### hdfs的优缺点

优点：

高容错性，适合批处理、适合大数据处理、支持流式文件访问、可构建在廉价机器上。

缺点：

不适合低延迟数据访问、不适合小文件存取、不适合并发写入、文件随机修改

### hdfs组件

NameNode(Master)

管理维护HDFS、接收客户端的请求、维护了两个重要的文件：edits（记录操作日志）、fsimage文件（HDFS的元信息）

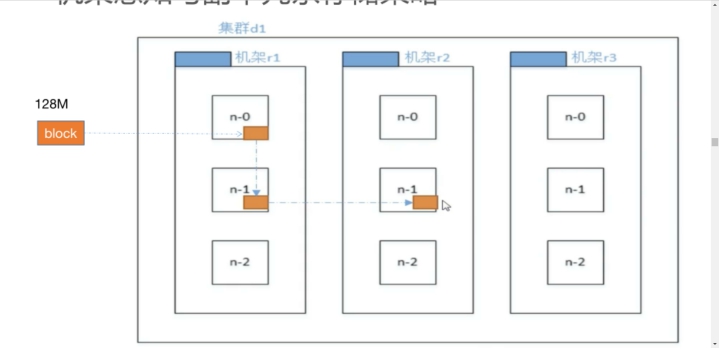
SecondaryNameNode

把edits 中最新的状态信息合并到fsimage文件中。

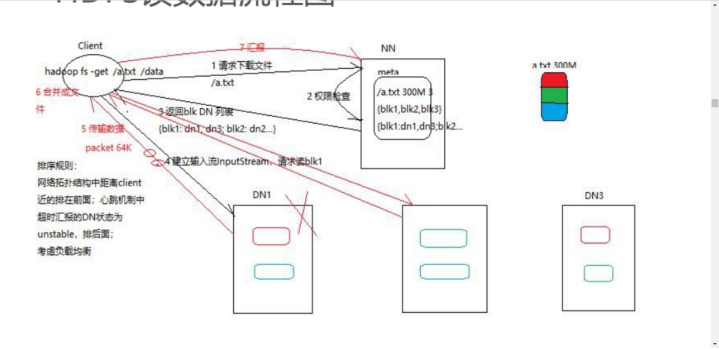
DataNode(Slave)

存储数据块、负责客户端对数据块的IO请求、DataNode定时和NameNode进行心跳通信，接受NameNode的指令。

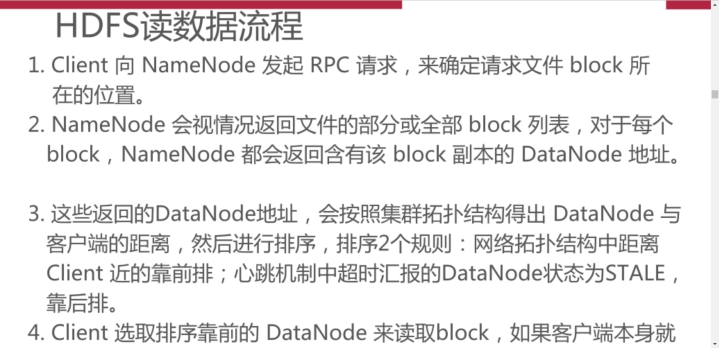
### 机架感知与副本冗余存错策略

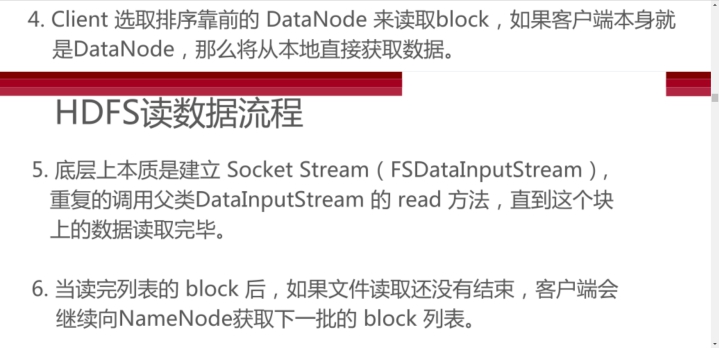


### HDFS文件读取过程



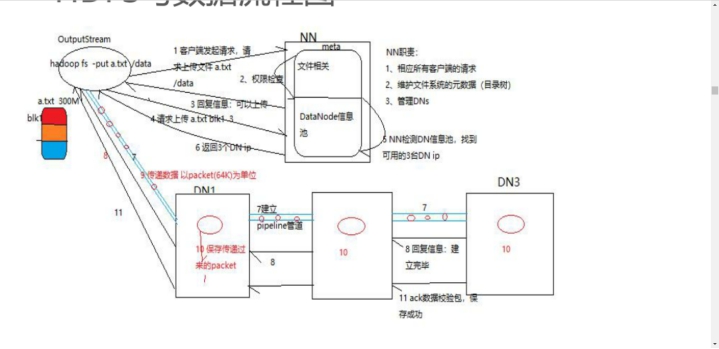
解析：







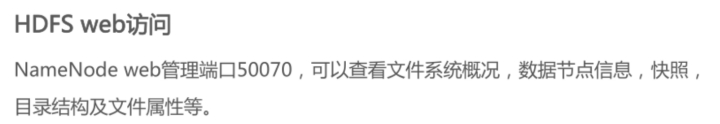
### HDFS文件写入流程

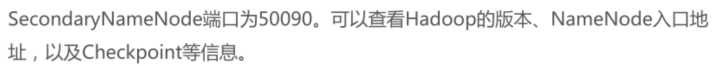


### HDFS三种操作方式

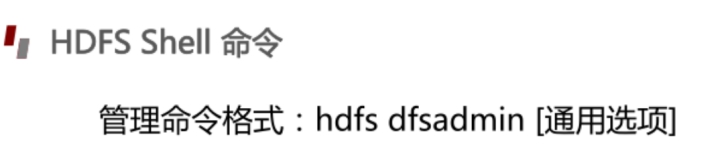


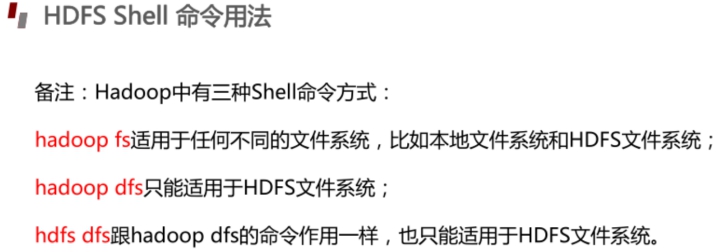
1.Web访问





2.





3.





联邦

