**集中化经分小二期集群使用规范**

1、HDFS适用于大文件存储，小文件数量不宜过多，小文件太多会加重namenode负担。日志集群BlockSize目前配置为512M，整体上要求平均文件大小不小于128M。建议应用侧定期清理历史业务数据，对业务数据进行合并，以减少小文件数量，减轻namenode压力。

2、日志集群目前namenode性能压力较大，建议定期关注namenode连接数，关注方法：

<http://active_namenode_ip:50070/jmx?qry=Hadoop:service=NameNode,name=RpcActivityForPort8020>

重点关注NumOpenConnections请求连接数、CallQueueLength callQueue队列长度两个指标，NumOpenConnections请求连接数建议维持在20000以下，CallQueueLength大部分时间应该为0。

namenode中rpc请求类型包括：a，客户端请求，主要包括读写hdfs数据，操作元数据；b，datanode请求，主要包括heartbeat、blockreport；c，管理员请求，dfsadmin各种操作，例如获取dn信息。对namenode产生压力的主要是a和b。建议应用侧进行HDFS上超大数据量操作之前，提前与平台侧沟通；避免劣质MR job（如10000+map/reduce、mr作业产生过大临时数据）等。

3、避免一次性删除超大规模HDFS数据，建议定期适量删除数据。如需大规模清理HDFS数据，建议与平台侧沟通。

4、对于新上业务的mr job，运行之前建议先预估一下数据量，若数据量较大，建议分拆多个mr job运行或者启用压缩。

5、关于集群客户端参数，建议业务侧根据自己的作业实际情况，做出适当调整，不能仅仅依赖于服务端的参数设置。如mapreduce.job.reduce.slowstart.completedmaps对于集群资源不足而map数量又较多的作业，建议调整为1；mapreduce.fileoutputcommitter.algorithm.version建议调整为2，yarn.scheduler.minimum-allocation-mb、mapreduce.map.memory.mb、mapreduce.reduce.memory.mb、mapred.min.split.size、mapred.map.tasks等等，建议根据作业实际数据量做出适当的调整。

6、合理设计HBase表rowkey，避免数据访问热点问题；

7、HBase建表时合理进行预分区，防止大数据量插入的热点问题，提高数据插入的效率。

8、合理控制分区数量，建议每个regionserver负载100-200个region。

9、由于清单集群HBase对查询的时间要求较高，如果启动大量的hadoop MR任务，会消耗大量集群资源，影响HBase查询性能，建议尽量不要在性能要求较高的在线HBase集群做类似操作。

10：建议业务高峰期段关闭HBase major compaction，避免影响集群性能。