

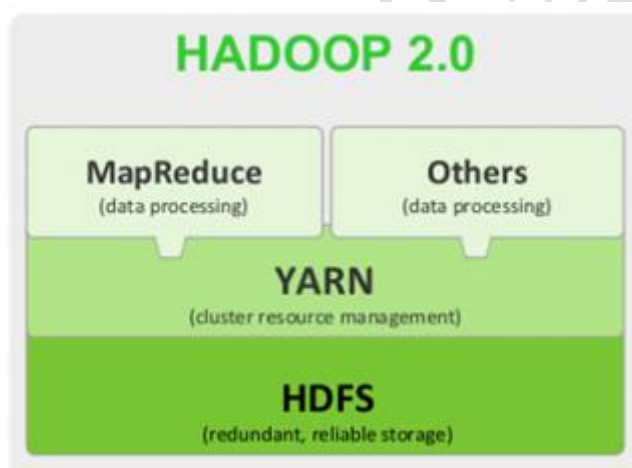
Hadoop 系统架构

侯在钱

目 录

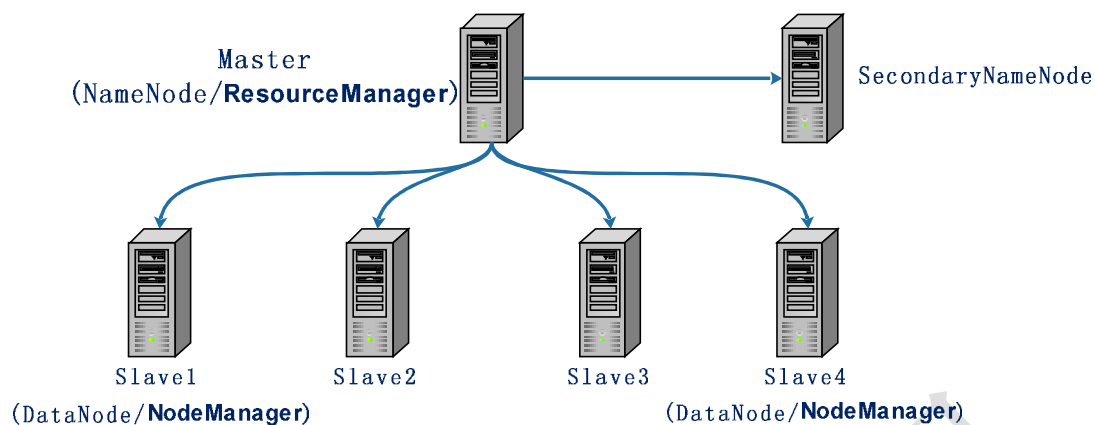
- 1. Hadoop 层次架构 1
- 2. Hadoop 集群架构 1

1. Hadoop 层次架构



2. Hadoop 集群架构

就像一个项目团队必须有一个项目 Leader 一样，Hadoop 系统中也需要一个 Leader 来管理和分配任务，这个 Leader 我们称之为主服务器 Master，其他计算机都为从服务器 Slave。一个 Hadoop 系统是由一个 Master 和多个 Slave 组成的集群系统。Hadoop 系统架构如下图：



Hadoop 有两个功能，一个是 HDFS，另一个是 MapReduce。

HDFS 用于分布式存储文件，系统将把数据分割成多个块存储在多个 Slave 上，而文件的目录和块信息（Metadata）存在 Master 上，在 HDFS 中 Master 叫 NameNode，Slave 叫 DataNode。

YARN 是一个分布式计算的资源管理、任务调度与监控的平台。在 YARN 上可以运行 MapReduce，Storm，Spark 等分布式计算模式。