

Hadoop 复习要点

选择题 20 分；填空题 10 分；简答题 40 分；编程题 20 分

选择题

机架感知理论
HDFS Federation 的优缺点
NameNode HA 的作用和特点
HDFS 的特点
Reducer 的作用
安全模式的作用特点
YARN 调度器
YARN 的架构和特点
大数据的基本特征
HDFS 的架构：例如 NameNode、DataNode 的职责等
Hadoop 各个服务的端口号
HDFS 回收站的操作和特点

填空题

YARN 的架构组成
MapReduce 的思想，任务组成
大数据的核心问题
HDFS 的 edits 文件和 fsimage 文件的区别
HDFS 的架构组成和作用
Hadoop 生态圈
Hadoop 安装好后的进程

简答题（注意：只有图形作为答案是会扣分的!）

Checkpoint 的工作流程。
HDFS 的体系结构。
理解并掌握整个 MapReduce 工作的流程。
MapReduce 的编程模型，理解各个 k 和 v 的值及其数据类型。
MapReduce 1 与 MapReduce 2 的区别
NameNode 的元信息具体包括哪些内容？
YARN 三种调度器，并简要说明其工作方法

编程题:

理解并默写 WordCount 全部代码。

HDFS Shell 基本操作命令，例如：创建、上传、下载、追加、统计、查看等等