
HDFS2.0

OutLine

HDFS 2.0

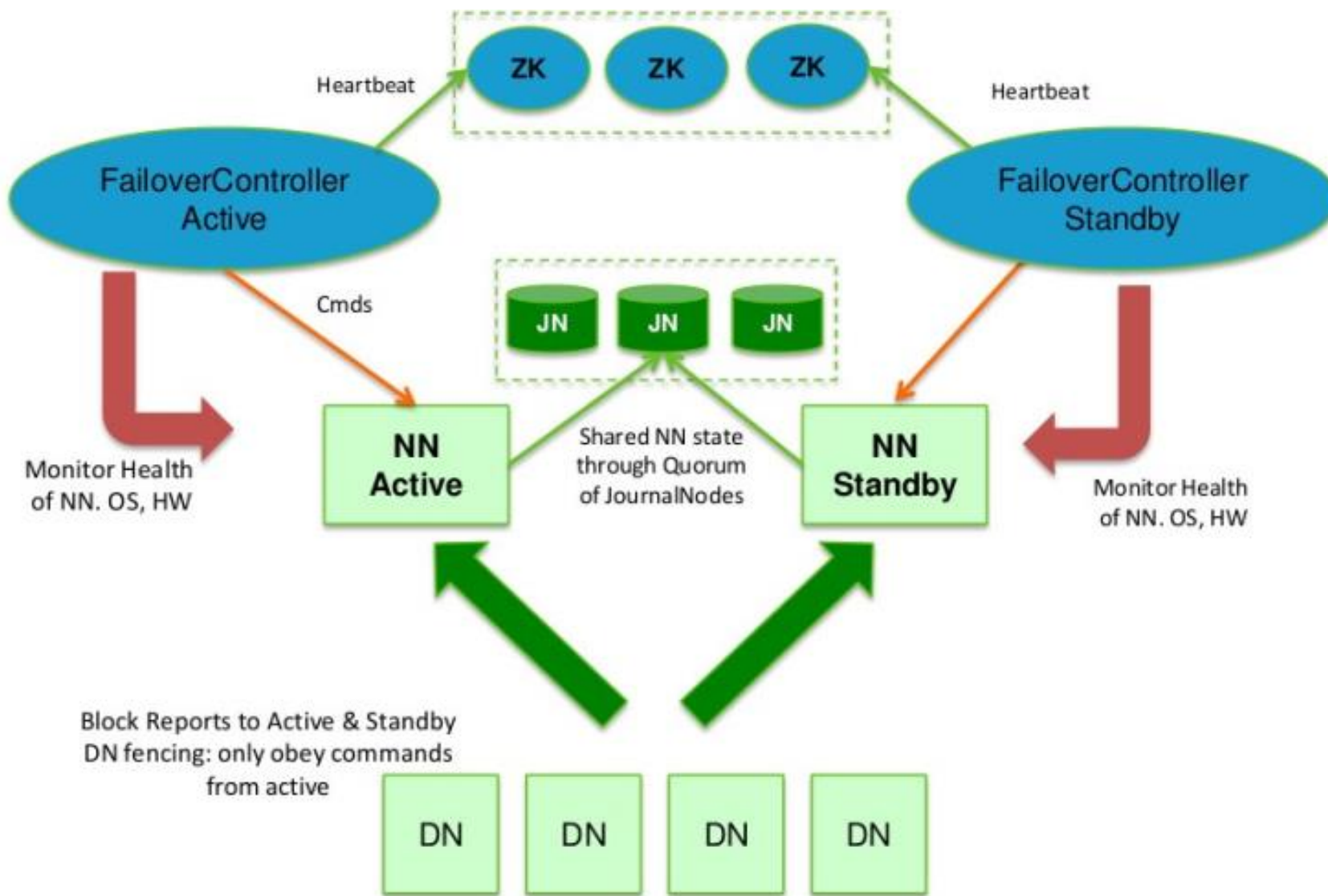
HDFS 2.0 的新特性

- NameNode HA
- NameNode Federation
- HDFS 快照
- HDFS 缓存
- HDFS ACL

NameNode HA

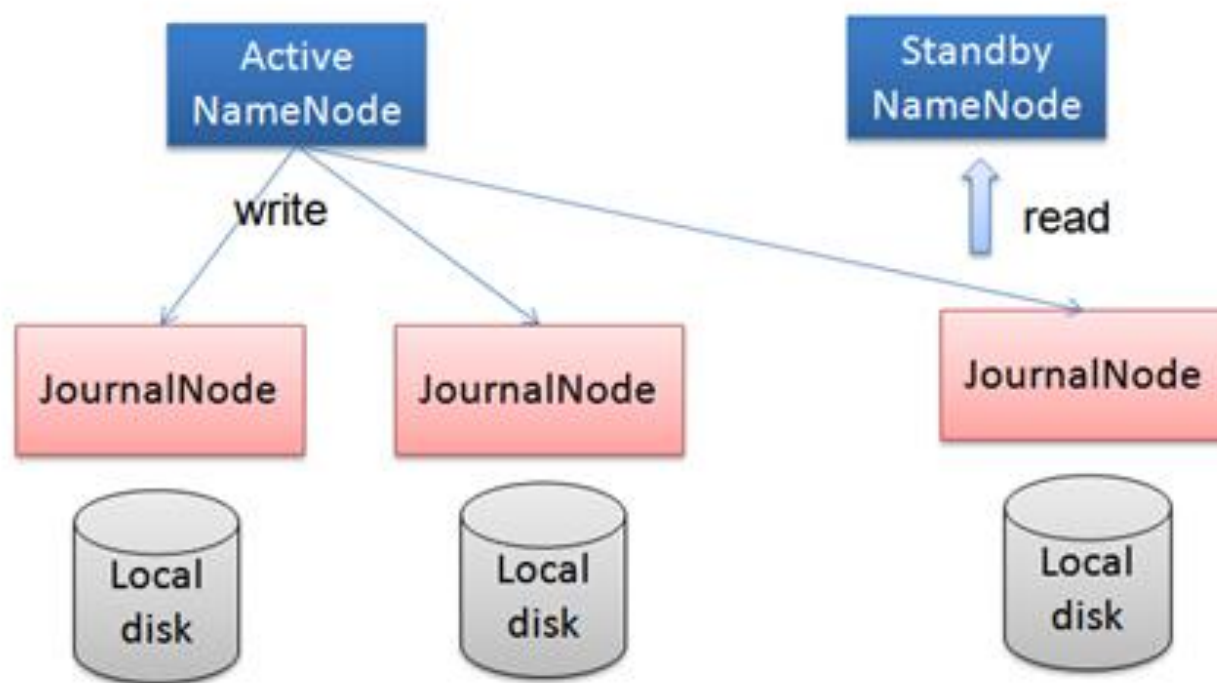
- **什么问题：** Hadoop 1.0中NameNode在整个HDFS中只有一个，存在单点故障风险，一旦NameNode挂掉，整个集群无法使用
- **解决方法：** HDFS的高可用性将通过在同一个集群中运行两个NameNode （ active NameNode & standby NameNode ） 来解决
- 在任何时间，只有一台机器处于Active状态；另一台机器是处于Standby状态
- Active NN负责集群中所有客户端的操作；
- Standby NN主要用于备用，它主要维持足够的状态，如果必要，可以提供快速的故障恢复。

NameNode HA



NameNode HA

- 同步问题：需要依赖JournalNodes守护进程，完成元数据的一致性



NameNode HA

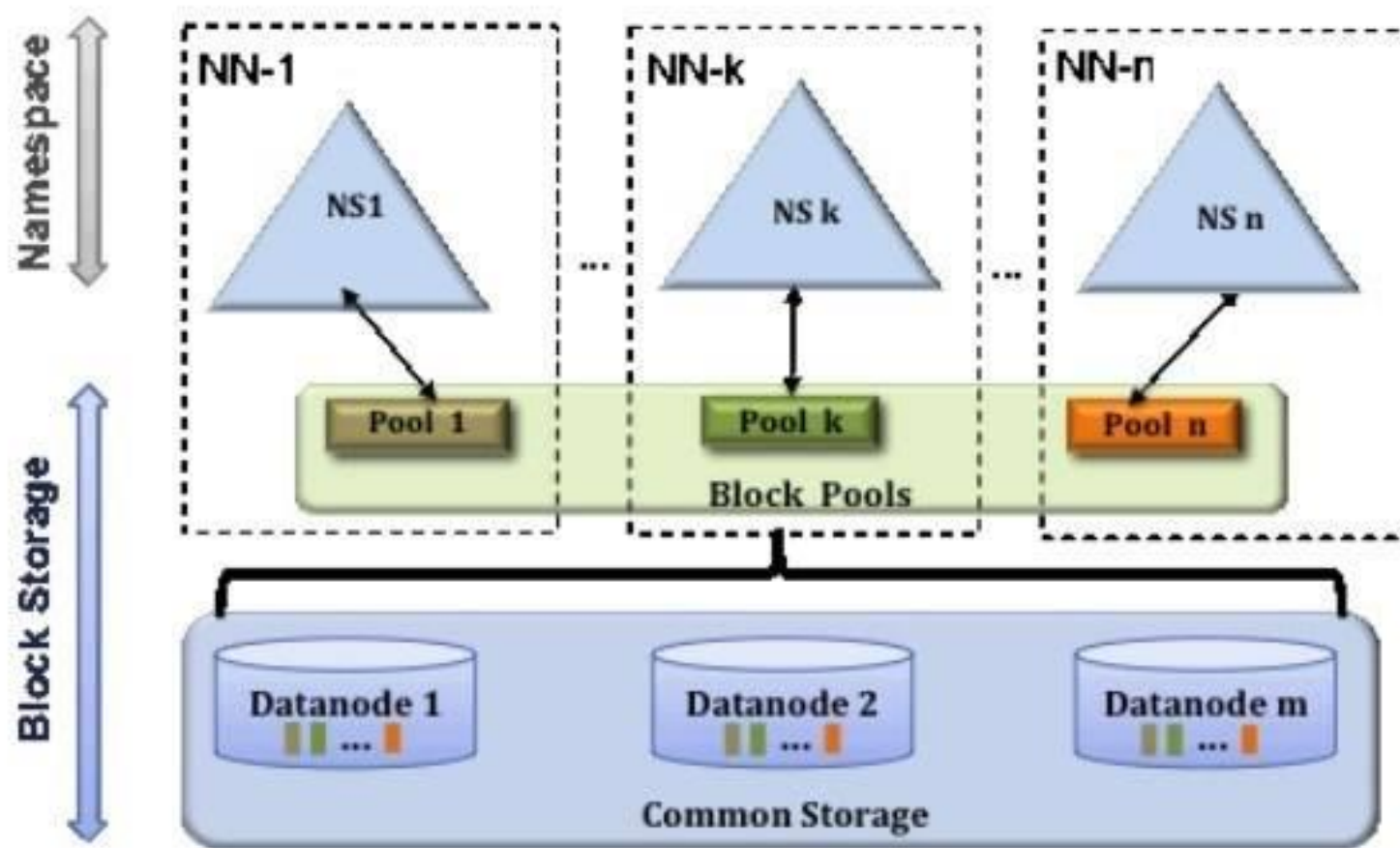
- 快速的故障恢复：心跳保证，Standby NN也需要保存集群中各个文件块的存储位置
- 避免分歧：任何情况下，NameNode只有一个Active状态，否则导致数据的丢失及其它不正确的结果
 - 如何做到？
 - 在任何时间，JNs只允许一个 NN充当writer。在故障恢复期间，将要变成Active 状态的NN将取得writer的角色，并阻止另外一个NN继续处于Active状态

NameNode HA

- 节点分配：
 - NameNode machines: 运行Active NN和Standby NN的机器需要相同的硬件配置
 - JournalNode machines: 也就是运行JN的机器。JN守护进程相对来说比较轻量，所以这些守护进程可以和其他守护线程（比如NN，YARN ResourceManager）运行在同一台机器上
 - 在一个集群中，最少要运行3个JN守护进程，这将使得系统有一定的容错能力。
- 在HA集群中，Standby NN也执行namespace状态的checkpoints，所以不必要运行Secondary NN、CheckpointNode和BackupNode；事实上，运行这些守护进程是错误的。

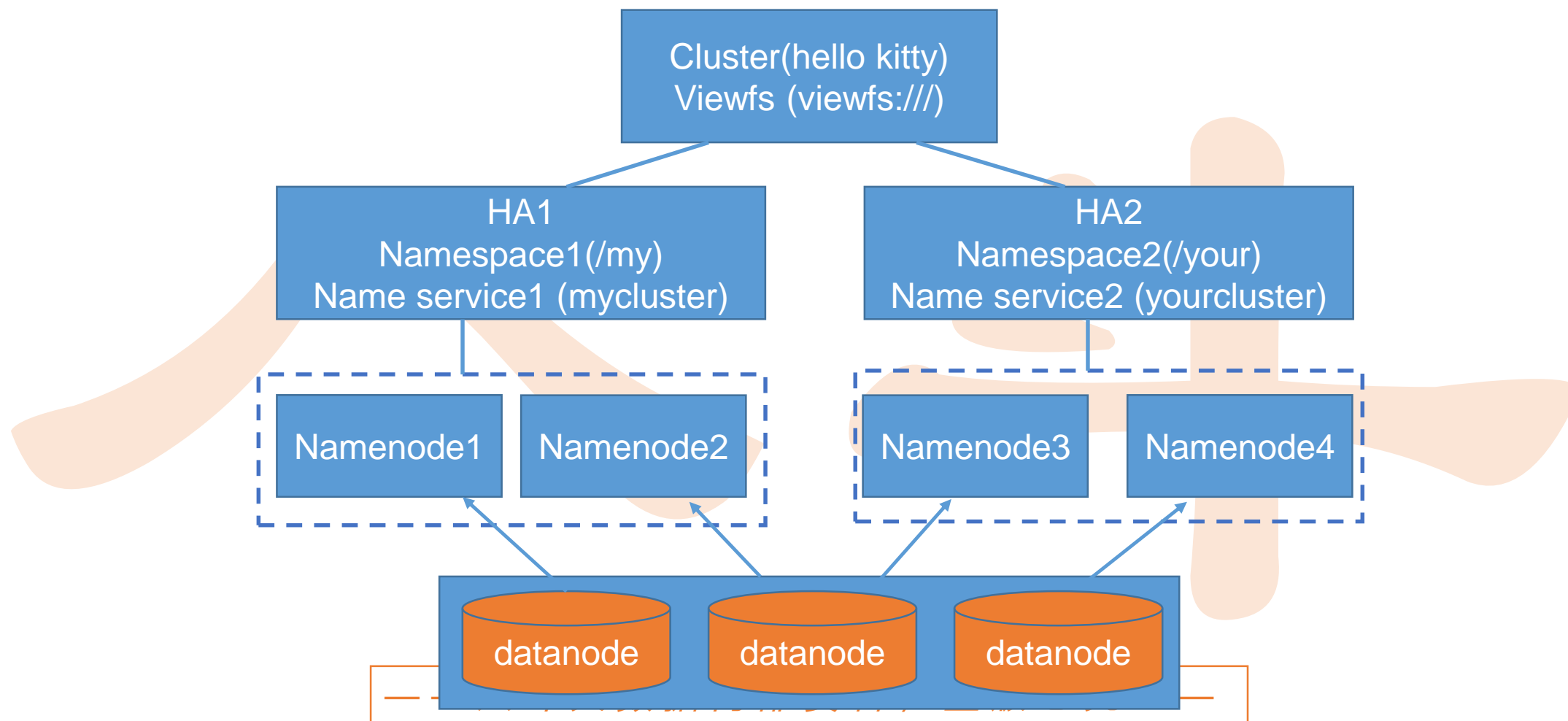
NameNode Federation

- 集群中提供多个NameNode，每个NameNode负责管理一部分DataNode



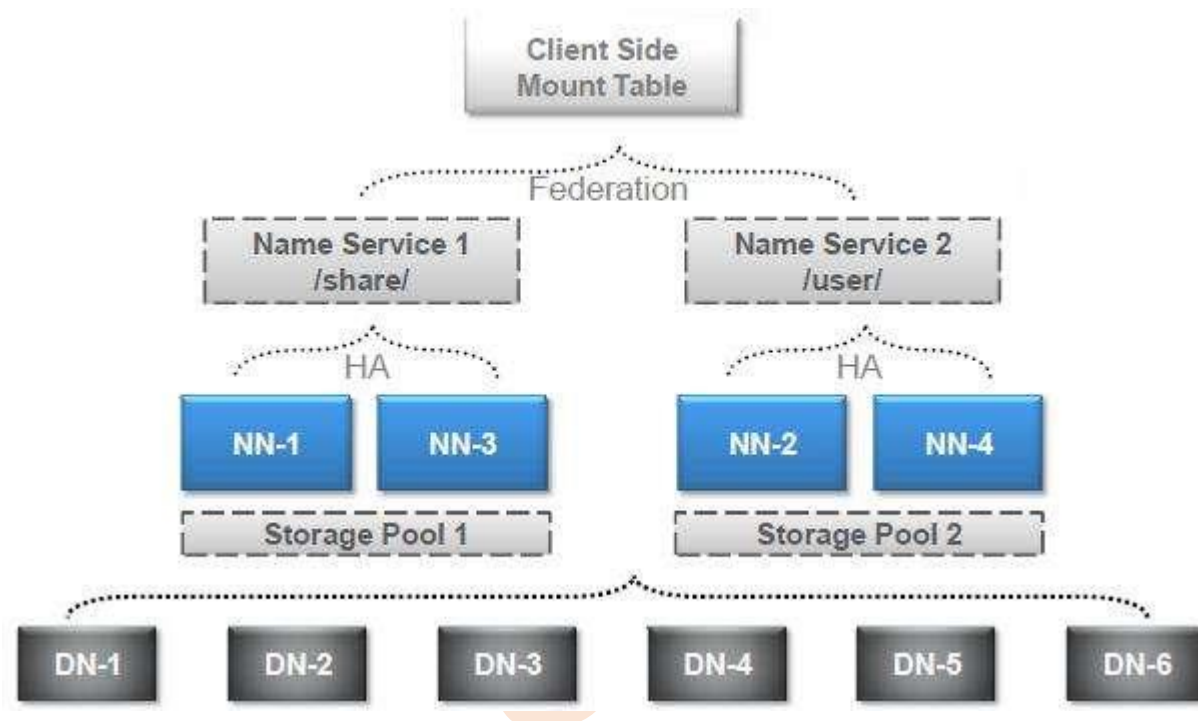
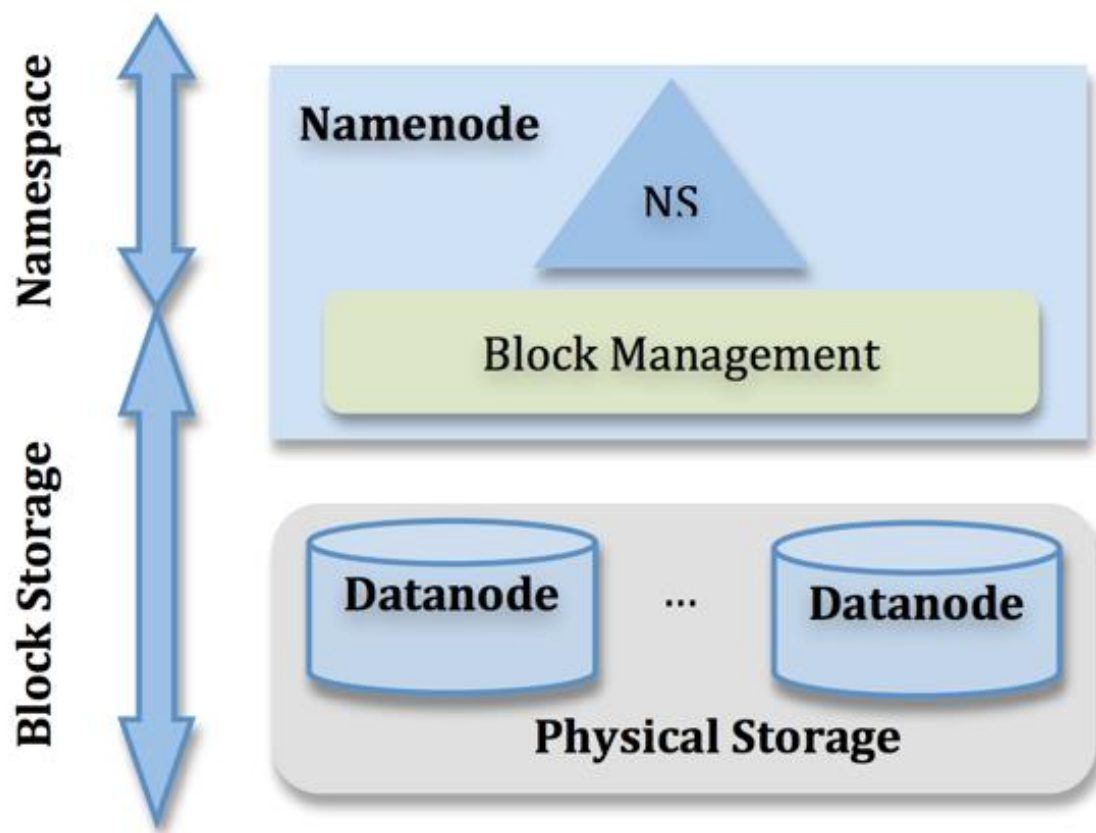
NameNode Federation

- 集群中提供多个NameNode, 每个NameNode负责管理一部分DataNode



NameNode Federation

- **好处：** 实现NameNode的横向扩展,使得Hadoop集群的规模可以达到上万台



HDFS 快照

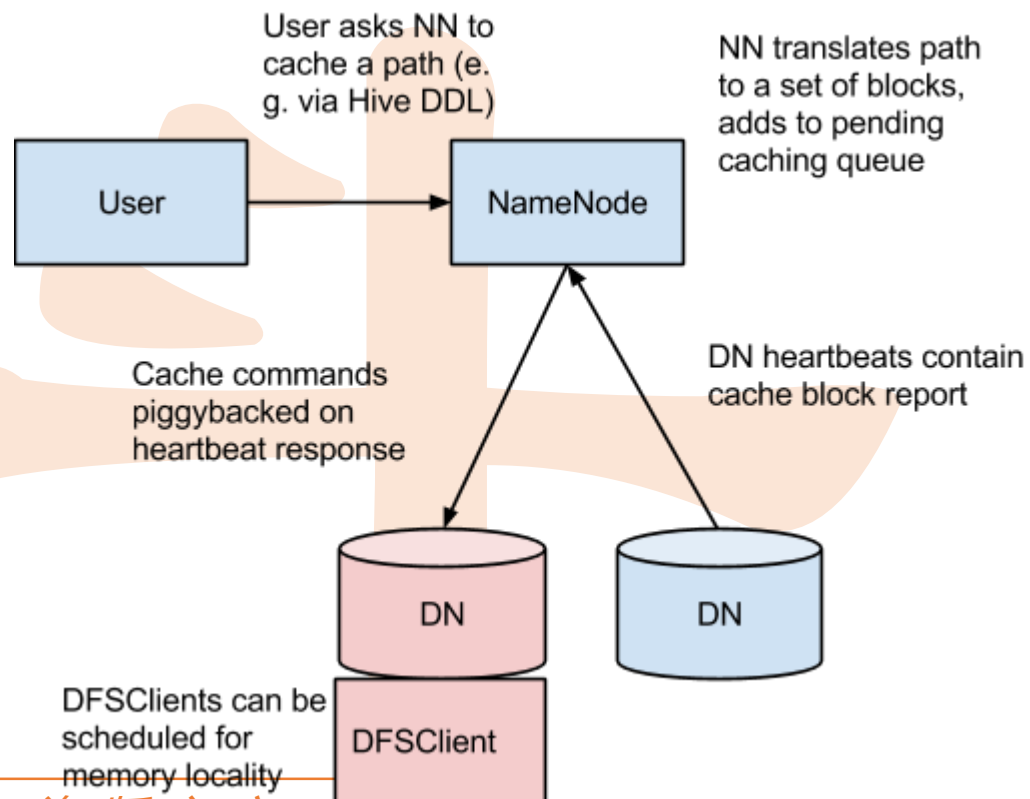
- HDFS快照是一个只读的基于时间点文件系统拷贝
- 快照可以是整个文件系统的也可以是一部分。
- 常用来作为数据备份，防止用户错误操作和容灾恢复。
- Snapshot 并不会影响HDFS 的正常操作：修改会按照时间的反序记录，这样可以直接读取到最新的数据。
- 快照数据是当前数据减去修改的部分计算出来的。
- 快照会存储在snapshottable的目录下。

HDFS 快照

- HDFS快照是对目录进行设定，是某个目录的某一个时刻的镜像
- 对于一个snapshotable文件夹，“.snapshot” 被用于进入他的快照 /foo 是一个 snapshotable目录， /foo/bar是一个/foo下面的文件目录， /foo有一个快照s0，那么路径就是：
： /foo/.snapshot/s0/bar
- `hdfs dfsadmin -allowSnapshot /user/spark`
- `hdfs dfs -createSnapshot /user/spark s0`
- `hdfs dfs -renameSnapshot /user/spark s0 s_init`
- `hdfs dfs -deleteSnapshot /user/spark s_init`
- `hdfs dfsadmin -disallowSnapshot /user/spark`

HDFS 缓存

- 允许用户指定要缓存的HDFS路径
- 明确的锁定可以阻止频繁使用的数据被从内存中清除
- 集中化缓存管理对于重复访问的文件很有用
- 可以换成目录或文件，但目录是非递归的



HDFS ACL

- Hadoop从2.4.0开始支持
- 目前HDFS的权限控制与Linux一致，包括用户、用户组、其他用户组三类权限，这种方式有很大局限性
- 首先参数上要开启基本权限和访问控制列表功能
 - `dfs.permissions.enabled`
 - `dfs.namenode.acls.enabled`
- 常用命令：
 - `hadoop fs -getfacl /input/acl`
 - `hdfs dfs -setfacl -m user:mapred:r-- /input/acl`
 - `hdfs dfs -setfacl -x user:mapred /input/acl`

Q & A

@八斗学院
