时间：20160412,

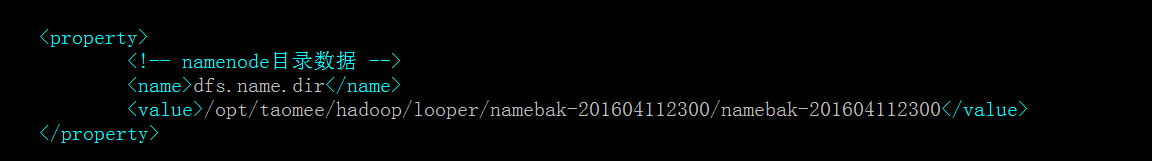
集群遇到的问题：某些数据被误删了。

处理过程:当时把集群给关闭了，因为199机器上面有hadoop元数据的镜像数据的备份，所以当时把主节点的镜像目录切换到相应的镜像备份目录下面(前一天的元数据目录)，此时集群恢复到这个元数据时刻的集群状态，也就是我们之前的误删数据操作暂时停止了。这样做的目的是可以恢复部分删除的数据，使得我们可以根据这种数据格式重写写MR，对之前的数据进行恢复。缺点是：导致这个时刻到误删数据操作的这段时间的数据都被删了，所以我们需要做的是将前一天的stat-server的basic数据给重新同步上来，广告的数据也需要找运营开发部的相关同事给其重新上传，同时老统计的那边的数据也需要重新上传上来。当天的数据都需要重新的计算。

具体过程:

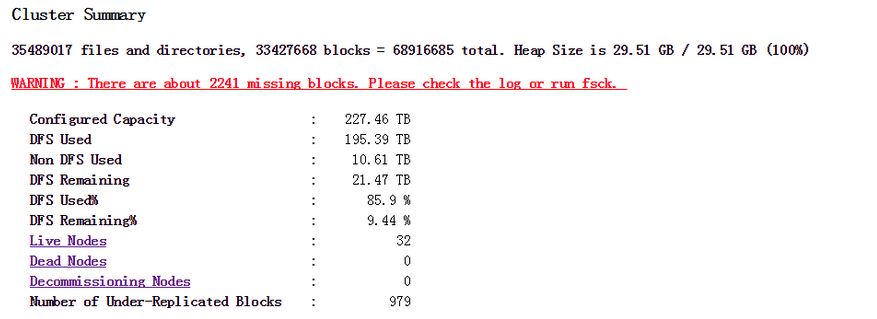
1. 切换主节点元数据的镜像

192.168.11.199的元数据路径切换成备份数据的路径(只需要改199机器的相关配置),将hdfs-site.xml里面的dfs.name.dir改下路径。



1. 将相关丢失的数据重新同步上来，新统计的basic数据改下stat-server的同步时间就可以了。广告数据/ads目录下的数据需要找相关运营开发部的同事，让他们重新上传即可。
2. 重新计算相关数据。
3. 处理掉误删的不能恢复的元数据。

因为有的数据已经被删除了，此时集群中的所有节点已经没有改数据了，此时查看页面如下所示：



这个时候执行fsck检查：

hadoop fsck -files  -blocks -locations -racks path

其中path只要填写你误删目录即可，不需要从根目录去检查，否则时间回很久。

然后针对上面的检查报告，执行：

hadoop fsck –delete path

5.根据目前被删目录还没被删掉的数据，按照相应格式与逻辑重新写MR，恢复相关被删掉的数据。