

Hadoop

-- HDFS和MapReduce

李铁
2012-11-06

主要内容

- **HDFS: 三类节点**NameNode, SecondNameNode, DataNode
- **MapReduce工作原理**, JobTracker, TaskTracker
- 运行WordCount例子
- Hadoop适合做细胞分割吗?
- Hadoop适合做什么?

HDFS-简介

Hadoop的分布式计算与分布式存储都采用了主/从 (Master/Slave) 结构.

HDFS的数据按块存储, 默认块大小是**64M** (Hadoop主要是为了处理大数据, 为了减少磁盘寻道时间, 那么数据传输时间>>寻道时间)

三类节点:

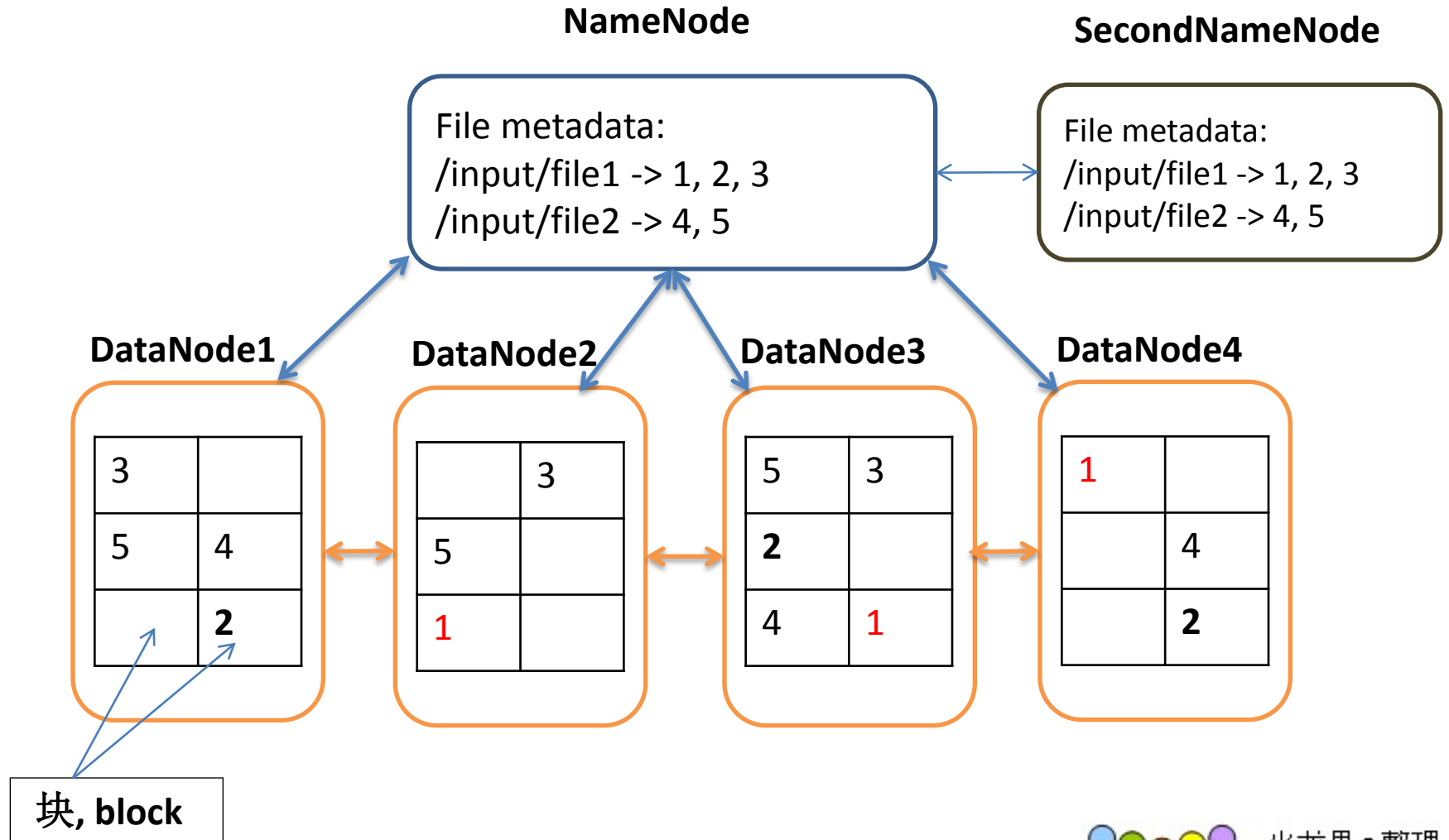
NameNode: HDFS的主端 (Master), 用于管理文件系统的命名空间, 维护整个文件系统树内所有文件和索引目录
(没有NameNode HDFS就无法使用, 找文件就找不到了! 万一NameNode失效了, 怎么办?)

SecondNameNode: NameNode的复本, 万一NameNode失效, 便于快速恢复NameNode. 一般不和NameNode在同一台机器上.

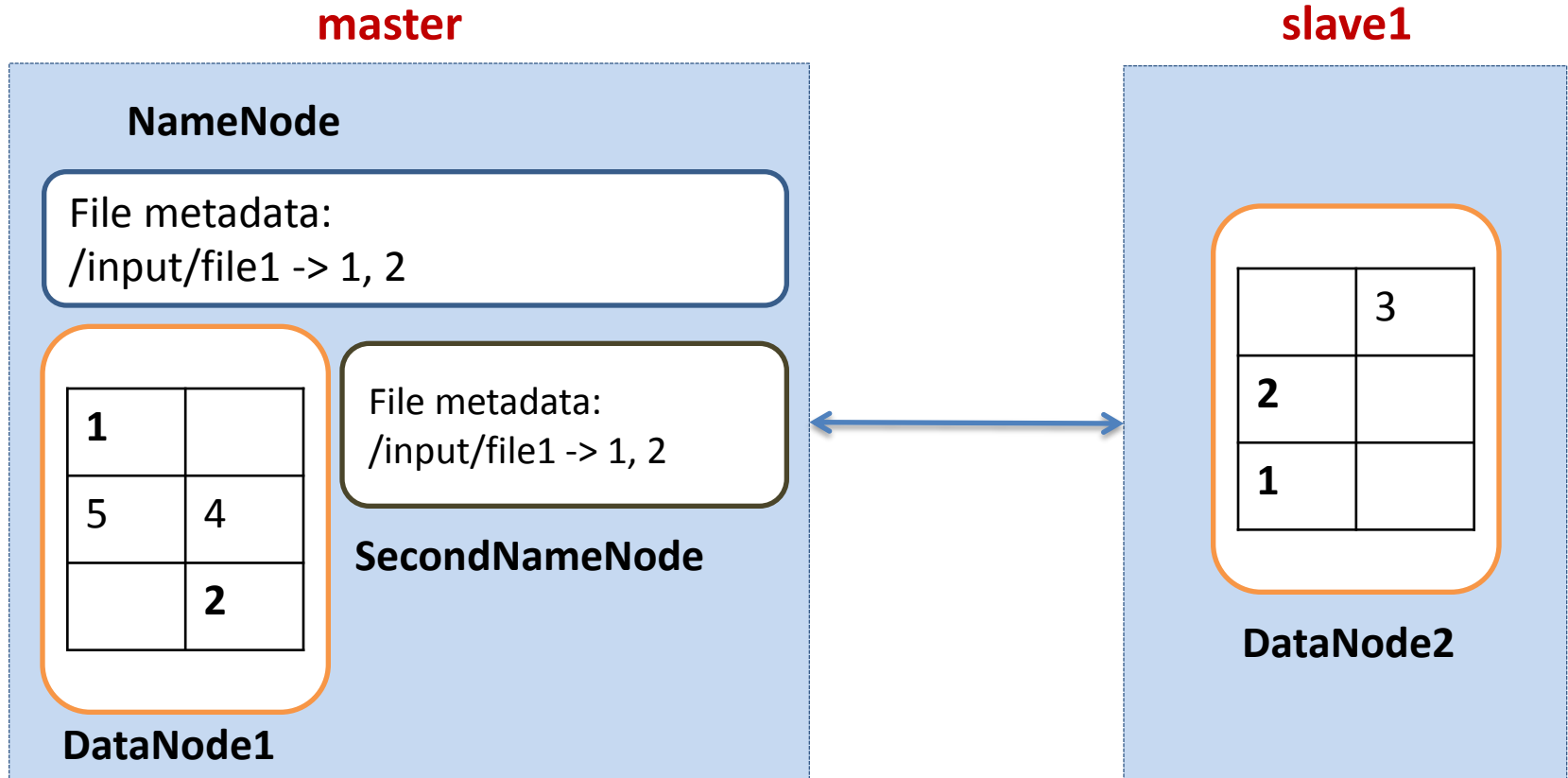
DataNode: slave端, 存储数据

三类节点一般在不同的机器上!

HDFS-三类节点



HDFS-本地小集群



命令:

- `start-dfs.sh` 启动HDFS, 运行NameNode, SecondNameNode, DataNode 进程
- `jps` 查看Java后台运行的进程 (jps: Java Virtual Machine)
- `ssh slave1` 远程连接slave1 (ssh: Secure Shell)

HDFS-文件信息

命令: `hadoop fsck /input/file1 -files -blocks -locations`
查看文件信息(fsck: File System Check)

```
/input/file1 113213521 bytes, 2 block(s): OK
0. blk_-2963066870703248498_1006 len=67108864 repl=2 [222.204.215.151:50010,
222.204.215.152:50010]
1. blk_5458412688561008755_1006 len=46104657 repl=2 [222.204.215.151:50010,
222.204.215.152:50010]
```

Status: HEALTHY

Total size: 113213521 B

Total dirs: 0

Total files: 1

Total blocks (validated): 2 (avg. block size 56606760 B)

Minimally replicated blocks: 2 (100.0 %)

Default replication factor: 2

Average block replication: 2.0 // 每一块的复本数

Number of data-nodes: 2 // 数据分布的数据节点数

MapReduce-简介

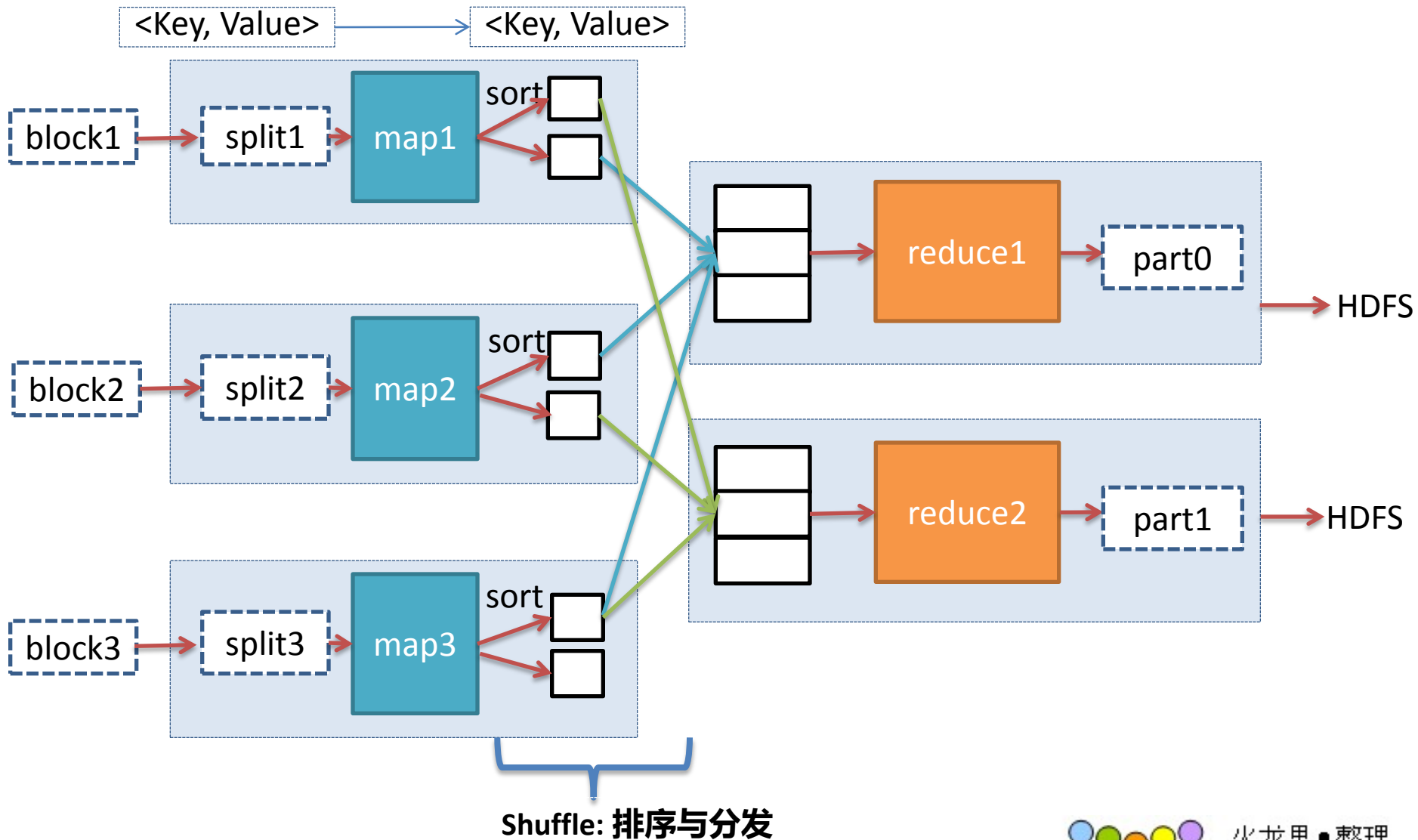
MapReduce是并行处理数据的计算框架. 要对分布式数据进行处理最简单的就只要实现两个方法, 即Map 和 Reduce.

Map主要是数据的获取, 处理, 分发

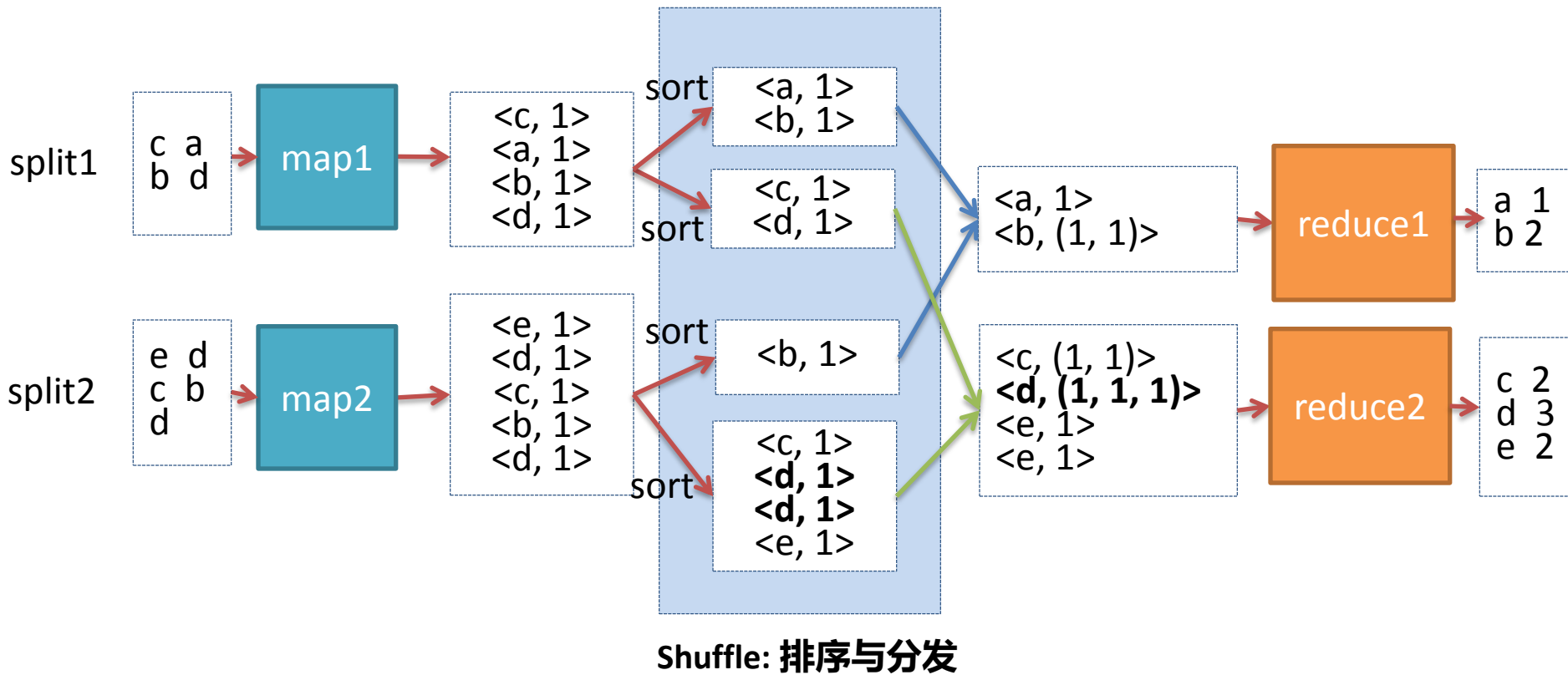
Reduce主要是数据的汇总

MapReduce的数据传输都是以<键, 值>对的形式来传输的.

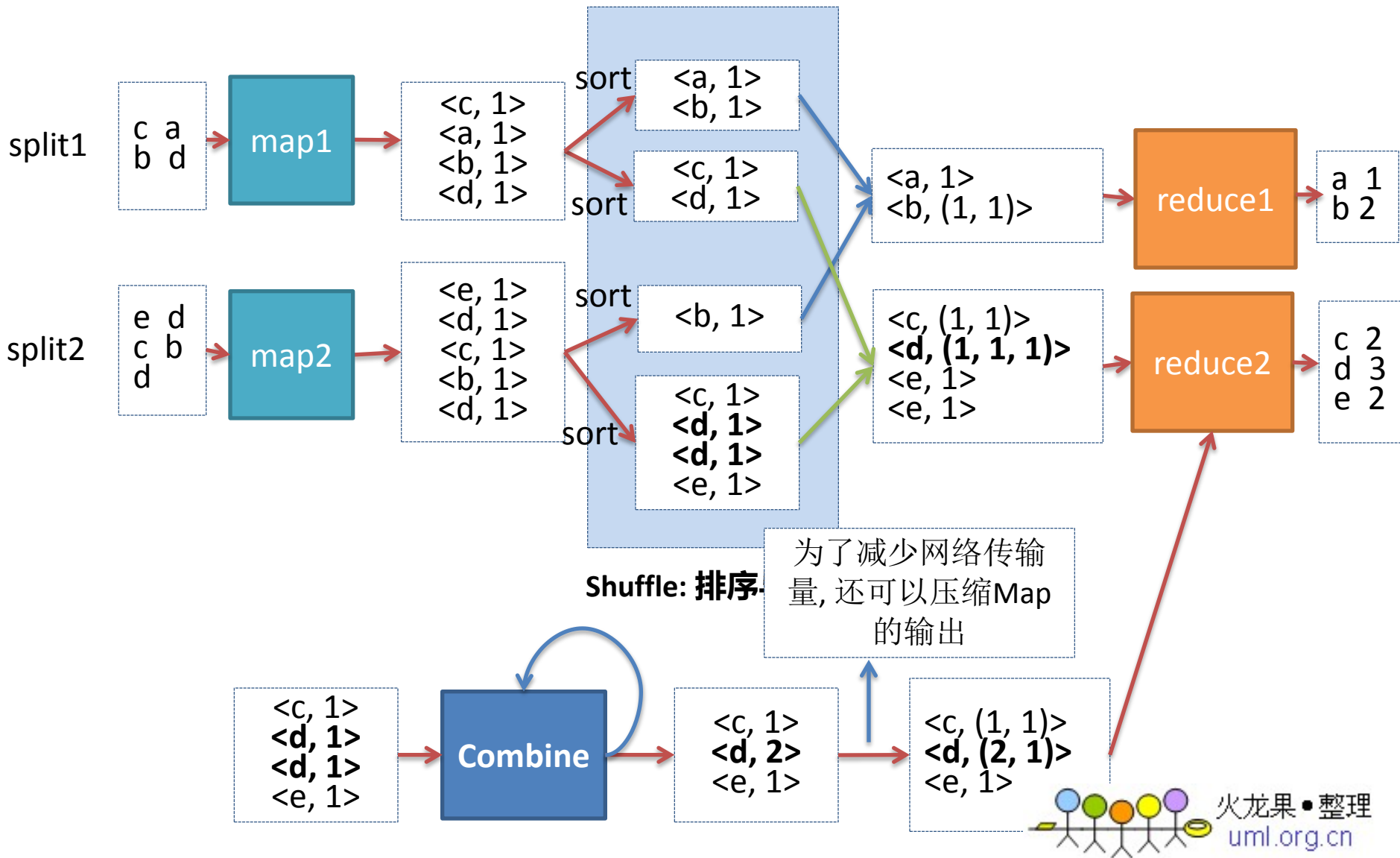
MapReduce-Map Reduce过程



MapReduce-WordCount例子



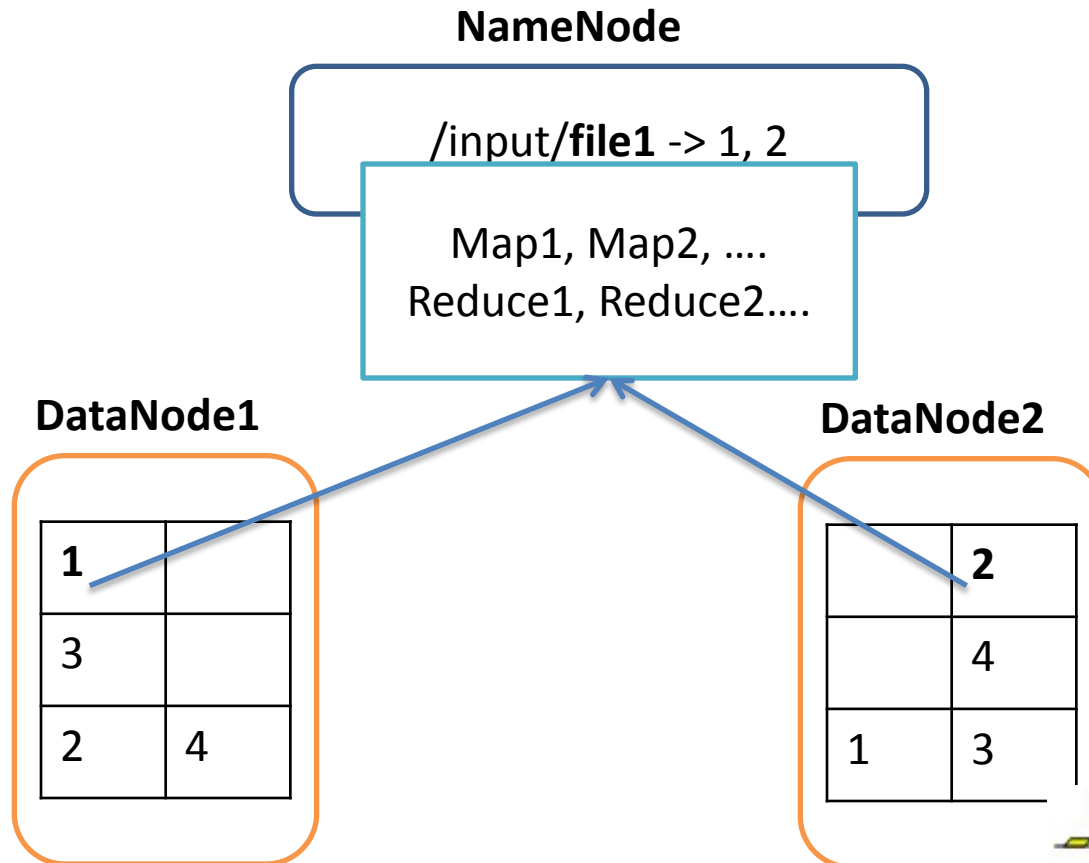
MapReduce-WordCount例子



MapReduce-数据本地化

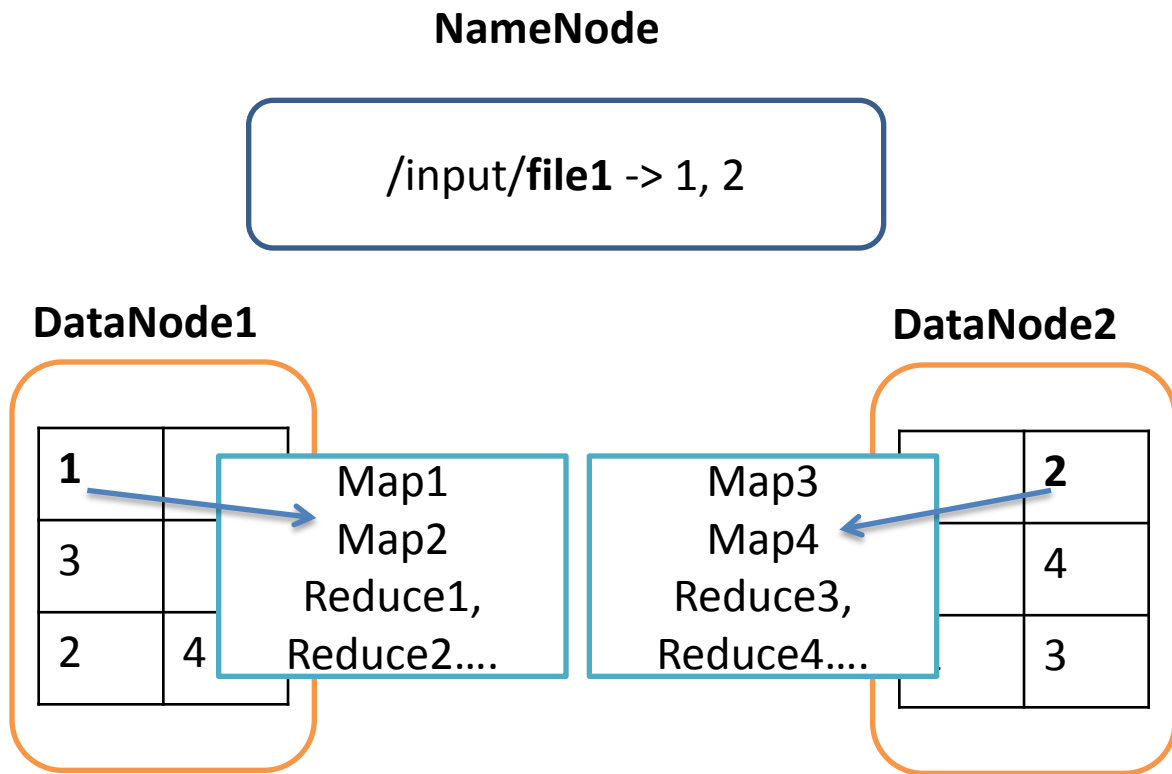
问题: Map, Reduce在哪里运行?

如果: Map在NameNode上运行, 那么所需要的数据需要从DataNode上传到NameNode上, 大数据的传输时间长...



MapReduce-数据本地化

hadoop思想: 移动计算, 实现数据本地化! 把Map, Reduce 移动数据结点上运行!



MapReduce-JobTracker&TaskTracker

客户端提交作业后, 主要由两类进程控制作业的运行:

JobTracker: 整个集群只有一个JobTracker. 为TaskTracker分配任务, 监测任务的运行情况, 调度任务.

(小集群通常运行在NameNode节点上, 大集群JobTracker单独一个节点)

TaskTracker: 运行在DataNode节点上, 每个节点一个TaskTracker进程. 每个TaskTracker可运行运行数个Map, Reduce进程.

MapReduce-JobTracker&TaskTracker

客户端提交作业

```
Hadoop jar WordCount.jar  
Wordcount /input/file1 /output
```

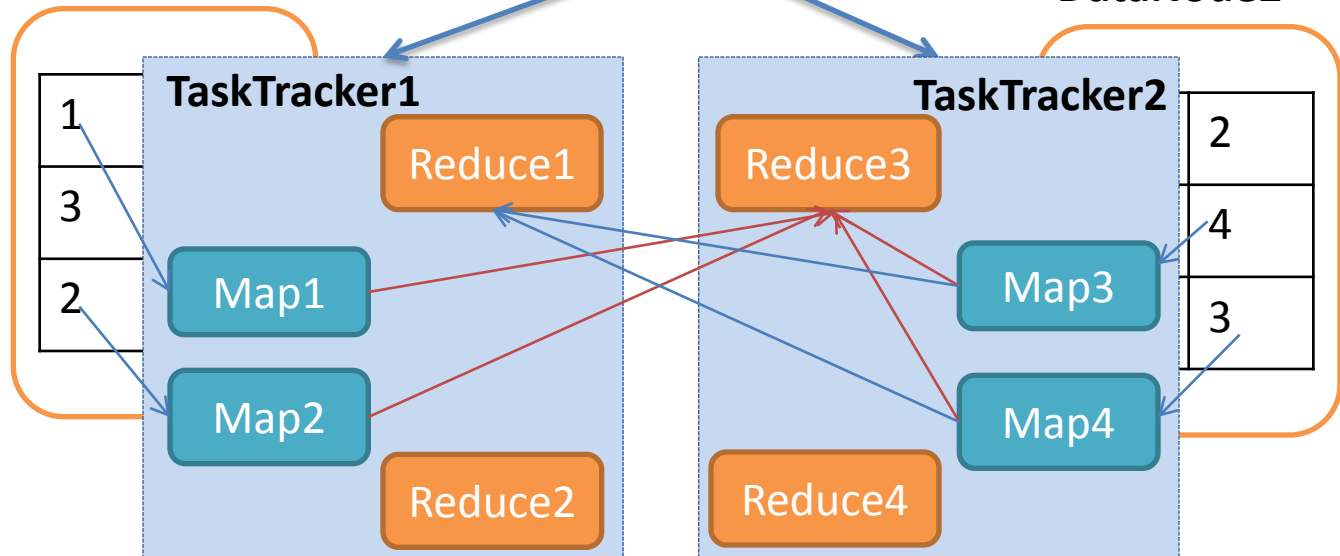
NameNode

/input/file1 -> 1, 2, 3, 4

JobTracker

DataNode1

DataNode2



MapReduce-JobTracker&TaskTracker

客户端提交作业

```
Hadoop jar WordCount.jar  
Wordcount /input/file1 /output
```

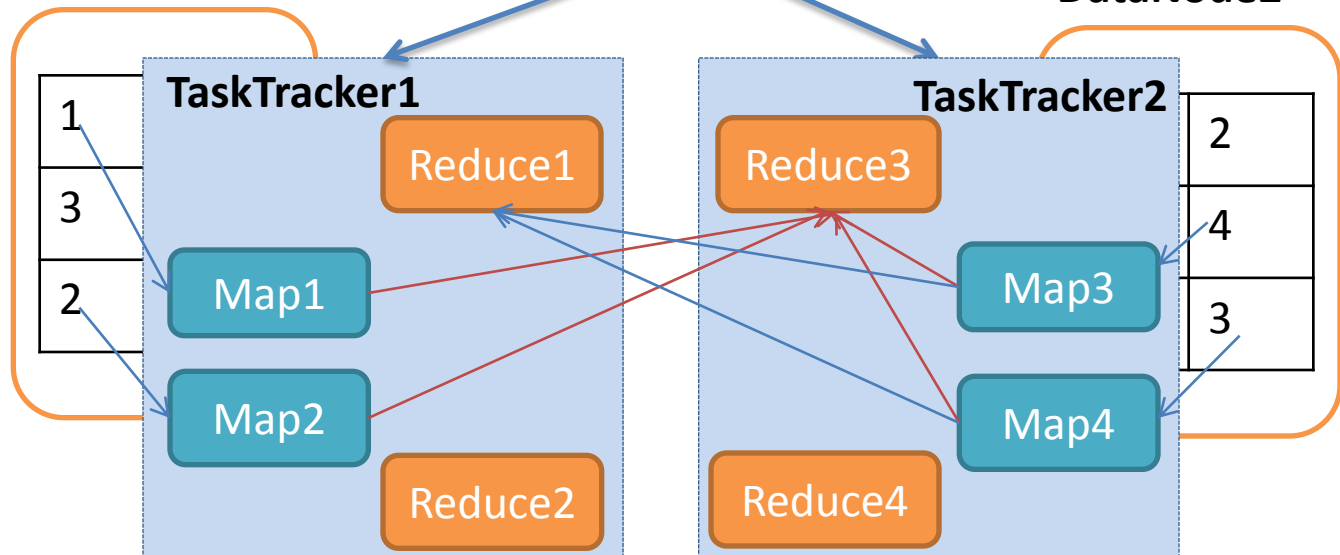
NameNode

/input/file1 -> 1, 2, 3, 4

JobTracker

DataNode1

DataNode2

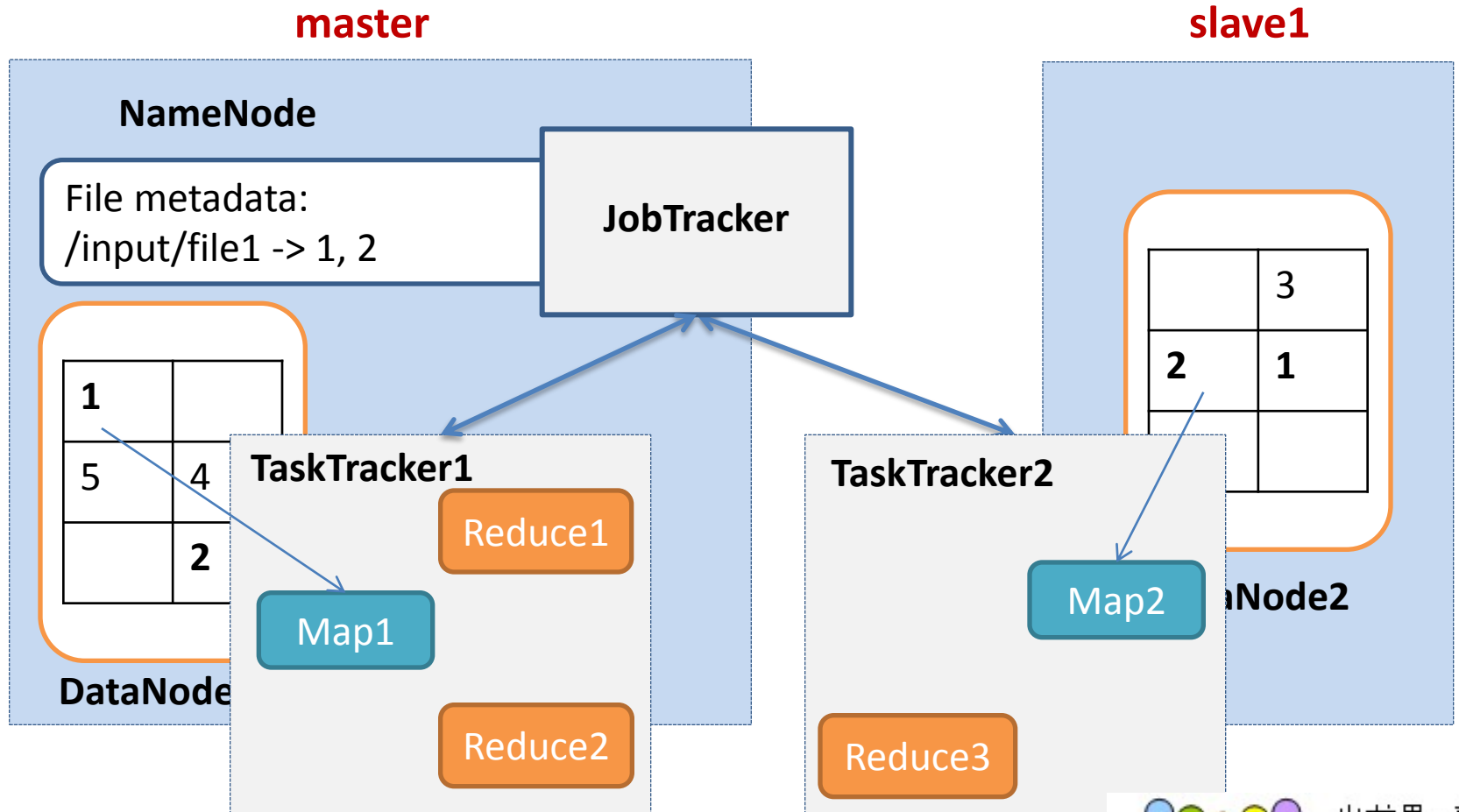


注意: Reduce并没有实现数据本地化!

MapReduce-运行WordCount

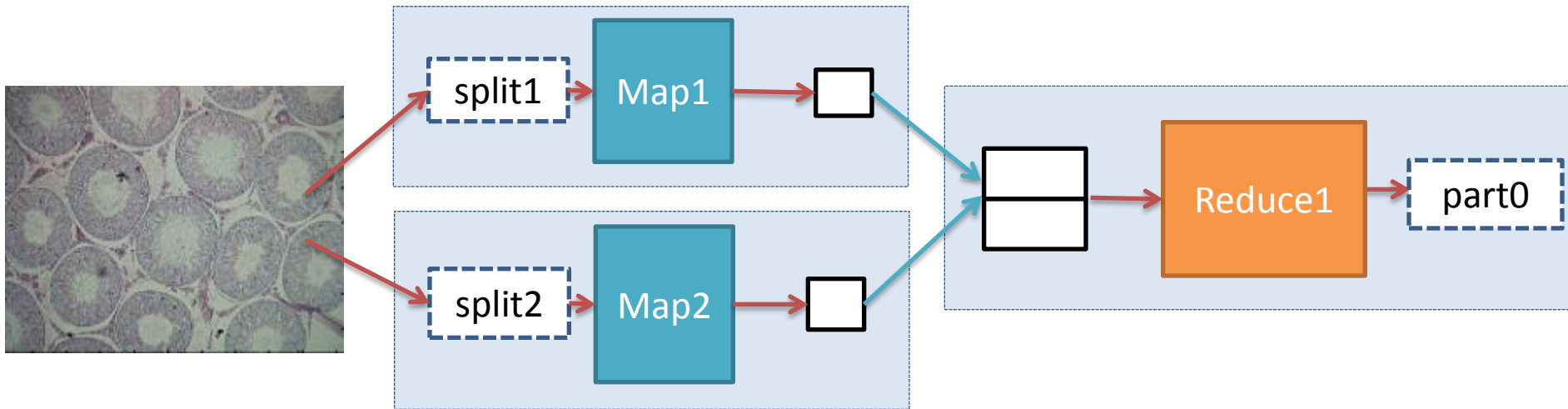
命令:

- `start-mapred.sh` 启动JobTracker, TaskTracker
- `hadoop jar WordCount.jar WordCount /input/file1 /output` 提交作业



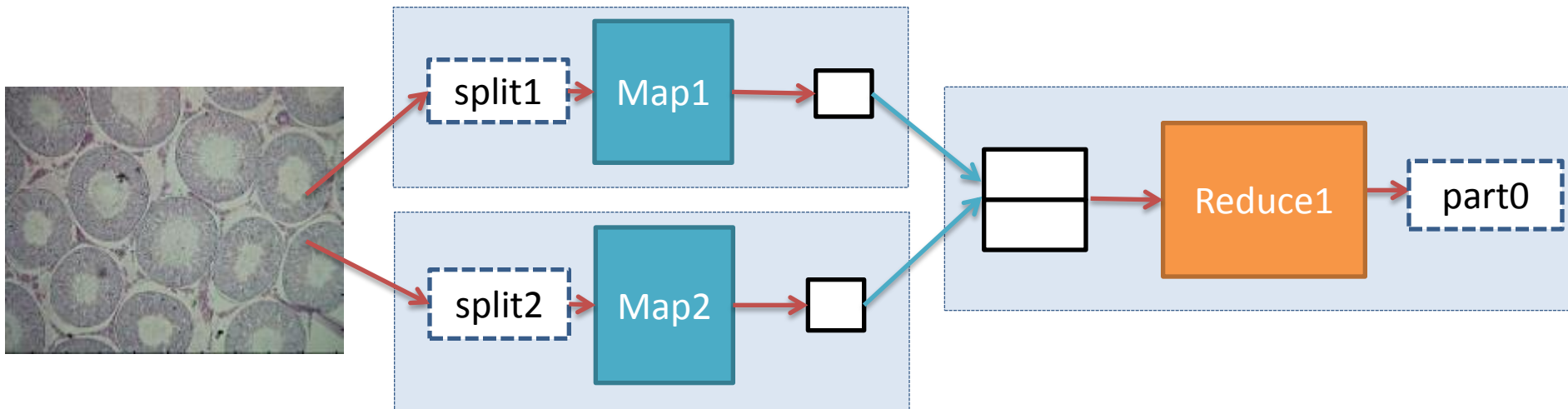
Hadoop适合做细胞分割吗?

简单例子: 统计细胞个数



Hadoop适合做细胞分割吗?

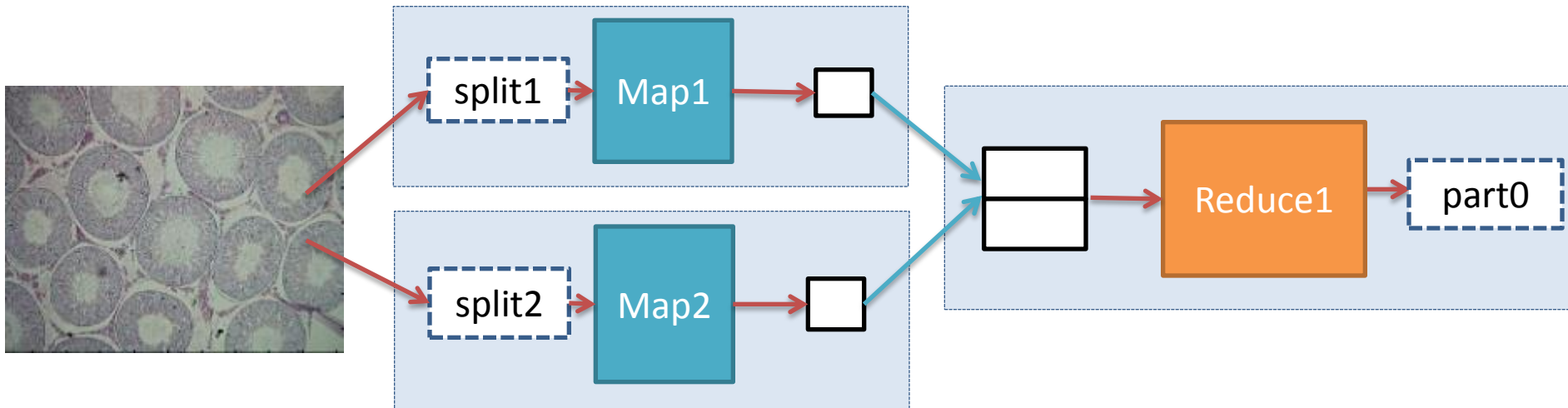
简单例子: 统计细胞个数



要解决的问题: 怎么分片? 一个图片才500KB-1MB. 单纯就文件的存储分片, 前面的200K为一个分片?

Hadoop适合做细胞分割吗？

简单例子: 统计细胞个数



要解决的问题: 怎么分片? 一个图片才500KB-1MB. 单纯就文件的存储分片, 前面的200K为一个分片?

方法: 一个图片对应于一个Map. 但是如果用hadoop来处理一个图片就不太合适了.

Hadoop适合做什么？

如果有大量的图片要进行处理, 分析, Hadoop是适合的. 每个图片一个Map进行分析.

“

一家创业型企业Skybox Imaging, 利用Hadoop来存储和处理高来自卫星捕捉的高分辨率图像, 并尝试将这些信息及图像与地理格局的变化相对应.

”

Hadoop适合做:

一是用HDFS存储, 二是用MapReduce海量级数据(TB级)并行计算

日志分析: 这是当前用的最多的一个用途, Baidu, facebook.

数据分析, 挖掘: (HBase, Hive) 对大数据进行并行计算分析, 如分析用户行为, 偏好(facebook, 淘宝, Ebay).

- MAC用户比Windows用户更大方

Williams acknowledges Hadoop can be a tough technology to learn, **but any blood, sweat and tears are worth** it to ensure his team really understands the data platform that underpins so much of eBay.

谢谢!

欢迎讨论!