

HDFS中NameNode和Secondary NameNode工作机制

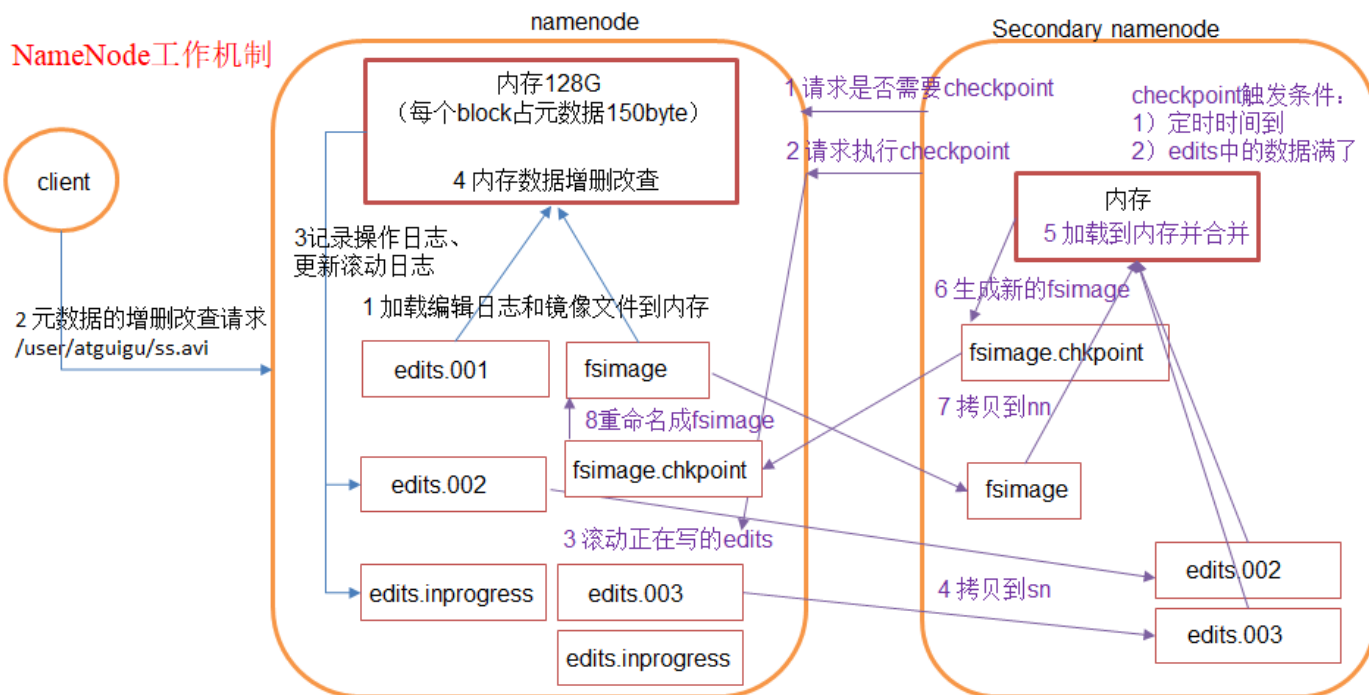
NameNode工作机制

0) 启动概述

namenode启动时，首先将映像文件（fsimage）载入内存，并执行编辑日志（edits）中的各项操作。一旦在内存中成功建立文件系统元数据的映像，则创建一个新的fsimage文件和一个空的编辑日志。此时，namenode开始监听datanode请求。但是此刻，namenode运行在安全模式，即namenode的文件系统对于客户端来说是只读的。

系统中的数据块的位置并不是由namenode维护的，而是以块列表的形式存储在datanode中。在系统的正常操作期间，namenode会在内存中保留所有块位置的映射信息。在安全模式下，各个datanode会向namenode发送最新的块列表信息，namenode了解到足够多的块位置信息之后，即可高效运行文件系统。

如果满足“最小副本条件”，namenode会在30秒钟之后就退出安全模式。所谓的最小副本条件指的是在整个文件系统中99.9%的块满足最小副本级别（默认值：dfs.replication.min=1）。在启动一个刚刚格式化的HDFS集群时，因为系统中还没有任何块，所以namenode不会进入安全模式。



1) 第一阶段：namenode启动（根据客户端的请求记录fsimage和edits，在内存中进行增删改查）

(1) 第一次启动namenode格式化后，创建HDFS镜像文件fsimage和编辑日志文件edits。如果不是第一次启动，直接加载编辑日志和镜像文件到内存。

(HDFS的镜像文件FsImage包含着集群所有文件的元数据信息；编辑日志edits类似“账本”记录数据操作)

(2) 客户端对元数据进行增删改的请求

(3) namenode记录操作日志，更新滚动日志（“记账”）到edits.002、edits.inprogress为接下来用的edits

(4) namenode在内存中对数据进行增删改查

2) 第二阶段: Secondary NameNode工作 (帮助NameNode具体操作edits和fsimage文件, NameNode只是在内存中执行增删改查)

(1) Secondary NameNode询问namenode是否需要checkpoint。直接带回namenode是否检查结果。

(checkpoint判断条件: ① 定时时间到, 默认1小时 ② edits中造作动作次数已满, 默认100万)

(2) Secondary NameNode请求执行checkpoint。

(3) namenode滚动正在写的edits日志 (将目前的edits.inprogress写入edits.003)

(4) 将滚动前的编辑日志 (edits.002、edits.003) 和镜像文件拷贝到Secondary NameNode

(5) Secondary NameNode将编辑日志和镜像文件加载到内存并合并。

(6) 生成新的镜像文件fsimage.chkpoint

(7) 拷贝fsimage.chkpoint到namenode

(8) namenode将fsimage.chkpoint重新命名成fsimage

3) web端访问SecondaryNameNode

(1) 启动集群

(2) 浏览器中输入: <http://hadoop102:50090/status.html>

(3) 查看SecondaryNameNode信息

hadoop102:50090/status.html

Hadoop Overview

Overview

Version	2.7.2
Compiled	2017-05-22T10:49Z by root from Unknown
NameNode Address	hadoop102:8020
Started	2017/7/20 下午5:21:29
Last Checkpoint	Never
Checkpoint Period	120 seconds
Checkpoint Transactions	1000000

Checkpoint Image URI

- file:///opt/module/hadoop-2.7.2/data/tmp/dfs/namesecondary

Checkpoint Editlog URI

- file:///opt/module/hadoop-2.7.2/data/tmp/dfs/namesecondary

4) chkpoint检查时间参数设置

(1) 通常情况下, SecondaryNameNode每隔一小时执行一次。

[hdfs-default.xml]

```
<property>
  <name>dfs.namenode.checkpoint.period</name>
  <value>3600</value>
</property>
```

(2) 一分钟检查一次操作次数, 当操作次数达到1百万时, SecondaryNameNode执行一次。

```
<property>
  <name>dfs.namenode.checkpoint.txns</name>
  <value>1000000</value>
<description>操作动作次数</description>
</property>

<property>
  <name>dfs.namenode.checkpoint.check.period</name>
  <value>60</value>
<description> 1分钟检查一次操作次数</description>
</property>
```

分类: [Hadoop](#)

标签: [Hadoop](#), [HDFS](#)

好文要顶

关注我

收藏该文



猫不夜行
关注 - 2
粉丝 - 16

+加关注

0

推荐

0

反对

« 上一篇: [HDFS数据流——读数据流程](#)

» 下一篇: [HDFS镜像文件fsimage和编辑日志文件edits](#)

posted @ 2019-07-19 17:35 猫不夜行 阅读(1133) 评论(0) 编辑 收藏

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论, 请 [登录](#) 或 [注册](#), [访问](#) 网站首页。