

## Hadoop NameNode元数据相关文件目录解析

下面所有的内容是针对Hadoop 2.x版本进行说明的，Hadoop 1.x和这里有点不一样。

在第一次部署好Hadoop集群的时候，我们需要在NameNode（NN）节点上格式化磁盘：

```
[wyp@wyp hadoop-2.2.0]$ $HADOOP_HOME/bin/hdfs namenode -format
```

格式化完成之后，将会在\$dfs.namenode.name.dir/current目录下如下的文件结构

```
current/  
|-- VERSION  
|-- edits_*  
|-- fsimage_0000000000008547077  
|-- fsimage_0000000000008547077.md5  
`-- seen_txid
```

其中的dfs.namenode.name.dir是在hdfs-site.xml文件中配置的，默认值如下：

```
<property>  
  <name>dfs.namenode.name.dir</name>  
  <value>file://${hadoop.tmp.dir}/dfs/name</value>  
</property>
```

hadoop.tmp.dir是在core-site.xml中配置的，默认值如下

```
<property>  
  <name>hadoop.tmp.dir</name>  
  <value>/tmp/hadoop-${user.name}</value>  
  <description>A base for other temporary directories.</description>  
</property>
```

dfs.namenode.name.dir属性可以配置多个目录，如/data1/dfs/name,/data2/dfs/name,/data3/dfs/name,....。各个目录存储的文件结构和内容都完全一样，相当于备份，这样做的好处是当其中一个目录损坏了，也不会影响到Hadoop的元数据，特别是当其中一个目录是NFS（网络文件

系统Network File System, NFS) 之上, 即使你这台机器损坏了, 元数据也得到保存。

下面对\$dfs.namenode.name.dir/current/目录下的文件进行解释。

1、VERSION文件是Java属性文件, 内容大致如下:

```
#Fri Nov 15 19:47:46 CST 2013
namespaceID=934548976
clusterID=CID-cdff7d73-93cd-4783-9399-0a22e6dce196
cTime=0
storageType=NAME_NODE
blockpoolID=BP-893790215-192.168.24.72-1383809616115
layoutVersion=-47
```

其中

(1)、namespaceID是文件系统的唯一标识符, 在文件系统首次格式化之后生成的;

(2)、storageType说明这个文件存储的是什么进程的数据结构信息(如果是DataNode, storageType=DATA\_NODE);

(3)、cTime表示NameNode存储时间的创建时间, 由于我的NameNode没有更新过, 所以这里的记录值为0, 以后对NameNode升级之后, cTime将会记录更新时间戳;

(4)、layoutVersion表示HDFS永久性数据结构的版本信息, 只要数据结构变更, 版本号也要递减, 此时的HDFS也需要升级, 否则磁盘仍旧是使用旧版本的数据结构, 这会导致新版本的NameNode无法使用;

(5)、clusterID是系统生成或手动指定的集群ID, 在-clusterid选项中可以使用它; 如下说明

a、使用如下命令格式化一个Namenode:

```
$ $HADOOP_HOME/bin/hdfs namenode -format [-clusterId <cluster_id>]
```

选择一个唯一的cluster\_id, 并且这个cluster\_id不能与环境中其他集群有冲突。如果没有提供cluster\_id, 则会自动生成一个唯一的ClusterID。

b、使用如下命令格式化其他Namenode:

```
$ $HADOOP_HOME/bin/hdfs namenode -format -clusterId <cluster_id>
```

c、升级集群至最新版本。在升级过程中需要提供一个ClusterID, 例如:

```
$ $HADOOP_PREFIX_HOME/bin/hdfs start namenode    W  
--config $HADOOP_CONF_DIR -upgrade -clusterId <cluster_ID>
```

如果没有提供ClusterID，则会自动生成一个ClusterID。

(6)、blockpoolID：是针对每一个Namespace所对应的blockpool的ID，上面的这个BP-893790215-192.168.24.72-1383809616115就是在我的ns1的namespace下的存储块池的ID，这个ID包括了其对应的NameNode节点的ip地址。

2、\$dfs.namenode.name.dir/current/seen\_txid非常重要，是存放transactionId的文件，format之后是0，它代表的是namenode里面的edits\_\*文件的尾数，namenode重启的时候，会按照seen\_txid的数字，循序从头跑edits\_0000001~到seen\_txid的数字。所以当你的hdfs发生异常重启的时候，一定要比对seen\_txid内的数字是不是你edits最后的尾数，不然会发生建置namenode时metaData的资料有缺少，导致误删Datanode上多余Block的资讯。

3、\$dfs.namenode.name.dir/current目录下在format的同时也会生成fsimage和edits文件，及其对应的md5校验文件。fsimage和edits是Hadoop元数据相关的重要文件，我打算在下一篇文章中fsimage和edits存储的内容和他们之间的关系进行说明，请关注本博客。

**本博客文章除特别声明，全部都是原创！**  
**原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。**  
**本文链接: [【】](#)（）**