|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| host | ip | journalNode | nameNode | dataNode | resourceManager | nodeManager | jobHistoryServer | hmaster | hregionserver |
| node1 | 47.119.115.107 | √ | √ | √ |  | √ | √ | √ | √ |
| node2 | 8.135.110.24 | √ | √ | √ | √ | √ |  | √ | √ |
| node3 | 8.135.118.28 | √ |  | √ | √ | √ |  |  | √ |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| host | ip | journalNode | nameNode | dataNode | resourceManager | nodeManager | jobHistoryServer | hmaster | hregionserver |
| hadoop14 | 10.83.64.192 | √ | √ | √ |  | √ | √ | √ | √ |
| hadoop15 | 10.83.64.96 | √ | √ | √ | √ | √ |  | √ | √ |
| hadoop16 | 10.83.64.211 | √ |  | √ | √ | √ |  |  | √ |

## 一、配置**ssh**免密登录

## 二、配置**xcall**、**xsync**

## 三、安装**Java**环境

## 四、安卓**zookeeper**集群

阿里云ECS启动zookeeper集群时，会出现：

|  |
| --- |
| As I'm leaving the listener thread, I won't be able to participate in leader election any longer: 8.135.118.28:3888 |

这时需要在zoo.cfg中增加

|  |
| --- |
| quorumListenOnAllIPs=true |

以及在阿里云控制台增加安全组

## 五、安装**hadoop**集群

**hdfs**

1. 三台起journalnode（起journalnode之前确保目录是空的）

|  |
| --- |
| hadoop-daemon.sh start journalnode |

1. 一台namenode格式化并启动

|  |
| --- |
| 1）hadoop namenode -format 要保证journalnode配置的数据目录为空  ==》Storage directory /tmp/hadoop/name has been successfully formatted  2）hadoop-daemon.sh start namenode  阿里云三个节点的/etc/hosts文件中，需要配置私网IP到主机名的映射。否则启动的时候提示找不到0.0.0.0:50070这个address |

1. 另一台namenode同步

|  |
| --- |
| hdfs namenode -bootstrapStandby |

4、格式化zkfc（找一台namenode节点）

|  |
| --- |
| hdfs zkfc -formatZK |

5、启动所有

|  |
| --- |
| start-dfs.sh |

**yarn**

1. 一台主节点

|  |
| --- |
| start-yarn.sh  yarn-daemon.sh start resourcemanager |

1. 另一台主节点

|  |
| --- |
| yarn-daemon.sh start resourcemanager |

1. 启动historyserver

|  |
| --- |
| mr-jobhistory-daemon.sh start historyserver |

|  |
| --- |
| **hdfs页面地址**  node1:50070 node2:50070 主副有自己各自页面  **yarn页面地址**  node2:8088 node3:8088 副的会跳到主页面  **Jobhistory页面地址**  node1:19888 |

yarn-site.xml

|  |
| --- |
| **nodemanager**  <!-- 每个NM有8个核心数可用 -->  yarn.nodemanager.resource.cpu-vcores 2  <!-- 每个NM有32GB内存可用 -->  yarn.nodemanager.resource.memory-mb 6000  **resourcemanager**  <!-- 每个container最少申请内存数 -->  yarn.scheduler.minimum-allocation-mb 1024  <!-- 每个container最多申请内存数 -->  yarn.scheduler.maximum-allocation-mb 4096 |

mapred-site.xml

|  |
| --- |
| <!-- MapTask内存大小限制 -->  mapreduce.map.memory.mb 2048  <!-- MapTask最大重试次数 -->  mapreduce.map.maxattempts 3  <!-- ReduceTask内存大小限制 -->  mapreduce.reduce.memory.mb 2048  <!-- ReduceTask最大重试次数 -->  mapreduce.reduce.maxattempts 3  **可能会出现java的OOM，需要加入下面的配置，值要小于**  **mapreduce.map.java.opts -Xmx2000m mapreduce.map.memory.mb**  <!-- 环形缓冲区大小 -->  mapreduce.task.io.sort.mb 200  <!-- 环形缓冲区溢写百分比 -->  mapreduce.map.sort.spill.percent 0.60  <!-- ApplicationMaster内存大小限制 -->  yarn.app.mapreduce.am.resource.mb 2048  yarn.app.mapreduce.am.command-opts -Xmx1000M |

## 六、安装**hive**

1. hive-env.xml 设置HADOOP\_HOME
2. hive-site.xml 设置mysql
3. hive-site.xml 设置自动创建

<property>

<name>datanucleus.schema.autoCreateAll</name>

<value>true</value>

</property>

1. 安装mysql，创建数据库
2. 初始化hive库中的表，Linux命令行执行 schematool -dbType mysql -initSchema
3. 启动hiveserver2 nohup hiveserver2 >> hiveserver2.log & hiveserver2.log在bin目录

## 安装hbase

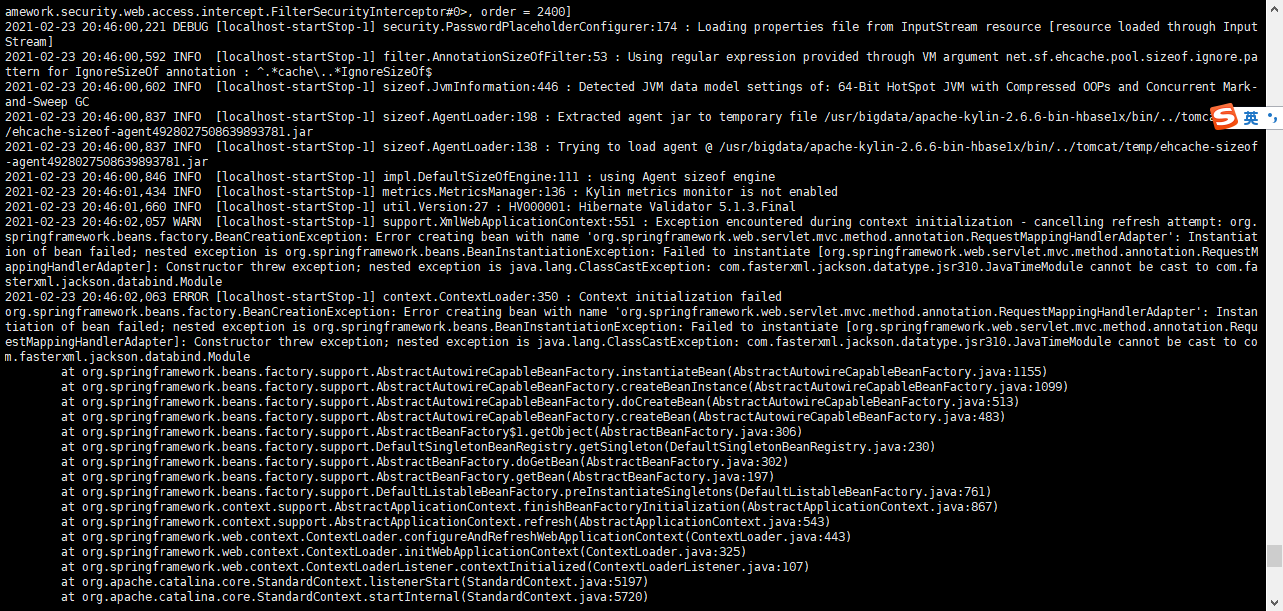
1. hbase-env.sh 设置JAVA\_HOME
2. hbase-site.sh设置bindAddress（每台机子设置自己的主机名）
3. 新建backup-masters文件（一台master写另一台master的主机名）
4. 修改regionservers文件（3台regionserver的主机名）
5. 建立core-site.xml、hdfs-site.xml的软连接

|  |
| --- |
| ln -s /usr/bigdata/hadoop-2.10.0/etc/hadoop/core-site.xml /usr/bigdata/hbase-1.6.0/conf/core-site.xml  ln -s /usr/bigdata/hadoop-2.10.0/etc/hadoop/hdfs-site.xml /usr/bigdata/hbase-1.6.0/conf/hdfs-site.xml |

1. 在一台master上启动start-hbase.sh
2. 访问web页面 node1:16010

## 安装kylin

1. ${KYLIN\_HOME}下创建hadoop\_conf文件夹，里面建立core-site.xml、hdfs-site.xml、yarn-site.xml、mapred-site.xml、hive-site.xml、hbase-site.xml的软连接
2. Kylin.properties设置kylin集群节点，beeline连接，hadoop\_conf目录
3. 执行kylin-env.sh，检查kylin环境，会在hdfs下面创建kylin目录
4. 每个节点执行 kylin.sh start
5. 出现问题：



解决方法：下载jackson-datatype-jsr310-2.10.3.jar包，放到${KYLIN\_HOME}/tomcat/webapps/kylin/WEB-INF/lib下面即可