

hdfs配置文件详解

- dfs.name.dir
 - NameNode 元数据存放位置
 - 默认值：使用core-site.xml中的hadoop.tmp.dir/dfs/name
- dfs.block.size
 - 对于新文件切分的大小，单位byte。默认是64M,建议是128M。每一个节点都要指定，包括客户端。
 - 默认值： 67108864
- dfs.data.dir
 - DataNode在本地磁盘存放block的位置，可以是以逗号分隔的目录列表，DataNode循环向磁盘中写入数据，每个DataNode可单独指定与其它DataNode不一样
 - 默认值： \${hadoop.tmp.dir}/dfs/data
- dfs.namenode.handler.count
 - NameNode用来处理来自DataNode的RPC请求的线程数量
 - 建议设置为DataNode数量的10%，一般在10~200个之间
 - 如设置太小，DataNode在传输数据的时候日志中会报告“connecton refused”信息
 - 在NameNode上设定
 - 默认值： 10
- dfs.datanode.handler.count
 - DataNode用来连接NameNode的RPC请求的线程数量
 - 取决于系统的繁忙程度
 - 设置太小会导致性能下降甚至报错
 - 在DataNode上设定
 - 默认值： 3
- dfs.datanode.max.xcievers
 - DataNode可以同时处理的数据传输连接数
 - 默认值： 256
 - 建议值： 4096
- dfs.permissions
 - 如果是true则检查权限，否则不检查（每一个人都可以存取文件）
 - 于NameNode上设定
 - 默认值： true
- dfs.datanode.du.reserved
 - 在每个卷上面HDFS不能使用的空间大小
 - 在每个DataNode上面设定
 - 默认值： 0
 - 建议为10737418240，即10G。需要结合MapReduce场景设置。
- dfs.datanode.failed.volumes.tolerated
 - DataNode可以容忍损坏的磁盘数量，超过这个数量DataNode将会离线，所有在这个节点上面的block将会被重新复制
 - 默认是0，但是在有多块磁盘的时候一般会增大这个值

- dfs.replication

- 在文件被写入的时候，每一块将要被复制多少份

- 默认是3份。建议3份

- 在客户端上设定

通常也需要在DataNode上设定

2、HDFS core-site.xml 参数配置

- fs.default.name

- 文件系统的名字。通常是NameNode的hostname与port

- 需要在每一个需要访问集群的机器上指定，包括集群中的节点

- 例如：hdfs://<your_namenode>:9000/

- fs.checkpoint.dir

- 以逗号分隔的文件夹列表，SecondNameNode用来存储checkpoint image文件

- 如果多于一个文件夹，那么都会被写入数据

- 需要在SecondNameNode上设定

- 默认值：\${hadoop.tmp.dir}/dfs/secondary

- hadoop.tmp.dir

- HDFS与本地磁盘的临时文件

默认是/tmp/hadoop-\${user.name}.需要在所有的节点中设定

- fs.trash.interval

- 当一个文件被删掉后，它会被放到用户目录的.Trash目录下，而不是立即删掉

- 经过此参数设置的分钟数之后，再删掉数据

- 默认是0，禁用此功能，建议1440（一天）

- io.file.buffer.size

- 设定在读写数据时的缓存大小，应该为硬件分页大小的2倍

- 默认是4096，建议为65536（64K）

3、设置log文件的大小和数量

- 修改core-site.xml中的参数

- hadoop.logfile.size

- hadoop.logfile.count

4、设置组件的日志级别

- 查看不同组件的日志级别

- hadoop daemonlog -getlevel host:port packageName

- 设置组件的日志级别

- hadoop daemonlog -setlevel host:port packageName level

- hadoop daemonlog -setlevel db74:50070 org.apache.hadoop ERROR

- DEBUG, INFO, ERROR, FATAL

- 端口为前台页面的端口，缺省为50070
- 组件名称(packageName)
- org.apache.hadoop.hdfs.server.namenode.NameNode
- org.apache.hadoop.hdfs.server.datanode.DataNode
- org.apache.hadoop.hdfs
- org.apache.hadoop
- org.apache.hadoop.mapred.JobTracker

重要程度由低到高依次为DEBUG < INFO < WARN < ERROR < FATAL。日志输出规则为：只输出级别不低于设定级别的日志信息。比如，级别设定为INFO，则INFO、WARN、ERROR和FATAL 级别的日志信息都会被输出，但级别比INFO 低的DEBUG 则不会被输出。DEBUG 为测试，INFO为默认，一般生产用，ERROR错误，

5、hdfs的进程节点

1.namenode

记录源数据的命名空间

数据分配到那些datanode保存

协调客户端对文件访问

2.datanode

负责所在物理节点的储存管理

一次写入,多次读取(不能修改)

文件由数据块组成,典型的块大小是64M

数据块尽量散步到各个节点

3.secondarynamenode (辅助)

当NameNode重启的时候，会合并硬盘上的fsimage文件和edits文件，得到完整的Metadata信息。这个fsimage文件可以看做是一个过时的Metadata信息文件（最新的Metadata修改信息在edits文件中）。

如果edits文件非常大，那么这个合并过程就非常慢，导致HDFS长时间无法启动，如果定时将edits文件合并到fsimage，那么重启NameNode就可以非常快。

SecondaryNameNode就做这个合并的工作。

6、hdfs的回收站功能

删除文件时，其实是放入回收站/trash，回收站里的文件可以快速恢复

可以设置一个时间阈值，当回收站里文件的存放时间超过这个阈值，就被彻底删除，并且释放占用的数据块

(开启回收站功能)

```
[hadoop@h1 ~]$ cd /usr/local/hadoop-1.2.1/conf
[hadoop@h1 conf]$ vi core-site.xml (添加下面一段,10080为保留时间,单位分钟)
```

```
<property>
<name>fs.trash.interval</name>
<value>10080</value>
<description>
Number of minutes between trashcheckpoints.
If zero, the trash feature is disabled
</description>
</property>
```

```
[hadoop@h1 hadoop-1.2.1]$ bin/start-all.sh (重启 回收站功能生效)
```

hdfs 常用端口和 hdfs-default配置文件参数的意义

刚配置完1.0.4发现对配置文件还是不太了解，就baidu了下，跟大家分享下。常用的端口配置1 HDFS端口

参数	描述	默认	配置文件	例子值
fs.default.name	namenode namenode RPC交互端口	8020	core-site.xml	hdfs://master:8020/
dfs.http.address	NameNode web管理端口	50070	hdfs-site.xml	0.0.0.0:50070
dfs.datanode.address	datanode 控制端口	50010	hdfs-site.xml	0.0.0.0:50010
dfs.datanode.ipc.address	datanode的RPC服务器地址和端口	50020	hdfs-site.xml	0.0.0.0:50020
dfs.datanode.http.address	datanode的HTTP服务器和端口	50075	hdfs-site.xml	0.0.0.0:50075

2 MR端口

参数	描述	默认	配置文件	例子值
mapred.job.tracker	job-tracker交互端口	8021	mapred-site.xml	hdfs://master:8021/
job	tracker的web管理端口	50030	mapred-site.xml	0.0.0.0:50030
mapred.task.tracker.http.address	task-tracker的HTTP端口	50060	mapred-site.xml	0.0.0.0:50060

3 其它端口

参数	描述	默认	配置文件	例子值
dfs.secondary.http.address	secondary NameNode web管理端口	50090	hdfs-site.xml	0.0.0.0:50090

4 hdfs-default.html

序号	参数名	参数值	参数说明
1	dfs.namenode.logging.level	info	输出日志类型
2	dfs.secondary.http.address	0.0.0.0:50090	备份名称节点的http协议访问地址与端口
3	dfs.datanode.address	0.0.0.0:50010	数据节点的TCP管理服务地址和端口
4	dfs.datanode.http.address	0.0.0.0:50075	数据节点的HTTP协议访问地址和端口
5	dfs.datanode.ipc.address	0.0.0.0:50020	数据节点的IPC服务访问地址和端口
6	dfs.datanode.handler.count	3	数据节点的服务连接处理线程数
7	dfs.http.address	0.0.0.0:50070	名称节点的http协议访问地址与端口
8	dfs.https.enable	false	支持https访问方式标识
9	dfs.https.need.client.auth	false	客户端指定https访问标识
10	dfs.https.server.keystore.resource	ssl-server.xml	Ssl密钥服务端的配置文件
11	dfs.https.client.keystore.resource	ssl-client.xml	Ssl密钥客户端的配置文件
12	dfs.datanode.https.address	0.0.0.0:50475	数据节点的HTTPS协议访问地址和端口
13	dfs.https.address	0.0.0.0:50470	名称节点的HTTPS协议访问地址和端口
14	dfs.datanode.dns.interface	default	数据节点采用IP地址标识
15	dfs.datanode.dns.nameserver	default	指定DNS的IP地址
16	dfs.replication.considerLoad	true	加载目标或不加载的标识
17	dfs.default.chunk.view.size	32768	浏览时的文件块大小设置为32K
18	dfs.datanode.du.reserved	0	每个卷预留的空闲空间数量

19	dfs.name.dir	\${hadoop.tmp.dir}/dfs/name	存贮在本地的名字节点数据镜象的目录,作为名字节点的冗余备份
20	dfs.name.edits.dir	\${dfs.name.dir}	存贮文件操作过程信息的存贮目录
21	dfs.web.ugi	webuser,webgroup	Web接口访问的用户名和组的帐户设定
22	dfs.permissions	true	文件操作时的权限检查标识。
23	dfs.permissions.supergroup	supergroup	超级用户的组名定义
24	dfs.block.access.token.enable	false	数据节点访问令牌标识
25	dfs.block.access.key.update.interval	600	升级访问钥时的间隔时间
26	dfs.block.access.token.lifetime	600	访问令牌的有效时间
27	dfs.data.dir	\${hadoop.tmp.dir}/dfs/data	数据节点的块本地存放目录
28	dfs.datanode.data.dir.perm	755	数据节点的存贮块的目录访问权限设置
29	dfs.replication	3	缺省的块复制数量
30	dfs.replication.max	512	块复制的最大数量
31	dfs.replication.min	1	块复制的最小数量
32	dfs.block.size	67108864	缺省的文件块大小为64M
33	dfs.df.interval	60000	磁盘空间统计间隔为6秒
34	dfs.client.block.write.retries	3	块写入出错时的重试次数
35	dfs.blockreport.intervalMsec	3600000	块的报告间隔时为1小时
36	dfs.blockreport.initialDelay	0	块顺序报告的间隔时间
37	dfs.heartbeat.interval	3	数据节点的心跳检测间隔时间
38	dfs.namenode.handler.count	10	名称节点的连接处理的线程数量
39	dfs.safemode.threshold.pct	0.999f	启动安全模式的阈值设定
40	dfs.safemode.extension	30000	当阈值达到量值后扩展的时限
41	dfs.balance.bandwidthPerSec	1048576	启动负载均衡的数据节点可利用带宽最大值为1M
42	dfs.hosts		可与名称节点连接的主机地址文件指定。
43	dfs.hosts.exclude		不充计与名称节点连接的主机地址文件设定
44	dfs.max.objects	0	文件数、目录数、块数的最大数量
45	dfs.namenode.decommission.interval	30	名称节点解除命令执行时的监测时间周期
46	dfs.namenode.decommission.nodes.per.interval	5	名称节点解除命令执行是否完检测次数
47	dfs.replication.interval	3	名称节点计算数据节点的复制工作的周期数。
48	dfs.access.time.precision	3600000	充许访问文件的时间精确到1小时
49	dfs.support.append	false	是否充许链接文件指定
50	dfs.namenode.delegation.key.update-interval	86400000	名称节点上的代理令牌的主key的更新间隔时间为24小时
51	dfs.namenode.delegation.token.max-lifetime	604800000	代理令牌的有效时间最大值为7天
52	dfs.namenode.delegation.token.renew-interval	86400000	代理令牌的更新时间为24小时
53	dfs.datanode.failed.volumes.tolerated	0	决定停止数据节点提供服务充许卷的出错次数。0次则任何卷出错都要停止数据节点