Hadoop

Hadoop

- Hadoop https://hadoop.apache.org/
- Hadoop https://hadoop.apache.org/releases.html

- 2C4G 8G
  - 172.16.0.17
  - 172.16.0.43
  - 172.16.0.180
- CentOS 7.5
  - root
- Java 1.8
  - `echo $JAVA_HOME`
- Hadoop:2.6.5
- `systemctl stop firewalld.service`

- Hadoop HDFS YARN
  - HDFS NameNode / DataNode
  - YARN ResourceManager /NodeManager
  - HDFS master/worker HDFS Namenode Datanodes Namenode Datanode
- hostname

```
hostnamectl --static set-hostname linux01
hostnamectl --static set-hostname linux02
hostnamectl --static set-hostname linux03
```

- hosts

```
vim /etc/hosts
172.16.0.17 linux01
172.16.0.43 linux02
172.16.0.180 linux03
```

- linux01

```
```

```
ssh-keygen -t rsa



      authorized_keys
cat /root/.ssh/id_rsa.pub >> /root/.ssh/authorized_keys



ssh localhost
```

- slave
  - SSH pem

```
ssh-copy-id -i ~/.ssh/id_rsa.pub -p 22 root@172.16.0.43       linux02    root
ssh-copy-id -i ~/.ssh/id_rsa.pub -p 22 root@172.16.0.180       linux03    root



  linux01
ssh linux02
ssh linux03
```

Hadoop

- Spark Flink
- 2.6.5 2.6.5
- https://hadoop.apache.org/docs/r2.6.5/hadoop-project-dist/hadoop-common/ClusterSetup.html
-

```
mkdir -p /data/hadoop/hdfs/name /data/hadoop/hdfs/data /data/hadoop/hdfs/tmp
```

- Hadoop http://apache.claz.org/hadoop/common/hadoop-2.6.5/
- linux01

```
cd /usr/local && wget http://apache.claz.org/hadoop/common/hadoop-2.6.5/hadoop-2.6.5.tar.gz
tar zxvf hadoop-2.6.5.tar.gz   191M
```

- **HADOOP_HOME**

- ansible playbookAnsible

```
vim /etc/profile

export HADOOP_HOME=/usr/local/hadoop-2.6.5
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin

source /etc/profile
```

linux01

```
   JAVA_HOME
vim $HADOOP_HOME/etc/hadoop/hadoop-env.sh

  25
export JAVA_HOME=${JAVA_HOME}

export JAVA_HOME=/usr/local/jdk1.8.0_191



vim $HADOOP_HOME/etc/hadoop/yarn-env.sh

       export JAVA_HOME=/usr/local/jdk1.8.0_191
```

- hadoop.tmp.dir == hadoop

```


vim $HADOOP_HOME/etc/hadoop/core-site.xml

<configuration>
    <property>
        <name>hadoop.tmp.dir</name>
        <value>file:/data/hadoop/hdfs/tmp</value>
    </property>
    <property>
        <name>io.file.buffer.size</name>
        <value>131072</value>
    </property>
```

```
    <!--
    <property>
        <name>fs.default.name</name>
        <value>hdfs://linux01:9000</value>
    </property>
    -->
    <property>
        <name>fs.defaultFS</name>
        <value>hdfs://linux01:9000</value>
    </property>
    <property>
        <name>hadoop.proxyuser.root.hosts</name>
        <value>*</value>
    </property>
    <property>
        <name>hadoop.proxyuser.root.groups</name>
        <value>*</value>
    </property>
</configuration>
```

-
  - datanode
-

```
vim $HADOOP_HOME/etc/hadoop/hdfs-site.xml

<configuration>
  <property>
    <name>dfs.replication</name>
    <value>2</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/data/hadoop/hdfs/name</value>
    <final>true</final>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:/data/hadoop/hdfs/data</value>
    <final>true</final>
  </property>
```

```
  <property>
    <name>dfs.webhdfs.enabled</name>
    <value>true</value>
  </property>
  <property>
    <name>dfs.permissions</name>
    <value>false</value>
  </property>
</configuration>
```

- YARN

```
   vim $HADOOP_HOME/etc/hadoop/mapred-site.xml

<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>

  <property>
    <name>mapreduce.map.memory.mb</name>
    <value>4096</value>
  </property>

  <property>
    <name>mapreduce.reduce.memory.mb</name>
    <value>8192</value>
  </property>

  <property>
    <name>mapreduce.map.java.opts</name>
    <value>-Xmx3072m</value>
  </property>

  <property>
    <name>mapreduce.reduce.java.opts</name>
    <value>-Xmx6144m</value>
```

```
  </property>

</configuration>
```

- yarn.resourcemanager.hostname == YARN ResourceManager
- yarn.nodemanager.aux-services == NodeManager mapreduce_shuffle MapRedu
- 32G

```
vim $HADOOP_HOME/etc/hadoop/yarn-site.xml



<configuration>
  <property>
    <name>yarn.resourcemanager.hostname</name>
    <value>linux01</value>
  </property>

  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>

  <property>
    <name>yarn.nodemanager.vmem-pmem-ratio</name>
    <value>2.1</value>
  </property>

  <property>
    <name>yarn.nodemanager.resource.memory-mb</name>
    <value>20480</value>
  </property>

  <property>
    <name>yarn.scheduler.minimum-allocation-mb</name>
    <value>2048</value>
  </property>

</configuration>
```

- slave

```

```

```
vim $HADOOP_HOME/etc/hadoop/slaves

        localhost
linux02
linux03
```

```
scp -r /usr/local/hadoop-2.6.5 root@linux02:/usr/local/

scp -r /usr/local/hadoop-2.6.5 root@linux03:/usr/local/
```

linux01

```
    HDFS
hdfs namenode -format
```

```
[root@linux01 hadoop-2.6.5]# hdfs namenode -format
18/12/17 17:47:17 INFO namenode.NameNode:  STARTUP_MSG:
/************************************************************
STARTUP_MSG:  Starting NameNode
STARTUP_MSG:    host = localhost/127.0.0.1
STARTUP_MSG:    args = [-format]
STARTUP_MSG:    version = 2.6.5
STARTUP_MSG:    classpath = /usr/local/hadoop-2.6.5/etc/hadoop:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/apacheds-kerberos-codec-2.0.0-M15.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-io-2.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/activation-1.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/netty-3.6.2.Final.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/slf4j-api-1.7.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/junit-4.11.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/curator-recipes-2.6.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jasper-compiler-5.5.23.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jets3t-0.9.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-lang-2.6.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-digester-1.8.jar:/usr/local/hadoop-
```

2.6.5/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/apacheds-i18n-2.0.0-M15.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/guava-11.0.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/gson-2.2.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jackson-jaxrs-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jettison-1.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jetty-6.1.26.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/api-util-1.0.0-M20.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/log4j-1.2.17.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-beanutils-core-1.8.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-httpclient-3.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-el-1.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/paranamer-2.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-collections-3.2.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jersey-server-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-net-3.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/hadoop-auth-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jasper-runtime-5.5.23.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jaxb-impl-2.2.3-1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/hamcrest-core-1.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/stax-api-1.0-2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-beanutils-1.7.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/curator-framework-2.6.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/xz-1.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jsr305-1.3.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jsp-api-2.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-compress-1.4.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/asm-3.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jsch-0.1.42.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-configuration-1.6.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-cli-1.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jackson-xc-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-logging-1.1.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/htrace-core-3.0.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jetty-util-6.1.26.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-math3-3.1.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/mockito-all-1.8.5.jar:/usr/local/hadoop-

2.6.5/share/hadoop/common/lib/jersey-json-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/zookeeper-3.4.6.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/httpclient-4.2.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/servlet-api-2.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/xmlenc-0.52.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/httpcore-4.2.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/api-asn1-api-1.0.0-M20.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/java-xmlbuilder-0.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/avro-1.7.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jaxb-api-2.2.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/commons-codec-1.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/jersey-core-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/snappy-java-1.0.4.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/curator-client-2.6.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/lib/hadoop-annotations-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/hadoop-common-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/hadoop-common-2.6.5-tests.jar:/usr/local/hadoop-
2.6.5/share/hadoop/common/hadoop-nfs-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs:/usr/local/hadoop-2.6.5/share/hadoop/hdfs/lib/commons-io-
2.4.jar:/usr/local/hadoop-2.6.5/share/hadoop/hdfs/lib/netty-3.6.2.Final.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jackson-mapper-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/commons-lang-2.6.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/commons-daemon-1.0.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jackson-core-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/guava-11.0.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jetty-6.1.26.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/log4j-1.2.17.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/commons-el-1.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/xercesImpl-2.9.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jersey-server-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jasper-runtime-5.5.23.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/protobuf-java-2.5.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jsr305-1.3.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/xml-apis-1.3.04.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jsp-api-2.1.jar:/usr/local/hadoop-2.6.5/share/hadoop/hdfs/lib/asm-
3.2.jar:/usr/local/hadoop-2.6.5/share/hadoop/hdfs/lib/commons-cli-1.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/commons-logging-1.1.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/htrace-core-3.0.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jetty-util-6.1.26.jar:/usr/local/hadoop-

2.6.5/share/hadoop/hdfs/lib/servlet-api-2.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/xmlenc-0.52.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/commons-codec-1.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/lib/jersey-core-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/hadoop-hdfs-2.6.5-tests.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/hadoop-hdfs-nfs-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/hdfs/hadoop-hdfs-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-io-2.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/activation-1.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/aopalliance-1.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/netty-3.6.2.Final.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jackson-mapper-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-lang-2.6.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jackson-core-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/guice-3.0.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/guava-
11.0.2.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/jackson-jaxrs-
1.9.13.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/jettison-1.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jetty-6.1.26.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/log4j-1.2.17.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-httpclient-3.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-collections-3.2.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jersey-server-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jaxb-impl-2.2.3-1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/stax-api-1.0-2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/protobuf-java-2.5.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/xz-1.0.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/jsr305-
1.3.9.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/jersey-client-
1.9.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/guice-servlet-3.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-compress-1.4.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/asm-3.2.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/commons-
cli-1.2.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/jersey-guice-
1.9.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/lib/jackson-xc-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-logging-1.1.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jetty-util-6.1.26.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/leveldbjni-all-1.8.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jersey-json-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/javax.inject-1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/zookeeper-3.4.6.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/servlet-api-2.5.jar:/usr/local/hadoop-

2.6.5/share/hadoop/yarn/lib/jaxb-api-2.2.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jline-0.9.94.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/commons-codec-1.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/lib/jersey-core-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-server-web-proxy-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-api-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-server-common-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-registry-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-server-nodemanager-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-client-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-common-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/yarn/hadoop-yarn-applications-unmanaged-am-launcher-
2.6.5.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/hadoop-yarn-server-tests-
2.6.5.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/hadoop-yarn-server-resourcemanager-
2.6.5.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/hadoop-yarn-server-
applicationhistoryservice-2.6.5.jar:/usr/local/hadoop-2.6.5/share/hadoop/yarn/hadoop-yarn-
applications-distributedshell-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/commons-io-2.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/aopalliance-1.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/netty-3.6.2.Final.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/jackson-mapper-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/junit-4.11.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/jackson-core-asl-1.9.13.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/guice-3.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/log4j-1.2.17.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/paranamer-2.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/jersey-server-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/hamcrest-core-1.3.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/protobuf-java-2.5.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/xz-1.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/guice-servlet-3.0.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/commons-compress-1.4.1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/asm-3.2.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/jersey-guice-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/leveldbjni-all-1.8.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/javax.inject-1.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/avro-1.7.4.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/jersey-core-1.9.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/lib/snappy-java-1.0.4.1.jar:/usr/local/hadoop-

```
2.6.5/share/hadoop/mapreduce/lib/hadoop-annotations-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-app-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-common-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-shuffle-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-jobclient-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-core-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-hs-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-hs-plugins-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar:/usr/local/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-client-jobclient-2.6.5-
tests.jar:/usr/local/hadoop-2.6.5/contrib/capacity-scheduler/*.jar
STARTUP_MSG:   build = https://github.com/apache/hadoop.git -r
e8c9fe0b4c252caf2ebf1464220599650f119997; compiled by 'sjlee' on 2016-10-02T23:43Z
STARTUP_MSG:   java = 1.8.0_191
************************************************************/
18/12/17 17:47:17 INFO namenode.NameNode: registered UNIX signal handlers for [TERM, HUP, INT]
18/12/17 17:47:17 INFO namenode.NameNode: createNameNode [-format]
Formatting using clusterid: CID-beba43b4-0881-48b4-8eda-5c3bca046398
18/12/17 17:47:17 INFO namenode.FSNamesystem: No KeyProvider found.
18/12/17 17:47:17 INFO namenode.FSNamesystem: fsLock is fair:true
18/12/17 17:47:17 INFO blockmanagement.DatanodeManager: dfs.block.invalidate.limit=1000
18/12/17 17:47:17 INFO blockmanagement.DatanodeManager: dfs.namenode.datanode.registration.ip-
hostname-check=true
18/12/17 17:47:17 INFO blockmanagement.BlockManager:
dfs.namenode.startup.delay.block.deletion.sec is set to 000:00:00:00.000
18/12/17 17:47:17 INFO blockmanagement.BlockManager: The block deletion will start around
2018 Dec 17 17:47:17
18/12/17 17:47:17 INFO util.GSet: Computing capacity for map BlocksMap
18/12/17 17:47:17 INFO util.GSet: VM type       = 64-bit
18/12/17 17:47:17 INFO util.GSet: 2.0% max memory 889 MB = 17.8 MB
18/12/17 17:47:17 INFO util.GSet: capacity      = 2^21 = 2097152 entries
18/12/17 17:47:17 INFO blockmanagement.BlockManager: dfs.block.access.token.enable=false
18/12/17 17:47:17 INFO blockmanagement.BlockManager: defaultReplication         = 2
18/12/17 17:47:17 INFO blockmanagement.BlockManager: maxReplication             = 512
18/12/17 17:47:17 INFO blockmanagement.BlockManager: minReplication             = 1
18/12/17 17:47:17 INFO blockmanagement.BlockManager: maxReplicationStreams      = 2
18/12/17 17:47:17 INFO blockmanagement.BlockManager: replicationRecheckInterval = 3000
18/12/17 17:47:17 INFO blockmanagement.BlockManager: encryptDataTransfer        = false
18/12/17 17:47:17 INFO blockmanagement.BlockManager: maxNumBlocksToLog          = 1000
```

```
18/12/17 17:47:17 INFO namenode.FSNamesystem: fsOwner             = root (auth:SIMPLE)
18/12/17 17:47:17 INFO namenode.FSNamesystem: supergroup          = supergroup
18/12/17 17:47:17 INFO namenode.FSNamesystem: isPermissionEnabled = false
18/12/17 17:47:17 INFO namenode.FSNamesystem: HA Enabled: false
18/12/17 17:47:17 INFO namenode.FSNamesystem: Append Enabled: true
18/12/17 17:47:17 INFO util.GSet: Computing capacity for map INodeMap
18/12/17 17:47:17 INFO util.GSet: VM type       = 64-bit
18/12/17 17:47:17 INFO util.GSet: 1.0% max memory 889 MB = 8.9 MB
18/12/17 17:47:17 INFO util.GSet: capacity      = 2^20 = 1048576 entries
18/12/17 17:47:17 INFO namenode.NameNode: Caching file names occuring more than 10 times
18/12/17 17:47:17 INFO util.GSet: Computing capacity for map cachedBlocks
18/12/17 17:47:17 INFO util.GSet: VM type       = 64-bit
18/12/17 17:47:17 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
18/12/17 17:47:17 INFO util.GSet: capacity      = 2^18 = 262144 entries
18/12/17 17:47:17 INFO namenode.FSNamesystem: dfs.namenode.safemode.threshold-pct =
0.9990000128746033
18/12/17 17:47:17 INFO namenode.FSNamesystem: dfs.namenode.safemode.min.datanodes = 0
18/12/17 17:47:17 INFO namenode.FSNamesystem: dfs.namenode.safemode.extension     = 30000
18/12/17 17:47:17 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
18/12/17 17:47:17 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and
retry cache entry expiry time is 600000 millis
18/12/17 17:47:17 INFO util.GSet: Computing capacity for map NameNodeRetryCache
18/12/17 17:47:17 INFO util.GSet: VM type       = 64-bit
18/12/17 17:47:17 INFO util.GSet: 0.029999999329447746% max memory 889 MB = 273.1 KB
18/12/17 17:47:17 INFO util.GSet: capacity      = 2^15 = 32768 entries
18/12/17 17:47:17 INFO namenode.NNConf: ACLs enabled? false
18/12/17 17:47:17 INFO namenode.NNConf: XAttrs enabled? true
18/12/17 17:47:17 INFO namenode.NNConf: Maximum size of an xattr: 16384
18/12/17 17:47:17 INFO namenode.FSImage: Allocated new BlockPoolId: BP-233285725-127.0.0.1-
1545040037972
18/12/17 17:47:18 INFO common.Storage: Storage directory /data/hadoop/hdfs/name has been
successfully formatted.
18/12/17 17:47:18 INFO namenode.FSImageFormatProtobuf: Saving image file
/data/hadoop/hdfs/name/current/fsimage.ckpt_0000000000000000000 using no compression
18/12/17 17:47:18 INFO namenode.FSImageFormatProtobuf: Image file
/data/hadoop/hdfs/name/current/fsimage.ckpt_0000000000000000000 of size 321 bytes saved in 0
seconds.
18/12/17 17:47:18 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid
>= 0
```

```
18/12/17 17:47:18 INFO util.ExitUtil: Exiting with status 0
18/12/17 17:47:18 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at localhost/127.0.0.1
************************************************************/
```

HDFS

- start-dfs.sh yes

```
        NameNode   SecondaryNameNode
        DataNode



     jps
21922 Jps
21603 NameNode
21787 SecondaryNameNode



     jps
19728 DataNode
19819 Jps
```

- `hdfs dfsadmin -report`

```
Configured Capacity: 0 (0 B)
Present Capacity: 0 (0 B)
DFS Remaining: 0 (0 B)
DFS Used: 0 (0 B)
DFS Used%: NaN%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
```

- `stop-dfs.sh`
- l `cd $HADOOP_HOME/logs`

YARN

```
start-yarn.sh
   jps       ResourceManager

       NodeManager

   stop-yarn.sh
```

- `netstat -tpnl | grep java`
-

```
tcp        0      0 172.16.0.17:9000        0.0.0.0:*               LISTEN      22932/java >>
NameNode
tcp        0      0 0.0.0.0:50070           0.0.0.0:*               LISTEN      22932/java >>
NameNode
tcp        0      0 0.0.0.0:50090           0.0.0.0:*               LISTEN      23125/java >>
SecondaryNameNode
tcp6       0      0 172.16.0.17:8030      :::*                    LISTEN      23462/java   >>
ResourceManager
tcp6       0      0 172.16.0.17:8031      :::*                    LISTEN      23462/java   >>
ResourceManager
tcp6       0      0 172.16.0.17:8032      :::*                    LISTEN      23462/java   >>
ResourceManager
tcp6       0      0 172.16.0.17:8033      :::*                    LISTEN      23462/java   >>
ResourceManager
tcp6       0      0 172.16.0.17:8088      :::*                    LISTEN      23462/java   >>
ResourceManager
```

- `netstat -tpnl | grep java`
-

```
tcp        0      0 0.0.0.0:50010           0.0.0.0:*               LISTEN      14545/java >>
DataNode
tcp        0      0 0.0.0.0:50020           0.0.0.0:*               LISTEN      14545/java >>
DataNode
tcp        0      0 0.0.0.0:50075           0.0.0.0:*               LISTEN      14545/java >>
DataNode
tcp6       0      0 :::8040                 :::*                    LISTEN      14698/java >>
```

```
NodeManager
tcp6       0      0 :::8042                 :::*                    LISTEN      14698/java >>
NodeManager
tcp6       0      0 :::13562                :::*                    LISTEN      14698/java >>
NodeManager
tcp6       0      0 :::37481                :::*                    LISTEN      14698/java >>
NodeManager
```

- HDFS NameNode http://linux01:50070
- YARN ResourceManager http://linux01:8088
- NodeManager-1 http://linux02:8042
- NodeManager-2 http://linux03:8042

- Mapreduce
  - `hadoop jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar pi 5 10`
-
  - hadoop HDFS wordcount HDFS
  - `hadoop fs -ls /`
    - `hadoop fs -ls hdfs://linux-05:9000/`
    - `hadoop fs -ls -R /`
    - hadoop HDFS
  - `hadoop fs -mkdir -p /tmp/zch/wordcount_input_dir`
  - `hadoop fs -put /opt/input.txt /tmp/zch/wordcount_input_dir`
  - `hadoop fs -ls /tmp/zch/wordcount_input_dir`
  - yarn `hadoop jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordcount /tmp/zch/wordcount_input_dir /tmp/zch/wordcount_output_dir`
  - `hadoop fs -ls /tmp/zch/wordcount_output_dir`
  - `hadoop fs -cat /tmp/zch/wordcount_output_dir/part-r-00000`
- H `yarn application -list`
- Hadoo `yarn application -kill  ApplicationId`

- MapReduce
- https://www.linode.com/docs/databases/hadoop/how-to-install-and-set-up-hadoop-cluster/
- http://www.cnblogs.com/Leo_wl/p/7426496.html
- https://blog.csdn.net/bingduanlbd/article/details/51892750
- https://blog.csdn.net/whdxjbw/article/details/81050597

Revision #1

Created 20 March 2020 15:26:42 by

Updated 14 August 2020 06:18:41 by