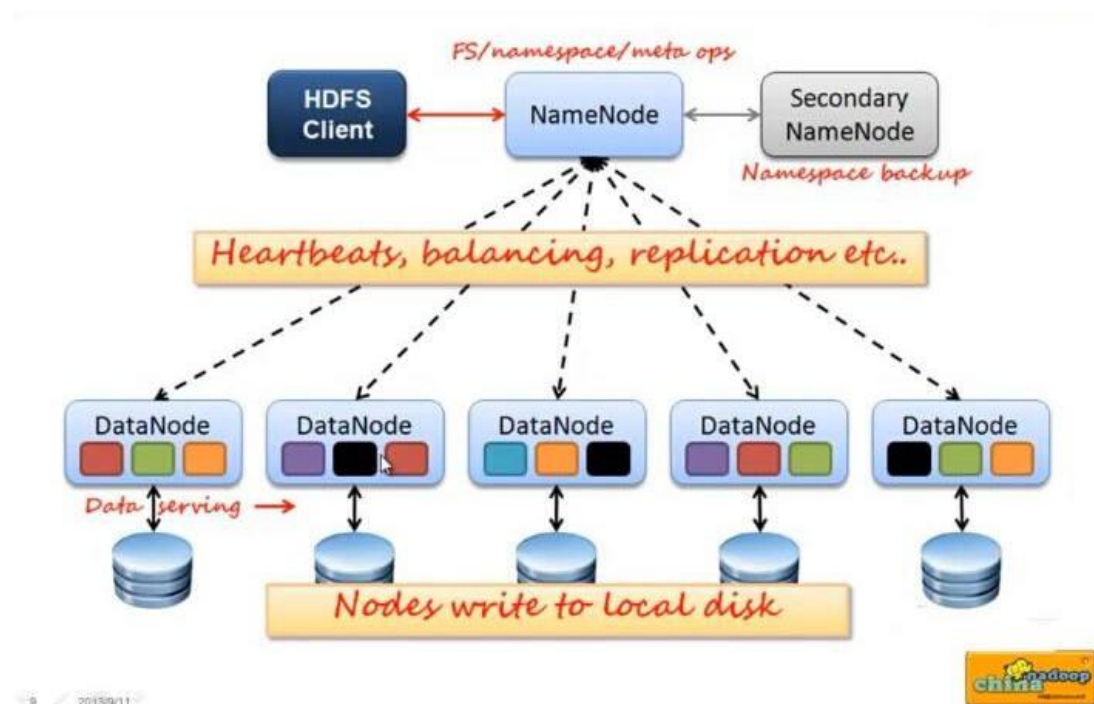
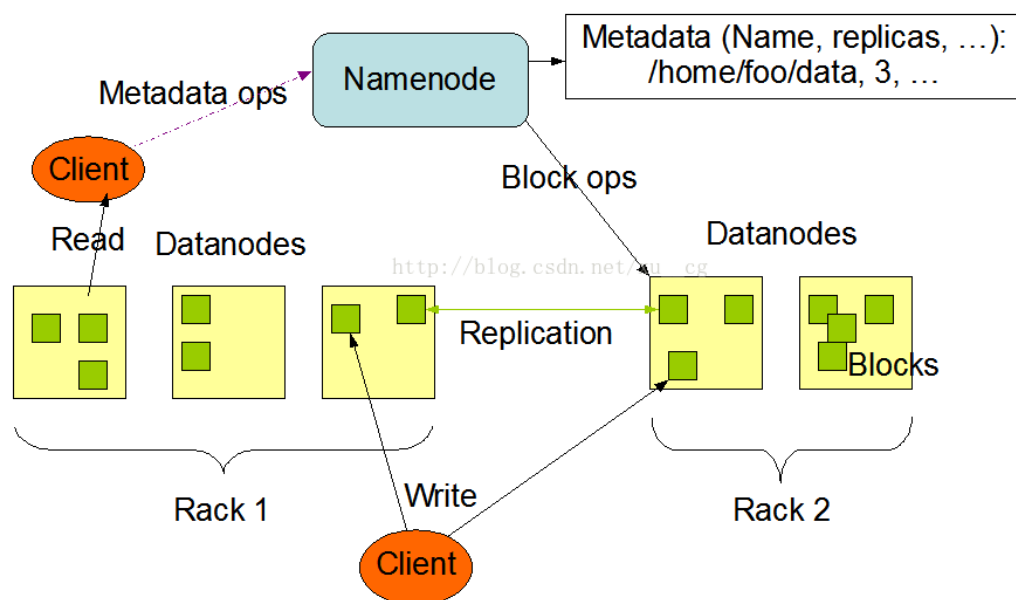


HDFS 的架构图之基础架构



HDFS Architecture



- 1、 NameNode 是一个中心服务器，单一节点（简化系统的设计和实现），负责管理文件系统的名字空间（namespace）以及客户端对文件的访问
- 2、 文件操作，namenode 是负责文件元数据的操作，datanode 负责处理文

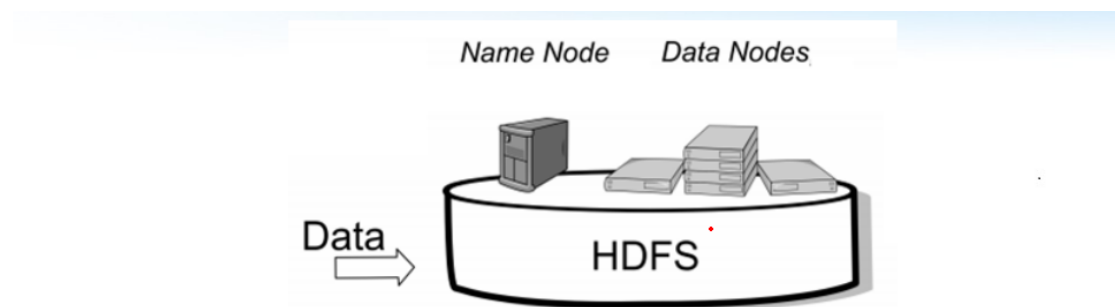
文件内容的读写请求，跟文件内容相关的数据流不经过 Namenode，只询问

它跟哪个 datanode 联系，否则 NameNode 会成为系统的瓶颈

3、副本存放在哪些 Datanode 上由 NameNode 来控制，根据全局情况作出块放置决定，读取文件时 NameNode 尽量让用户先读取最近的副本（网络拓扑图），降低读取网络开销和读取延时

4、NameNode 全权管理数据库的复制（永远保证有三个副本），它周期性的从集群中的每个 DataNode 接收心跳信号和状态报告，接收到心跳信号意味着 DataNode 节点工作正常，块状态报告包含了一个该 DataNode 上所有的数据列表

NameNode 与 Datanode 的总结概述



NameNode	DataNode
存储元数据	存储文件内容
元数据保存在内存中	文件内容保存在磁盘
保存文件、block、DataNode之间的映射关系	维护了block id到DataNode本地文件的映射关系

文件划了几个 block 块，存在那些个 DataNode 上