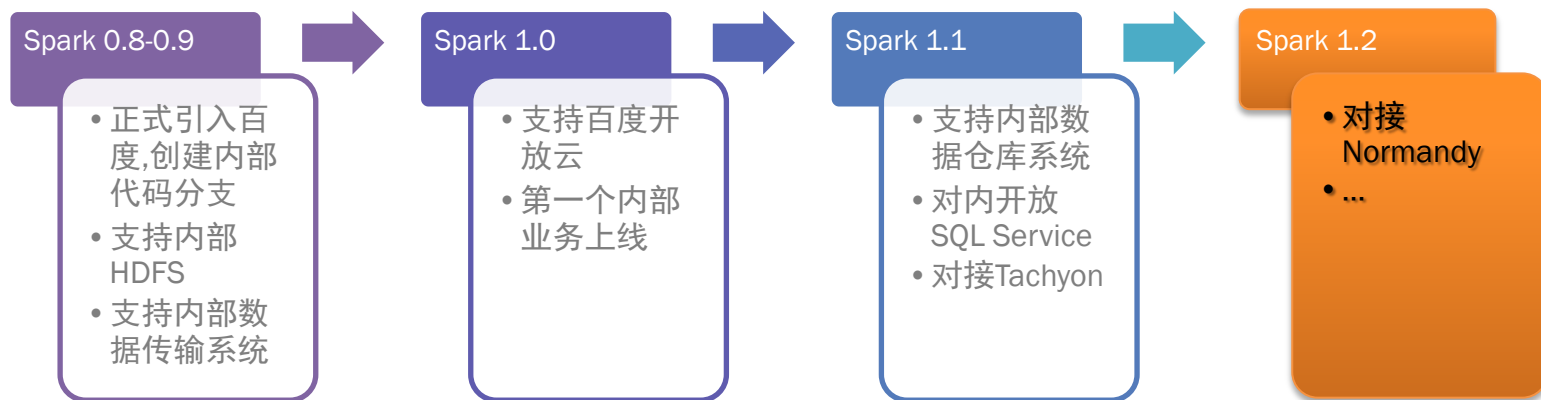
- ✓ 支持百度公有云产品BMR, 对外提供大数据计算服务
- ✓ 第一个内部业务上线

Spark In Baidu



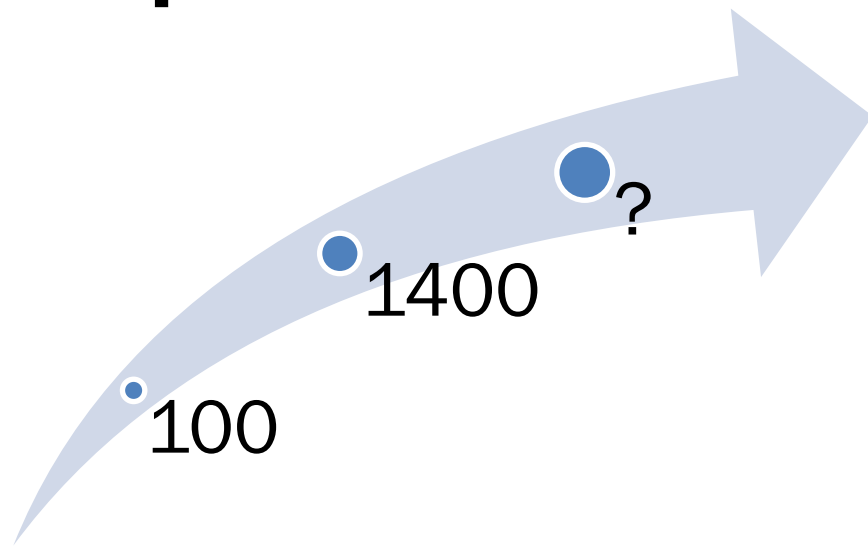
- ✓ UDW 数据接入, Spark SQL 直接读写 UDW 表格
- ✓ CloudQuery, 高速 UDW 查询引擎
- ✓ 整合Tachyon做透明Cache层

Spark In Baidu



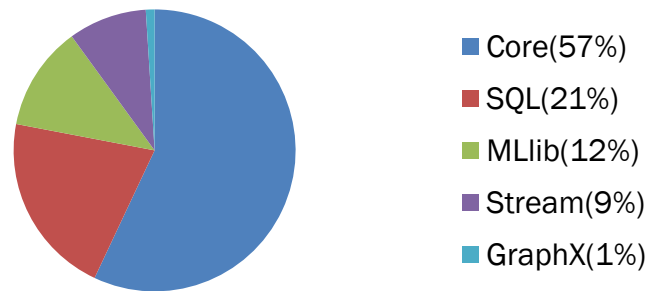
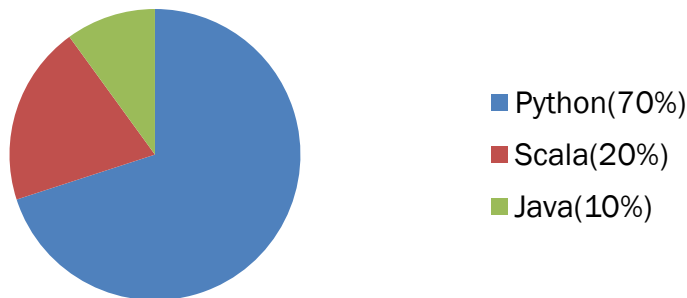
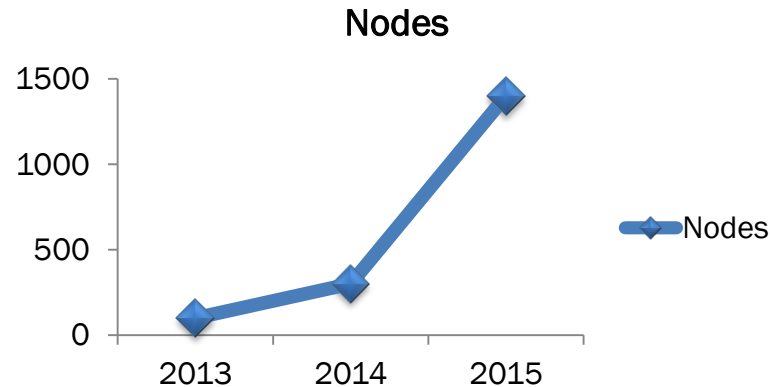
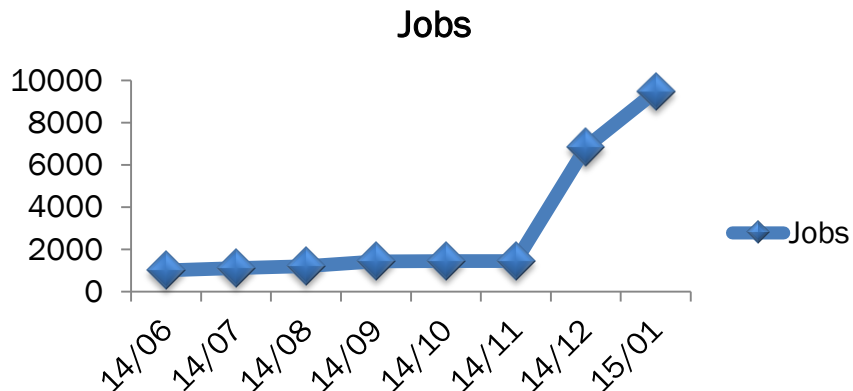
✓ 对接 Normandy 调度器，统一调度，实现计算混布

Spark In Baidu



- ✓ 在南方和北方多集群，当前集群规模1400台（30000~Cores, 100~TB Memory）
- ✓ 日提交 App: 200~300
- ✓ 业务方：凤巢，大搜索，直达号，百度大数据，百度知道、云安全等等

Spark In Baidu



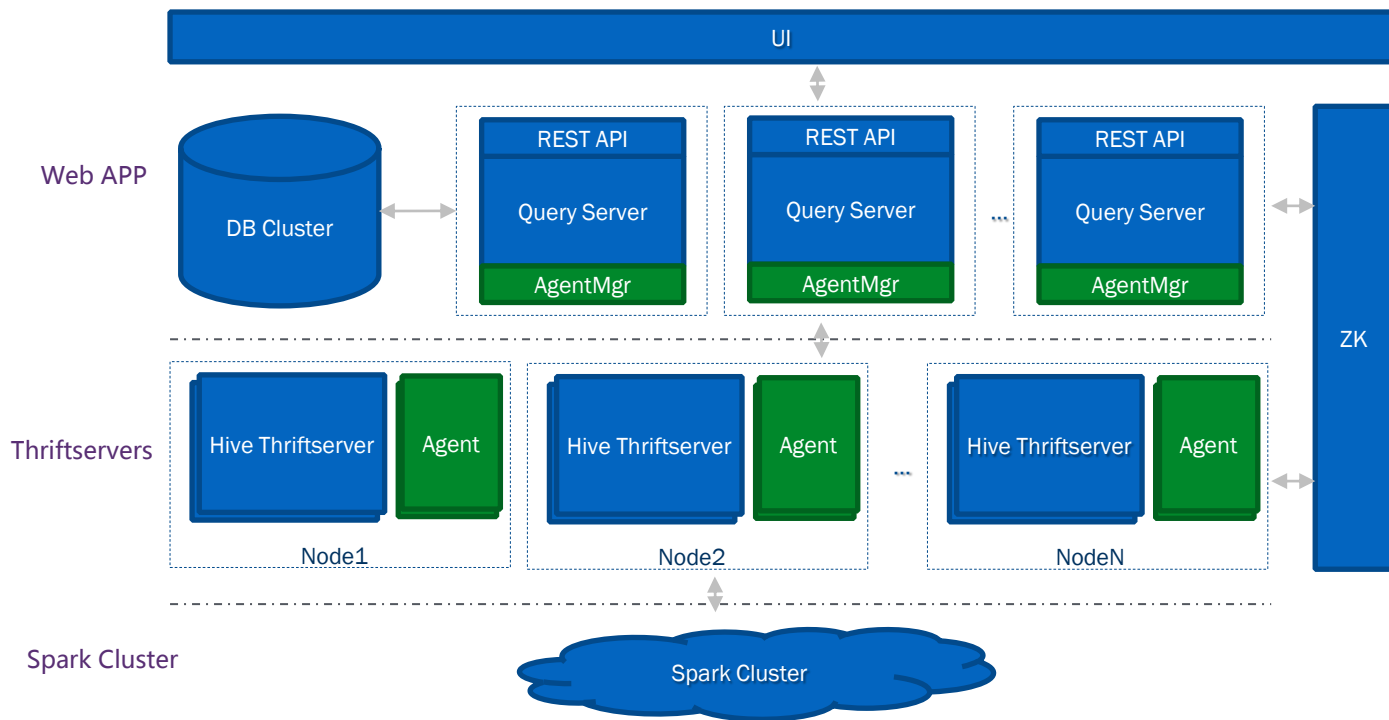
工程（产品）实践案例

- ✓ Adhoc Query On Spark SQL
- ✓ Spark And Tachyon
- ✓ Spark SQL And Mllib
- ✓ Spark And ElasticSearch
- ✓ Spark And Real-time Security
- ✓ 百度开放云:Spark On BMR

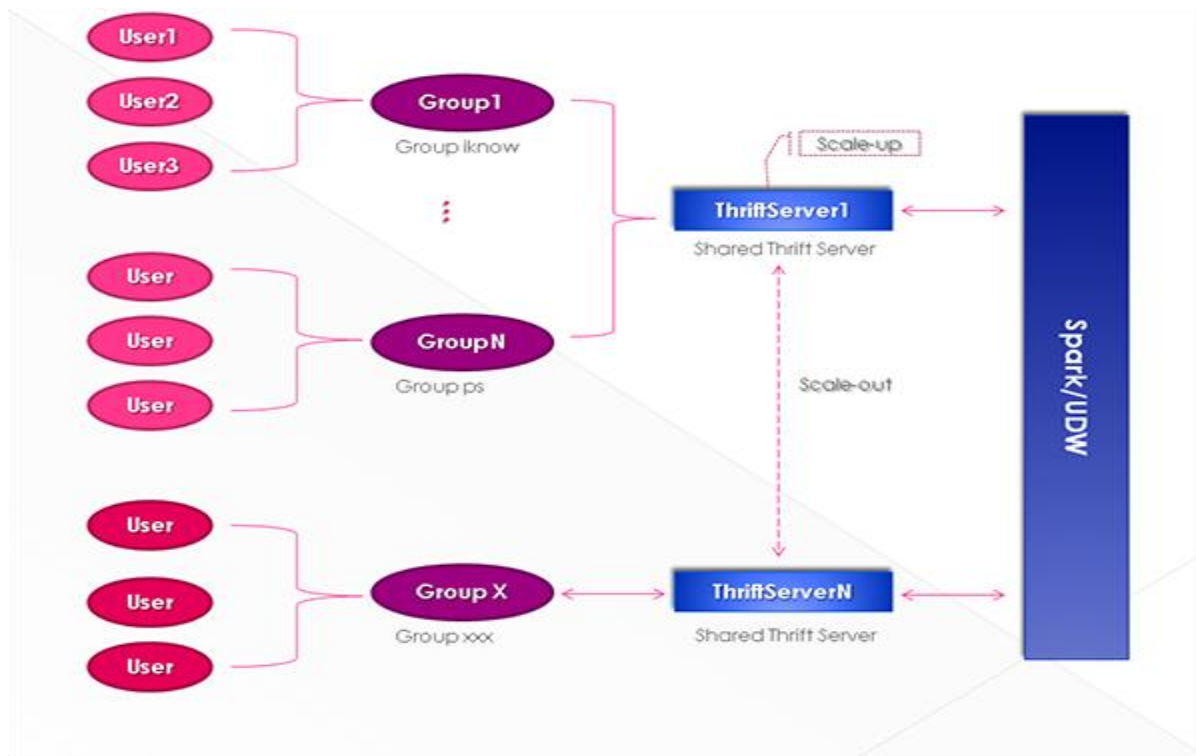
Adhoc Query On Spark SQL

- Cloud Query是一个基于Spark的SQL查询系统，用户可以通过Cloud Query以类SQL的方式快速、高效、无缝地查询公司UDW上的数据。

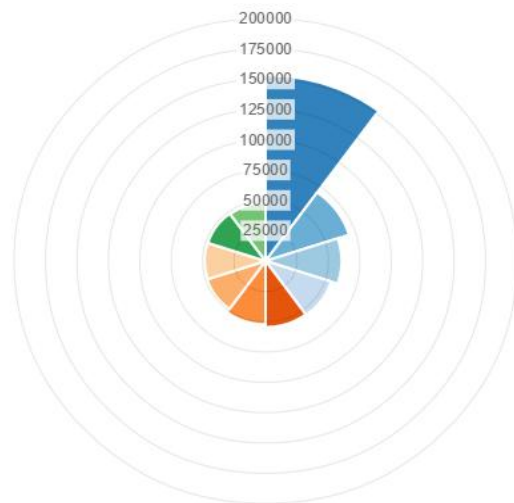
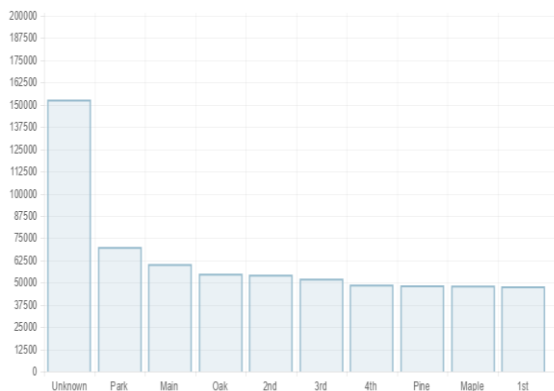
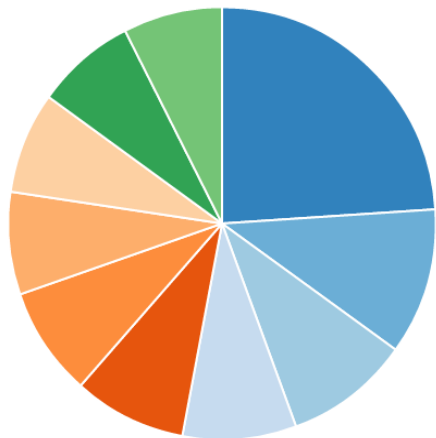
Adhoc Query On Spark SQL(1)



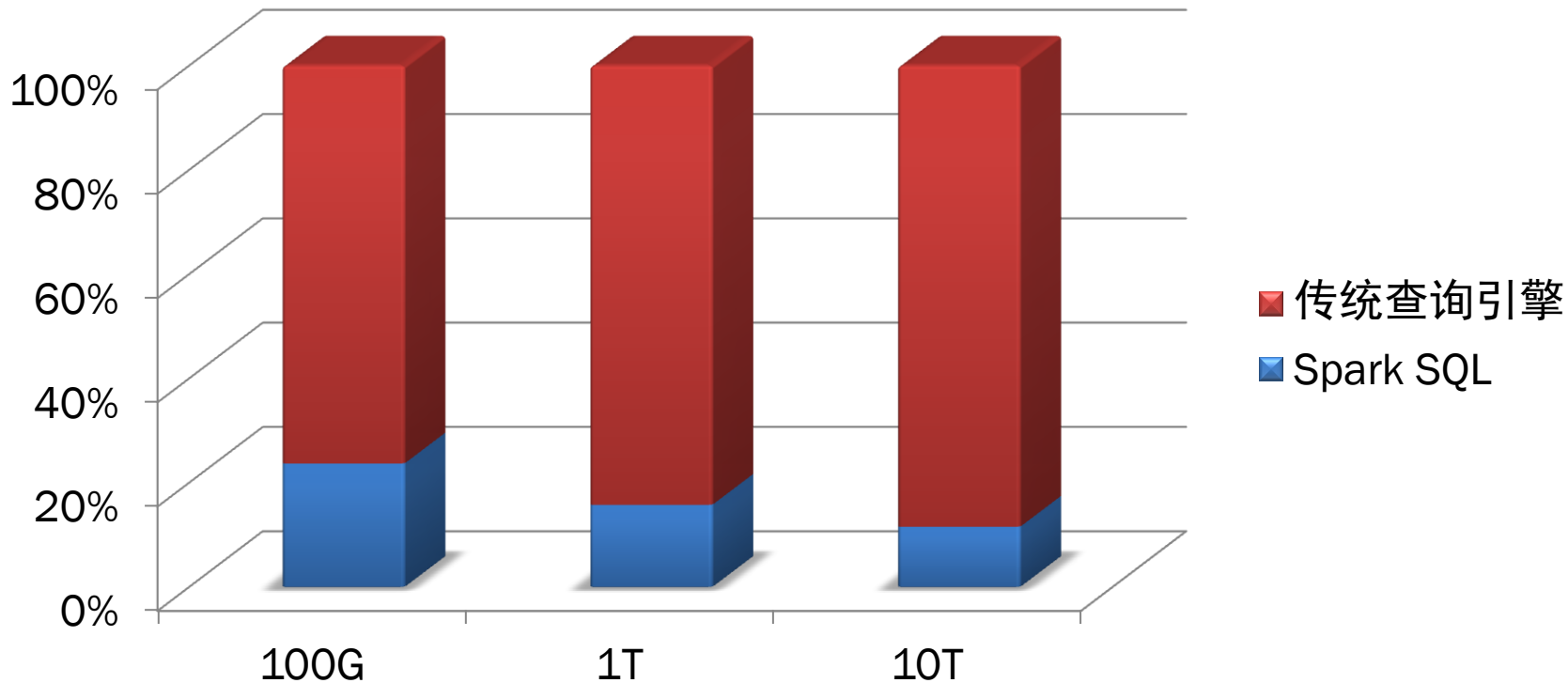
Adhoc Query On Spark SQL(2)



Adhoc Query On Spark SQL(3)



Adhoc Query On Spark SQL(4)



传统的SQL 查询引擎和Spark SQL在不同数据规模下的查询耗时情况对比

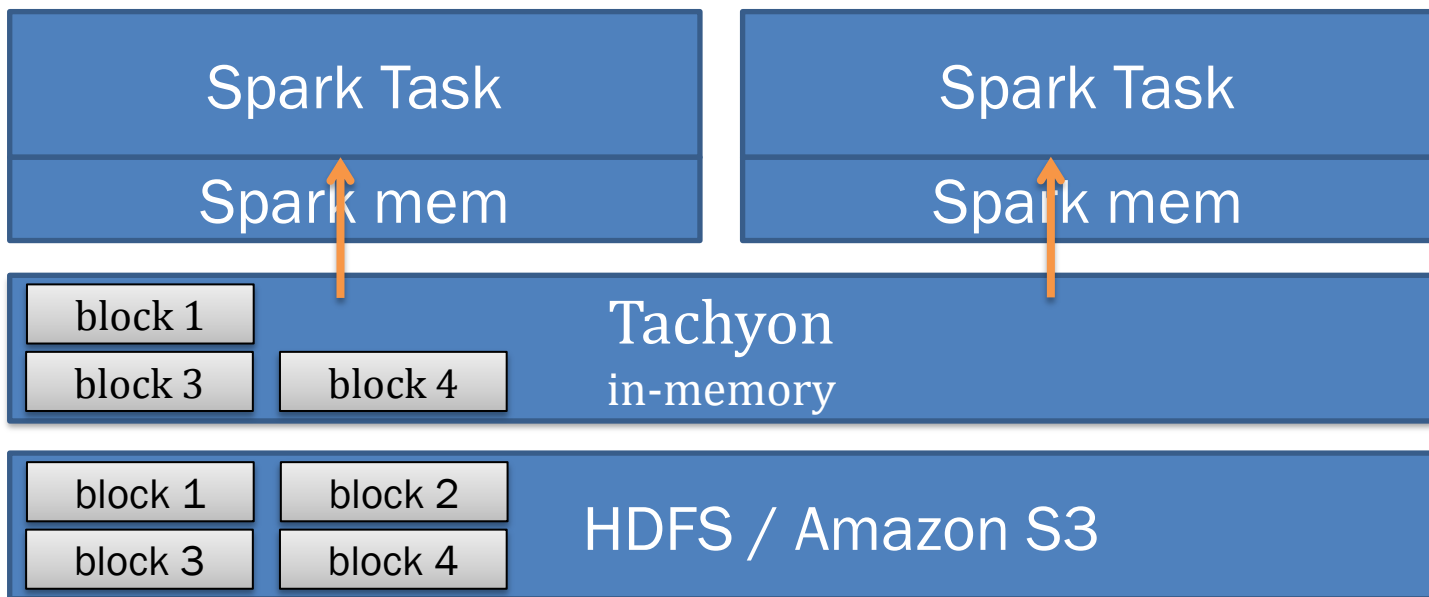
Spark And Tachyon

- Spark存在的一些问题

- 数据计算的过程中，不同的Task之间有共享数据的需求，当需要与文件系统交互时，数据的读写速度就成为了计算的瓶颈；
- 在数据计算的过程中，如果某个JVM进程Crash,那么cache在JVM里面的数据就会丢失；
- Spark 的execution engine & storage engine在相同的节点上位于同一个JVM进程内，这可能会带来两个问题：
 - 1) 增加GC的负担，
 - 2) 相同的数据block可能会被重复cache.

Spark And Tachyon(1)

通过引入Tachyon来改进和优化Spark的计算



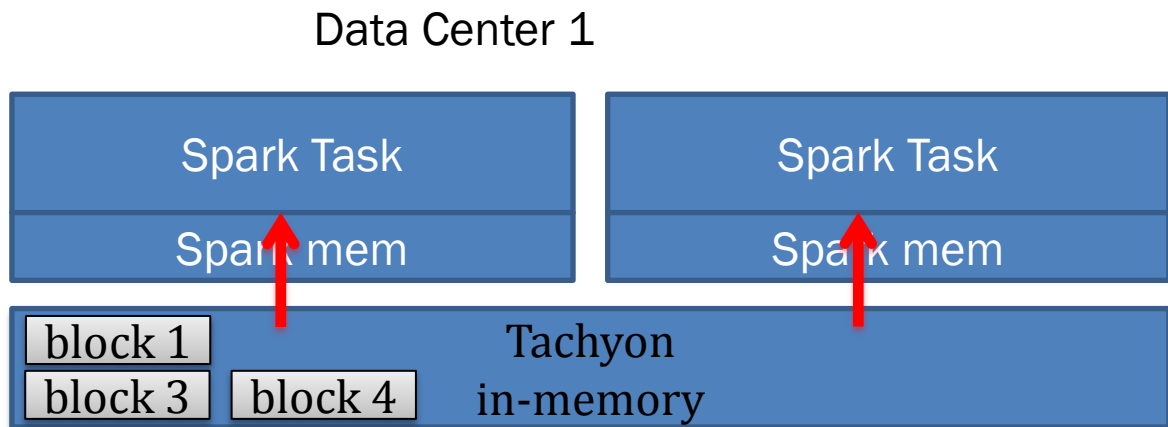
Spark And Tachyon(2)

- Tachyon在Baidu应用场景下面临的一个问题
 - 数据节点和计算节点可能不在同一个数据中心，会带来非常大的数据访问延迟；
 - 尤其对于前面介绍的ad-hoc query的案例 来说，这个问题可能是系统面临的最主要的性能问题；

Spark And Tachyon(3)

- 解决方案
 - 使用 Tachyon 做为 transparent cache layer
 - Cold query: 从远程存储节点读取数据
 - Warm\hot query: 直接从Tachyon 读取
 - 在百度, 当前有一个大概50台机器的集群部署了Spark和Tachyon的服务:
 - 提供基于Spark SQL 的ad-hoc 查询服务
 - 用Tachyon来 做为 transparent cache layer
 - 我们在Warm\hot query上获得了10X +的性能提升

Spark And Tachyon(4)



Data Center 2

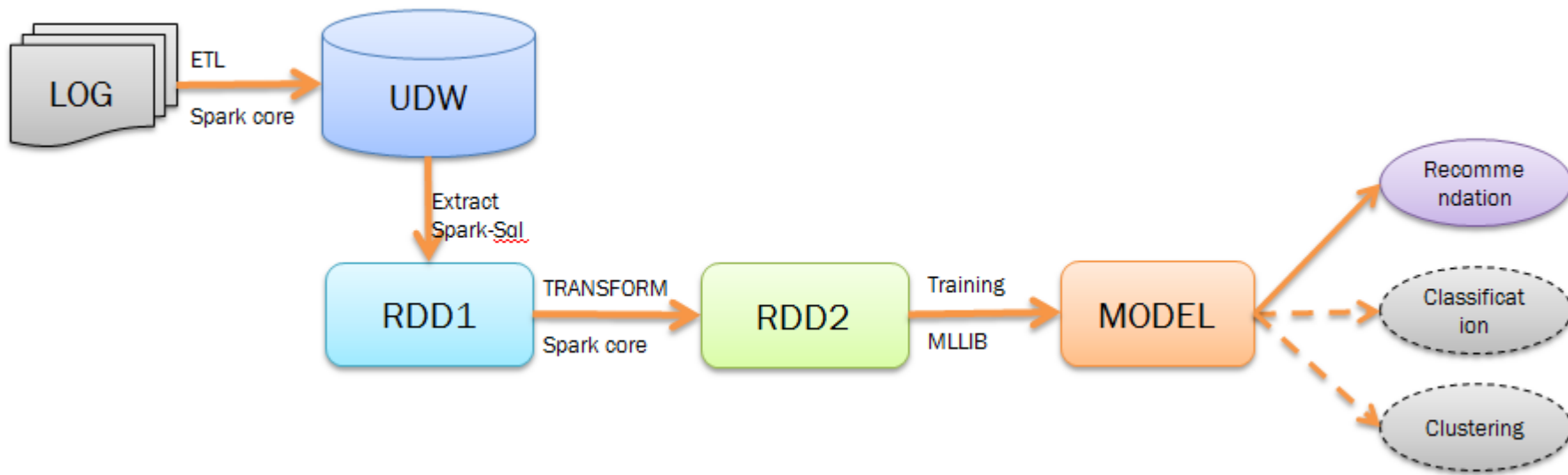
Under File System (HDFS)

某一组实测数据：

- Read from remote data center: ~ 100 ~ 150 seconds
- Read from Tachyon remote node: 10 ~ 15 sec
- Read from Tachyon local node: ~ 5 sec

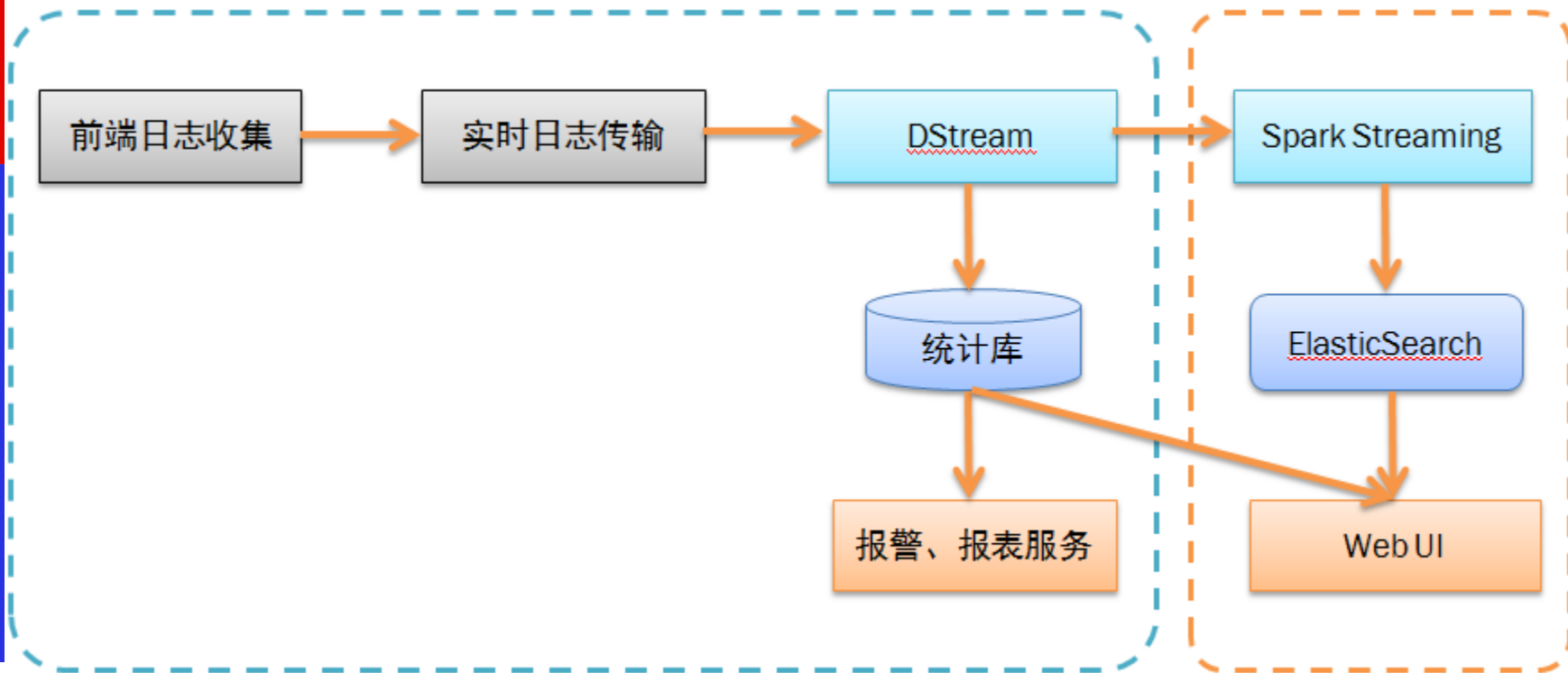
Spark SQL And MLlib

- 推荐系统一站式解决方案



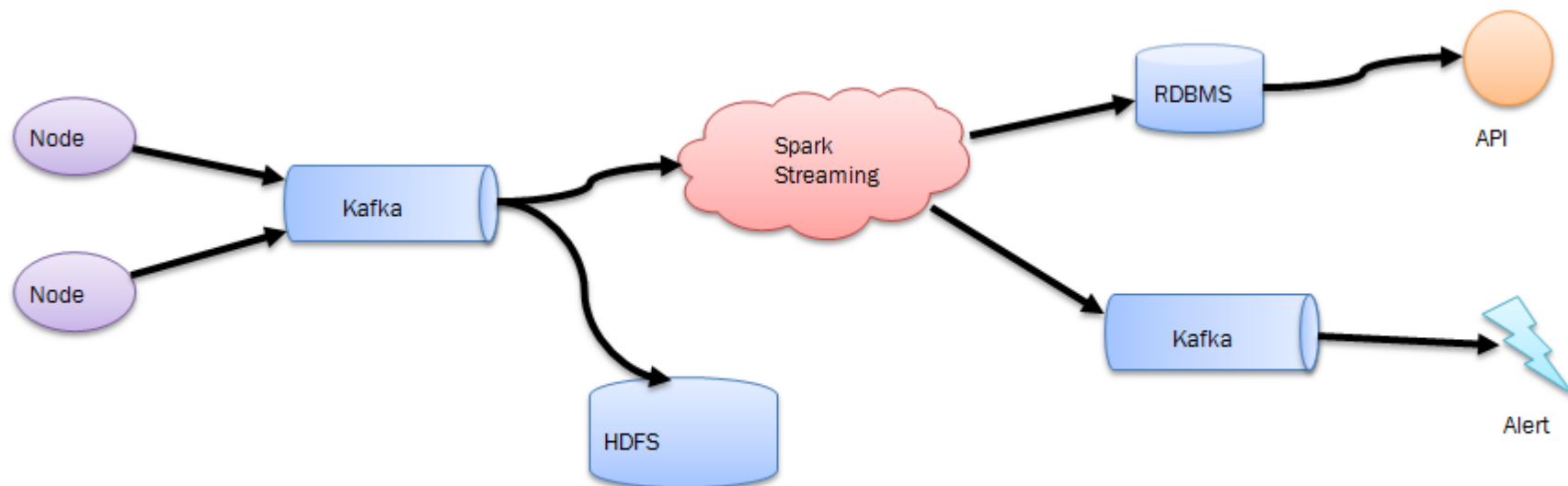
Spark And ElasticSearch

- 流式Tracing系统



Spark And Real-time Security

- 在线入侵检测



百度开放云：Spark On BMR

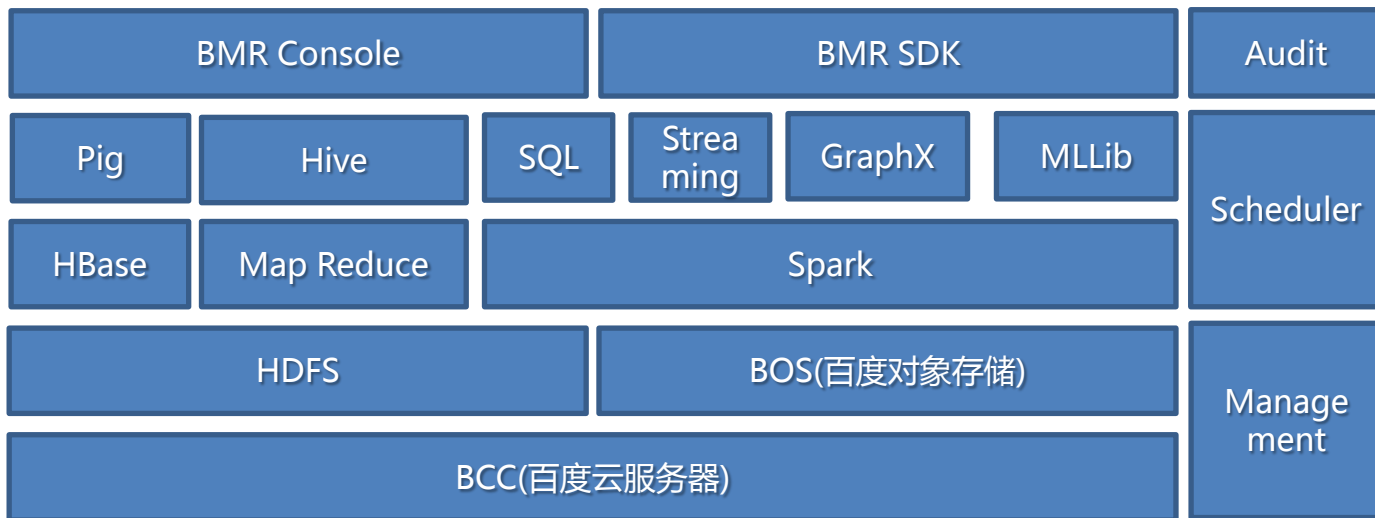
- What's BMR?

- ✓ 百度开放云内面向企业和开发者的数据分析服务产品
- ✓ 基于百度多年大数据处理分析经验和技術积累
- ✓ 提供弹性的基于Hadoop&Spark的计算集群服务
- ✓ 降低大数据分析技术门槛，助力客户提升自己的海量数据分析和挖掘的能力以及业务竞争力

<http://cloud.baidu.com/product/bmr.html>

Spark On BMR(1)

- 大数据全栈解决方案

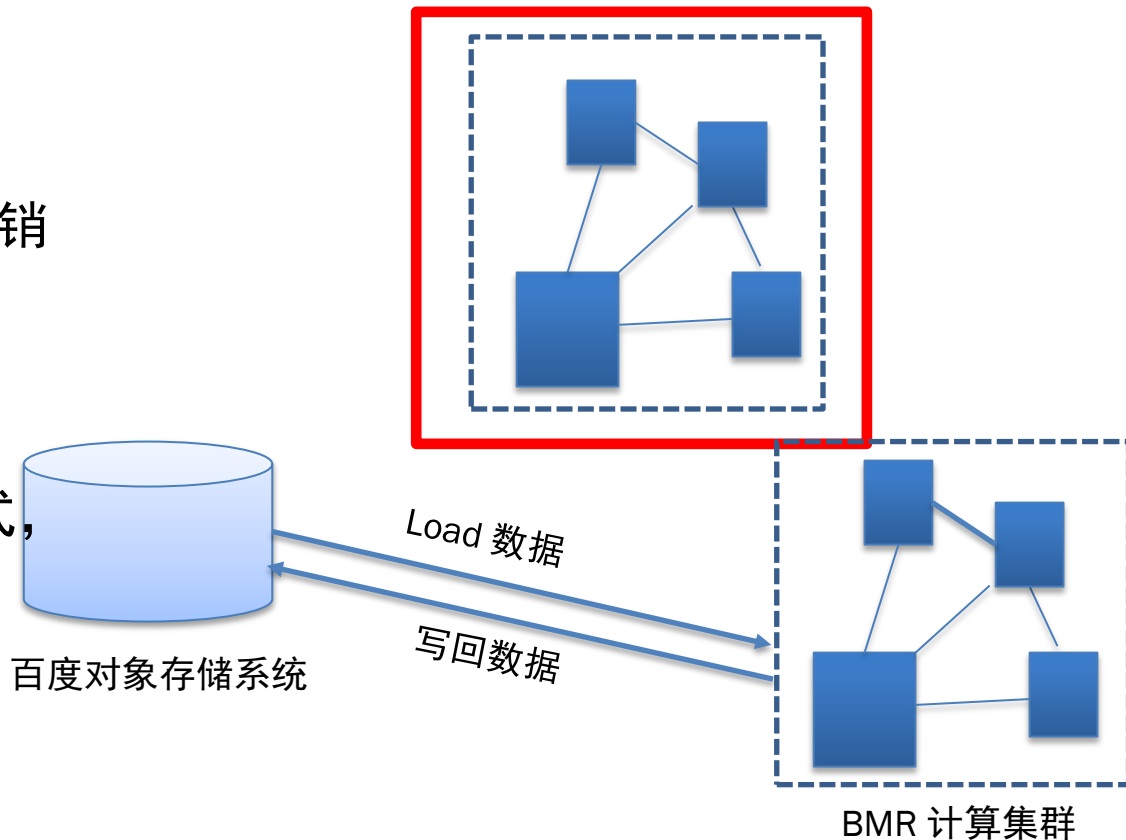


Spark On BMR(2)

- 随用随起

使用即创建，集群空闲即销毁：

- 最大限度节省预算
- 可选 Long Running 模式，
不会自动销毁
- 多种套餐可选

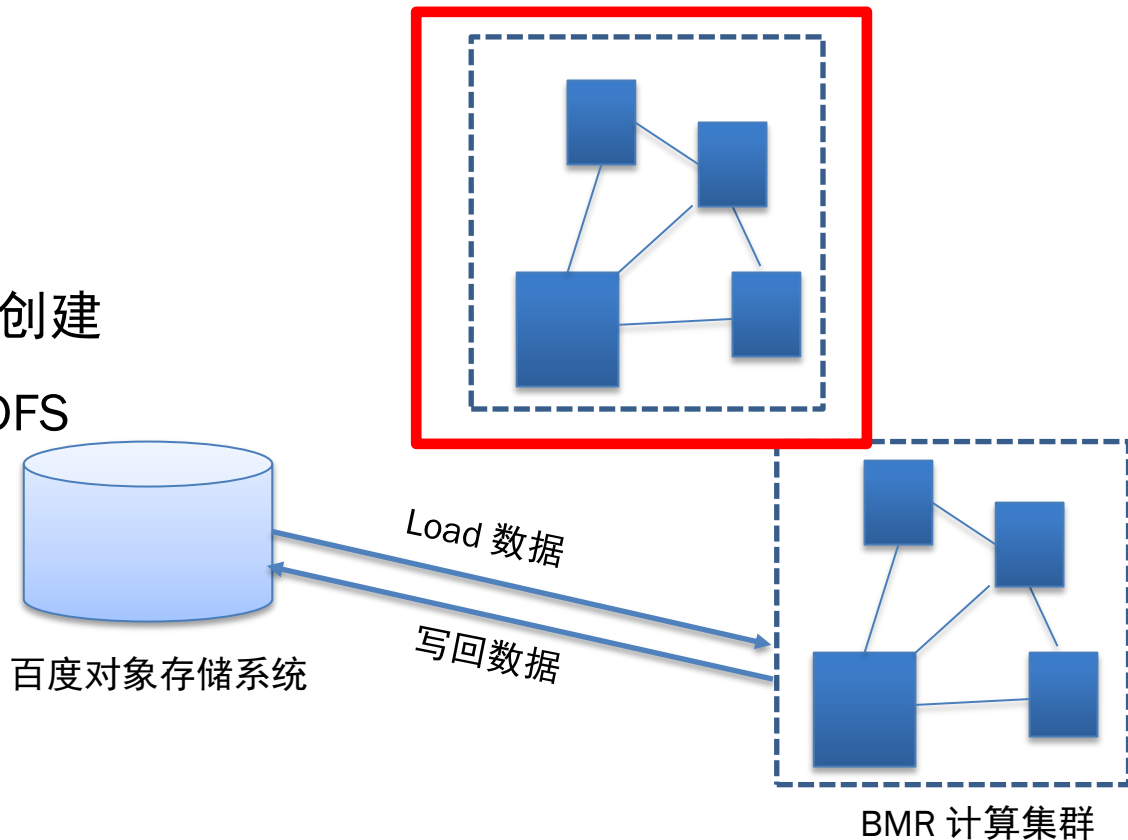


Spark On BMR(3)

- 快速创建

集群快速创建：

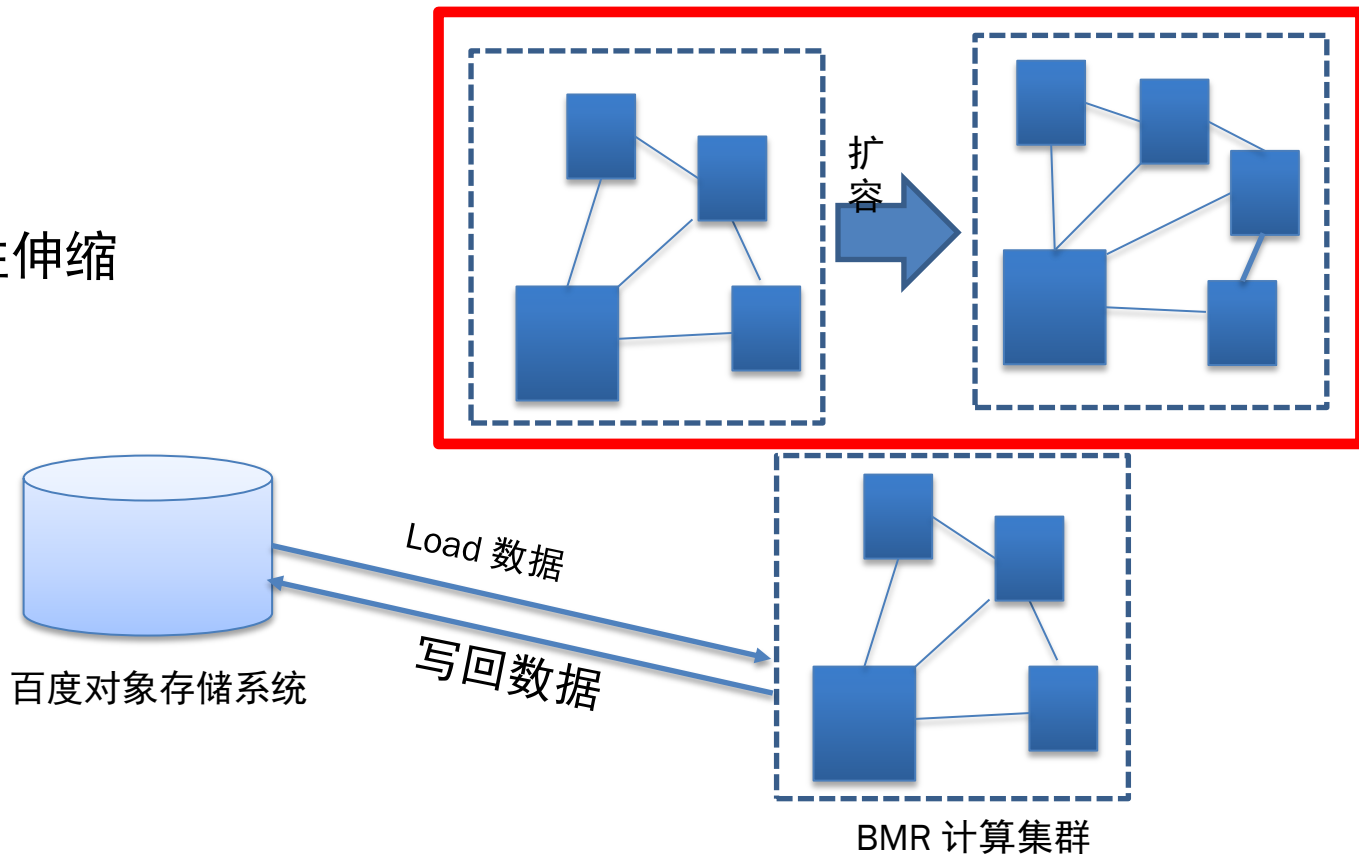
- 3~5分钟即可完成集群创建
- 包含完整的 Spark + HDFS
+ Yarn



Spark On BMR(4)

- 弹性伸缩

节点规模可弹性伸缩



Spark On BMR(5)

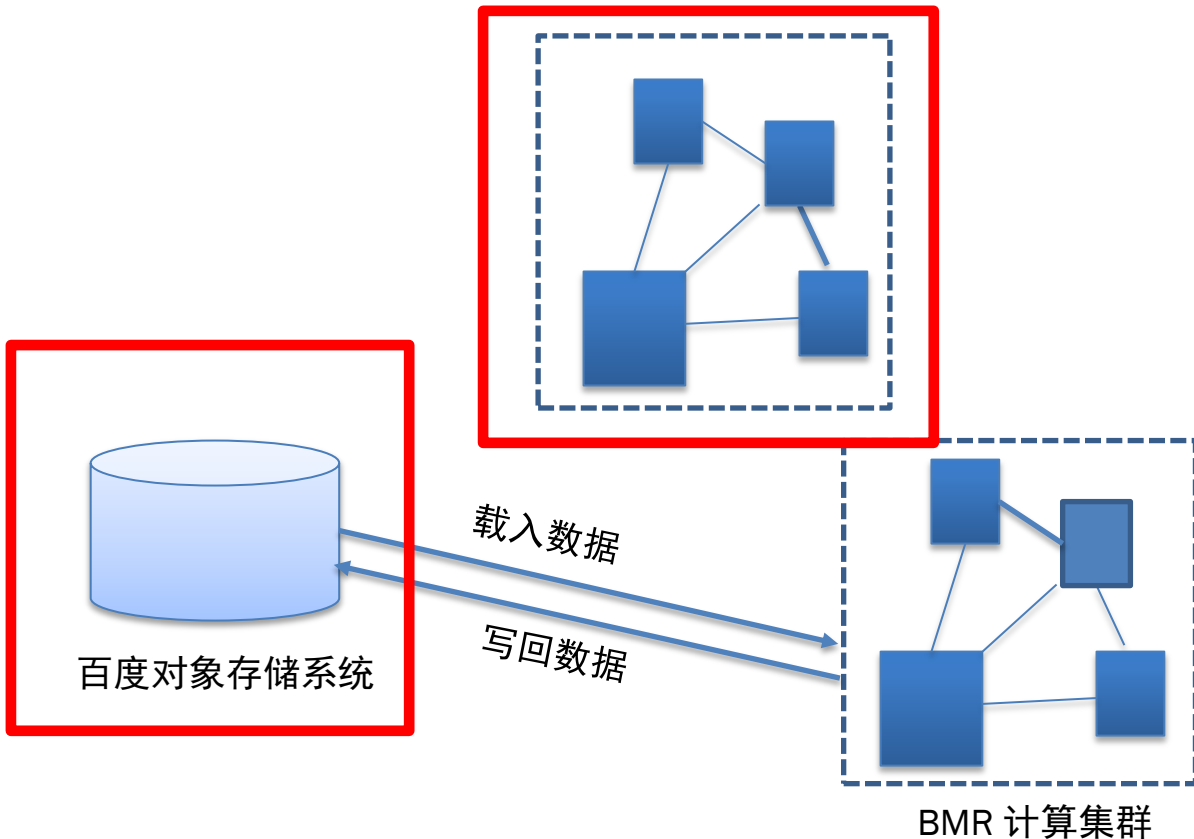
- 安全、可靠

用户间网络完全隔离：

- 虚拟的独立网络

百度对象存储：

- 用户的输入、输出
- 集群本身的备份信息





我们会致力于Spark在百度的传播与推广！

- ✓ Spark Lessons :Wiki,Video, Sandbox
- ✓ Technical Day with on-site QA
- ✓ Spark learning group
- ✓ Scala language learning group
- ✓ ...

THANKS



百度开放云

<http://cloud.baidu.com>

@kaiwang

<http://weibo.com/kaiw>

2015/06/27