

11.1 Configure Hadoop and start cluster services using Ansible Playbook.

Step1: Checking the Hadoop is installed in Namenode.

```
root@NameNode:~  
[root@NameNode ~]# hadoop  
-bash: hadoop: command not found  
[root@NameNode ~]#
```

Step2: Checking the Hadoop is installed in Datanode.

```
root@DataNode1:~  
[root@DataNode1 ~]# hadoop  
-bash: /usr/bin/hadoop: No such file or directory  
[root@DataNode1 ~]#
```

Here, Hadoop was not installed in both Namenode and Datanode.

Step3: Lets configure the Hadoop in both Namenode and Datanode

- ➔ Created a playbook named play.yml to configure Hadoop cluster.
- ➔ Created another playbook named delete_play.yml to destroy Hadoop cluster.

```
root@localhost:~/hadoop  
[root@ansible_controller hadoop]# ls  
core-site.xml  delete_play.yml  hadoop-1.2.1-1.x86_64.rpm  hdfs-site.xml  jdk-8u171-linux-x64.rpm  play.yml  play.yml.bak  test  
[root@ansible_controller hadoop]#
```

[illegible]

```

TASK [Checking hadoop] *****
fatal: [192.168.99.131]: FAILED! => ("changed": false, "cmd": ["rpm", "-q", "hadoop"], "delta": "0:00:00.291015", "end": "2021-02-12 07:56:06.945494", "msg": "non-zero return code", "rc": 1, "start": "2021-02-12 07:56:06.654475", "stderr": "", "stderr_lines": [], "stdout": "package hadoop is not installed", "stdout_lines": ["package hadoop is not installed"])

TASK [Installing hadoop] *****
changed: [192.168.99.131] => ("changed": true, "cmd": ["rpm", "-i", "hadoop-1.2.1-1.x86_64.rpm", "--force"], "delta": "0:00:13.114499", "end": "2021-02-12 07:56:20.571484", "rc": 0, "start": "2021-02-12 07:56:07.456985", "stderr": "/sbin/ldconfig: /lib64/libhadoop.so.1 is not a symbolic link\n/sbin/ldconfig: /lib64/libhdfs.so.0 is not a symbolic link", "stderr_lines": ["/sbin/ldconfig: /lib64/libhadoop.so.1 is not a symbolic link", "", "/sbin/ldconfig: /lib64/libhdfs.so.0 is not a symbolic link"], "stdout": "", "stdout_lines": [])

TASK [Configuring Hadoop Datanode] *****
changed: [192.168.99.131] => (item=hdfs-site.xml) => ("ansible_loop_var": "item", "changed": true, "checksum": "6647debb0b424dc9553f4438c67b11b8251800c3", "dest": "/etc/hadoop/hdfs-site.xml", "gid": 0, "group": "root", "item": "hdfs-site.xml", "md5sum": "4c24b225e0635e604c85fb800588bde0", "mode": "0644", "owner": "root", "secontext": "system_u:object_r:etc_t:s0", "size": 256, "src": "/root/.ansible/tmp/ansible-tmp-1613134581.2108843-23469-136408970980490/source", "state": "file", "uid": 0)
changed: [192.168.99.131] => (item=core-site.xml) => ("ansible_loop_var": "item", "changed": true, "checksum": "5f9648bbe3883e33dea7d7a8ec85449cb6a123a", "dest": "/etc/hadoop/core-site.xml", "gid": 0, "group": "root", "item": "core-site.xml", "md5sum": "f0e15922d9ee73ba6543015d1f67926f", "mode": "0644", "owner": "root", "secontext": "system_u:object_r:etc_t:s0", "size": 271, "src": "/root/.ansible/tmp/ansible-tmp-1613134583.66671-23469-108574970213221/source", "state": "file", "uid": 0)

TASK [Enable Firewall to port 9001] *****
changed: [192.168.99.131] => ("changed": true, "msg": "Permanent and Non-Permanent(immediate) operation, Changed port 9001/tcp to enabled")

RUNNING HANDLER [Creating Datanode Directory] *****
changed: [192.168.99.131] => ("changed": true, "gid": 0, "group": "root", "mode": "0755", "owner": "root", "path": "/hadoop_DN", "secontext": "unconfined_u:object_r:default_t:s0", "size": 6, "state": "directory", "uid": 0)

RUNNING HANDLER [start hadoop] *****
changed: [192.168.99.131] => ("changed": true, "cmd": ["hadoop-daemon.sh", "start", "datanode"], "delta": "0:00:01.174588", "end": "2021-02-12 07:56:31.540485", "rc": 0, "start": "2021-02-12 07:56:30.365897", "stderr": "", "stderr_lines": [], "stdout": "starting datanode, logging to /var/log/hadoop/root/hadoop-root-datanode-DataNode1.out", "stdout_lines": ["starting datanode, logging to /var/log/hadoop/root/hadoop-root-datanode-DataNode1.out"])

PLAY RECAP *****
192.168.99.129      : ok=9   changed=5   unreachable=0   failed=0   skipped=0   rescued=1   ignored=0
192.168.99.131     : ok=8   changed=5   unreachable=0   failed=0   skipped=0   rescued=1   ignored=0

[root@ansible_controller hadoop]#

```

Step4: Checking the Hadoop is perfectly configured for Namenode or not.

```

root@NameNode:/etc/hadoop

[root@NameNode ~]# jps
3856 Jps
3651 NameNode
[root@NameNode ~]# cd /etc/hadoop/
[root@NameNode hadoop]# ls
capacity-scheduler.xml  fair-scheduler.xml  hadoop-policy.xml  mapred-queue-acls.xml  slaves  taskcontroller.cfg
configuration.xml      hadoop-env.sh       hdfs-site.xml     mapred-site.xml       ssl-client.xml.example  task-log4j.properties
core-site.xml          hadoop-metrics2.properties  log4j.properties  masters               ssl-server.xml.example

[root@NameNode hadoop]#

```

```

root@NameNode:/etc/hadoop

[root@NameNode hadoop]# cat hdfs-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>


<!-- Put site-specific property overrides in this file. -->

<configuration>

<property>
<name>dfs.name.dir</name>
<value>/hadoop_NN</value>
</property>

</configuration>


```

 root@NameNode:/etc/hadoop

```
[root@NameNode hadoop]# cat core-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://192.168.99.129:9001</value>
</property>
</configuration>
[root@NameNode hadoop]#
```


 root@NameNode:/etc/hadoop

```
[root@NameNode hadoop]# hadoop dfsadmin -report
Configured Capacity: 6641680384 (6.19 GB)
Present Capacity: 4555804672 (4.24 GB)
DFS Remaining: 4555796480 (4.24 GB)
DFS Used: 8192 (8 KB)
DFS Used%: 0%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0


-----
Datanodes available: 1 (1 total, 0 dead)

Name: 192.168.99.131:50010
Decommission Status : Normal
Configured Capacity: 6641680384 (6.19 GB)
DFS Used: 8192 (8 KB)
Non DFS Used: 2085875712 (1.94 GB)
DFS Remaining: 4555796480 (4.24 GB)
DFS Used%: 0%
DFS Remaining%: 68.59%
Last contact: Fri Feb 12 08:09:24 EST 2021
```


Step5: Checking the Hadoop is perfectly configured for Datanode or not.

 root@DataNode1:/etc/hadoop

```
[root@DataNode1 hadoop]# jps
23336 DataNode
23646 Jps
```

 root@DataNode1:/etc/hadoop

```
[root@DataNode1 hadoop]# ls
capacity-scheduler.xml  fair-scheduler.xml      hadoop-policy.xml  mapred-queue-acls.xml  slaves          taskcontroller.cfg
configuration.xml       hadoop-env.sh           hdfs-site.xml     mapred-site.xml       ssl-client.xml.example  task-log4j.properties
core-site.xml           hadoop-metrics2.properties  log4j.properties  masters               ssl-server.xml.example
[root@DataNode1 hadoop]#
```

 root@DataNode1:/etc/hadoop


```
[root@DataNode1 hadoop]# cat hdfs-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>

<property>
<name>dfs.data.dir</name>
<value>/hadoop_DN</value>
</property>

</configuration>
```

 root@DataNode1:/etc/hadoop

```
[root@DataNode1 hadoop]# cat core-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://192.168.99.129:9001</value>
</property>
</configuration>
[root@DataNode1 hadoop]#
```

```
root@DataNode1:/etc/hadoop

[root@DataNode1 hadoop]# hadoop dfsadmin -report
Configured Capacity: 6641680384 (6.19 GB)
Present Capacity: 4555677696 (4.24 GB)
DFS Remaining: 4555669504 (4.24 GB)
DFS Used: 8192 (8 KB)
DFS Used%: 0%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0

-----
Datanodes available: 1 (1 total, 0 dead)

Name: 192.168.99.131:50010
Decommission Status : Normal
Configured Capacity: 6641680384 (6.19 GB)
DFS Used: 8192 (8 KB)
Non DFS Used: 2086002688 (1.94 GB)
DFS Remaining: 4555669504 (4.24 GB)
DFS Used%: 0%
DFS Remaining%: 68.59%
Last contact: Fri Feb 12 08:11:45 EST 2021
```

Step6: Delete the setup using delete_play.yml

```
root@localhost:~/hadoop

[root@ansible_controller hadoop]# ansible-playbook -v delete_play.yml
Using /etc/ansible/ansible.cfg as config file

PLAY [Deleting Datanode] *****
TASK [Gathering Facts] *****
ok: [192.168.99.131]

TASK [stop hadoop] *****
fatal: [192.168.99.131]: FAILED! => ("changed": false, "cmd": "hadoop-daemon.sh stop datanode", "msg": "[Errno 2] No such file or directory: b'hadoop-daemon.sh': b'hadoop-daemon.sh'", "rc": 2)
...ignoring

TASK [Uninstalling packages] *****
ok: [192.168.99.131] => ("changed": false, "msg": "Nothing to do", "rc": 0, "results": [])

TASK [Deleting Datanode Directory] *****
ok: [192.168.99.131] => (item=/hadoop/DN) => ("ansible_loop_var": "item", "changed": false, "item": "/hadoop/DN", "path": "/hadoop/DN", "state": "absent")
ok: [192.168.99.131] => (item=/etc/hadoop) => ("ansible_loop_var": "item", "changed": false, "item": "/etc/hadoop", "path": "/etc/hadoop", "state": "absent")

PLAY [Deleting Namenode] *****
TASK [Gathering Facts] *****
ok: [192.168.99.129]

TASK [stop hadoop] *****
fatal: [192.168.99.129]: FAILED! => ("changed": false, "cmd": "hadoop-daemon.sh stop namenode", "msg": "[Errno 2] No such file or directory: b'hadoop-daemon.sh': b'hadoop-daemon.sh'", "rc": 2)
...ignoring


TASK [Uninstalling packages] *****
ok: [192.168.99.129] => ("changed": false, "msg": "Nothing to do", "rc": 0, "results": [])

TASK [Deleting Namenode Directory] *****
changed: [192.168.99.129] => (item=/hadoop/NN) => ("ansible_loop_var": "item", "changed": true, "item": "/hadoop/NN", "path": "/hadoop/NN", "state": "absent")
changed: [192.168.99.129] => (item=/etc/hadoop) => ("ansible_loop_var": "item", "changed": true, "item": "/etc/hadoop", "path": "/etc/hadoop", "state": "absent")


PLAY RECAP *****
192.168.99.129      : ok=4    changed=1    unreachable=0    failed=0    skipped=0    rescued=0    ignored=1
192.168.99.131      : ok=4    changed=0    unreachable=0    failed=0    skipped=0    rescued=0    ignored=1

[root@ansible_controller hadoop]#
```

Step7: Checking the setup is deleted or not.

 root@NameNode:~

```
[root@NameNode ~]# jps  
4765 Jps  
[root@NameNode ~]# █
```

 root@DataNode1:~

```
[root@DataNode1 ~]# jps  
24494 Jps  
[root@DataNode1 ~]# █
```

Check the playbooks and files in GitHub.

GitHub Link:

https://github.com/Anuddeeph/Configure_Hadoop_using_Ansible.git