

它出现在Hadoop1.x版本中，又称辅助NameNode，在Hadoop2.x以后的版本中此角色消失。如果充当datanode节点的一台机器宕机或者损害，其数据不会丢失，因为备份数据还存在于其他的datanode中。但是，如果充当namenode节点的机器宕机或损害导致文件系统无法使用，那么文件系统上的所有文件将会丢失，因为我们不知道如何根据datanode的块重建文件。因此，对namenode实现容错非常重要。Hadoop提供了两种机制实现高容错性。

第一种机制是备份那些组成文件系统元数据持久状态的文件。Hadoop可以通过配置使namenode在多个文件系统上保存元数据的持久化状态。这些写操作是实时同步的，是原子操作。一般的配置是，将持久状态写入本地磁盘的同时，写入一个远程挂载的网络文件系统(NFS)。

另一种可行的方法是运行一个辅助namenode，但是它不能用作namenode，这个辅助namenode，在hadoop1.x中被称为secondary namenode，在hadoop2.x中，利用高可用（HA）解决单点故障问题。

Secondary namenode，以下简称SN，其重要作用是定期将编辑日志和元数据信息合并，防止编辑日志文件过大，并且能保证其信息与namenode信息保持一致。SN一般在另一台单独的物理计算机上运行，因为它需要占用大量CPU时间来与namenode进行合并操作，一般情况是单独开一个线程来执行操作过程。但是，SN保存的信息永远是滞后于namenode，所以在namenode失效时，难免会丢失部分数据。在这种情况下，一般把存储在NFS上的namenode元数据复制到SN并作为新的namenode。SN不是namenode的备份，可以作为备份。SN主要工作是帮助NN合并edits和fsimage，减少namenode的启动时间。

它不是NameNode的备份，但可以做备份，其主要工作是帮助NameNode合并editslog，减少NameNode的启动时间。SecondaryNameNode执行合并的时机决定于：

- (1) 配置文件设置的时间间隔fs.checkpoint.period，默认为3600秒。
- (2) 配置文件设置edits log大小fs.checkpoint.size,规定edits文件的最大值默认是64MB。

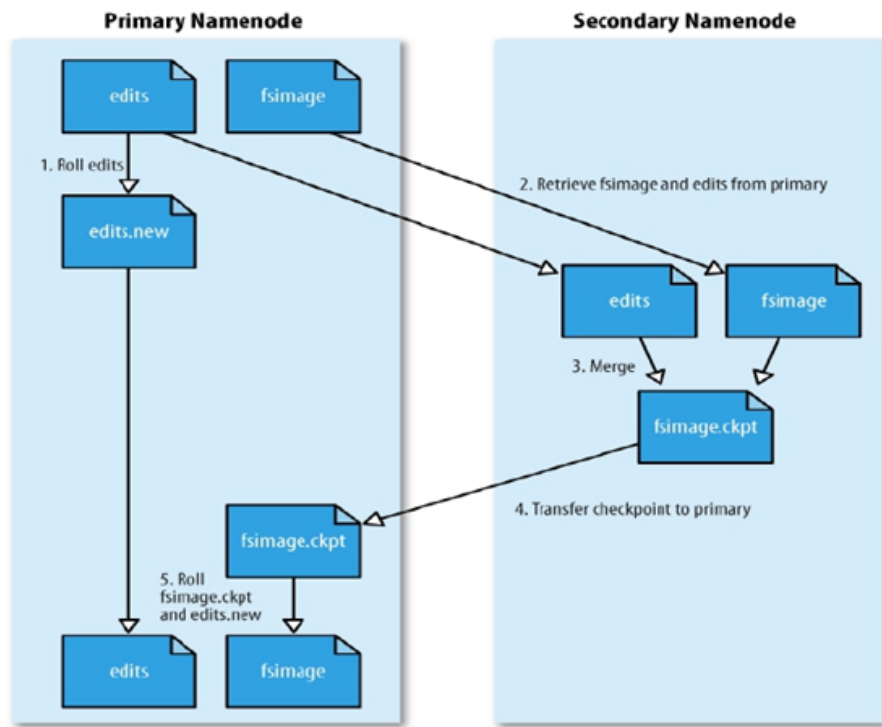


图1.6 SecondaryNameNode合并流程

如上图，当namenode运行了3600s后，SN取出fsimage和edits，合并，更新fsimage，命名为fsimage.ckpt，将fsimage.ckpt文件传入namenode中，合并过程中，客户端会继续上传文件。同时，namenode会创建新的edits.new文件，将合并过程中，产生的日志存入edits.new，namenode将 fsimage.ckpt,更名为fsimage，edits.new更名为edits。

如果在合并过程中，namenode损坏，那么，丢失了在合并过程中产生的edits.new,因此namenode失效时，难免会丢失部分数据。

好文要顶

关注我

收藏该文

高登军

关注 - 1

粉丝 - 0

+ 加关注

0

0

推荐

反对