



插上翅膀的大象 基于Spark on Yarn的淘宝数据挖掘平台

淘宝技术部——数据挖掘与计算
明风



目录

- 为什么选择Spark On Yarn
- Spark On Yarn的原理和框架
- 淘宝在Spark On Yarn上做的工作
- 基于Spark on Yarn的数据挖掘平台架构
- 案例性能



Hadoop在数据挖掘遇到的问题

- 多次迭代
- 中间数据的序列化和反序列化
- 简单的MR模式 vs 复杂的机器学习算法
- OO编程 vs 函数式风格
- 图计算能力



Why Spark

RDD

- 内存计算
- 快速迭代
- DAG

Scala

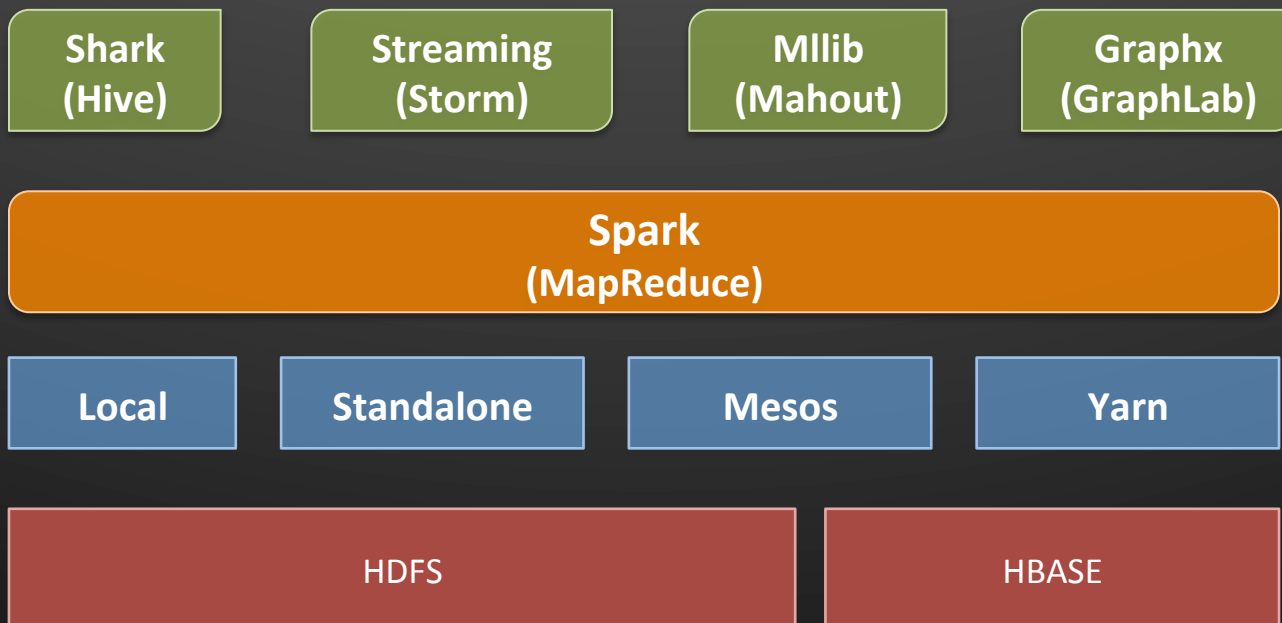
- FP编程
- Actor模型
- 并发能力

Hadoop

- MapReduce
- HDFS访问

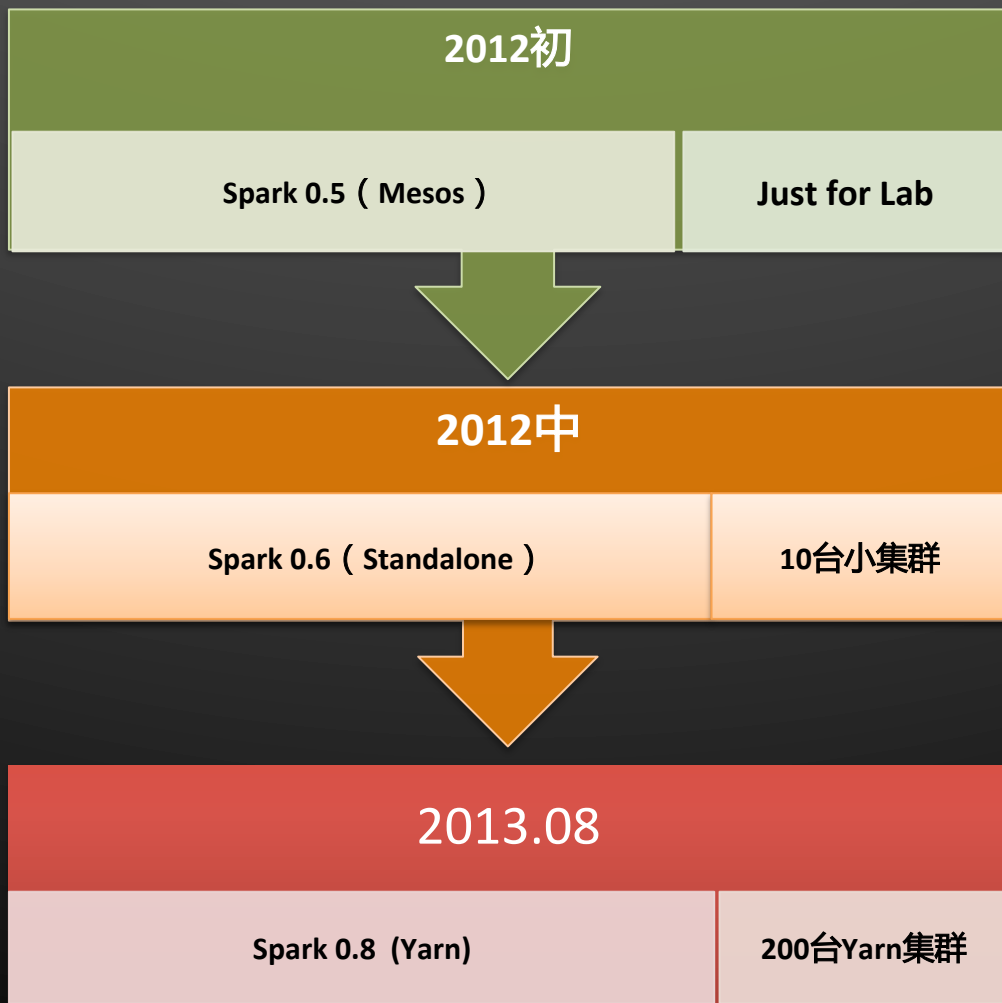


Spark的生态圈





淘宝的Spark之路

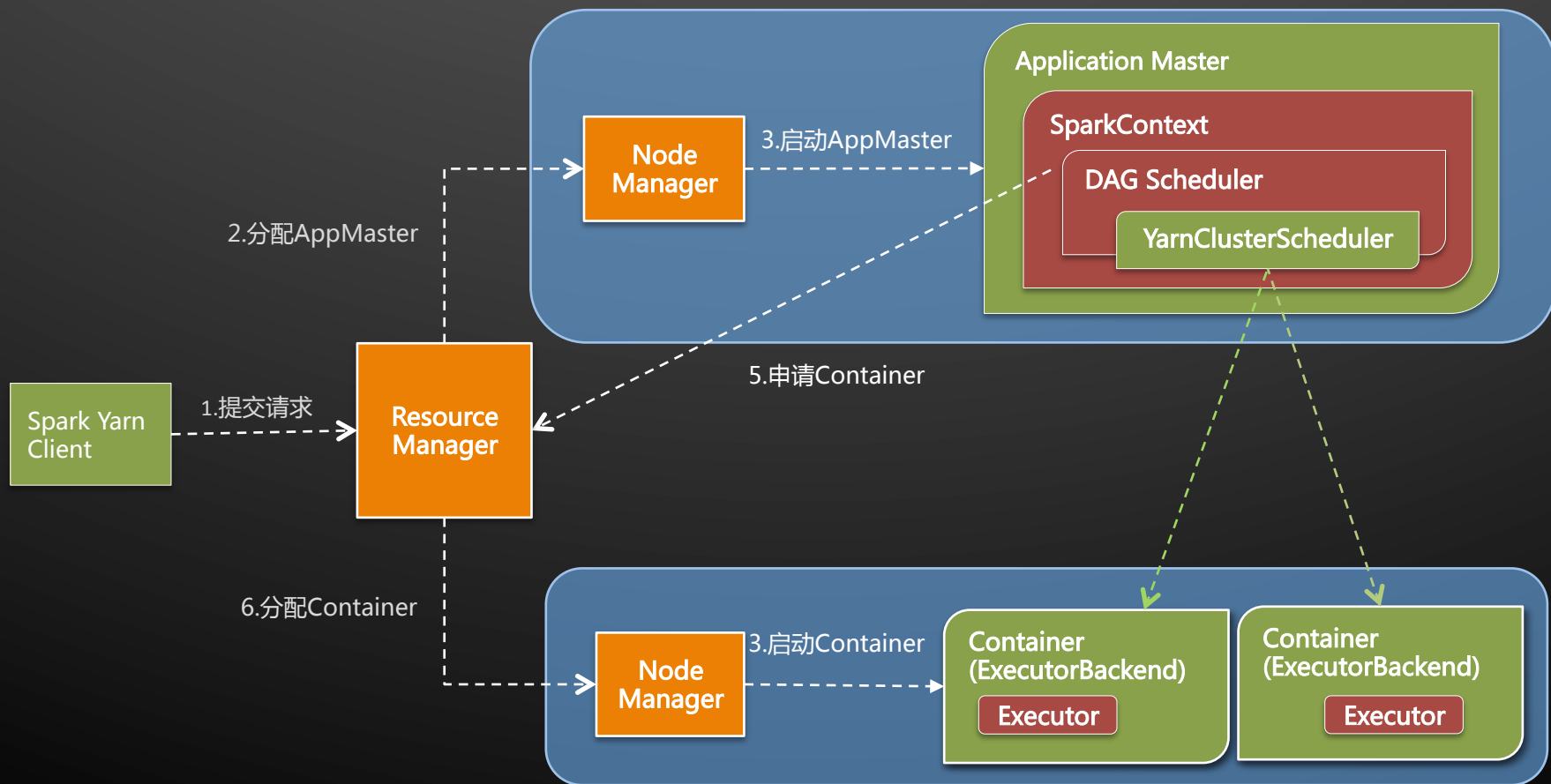


Yarn版本:
0.23.7

阿里云梯1目前规模:
5000*2

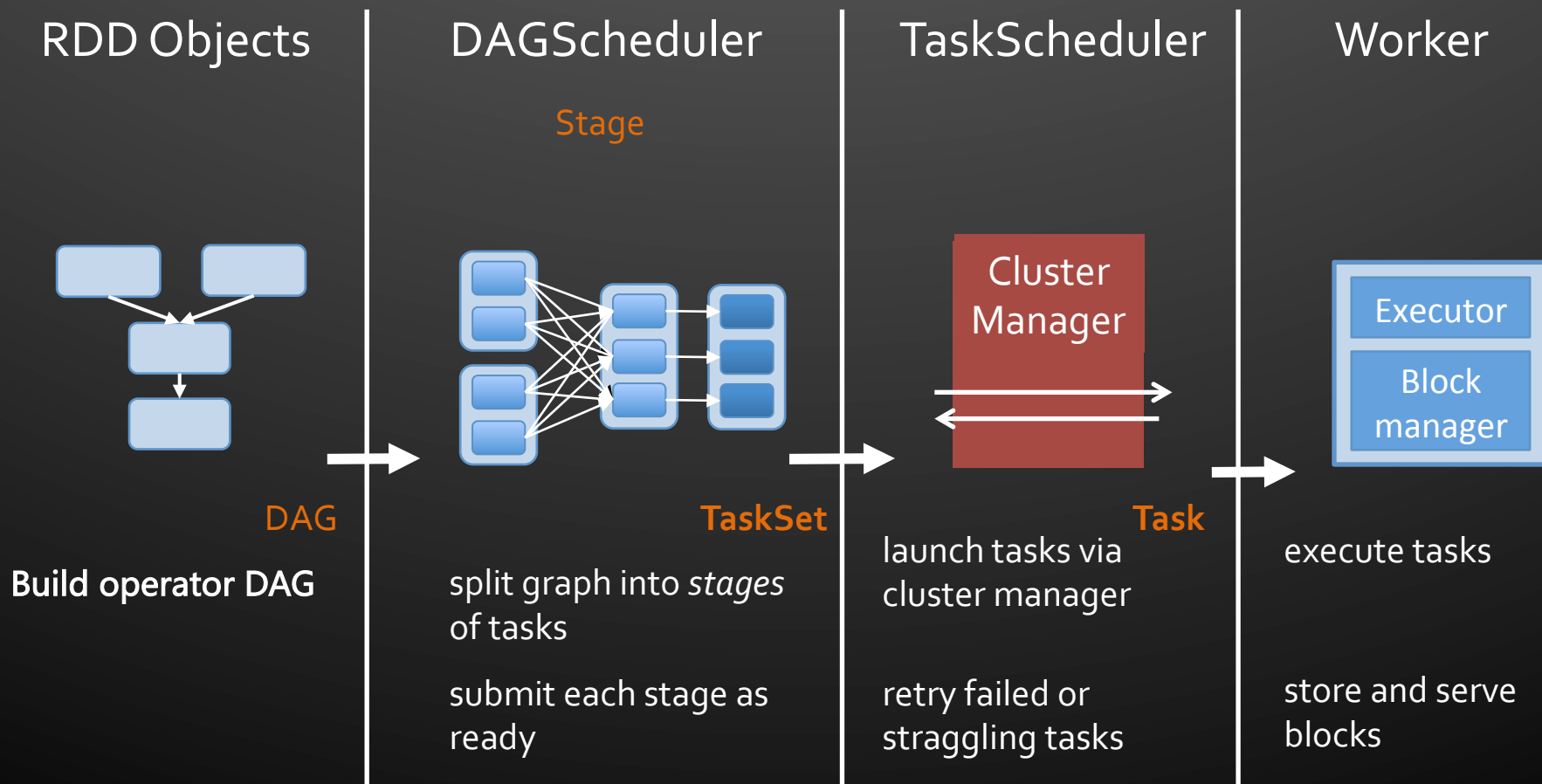


Spark on Yarn的框架





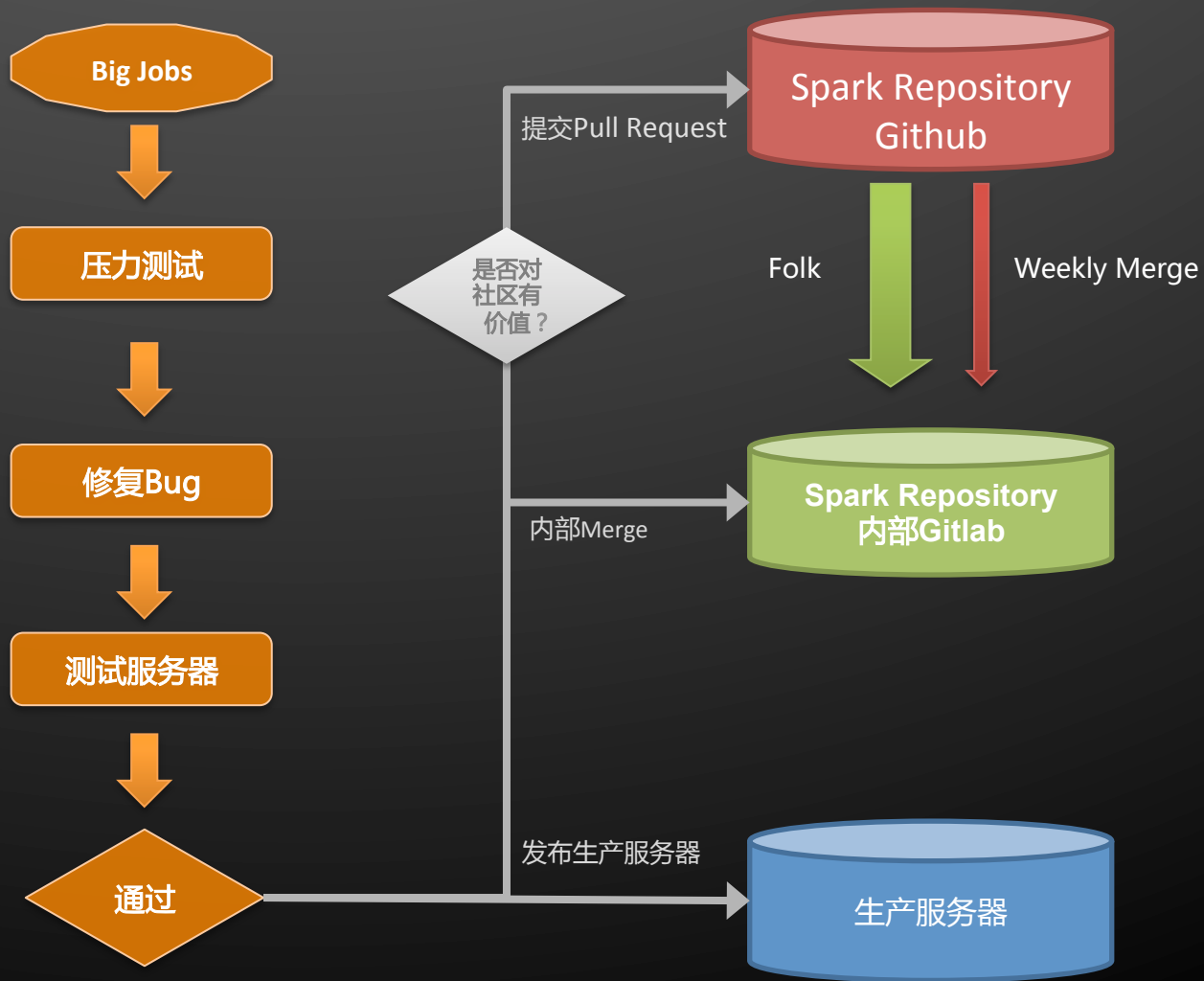
调度过程



Introduction to Spark Internals @Matei



我们的社区跟进模式





我们做了什么？

- 公开的

1. Pull Request 681: Remove active job from idToActiveJob when job finished or aborted
2. Pull Request 689: Jobs are always marked as SUCCEEDED even it's actually failed on Yarn
3. Pull Request 757: ResultTask's serialization forget about handling "generation" field, while ShuffleMapTask does

- 未公开的（和云梯Yarn团队密切相关）

1. 增加用户权限管理
2. 工作jar包缓存机制
3. 自动配置Spark临时缓存目录
4. 封装了一个运行脚本，使用资源文件配置Spark作业性能参数
5. 添加一个Syslog来存放完整日志输出
6. Executor监控页面增加跳转到日志的链接

.....



封装的启动命令

**spark_on_yarn **

--queue cug-queue-spark \

--jar /home/spark/mllib/libs/spark-mllib_2.9.3-0.7.0-SNAPSHOT.jar \

--class spark.mllib.recommendation.ALS \

--args yarn-standalone \

--args /group/spark/als/test.data \

--args 3 \

--args 10 \

--args /group/spark/als/output \

--num-workers 10 \

--worker-memory 24g \

--worker-cores 3 \

--master-memory 5g

--conf ~/spark-config.properties

spark-config.properties

spark.akka.frameSize=50

spark.akka.timeout=60

spark.storage.memoryFraction=0.5

spark.akka.askTimeout=60

spark.worker.timeout=60



实现的算法

聚类

Canopy

KMeans

关联规则

FPGrowth

图算法

PageRank

分类

MLR(Softmax
Regression)

协同过滤

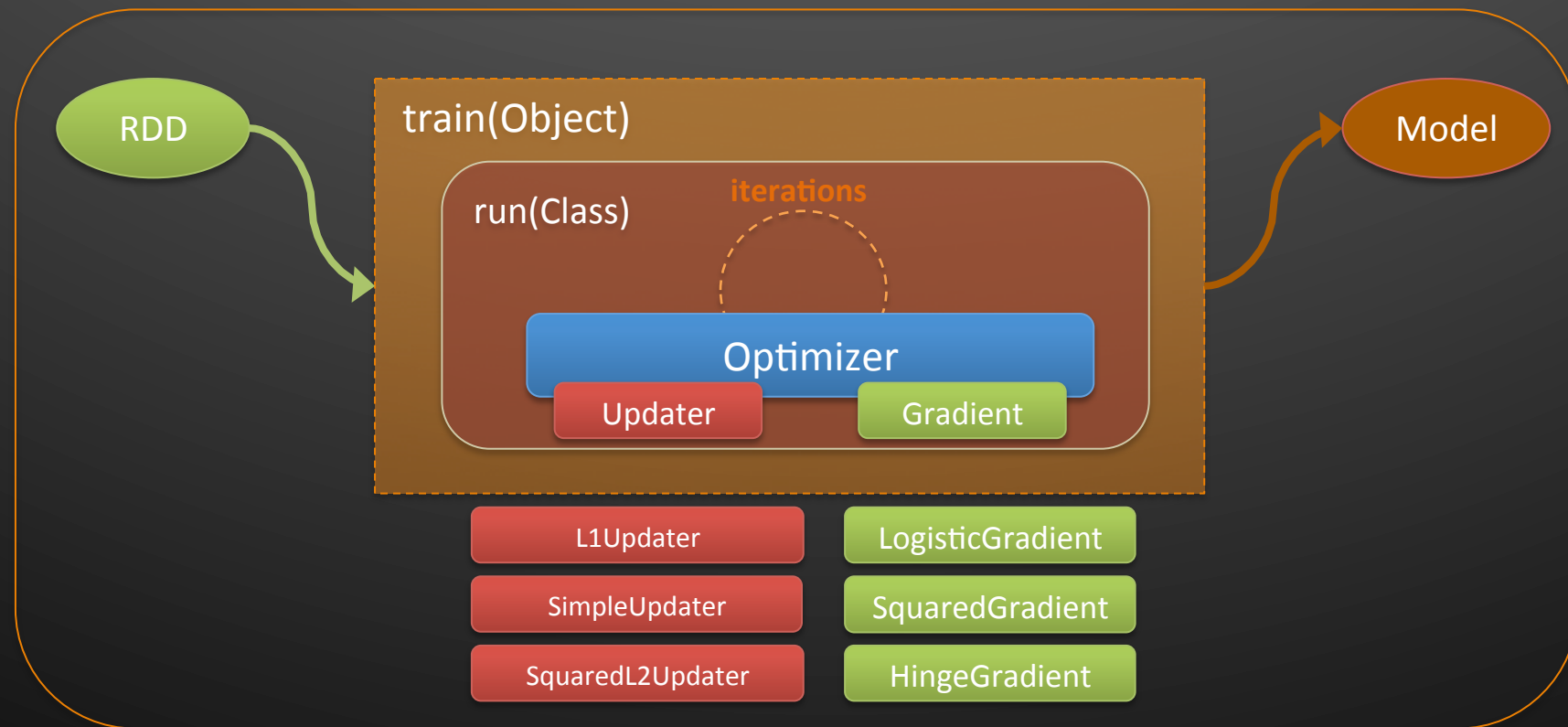
Simrank

ItemBased
(BWC)

UserBased



Mllib的架构图



ALS

KMeans

SVM

LinearRegression

RidgeRegression

LogisticRegression

LASSO

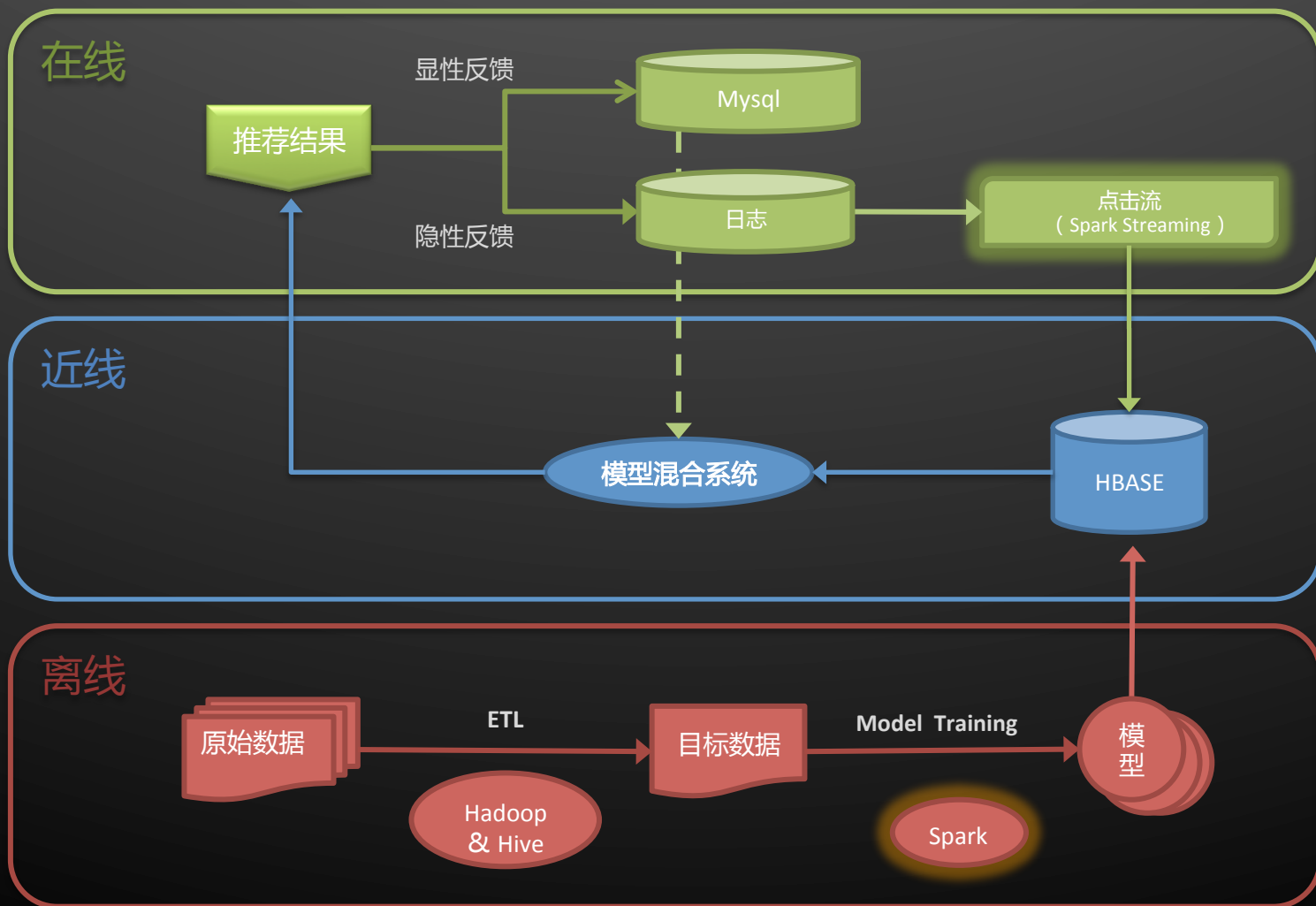
@尹绪森

Jblas(a fast linear algebra library for Java)

gFortran 4.4



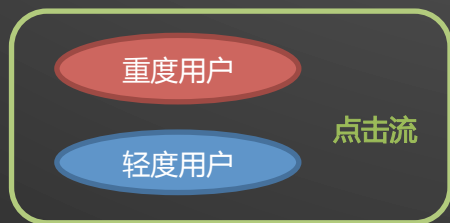
淘宝推荐系统架构





基于兴趣图谱推荐

用户短期行为

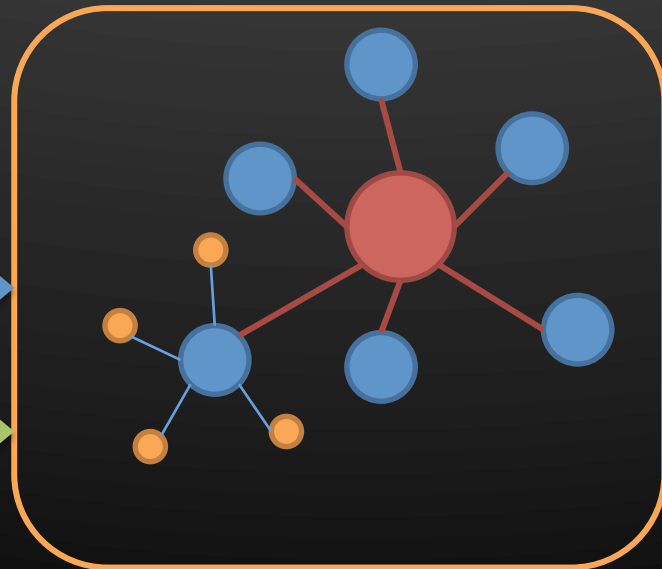


用户长期行为

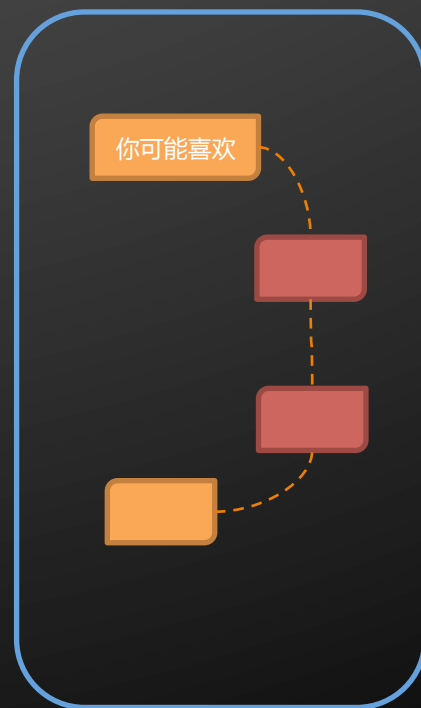


Spark Streaming

GraphX

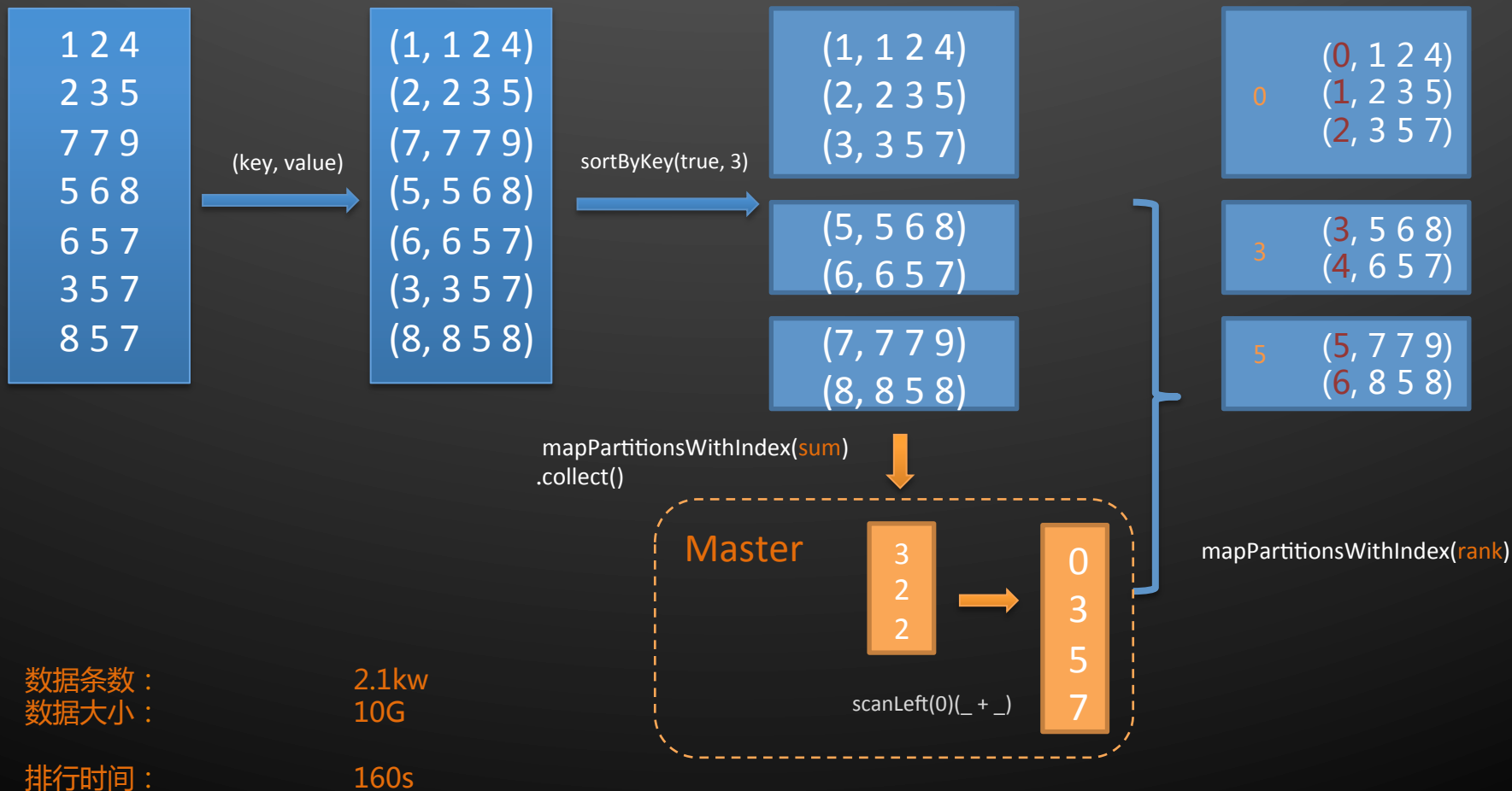


用户推荐Feeds流
(兴趣部分)





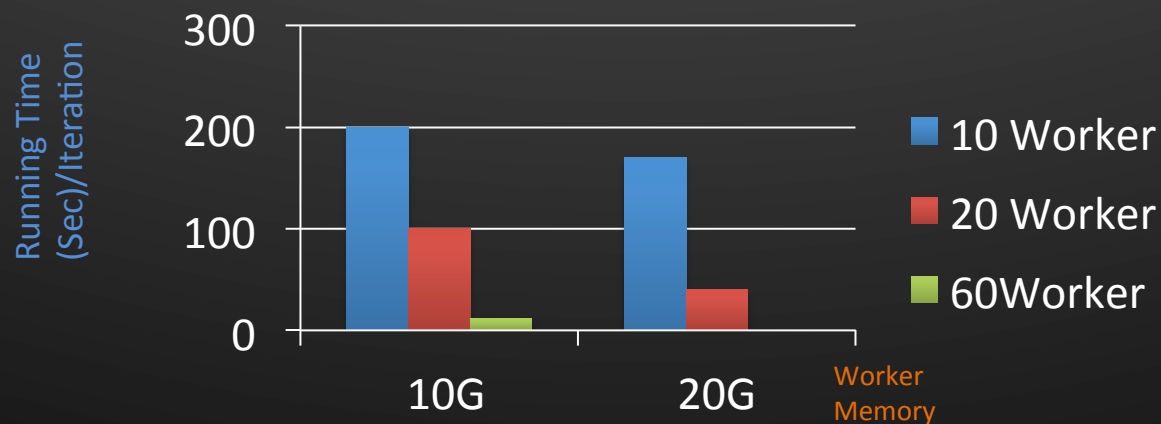
Case 1: Parallel Rank





Case 2: KMeans

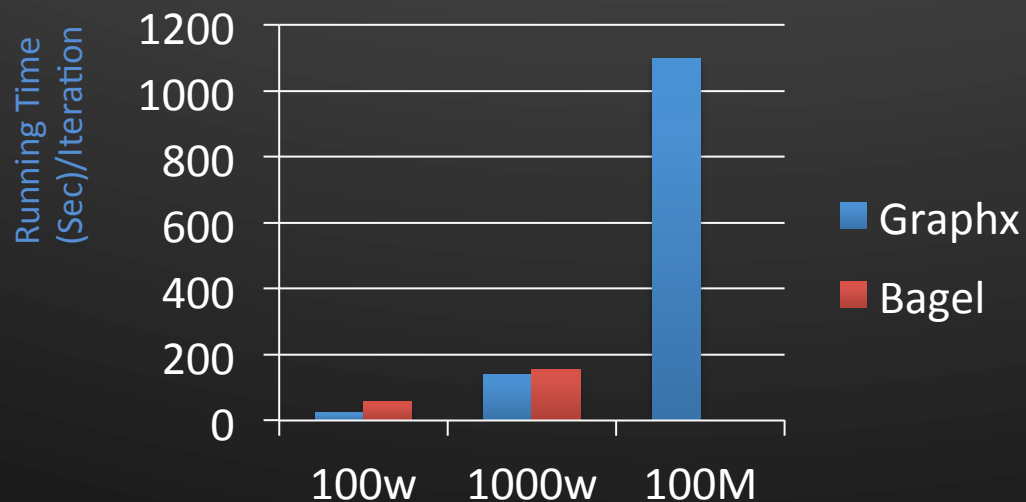
数据条数 : 2.6kw
数据大小 : 10G
中心点 : 80





Case 3 : PageRank

Worker: 50
Worker-Memory: 40g
Worker-Core: 2
Parallel Num : 200





Spark On Yarn目前存在的问题

- 多生态作业竞争
- 机器内存性能
- 粗粒度的资源预申请
- 内存把控能力



加入我们

- 我们需要
 - 推荐系统专家
 - 推荐算法工程师
 - 开发工程师 (Scala , Java)
- 微博: @明风Andy
- 来往





Q & A

谢谢