DataNode decommission分析

DataNode实际上是存储数据，在Decommission Datanode的时候需要把Datanode上的数据迁移到别的机器上。Decommission Data的步骤如下：

1）在NameNode上，需要把Decommision的datanode上的机器名加入到dfs.hosts.exclude(该配置项在hdfs-site.xml)所指定的文件中，告诉NameNode哪些DataNode需要被Decommission

需要把Decommissio的机器写到/etc/hadoop/conf/dfs.exclude中去

<property>

<name>dfs.hosts.exclude</name>

<value>/etc/hadoop/conf/dfs.exclude</value>

</property>

2）用如下的命令，启动Decommission

$hdfs dfsadmin -refreshNodes

接下来就可以在NameNode的UI上看到对应的Datanode会处在Decommissioning状态。完成后，对应的Datanode会处在Decommissioned状态

参考文献：

http://blog.csdn.net/jewes/article/details/40090779