安装步骤

1. 下载安装jdk
2. Hadoop下载地址 <https://dlcdn.apache.org/hadoop/core/>
3. 3台机器（8核16G）

192.168.29.185 hadoop1 zookeeper、datanod、journalnode、nodemanage节点

192.168.29.186 hadoop2 zookeeper、namenode、zxfc、datanode、journalnode、nodemanage、resourcemanager节点

192.168.29.61 hadoop3 zookeeper、namenode、zxfc、datanode、journalnode、nodemanage、resourcemanager节点

节点规划说明：  
  
 Zookeeper集群: 需要至少3个节点，并且节点数为奇数个，可以部署在任意独立节点上，NameNode及ResourceManager依赖zookeeper进行主备选举和切换  
   
 NameNode: 至少需要2个节点，一主多备，可以部署在任意独立节点上，用于管理HDFS的名称空间和数据块映射，依赖zookeeper和zkfc实现高可用和自动故障转移，并且依赖journalnode实现状态同步  
   
 ZKFailoverController: 即zkfc，在所有NameNode节点上启动，用于监视和管理NameNode状态，参与故障转移  
   
 Journalnode: 至少需要3个节点，并且节点数为奇数个，可以部署在任意独立节点上，用于主备NameNode状态信息同步  
   
 ResourceManager: 至少需要2个节点，一主多备，可以部署在任意独立节点上，依赖zookeeper实现高可用和自动故障转移，用于资源分配和调度  
   
 DataNode: 至少需要3个节点，因为hdfs默认副本数为3，可以部署在任意独立节点上，用于实际数据存储  
   
 NodeManage: 部署在所有DataNode节点上，用于节点资源管理和监控

1. 放到hadoop1中的/data/hadoop目录中

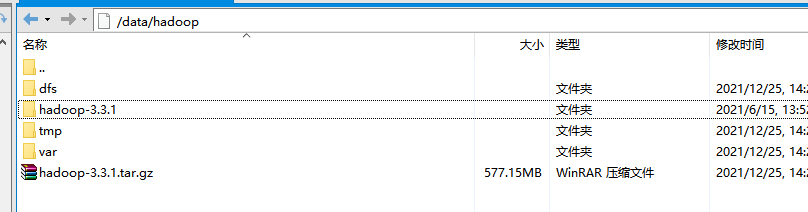
新建以下目录

mkdir -p /data/hadoop

mkdir -p /data/hadoop/tmp

mkdir -p /data/hadoop/dfs/data

mkdir -p /data/hadoop/dfs/name



1. 修改 /data/hadoop/hadoop-3.3.1/etc/hadoop/core-site.xml文件

在<configuration>节点内加入配置

<property>

<name>hadoop.tmp.dir</name>

<value>/data/hadoop/tmp</value>

<description>Abase for other temporary directories.</description>

</property>

<property>

<name>fs.defaultFS</name>

<value>hdfs://hadoop1:9000</value>

</property>

<property>

<name>io.file.buffer.size</name>

<value>131072</value>

<description>Size of read/write buffer used in SequenceFiles</description>

</property>

<property>

<name>ha.zookeeper.quorum</name>

<value>sfserver:2181,ywzfz:2181,hadoop3:2181</value>

<description>zookeeper地址</description>

</property>

<property>

<name>ha.zookeeper.session-timeout.ms</name>

<value>1000</value>

<description>hadoop链接zookeeper的超时时长设置ms</description>

</property>

1. 修改hadoop-env.sh

设置java home路径

export JAVA\_HOME=/data/jdk1.8.0\_121

export HADOOP\_PID\_DIR=/data/hadoop/tmp/pids

export HDFS\_NAMENODE\_USER=zhyx

export HDFS\_DATANODE\_USER=zhyx

export HDFS\_SECONDARYNAMENODE\_USER=zhyx

export YARN\_RESOURCEMANAGER\_USER=zhyx

export YARN\_NODEMANAGER\_USER=zhyx

export HDFS\_ZKFC\_USER=zhyx

export HDFS\_JOURNALNODE\_USER=zhyx

export HDFS\_SHELL\_EXECNAME=zhyx

1. 修改hdfs-site.xml
2. 新建并且修改mapred-site.xml
3. 修改works文件
4. 修改yarn-site.xml文件
5. 启动hadoop

在85上初始化

cd /data/hadoop/hadoop-3.3.1/bin

./hdfs --daemon start journalnode

./hadoop namenode -format

主节点初始化namenode后需要在从节点的namenode机器上执行同步namenode配置

./hdfs namenode -bootstrapStandby

在zk的leader上执行(启动journalnode时已经格式化了)

./hdfs zkfc -formatZK

初始化成功后启动

cd /data/hadoop/hadoop-3.3.1/sbin

这个启动脚本是通过ssh对多个节点的namenode、datanode、journalnode，resource以及zkfc进程进行批量启动的。

./start-all.sh

// mr-jobhistory-daemon.sh start historyserver

// 使用 mapred --daemon start historyserver 代替

hdfs主备切换

查看hdfs状态

*hdfs haadmin -getAllServiceState*

sfserver-dev:8020 standby

sfserver:8020 active

其中sfserver-dev 为nn1 sfserver为nn2

尽量使用

hdfs haadmin -failover --forcefence --forceactive NameNode2 NameNode1

来做主备切换

此处 “NameNode2 NameNode1” 的顺序表示 active 状态由 NameNode2 转换到 NameNode1 上（虽然 NameNode2 在转化前也是 standby 状态）。

此时如果使用

hdfs haadmin -transitionToActive --forcemanual nn1 时，会报错

必须使用

hdfs haadmin -transitionToStandby --forcemanual nn2

将nn2节点转为standby ,这样一来nn1会自动切换为active

*hdfs haadmin -getAllServiceState*

sfserver-dev:8020 active

sfserver:8020 standby

但是需要注意的是以上这两个命令不会尝试运行任何的 fence，因此不应该经常使用

yarn主备切换

yarn rmadmin -getAllServiceState

sfserver-dev:8033 standby

sfserver:8033 active

其中sfserver-dev 为rm1 sfserver为rm2

./yarn rmadmin -transitionToActive --forcemanual rm1

./yarn rmadmin -transitionToStandby --forcemanual rm2

查看zk选举情况

get -s /yarn-leader-election/yrc/ActiveStandbyElectorLock

get -s /hadoop-ha/myha01/ActiveStandbyElectorLock