

MCIS 6263 – Big Data Assignment 1 : Hadoop

Name : Saranya Balasubramaniyan

Student ID : 999901316

Contents

Oracle VirtualBox	3
Advantages of VirtualBox.....	3
Assignment.....	3
Connecting to the Virtual Machine using Oracle VirtualBox	3
Terminal in Virtual Machine – AntiX Hadoop Clone1	4
Switching to user: Hadoop.....	4
Find and Open: hdfs-site.xml	5
Find and Open: core-site.xml.....	10
Find and Open: mapred-site.xml	13
Find and Open: yarn-site.xml.....	18
References	25

Oracle VirtualBox

Oracle VM VirtualBox is a free and open-source hosted hypervisor for x86 virtualization, developed by Oracle Corporation

VirtualBox can be used to load multiple guest OSes under a single host operating system(host OS). Each guest can be started, paused and stopped independently within its own virtual machine(VM). The host OS and guest OSs and applications can communicate with each other through a number of mechanisms including a common clipboard and a virtualized network facility.

Advantages of VirtualBox

- Running multiple operating systems simultaneously
- Easier software installations
- Testing and disaster recovery
- Infrastructure consolidation

Assignment

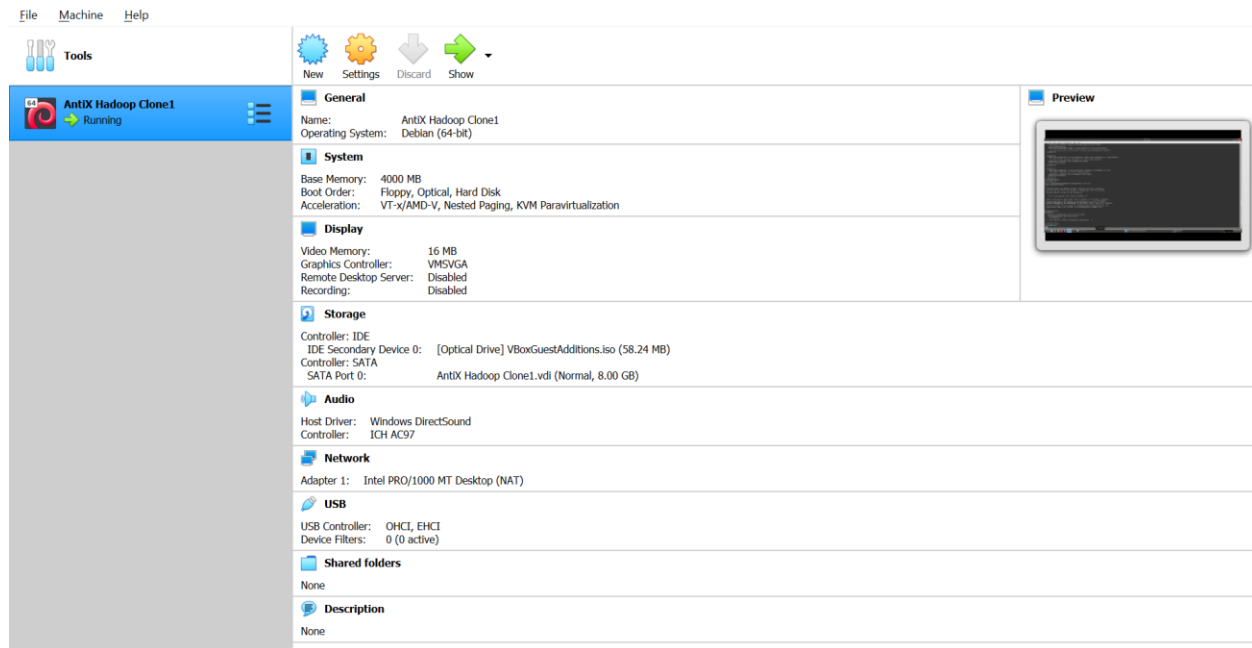
Connecting to the Virtual Machine using Oracle VirtualBox

Name of the Virtual Machine: AntiX Hadoop Clone1

Operating system: Debian

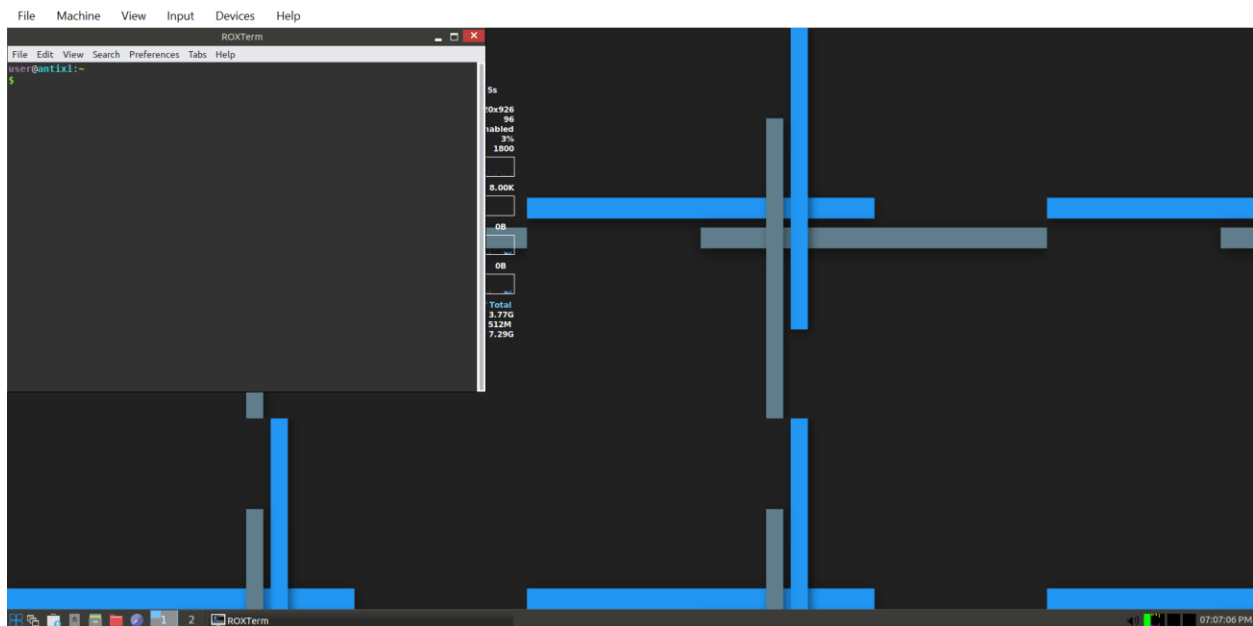
Base Memory: 4000MB

The Oracle Virtual Box showcasing the Virtual machine used for the assignment is shown below,



Terminal in Virtual Machine – AntiX Hadoop Clone1

Below image is a snapshot of the Virtual Machine – AntiX Hadoop Clone1 where the Terminal is opened.



Switching to user: Hadoop

The below image showcases the switching of user account to user: hadoop

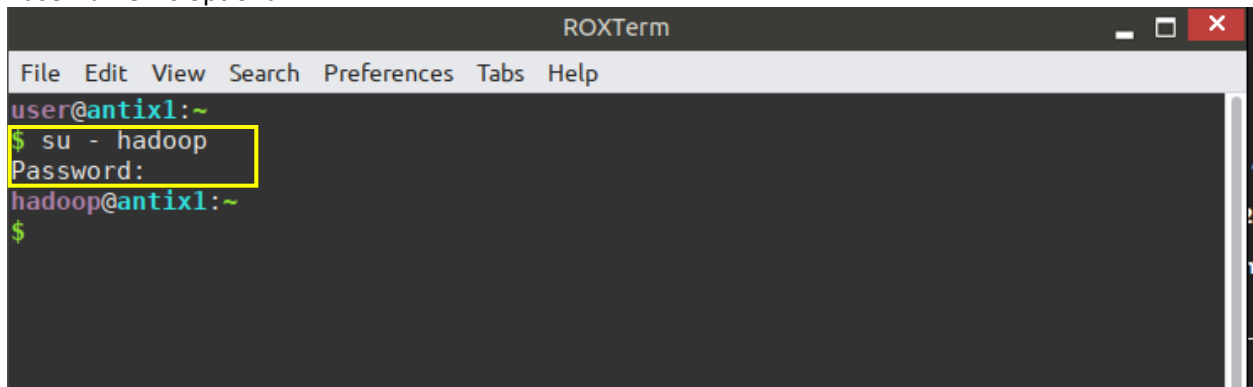
Command used: **su – hadoop**

Password would be requested

Command su : The su command is used to switch to another user, in other words change user ID during a normal login session.

Syntax: \$ su - <username>

<username> is optional



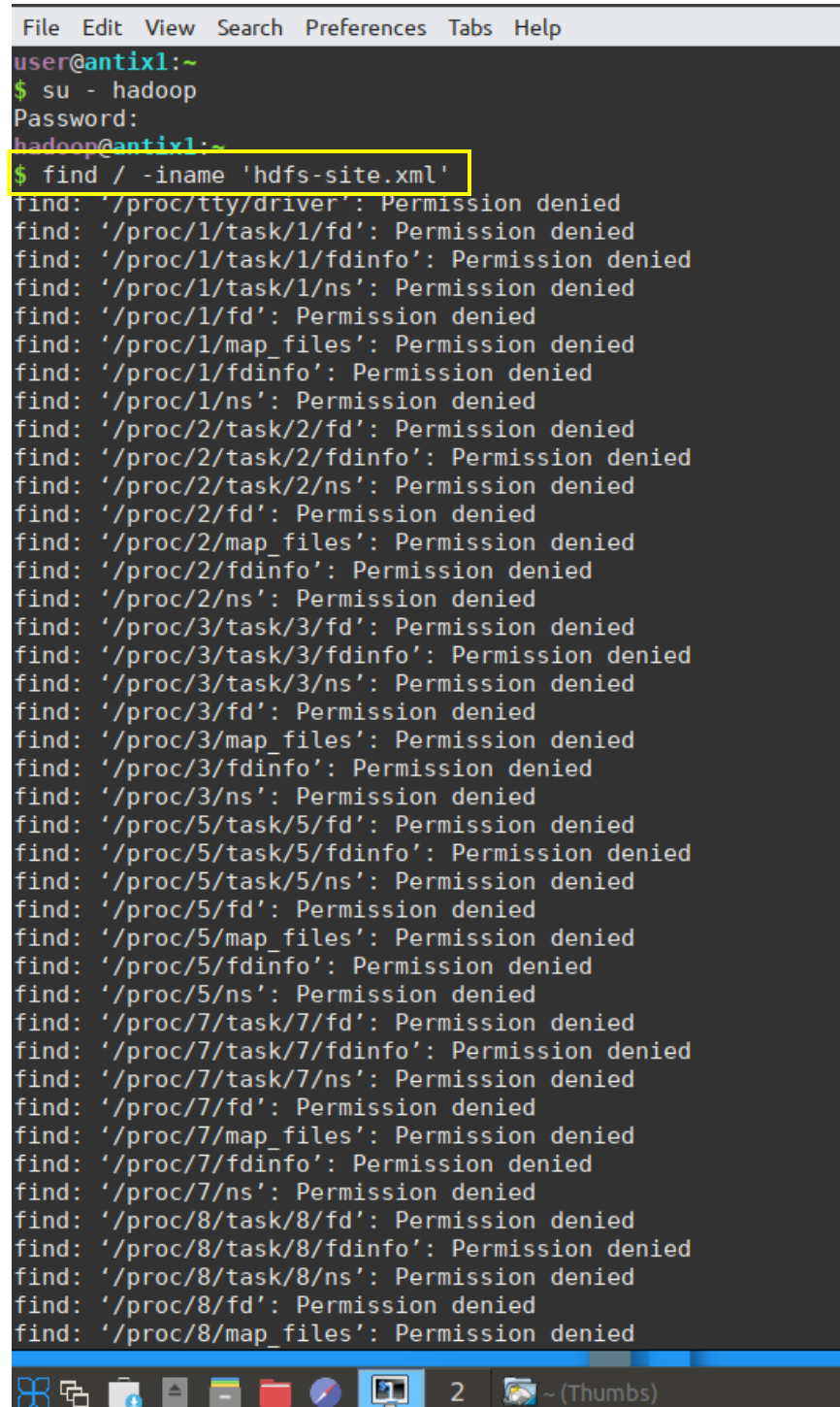
Find and Open: hdfs-site.xml

To find a file in a Linux terminal, find command can be used to locate a file.

Syntax: \$ find [where to start searching from] [-options] [what to find]

Command used: find / -iname 'hdfs-site.xml'

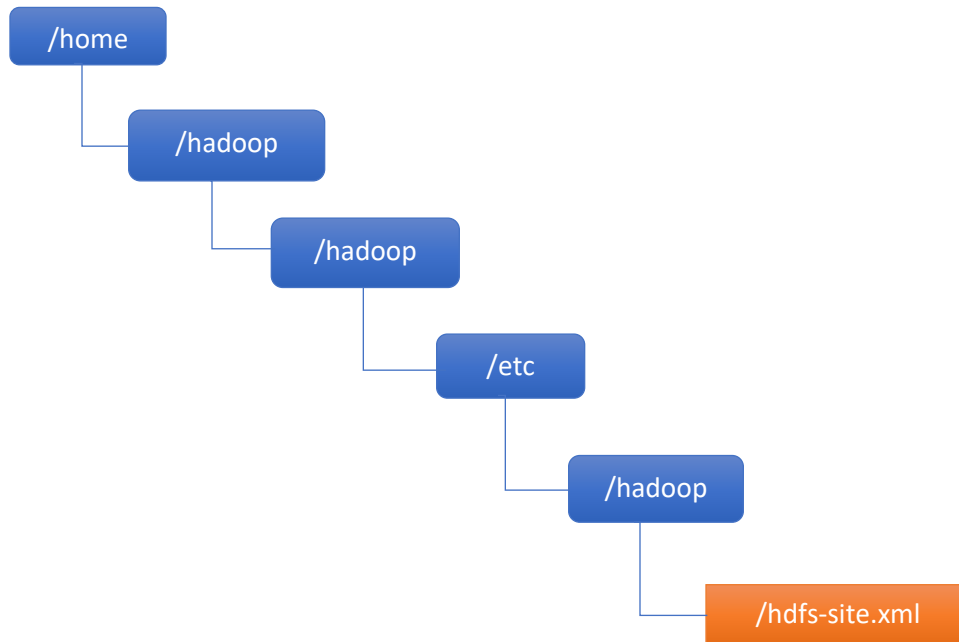
Option iname is used to specify that the name of the file is provided in the “what to find” section



The screenshot shows a terminal window with a menu bar (File, Edit, View, Search, Preferences, Tabs, Help) and a title bar (user@antix1:~). The user is logged in as 'user@antix1:~'. They run the command '\$ su - hadoop' and enter the password. The prompt changes to 'hadoop@antix1:~'. The user then runs the command '\$ find / -iname 'hdfs-site.xml'', which is highlighted with a yellow box. The output of the command is a long list of 'Permission denied' messages for various files and directories, including /proc/tty/driver, /proc/1/task/1/fd, /proc/1/task/1/fdinfo, /proc/1/task/1/ns, /proc/1/fd, /proc/1/map_files, /proc/1/fdinfo, /proc/1/ns, /proc/2/task/2/fd, /proc/2/task/2/fdinfo, /proc/2/task/2/ns, /proc/2/fd, /proc/2/map_files, /proc/2/fdinfo, /proc/2/ns, /proc/3/task/3/fd, /proc/3/task/3/fdinfo, /proc/3/task/3/ns, /proc/3/fd, /proc/3/map_files, /proc/3/fdinfo, /proc/3/ns, /proc/5/task/5/fd, /proc/5/task/5/fdinfo, /proc/5/task/5/ns, /proc/5/fd, /proc/5/map_files, /proc/5/fdinfo, /proc/5/ns, /proc/7/task/7/fd, /proc/7/task/7/fdinfo, /proc/7/task/7/ns, /proc/7/fd, /proc/7/map_files, /proc/7/fdinfo, /proc/7/ns, /proc/8/task/8/fd, /proc/8/task/8/fdinfo, /proc/8/task/8/ns, /proc/8/fd, and /proc/8/map_files.

The hdfs-site.xml is found and the path where the file is located is displayed as a part of the result of the find command.

The file is found in the location,



The file location is highlighted in the below image.

File Machine View Input Devices Help

```
File Edit View Search Preferences Tabs Help
find: '/proc/5259/fdinfo': Permission denied
find: '/proc/5259/ns': Permission denied
find: '/proc/20423/task/20423/fd': Permission denied
find: '/proc/20423/task/20423/fdinfo': Permission denied
find: '/proc/20423/task/20423/ns': Permission denied
find: '/proc/20423/fd': Permission denied
find: '/proc/20423/map_files': Permission denied
find: '/proc/20423/fdinfo': Permission denied
find: '/proc/20423/ns': Permission denied
find: '/tmp/libgksu-a3EXdy': Permission denied
find: '/tmp/libgksu-WDvJYp': Permission denied
find: '/lost+found': Permission denied
find: '/root': Permission denied
/home/hadoop/hadoop/etc/hadoop/hdfs-site.xml
find: '/home/user/.config/dconf': Permission denied
find: '/home/user/.config/geany/filedefs': Permission denied
find: '/home/user/.config/geany/templates': Permission denied
find: '/home/user/.config/geany/tags': Permission denied
find: '/home/user/.config/gtk-2.0': Permission denied
find: '/home/user/.run': Permission denied
find: '/home/user/.local/share/nano': Permission denied
find: '/home/user/.fltk': Permission denied
find: '/home/user/.mozilla': Permission denied
find: '/home/user/.dbus': Permission denied
find: '/home/user/.thumbnails': Permission denied
find: '/home/user/.cache': Permission denied
find: '/run/user/1000': Permission denied
find: '/run/sudo': Permission denied
find: '/run/systemd/inaccessible/dir': Permission denied
find: '/run/cups/certs': Permission denied
find: '/var/spool/cron/crontabs': Permission denied
find: '/var/spool/cups': Permission denied
find: '/var/spool/rsyslog': Permission denied
find: '/var/cache/apt/archives/partial': Permission denied
find: '/var/cache/ldconfig': Permission denied
find: '/var/lib/apt/lists/partial': Permission denied
find: '/var/lib/polkit-1': Permission denied
find: '/var/lib/connman/ethernet_080027c95d15_cable': Permission denied
find: '/var/lib/connman/ethernet_080027d5ee35_cable': Permission denied
find: '/var/lib/connman/ethernet_080027e586bd_cable': Permission denied
find: '/var/log/unattended-upgrades': Permission denied
find: '/etc/ssl/private': Permission denied
find: '/etc/cups/ssl': Permission denied
find: '/etc/polkit-1/localauthority': Permission denied
hadoop@antix1:~
$
```

2 ~ (Thumbs)

To open a text file cat command can be used,

Syntax: \$ cat <file_name>

Command used: cat /home/hadoop/hadoop/etc/hadoop/hdfs-site.xml

This displays the contents of the hdfs-site.xml file in the Terminal.

hdfs-site.xml: The hdfs-site.xml file contains the configuration settings for HDFS daemons, the NameNode, the Secondary NameNode and the DataNodes to specify default block replication and permission checking on HDFS. The actual number of replications can also be specified when the file is created.

In the file are the below configurations,

Name: dfs.replication, Value : 1

Default block replication. The actual number of replications can be specified when the file is created. The default is used if replication is not specified in create time.

Name: dfs.namenode.name.dir, Value: /home/Hadoop/hadoopdata/hdfs/namenode

Determines where on the local filesystem the DFS name node should store the name table(fsimage). If this is a comma-delimited list of directories then the name table is replicated in all of the directories, for redundancy.

Name: dfs.datanode.data.dir, Value: /home/Hadoop/hadoopdata/hdfs/datanode

Determines where on the local filesystem an DFS data node should store its blocks. If this is a comma-delimited list of directories, then data will be stored in all named directories, typically on different devices. Directories that do not exist are ignored.

File Edit View Search Preferences Tabs Help

```
find: '/var/lib/polkit-1': Permission denied
find: '/var/lib/connman/ethernet_080027c95d15_cable': Permission denied
find: '/var/lib/connman/ethernet_080027d5ee35_cable': Permission denied
find: '/var/lib/connman/ethernet_080027e586bd_cable': Permission denied
find: '/var/log/unattended-upgrades': Permission denied
find: '/etc/ssl/private': Permission denied
find: '/etc/cups/ssl': Permission denied
find: '/etc/polkit-1/localauthority': Permission denied
```

hadoop@antix1:~

```
$ cat /home/hadoop/hadoop/etc/hadoop/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.namenode.name.dir</name>
```

```
<value>/home/hadoop/hadoopdata/hdfs/namenode</value>
```

```
<final>true</final>
```

```
</property>
```

```
<property>
```

```
<name>dfs.datanode.data.dir</name>
```

```
<value>/home/hadoop/hadoopdata/hdfs/datanode</value>
```

```
<final>true</final>
```

```
</property>
```

```
</configuration>
```

hadoop@antix1:~

```
$
```

Find and Open: core-site.xml

To find a file in a Linux terminal, find command can be used to locate a file.

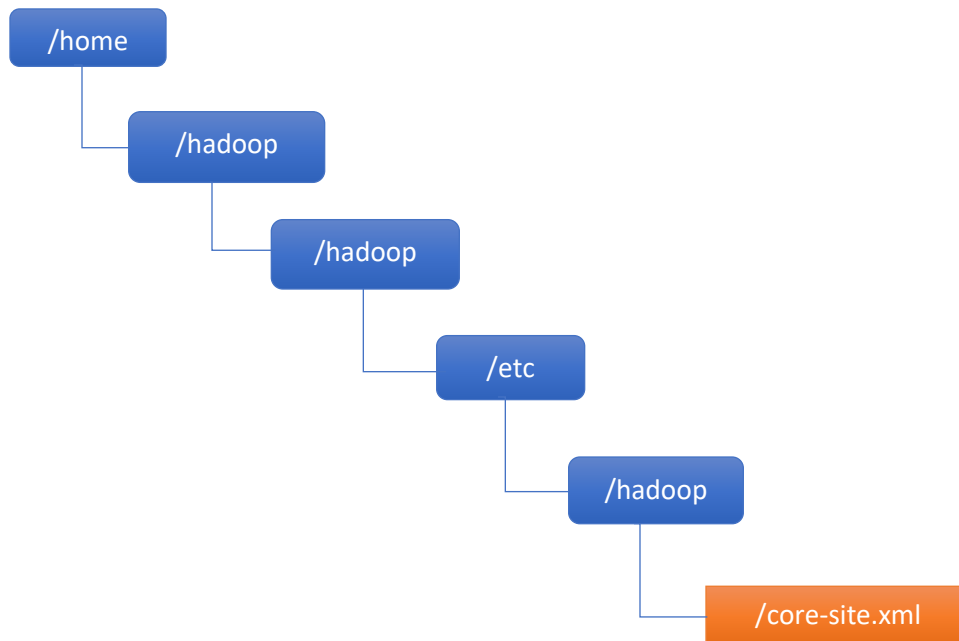
Syntax: \$ find [where to start searching from] [-options] [what to find]

Command used: find / -iname 'core-site.xml'

Option iname is used to specify that the name of the file is provided in the “what to find” section

The core-site.xml is found and the path where the file is located is displayed as a part of the result of the find command.

The file is found in the location,



```
File Edit View Search Preferences Tabs Help
<final>true</final>
</property>
</configuration>
hadoop@antix1:~
$ find / -iname 'core-site.xml'
find: '/proc/tty/driver': Permission denied
find: '/proc/1/task/1/fd': Permission denied
find: '/proc/1/task/1/fdinfo': Permission denied
find: '/proc/1/task/1/ns': Permission denied
find: '/proc/1/fd': Permission denied
find: '/proc/1/map_files': Permission denied
find: '/proc/1/fdinfo': Permission denied
find: '/proc/1/ns': Permission denied
find: '/proc/2/task/2/fd': Permission denied
find: '/proc/2/task/2/fdinfo': Permission denied
find: '/proc/2/task/2/ns': Permission denied
find: '/proc/2/fd': Permission denied
find: '/proc/2/map_files': Permission denied
find: '/proc/2/fdinfo': Permission denied
find: '/proc/2/ns': Permission denied
find: '/proc/3/task/3/fd': Permission denied
find: '/proc/3/task/3/fdinfo': Permission denied
find: '/proc/3/task/3/ns': Permission denied
find: '/proc/3/fd': Permission denied
find: '/proc/3/map_files': Permission denied
find: '/proc/3/fdinfo': Permission denied
find: '/proc/3/ns': Permission denied
find: '/proc/5/task/5/fd': Permission denied
find: '/proc/5/task/5/fdinfo': Permission denied
find: '/proc/5/task/5/ns': Permission denied
find: '/proc/5/fd': Permission denied
find: '/proc/5/map_files': Permission denied
find: '/proc/5/fdinfo': Permission denied
find: '/proc/5/ns': Permission denied
find: '/proc/7/task/7/fd': Permission denied
find: '/proc/7/task/7/fdinfo': Permission denied
find: '/proc/7/task/7/ns': Permission denied
find: '/proc/7/fd': Permission denied
find: '/proc/7/map_files': Permission denied
find: '/proc/7/fdinfo': Permission denied
find: '/proc/7/ns': Permission denied
find: '/proc/8/task/8/fd': Permission denied
find: '/proc/8/task/8/fdinfo': Permission denied
find: '/proc/8/task/8/ns': Permission denied
find: '/proc/8/fd': Permission denied
find: '/proc/8/map_files': Permission denied
```

To open a text file cat command can be used,

Syntax: \$ cat <file_name>

Command used: cat /home/hadoop/hadoop/etc/hadoop/core-site.xml

This displays the contents of the core-site.xml file in the Terminal.

core-site.xml

The core-site.xml file informs Hadoop daemon where NameNode runs in the cluster. It contains the configuration settings for Hadoop Core such as I/O settings that are common to HDFS and MapReduce

Name: fs.default.name, value: hdfs://localhost:9000

The name of the default file system. A URI whose scheme and authority determine the FileSystem implementation. The uri's scheme determines the config property (fs.SCHEME.impl) naming the FileSystem implementation class. The uri's authority is used to determine the host, port, etc. for a filesystem.

```
hadoop@antix1:~  
$ cat /home/hadoop/hadoop/etc/hadoop/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>  
hadoop@antix1:~  
$
```



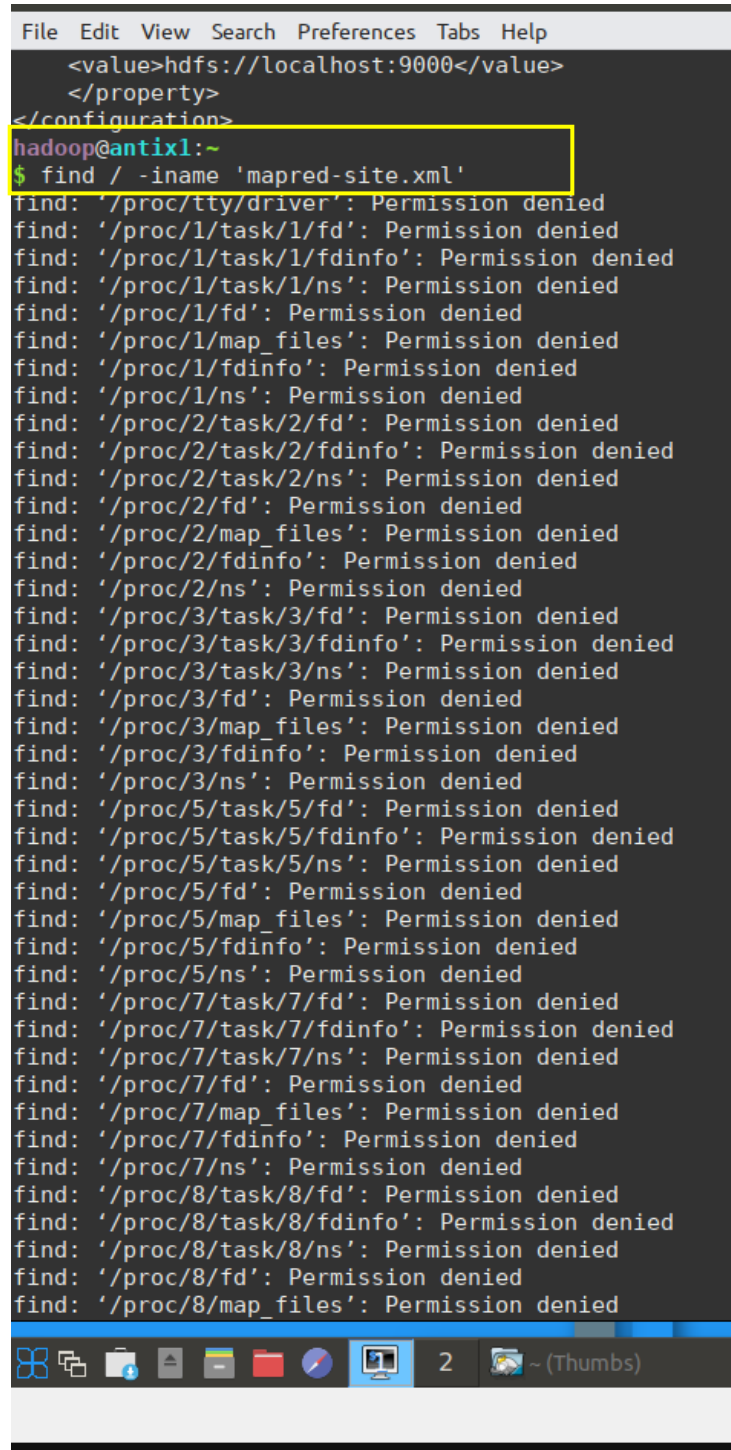
Find and Open: mapred-site.xml

To find a file in a Linux terminal, find command can be used to locate a file.

Syntax: \$ find [where to start searching from] [-options] [what to find]

Command used: find / -iname 'mapred-site.xml'

Option iname is used to specify that the name of the file is provided in the “what to find” section



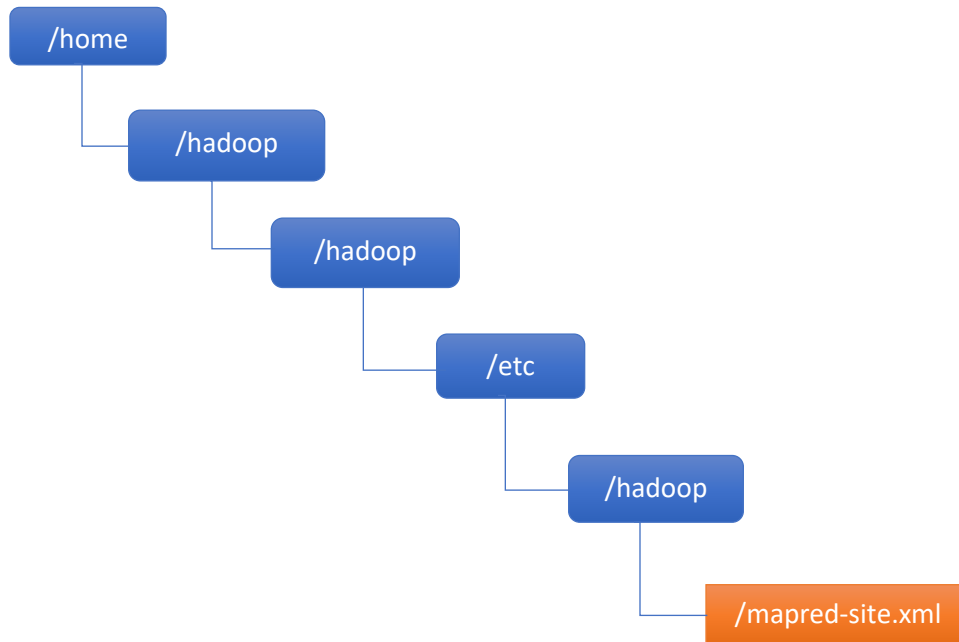
The screenshot shows a terminal window with a menu bar (File, Edit, View, Search, Preferences, Tabs, Help) and a toolbar. The terminal content is as follows:

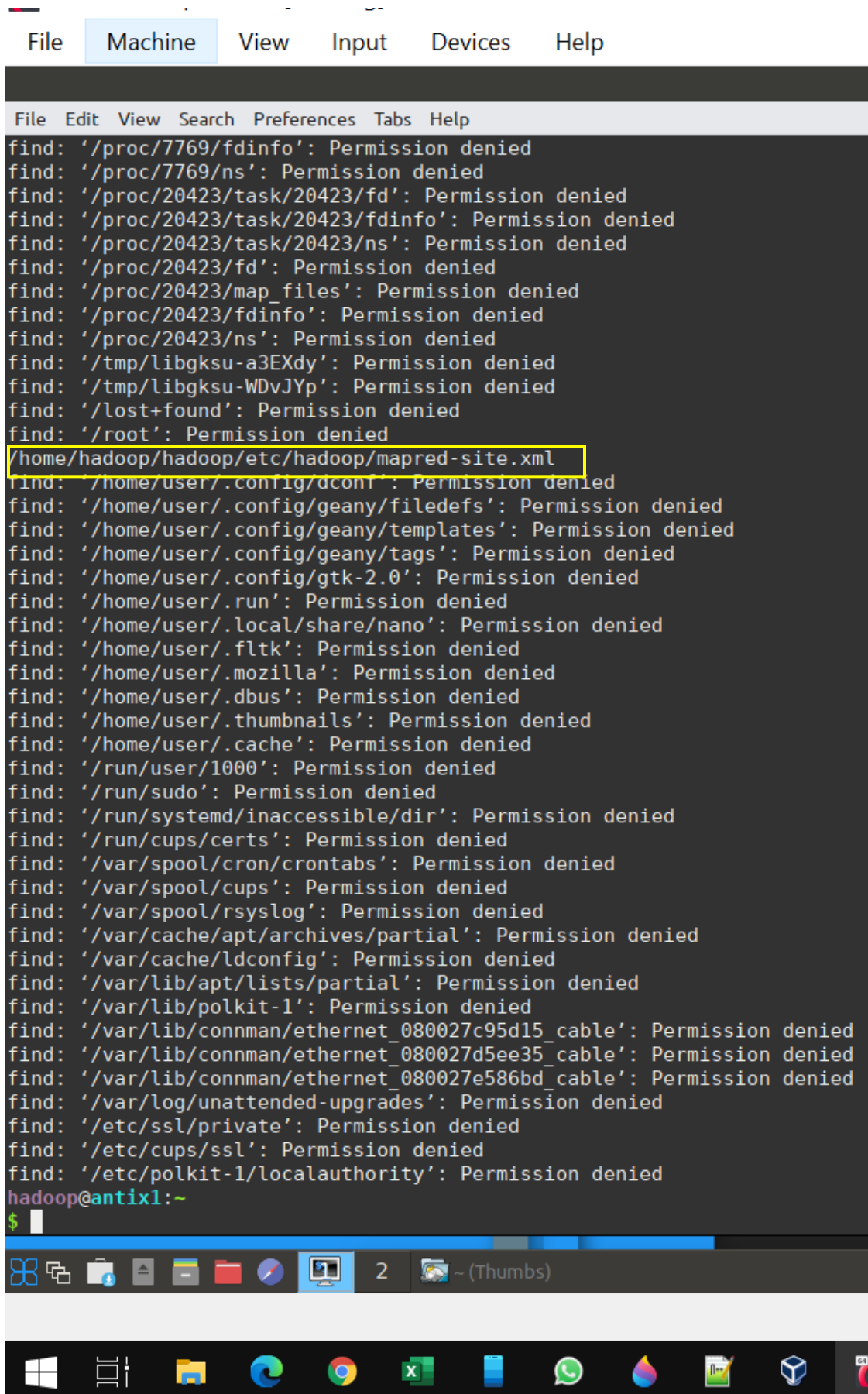
```
<value>hdfs://localhost:9000</value>
</property>
</configuration>
hadoop@antix1:~
$ find / -iname 'mapred-site.xml'
```

The output of the command is a long list of paths, each followed by "Permission denied". The paths are organized by process ID (proc/1, proc/2, proc/3, proc/5, proc/7, proc/8) and task ID (task/1, task/2, task/3, task/5, task/7, task/8). The paths include /proc/tty/driver, /proc/1/task/1/fd, /proc/1/task/1/fdinfo, /proc/1/task/1/ns, /proc/1/fd, /proc/1/map_files, /proc/1/fdinfo, /proc/1/ns, /proc/2/task/2/fd, /proc/2/task/2/fdinfo, /proc/2/task/2/ns, /proc/2/fd, /proc/2/map_files, /proc/2/fdinfo, /proc/2/ns, /proc/3/task/3/fd, /proc/3/task/3/fdinfo, /proc/3/task/3/ns, /proc/3/fd, /proc/3/map_files, /proc/3/fdinfo, /proc/3/ns, /proc/5/task/5/fd, /proc/5/task/5/fdinfo, /proc/5/task/5/ns, /proc/5/fd, /proc/5/map_files, /proc/5/fdinfo, /proc/5/ns, /proc/7/task/7/fd, /proc/7/task/7/fdinfo, /proc/7/task/7/ns, /proc/7/fd, /proc/7/map_files, /proc/7/fdinfo, /proc/7/ns, /proc/8/task/8/fd, /proc/8/task/8/fdinfo, /proc/8/task/8/ns, /proc/8/fd, and /proc/8/map_files.

The mapred-site.xml is found and the path where the file is located is displayed as a part of the result of the find command.

The file is found in the location,





To open a text file cat command can be used,

Syntax: \$ cat <file_name>

Command used: cat /home/hadoop/hadoop/etc/hadoop/mapred-site.xml

This displays the contents of the mapred-site.xml file in the Terminal.

mapred-site.xml

The mapred-site.xml file contains the configuration settings for MapReduce daemons; the job tracker and the task-trackers.

The properties defined in the file found are,

Name: **mapreduce.framework.name**, **Value:** **yarn**

The runtime framework for executing MapReduce jobs. Can be one of local, classic or yarn.


```
File Edit View Search Preferences Tabs Help
find: '/run/user/1000': Permission denied
find: '/run/sudo': Permission denied
find: '/run/systemd/inaccessible/dir': Permission denied
find: '/run/cups/certs': Permission denied
find: '/var/spool/cron/crontabs': Permission denied
find: '/var/spool/cups': Permission denied
find: '/var/spool/rsyslog': Permission denied
find: '/var/cache/apt/archives/partial': Permission denied
find: '/var/cache/ldconfig': Permission denied
find: '/var/lib/apt/lists/partial': Permission denied
find: '/var/lib/polkit-1': Permission denied
find: '/var/lib/connman/ethernet_080027c95d15_cable': Permission denied
find: '/var/lib/connman/ethernet_080027d5ee35_cable': Permission denied
find: '/var/lib/connman/ethernet_080027e586bd_cable': Permission denied
find: '/var/log/unattended-upgrades': Permission denied
find: '/etc/ssl/private': Permission denied
find: '/etc/cups/ssl': Permission denied
find: '/etc/polkit-1/localauthority': Permission denied
hadoop@antix1:~
$ cat /home/hadoop/hadoop/etc/hadoop/mapred-site.xml
<?xml version="1.0"?>
<?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>mapreduce.framework.name</name>
  <value>yarn</value>
  </property>
</configuration>
hadoop@antix1:~
$
```

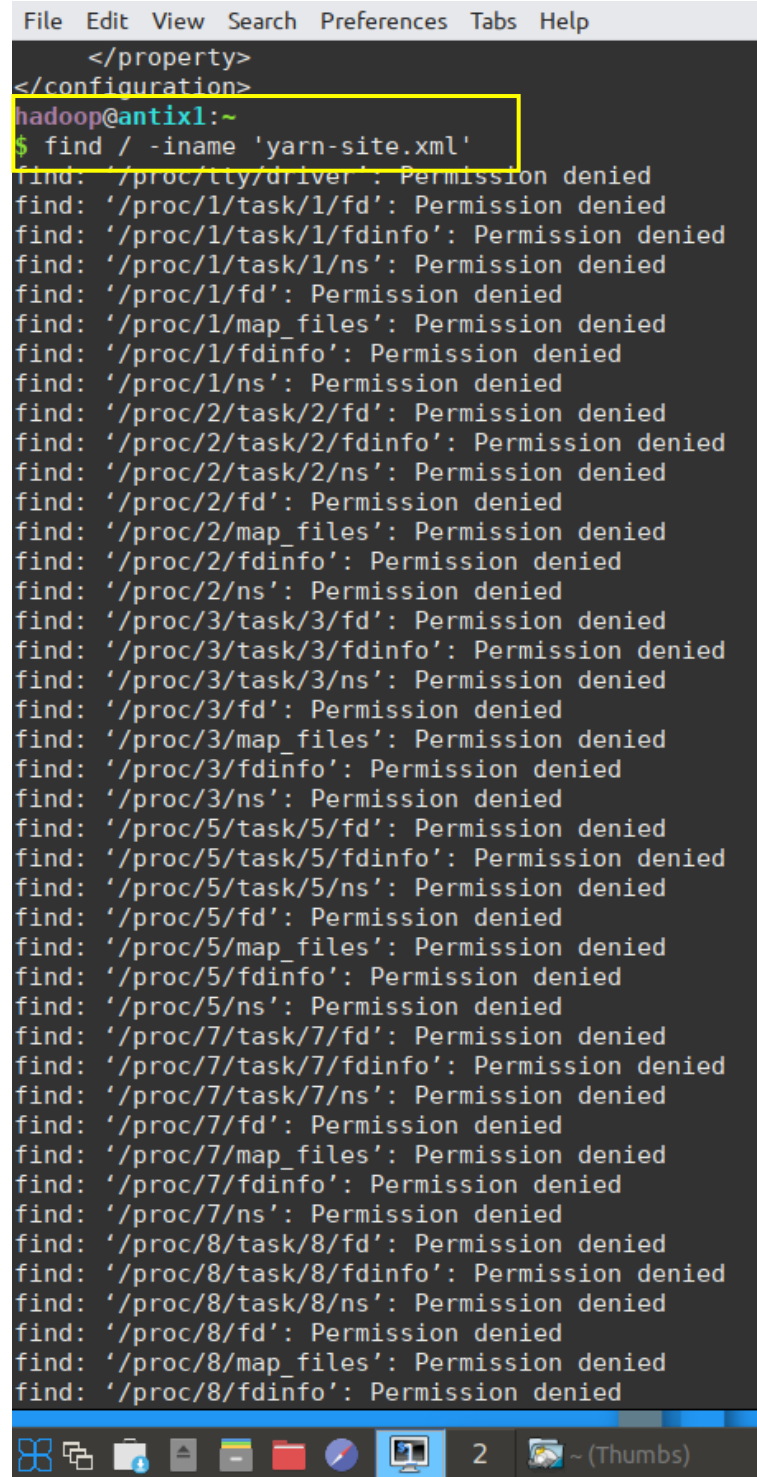
Find and Open: yarn-site.xml

To find a file in a Linux terminal, find command can be used to locate a file.

Syntax: \$ find [where to start searching from] [-options] [what to find]

Command used: find / -iname 'yarn-site.xml'

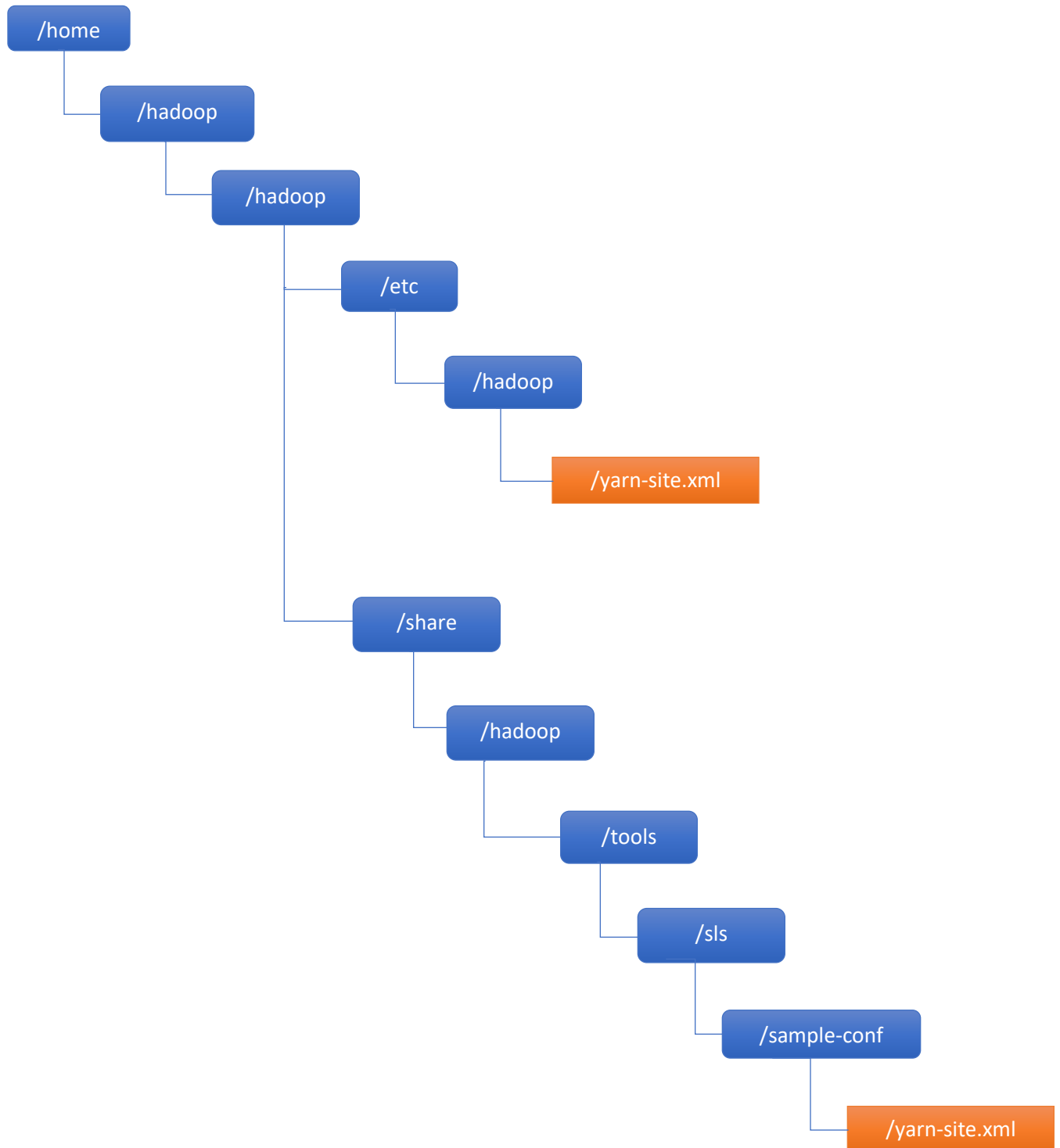
Option iname is used to specify that the name of the file is provided in the “what to find” section



```
File Edit View Search Preferences Tabs Help
</property>
</configuration>
hadoop@antix1:~
$ find / -iname 'yarn-site.xml'
find: '/proc/tty/driver': Permission denied
find: '/proc/1/task/1/fd': Permission denied
find: '/proc/1/task/1/fdinfo': Permission denied
find: '/proc/1/task/1/ns': Permission denied
find: '/proc/1/fd': Permission denied
find: '/proc/1/map_files': Permission denied
find: '/proc/1/fdinfo': Permission denied
find: '/proc/1/ns': Permission denied
find: '/proc/2/task/2/fd': Permission denied
find: '/proc/2/task/2/fdinfo': Permission denied
find: '/proc/2/task/2/ns': Permission denied
find: '/proc/2/fd': Permission denied
find: '/proc/2/map_files': Permission denied
find: '/proc/2/fdinfo': Permission denied
find: '/proc/2/ns': Permission denied
find: '/proc/3/task/3/fd': Permission denied
find: '/proc/3/task/3/fdinfo': Permission denied
find: '/proc/3/task/3/ns': Permission denied
find: '/proc/3/fd': Permission denied
find: '/proc/3/map_files': Permission denied
find: '/proc/3/fdinfo': Permission denied
find: '/proc/3/ns': Permission denied
find: '/proc/5/task/5/fd': Permission denied
find: '/proc/5/task/5/fdinfo': Permission denied
find: '/proc/5/task/5/ns': Permission denied
find: '/proc/5/fd': Permission denied
find: '/proc/5/map_files': Permission denied
find: '/proc/5/fdinfo': Permission denied
find: '/proc/5/ns': Permission denied
find: '/proc/7/task/7/fd': Permission denied
find: '/proc/7/task/7/fdinfo': Permission denied
find: '/proc/7/task/7/ns': Permission denied
find: '/proc/7/fd': Permission denied
find: '/proc/7/map_files': Permission denied
find: '/proc/7/fdinfo': Permission denied
find: '/proc/7/ns': Permission denied
find: '/proc/8/task/8/fd': Permission denied
find: '/proc/8/task/8/fdinfo': Permission denied
find: '/proc/8/task/8/ns': Permission denied
find: '/proc/8/fd': Permission denied
find: '/proc/8/map_files': Permission denied
find: '/proc/8/fdinfo': Permission denied
```

The yarn-site.xml is found in 2 locations and the path where the file is located is displayed as a part of the result of the find command.

The file is found in the location,



```
File Edit View Search Preferences Tabs Help
find: '/proc/5259/ns': Permission denied
find: '/proc/7769/task/7769/fd': Permission denied
find: '/proc/7769/task/7769/fdinfo': Permission denied
find: '/proc/7769/task/7769/ns': Permission denied
find: '/proc/7769/fd': Permission denied
find: '/proc/7769/map_files': Permission denied
find: '/proc/7769/fdinfo': Permission denied
find: '/proc/7769/ns': Permission denied
find: '/proc/8467/task/8467/fd': Permission denied
find: '/proc/8467/task/8467/fdinfo': Permission denied
find: '/proc/8467/task/8467/ns': Permission denied
find: '/proc/8467/fd': Permission denied
find: '/proc/8467/map_files': Permission denied
find: '/proc/8467/fdinfo': Permission denied
find: '/proc/8467/ns': Permission denied
find: '/proc/20423/task/20423/fd': Permission denied
find: '/proc/20423/task/20423/fdinfo': Permission denied
find: '/proc/20423/task/20423/ns': Permission denied
find: '/proc/20423/fd': Permission denied
find: '/proc/20423/map_files': Permission denied
find: '/proc/20423/fdinfo': Permission denied
find: '/proc/20423/ns': Permission denied
find: '/tmp/libgksu-a3EXdy': Permission denied
find: '/tmp/libgksu-WDvJYp': Permission denied
find: '/lost+found': Permission denied
find: '/root': Permission denied
/home/hadoop/hadoop/share/hadoop/tools/sls/sample-conf/yarn-site.xml
/home/hadoop/hadoop/etc/hadoop/yarn-site.xml
find: '/home/user/.config/dconf': Permission denied
find: '/home/user/.config/geany/filedefs': Permission denied
find: '/home/user/.config/geany/templates': Permission denied
find: '/home/user/.config/geany/tags': Permission denied
find: '/home/user/.config/gtk-2.0': Permission denied
find: '/home/user/.run': Permission denied
find: '/home/user/.local/share/nano': Permission denied
find: '/home/user/.fltk': Permission denied
find: '/home/user/.mozilla': Permission denied
find: '/home/user/.dbus': Permission denied
find: '/home/user/.thumbnails': Permission denied
find: '/home/user/.cache': Permission denied
find: '/run/user/1000': Permission denied
find: '/run/sudo': Permission denied
find: '/run/systemd/inaccessible/dir': Permission denied
find: '/run/cups/certs': Permission denied
find: '/var/spool/cron/crontabs': Permission denied
find: '/var/spool/cups': Permission denied
```

To open a text file cat command can be used,

Syntax: \$ cat <file_name>

Command used: cat /home/hadoop/hadoop/etc/hadoop/yarn-site.xml

This displays the contents of the yarn-site.xml file in the Terminal.

yarn-site.xml


This file contains configuration information that overrides the default values for YARN parameters. Overrides of the default values for core configuration properties are stored in the Default YARN parameters file.

The properties that are defined in this configuration file are,

Name: yarn.nodemanager.aux-services **Value:** mapreduce_shuffle

A comma separated list of services where service name should only contain a-zA-Z0-9_ and can not start with numbers

```
File Edit View Search Preferences Tabs Help
find: '/proc/8467/ns': Permission denied
find: '/proc/20423/task/20423/fd': Permission denied
find: '/proc/20423/task/20423/fdinfo': Permission denied
find: '/proc/20423/task/20423/ns': Permission denied
find: '/proc/20423/fd': Permission denied
find: '/proc/20423/map_files': Permission denied
find: '/proc/20423/fdinfo': Permission denied
find: '/proc/20423/ns': Permission denied
find: '/tmp/libgksu-a3EXdy': Permission denied
find: '/tmp/libgksu-WDvJYp': Permission denied
find: '/lost+found': Permission denied
find: '/root': Permission denied
/home/hadoop/hadoop/share/hadoop/tools/sls/sample-conf/yarn-site.xml
/home/hadoop/hadoop/etc/hadoop/yarn-site.xml
find: '/home/user/.config/dconf': Permission denied
find: '/home/user/.config/geany/filedefs': Permission denied
find: '/home/user/.config/geany/templates': Permission denied
find: '/home/user/.config/geany/tags': Permission denied
find: '/home/user/.config/gtk-2.0': Permission denied
find: '/home/user/.run': Permission denied
find: '/home/user/.local/share/nano': Permission denied
find: '/home/user/.fltk': Permission denied
find: '/home/user/.mozilla': Permission denied
find: '/home/user/.dbus': Permission denied
find: '/home/user/.thumbnails': Permission denied
find: '/home/user/.cache': Permission denied
find: '/run/user/1000': Permission denied
find: '/run/sudo': Permission denied
find: '/run/systemd/inaccessible/dir': Permission denied
find: '/run/cups/certs': Permission denied
find: '/var/spool/cron/crontabs': Permission denied
find: '/var/spool/cups': Permission denied
find: '/var/spool/rsyslog': Permission denied
find: '/var/cache/apt/archives/partial': Permission denied
find: '/var/cache/ldconfig': Permission denied
find: '/var/lib/apt/lists/partial': Permission denied
find: '/var/lib/polkit-1': Permission denied
find: '/var/lib/connman/ethernet_080027c95d15_cable': Permission denied
find: '/var/lib/connman/ethernet_080027d5ee35_cable': Permission denied
find: '/var/lib/connman/ethernet_080027e586bd_cable': Permission denied
find: '/var/log/unattended-upgrades': Permission denied
find: '/etc/ssl/private': Permission denied
find: '/etc/cups/ssl': Permission denied
find: '/etc/polkit-1/localauthority': Permission denied
hadoop@antix1:~
$
```



The other yarn-site.xml file has the below properties configured,

Name: yarn.resourcemanager.scheduler.class, Value:

org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler

The class to use as the resource scheduler.

Name: yarn.resourcemanager.webapp.address, Value: localhost:18088

The http address of the RM web application.

Name: yarn.resourcemanager.resource-tracker.address, Value: localhost:18031

Name: yarn.resourcemanager.scheduler.address, Value: localhost:18030

The address of the scheduler interface.

Name: yarn.resourcemanager.address, Value: localhost:18032

The address of the applications manager interface in the RM.

Name: yarn.resourcemanager.admin.address, Value: localhost:18030

The address of the RM admin interface.

Name: Hadoop.security.token.service.use_ip Value: false

Set to avoid ip check

Name: yarn.scheduler.capacity.maximum-applications, Value: 1000

Maximum number of applications that can run at the same for a specific user. This overrides the global setting

Name: yarn.scheduler.fair.preemption, Value:true

Whether to use preemption. Defaults to false.

Name: yarn.scheduler.fair.assignmultiple, Value:true

Whether to allow multiple container assignments in one heartbeat. Defaults to false

```
File Edit View Search Preferences Tabs Help
find: '/etc/cups/ssl': Permission denied
find: '/etc/polkit-1/localauthority': Permission denied
hadoop@antix1:~
$ cat /home/hadoop/hadoop/share/hadoop/tools/sls/sample-conf/yarn-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.
-->
<configuration>
  <property>
    <name>yarn.resourcemanager.scheduler.class</name>
    <value>org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler</value>
    <!-- <value>org.apache.hadoop.yarn.server.resourcemanager.scheduler.fifo.FifoScheduler</value> -->
    <!-- <value>org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler</value> -->
  </property>

  <property>
    <description>The address of the RM web application.</description>
    <name>yarn.resourcemanager.webapp.address</name>
    <value>localhost:18088</value>
  </property>

  <property>
    <name>yarn.resourcemanager.resource-tracker.address</name>
    <value>localhost:18031</value>
  </property>

  <property>
    <description>The address of the scheduler interface.</description>
    <name>yarn.resourcemanager.scheduler.address</name>
    <value>localhost:18030</value>
  </property>

  <property>
```



```
<description>The address of the scheduler interface.</description>
<name>yarn.resourcemanager.scheduler.address</name>
<value>localhost:18030</value>
</property>

<property>
<description>The address of the applications manager interface in the RM.</description>
<name>yarn.resourcemanager.address</name>
<value>localhost:18032</value>
</property>

<property>
<description>The address of the RM admin interface.</description>
<name>yarn.resourcemanager.admin.address</name>
<value>localhost:18033</value>
</property>

<property>
<description>Set to false, to avoid ip check</description>
<name>hadoop.security.token.service.use_ip</name>
<value>false</value>
</property>

<property>
<name>yarn.scheduler.capacity.maximum-applications</name>
<value>1000</value>
<description>Maximum number of applications in the system which
    can be concurrently active both running and pending</description>
</property>

<property>
<description>Whether to use preemption. Note that preemption is experimental
    in the current version. Defaults to false.</description>
<name>yarn.scheduler.fair.preemption</name>
<value>true</value>
</property>

<property>
<description>Whether to allow multiple container assignments in one
    heartbeat. Defaults to false.</description>
<name>yarn.scheduler.fair.assignmultiple</name>
<value>true</value>
</property>
</configuration>
hadoop@antix1:~
$
```

References

<https://en.wikipedia.org/wiki/VirtualBox>

<https://www.virtualbox.org/manual/ch01.html#virt-why-useful>
<https://hadoop.apache.org/docs/r2.4.1/hadoop-project-dist/hadoop-hdfs/hdfs-default.xml>
<https://www.edureka.co/blog/explaining-hadoop-configuration/#:~:text=The%20core%2Dsite.,The%20hdfs%2Dsite.>
<https://hadoop.apache.org/docs/current/hadoop-yarn/hadoop-yarn-site/FairScheduler.html>