

Screenshots:

Instance created using lxc

Small instance

```
root@nfs: ~  
cc@hw8rohitl:~$ lxc launch ubuntu:16.04  
Creating the container  
Container name is: true-hookworm  
Starting true-hookworm  
cc@hw8rohitl:~$ sudo lxc launch ubuntu:16.04 nfs -c security.privileged=true -c raw.apparmor="mount fstype=rpc_pipefs, m  
ount fstype=nfsd, mount fstype=nfs,"  
Creating nfs  
Starting nfs  
cc@hw8rohitl:~$ sudo lxc exec nfs /bin/bash  
root@nfs:~#
```

```
root@nfs: ~  
ount fstype=nfsd, mount fstype=nfs,"  
Creating nfs  
Starting nfs  
cc@hw8rohitl:~$ sudo lxc exec nfs /bin/bash  
root@nfs:~# df  
Filesystem            1K-blocks    Used Available Use% Mounted on  
lxd/containers/nfs    50527104 313216  50213888    1% /  
none                   492         0        492    0% /dev  
udev                  65871424     0 65871424    0% /dev/fuse  
tmpfs                   100         0         100    0% /dev/lxd  
tmpfs                   100         0         100    0% /dev/.lxd-mounts  
tmpfs                  65892276     0 65892276    0% /dev/shm  
tmpfs                  65892276   8352 65883924    1% /run  
tmpfs                   5120         0        5120    0% /run/lock  
tmpfs                  65892276     0 65892276    0% /sys/fs/cgroup  
root@nfs:~#
```

```
cc@hw8rohitl: ~  
error: not found  
cc@hw8rohitl:~$ lxc config set nfs limits.cpu 2  
cc@hw8rohitl:~$ lxc config set nfs limits.cpu 4  
cc@hw8rohitl:~$ lxc config set nfs limits.memory 8000MB  
cc@hw8rohitl:~$
```

Large instance:

```
root@largeinstance: ~  
cc@hw8rohitl:~$ lxc launch ubuntu:16.04 largeinstance  
Creating largeinstance  
Starting largeinstance  
cc@hw8rohitl:~$ sudo lxc exec largeinstance /bin/bash  
root@largeinstance:~#
```

cc@hw8rohitl: ~

```
cc@hw8rohitl:~$ lxc config set largeinstance limits.cpu 16
cc@hw8rohitl:~$ lxc config set largeinstance limits.memory 32000MB
cc@hw8rohitl:~$
```

Hadoop Installation steps:

1. Java installation

cc@hw8rohitl: ~/HW8/jdk-11

```
javac 1.8.0_191
cc@hw8rohitl:~/HW8$ ls
jdk-11  openjdk-11_linux-x64_bin.tar.gz
cc@hw8rohitl:~/HW8$ cd jdk-11/
cc@hw8rohitl:~/HW8/jdk-11$
```

- 2.

cc@hw8rohitl: ~/HW8/jdk-11

```
cc@hw8rohitl:~/HW8/jdk-11$ sudo apt install openssh-server
Reading package lists... Done
Building dependency tree
Reading state information... Done
openssh-server is already the newest version (1:7.2p2-4ubuntu2.8).
0 upgraded, 0 newly installed, 0 to remove and 20 not upgraded.
cc@hw8rohitl:~/HW8/jdk-11$ sudo addgroup hadoop
Adding group `hadoop' (GID 1002) ...
Done.
cc@hw8rohitl:~/HW8/jdk-11$
```

```

cc@hw8rohitl: ~/HW8
cc@hw8rohitl:~/HW8$ ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/home/cc/.ssh/id_rsa):
/home/cc/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/cc/.ssh/id_rsa.
Your public key has been saved in /home/cc/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:0AMaRogPlaVGQPSZgfzoDQtkhsUBuxJV8Byc9LX8VjY cc@hw8rohitl
The key's randomart image is:
+---[RSA 2048]-----+
|BX0%O.. .|
|+X==*= = .|
|*O=++ O = E|
|.*.. . o o .|
|= + S O|
|.O . .|
|_|
+---[SHA256]-----+
cc@hw8rohitl:~/HW8$ ls
jdk-11 openjdk-11_linux-x64_bin.tar.gz
cc@hw8rohitl:~/HW8$ sudo cat /home/cc/.ssh/id_rsa.pub >> /home/cc/.ssh/authorized_keys
cc@hw8rohitl:~/HW8$

```

```

cc@hw8rohitl: ~/HW8/hadoop-3.2.0/etc/hadoop
hadoop-env.sh          httpfs-site.xml        mapred-queues.xml.template  yarn-env.sh
hadoop-metrics2.properties  kms-acls.xml          mapred-site.xml            yarnservice-log4j.properties
hadoop-policy.xml       kms-env.sh            shellprofile.d             yarn-site.xml
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ cat core-site.xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the license at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the license for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

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

<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>
</configuration>

```

cc@hw8rohitl: ~/HW8/hadoop-3.2.0/etc/hadoop

```
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ vi hdfs-site.xml
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ cat hdfs-site.xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

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

<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
<property>
<name>dfs.permission</name>
<value>>false</value>
</property>
```

cc@hw8rohitl: ~/HW8/hadoop-3.2.0/etc/hadoop

```
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ vi yarn-site.xml
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ cat yarn-site.xml
<?xml version="1.0"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<configuration>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
</configuration>
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$
```

cc@hw8rohitl: ~/HW8/hadoop-3.2.0

```
cc@hw8rohitl:~/HW8/hadoop-3.2.0$ hadoop jar Hadoopsort.jar HadoopSort data60gb4smallinstance.in output
Time taken to sort 1126345
19/04/28 15:07:00 INFO client.RMPProxy: Connecting to ResourceManager at hadoop-f/192.168.2.30:8031
19/04/28 15:07:00 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed.
ol interface and execute your application with ToolRunner to remedy this.
19/04/28 15:07:00 INFO input.FileInputFormat: Total input files to process : 1
19/04/28 15:07:00 INFO mapreduce.JobSubmitter: number of splits:118
19/04/28 15:07:01 INFO Configuration.deprecation: yarn.resourcemanager.system-metrics-publisher.enabled :
nstead, use yarn.system-metrics-publisher.enabled
19/04/28 15:07:01 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1524521941871_0401
19/04/28 15:07:01 INFO impl.YarnClientImpl: Submitted application application_1524521941871_0401
19/04/28 15:07:01 INFO mapreduce.Job: The url to track the job: http://hadoop-f:8088/proxy/application_15
3/
19/04/28 15:07:01 INFO mapreduce.Job: Running job: job_1524521941871_0409
19/04/28 15:07:09 INFO mapreduce.Job: Job job_1524521941871_0404 running in uber mode : false
19/04/28 15:07:09 INFO mapreduce.Job: map 1% reduce 0%
19/04/28 15:07:24 INFO mapreduce.Job: map 4% reduce 0%
19/04/28 15:07:26 INFO mapreduce.Job: map 5% reduce 0%
19/04/28 15:07:28 INFO mapreduce.Job: map 8% reduce 0%
19/04/28 15:07:31 INFO mapreduce.Job: map 9% reduce 0%
19/04/28 15:07:33 INFO mapreduce.Job: map 10% reduce 0%
19/04/28 15:07:34 INFO mapreduce.Job: map 17% reduce 0%
19/04/28 15:07:35 INFO mapreduce.Job: map 19% reduce 0%
```

Spark Installation:

Scala installation

cc@hw8rohitl: ~/HW8/hadoop-3.2.0/etc/hadoop

```
fdf: command not found
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ sudo apt-get install scala
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following packages were automatically installed and are no longer required:
  libevdev2 libgsoap8 libgudev-1.0-0 libice6 libinput-bin libinput10 libmtdev1
  libqt5network5 libqt5opengl5 libqt5printsupport5 libqt5widgets5 libqt5x11extra
  libwacom-common libwacom2 libxcb-icccm4 libxcb-image0 libxcb-keysyms1 libxcb-
  libxcb-shape0 libxcb-util1 libxcb-xkb1 libxkbcommon-x11-0 libxmu6 libxt6
Use 'sudo apt autoremove' to remove them.
The following additional packages will be installed:
  libhawtjni-runtime-java libjansi-java libjansi-native-java libjline2-java scala
  scala-xml
Suggested packages:
  scala-doc
The following NEW packages will be installed:
  libhawtjni-runtime-java libjansi-java libjansi-native-java libjline2-java scala
  scala-parser-combinators scala-xml
0 upgraded, 8 newly installed, 0 to remove and 9 not upgraded.
Need to get 23.6 MB of archives.
After this operation, 28.1 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://nova.clouds.archive.ubuntu.com/ubuntu xenial/universe amd64 libhaw
Get:2 http://nova.clouds.archive.ubuntu.com/ubuntu xenial/universe amd64 libjans
Get:3 http://nova.clouds.archive.ubuntu.com/ubuntu xenial/universe amd64 libjans
```

cc@hw8rohitl: ~/HW8/hadoop-3.2.0/etc/hadoop

```
Unpacking scala-library (2.11.6-6) ...
Selecting previously unselected package scala-parser-combinators.
Preparing to unpack .../scala-parser-combinators_1.0.3-3_all.deb ...
Unpacking scala-parser-combinators (1.0.3-3) ...
Selecting previously unselected package scala-xml.
Preparing to unpack .../scala-xml_1.0.3-3_all.deb ...
Unpacking scala-xml (1.0.3-3) ...
Selecting previously unselected package scala.
Preparing to unpack .../scala_2.11.6-6_all.deb ...
Unpacking scala (2.11.6-6) ...
Setting up libhawtjni-runtime-java (1.10-1) ...
Setting up libjansi-native-java (1.0-4) ...
Setting up libjansi-java (1.4-3) ...
Setting up libjline2-java (2.11-4) ...
Setting up scala-library (2.11.6-6) ...
Setting up scala-parser-combinators (1.0.3-3) ...
Setting up scala-xml (1.0.3-3) ...
Setting up scala (2.11.6-6) ...
update-alternatives: using /usr/share/scala-2.11/bin/scala to provide /usr/bin/scala (s
cc@hw8rohitl:~/HW8/hadoop-3.2.0/etc/hadoop$ ls
capacity-scheduler.xml      hadoop-user-functions.sh.example  kms-log4j.properties
configuration.xml           hdfs-site.xml                    kms-site.xml
container-executor.cfg      httpfs-env.sh                    log4j.properties
core-site.xml               httpfs-log4j.properties          mapred-env.cmd
hadoop-env.cmd              httpfs-signature.secret          mapred-env.sh
hadoop-env.sh               httpfs-site.xml                  mapred-queues.xml.template
hadoop-metrics2.properties kms-acls.xml                      mapred-site.xml
hadoop-policy.xml           kms-env.sh                        shellprofile.d
```

```
cc@hw8rohitl:~$ cd HW8/
cc@hw8rohitl:~/HW8$ cd hadoop-3.2.0/
cc@hw8rohitl:~/HW8/hadoop-3.2.0$ wget "https://www.apache.org/dyn/closer.lua/spark/spark-2.4.2/spark-2.4.2-bin-h
.tgz"
--2019-05-04 04:47:44-- https://www.apache.org/dyn/closer.lua/spark/spark-2.4.2/spark-2.4.2-bin-hadoop2.7.tgz
Resolving www.apache.org (www.apache.org)... 40.79.78.1, 95.216.24.32, 2a01:4f9:2a:185f::2
Connecting to www.apache.org (www.apache.org)|40.79.78.1|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: unspecified [text/html]
Saving to: 'spark-2.4.2-bin-hadoop2.7.tgz'

spark-2.4.2-bin-hadoop2.7.tgz      [ <=> ] 35.70K --.-KB/s in 0
2019-05-04 04:47:44 (1.44 MB/s) - 'spark-2.4.2-bin-hadoop2.7.tgz' saved [36560]

cc@hw8rohitl:~/HW8/hadoop-3.2.0$
```