Hands-On Exercise: Installing Cloudera Manager Server

For this installation, you will install Cloudera Manager Server on the `cmhost` machine. Before installing Cloudera Manager, you will configure an external database (MySQL) to be used by Cloudera Manager and some of the services, which you will install in the next exercise.

Verify Environment Configuration

IMPORTANT: Complete all steps in this exercise on `cmhost`

In this section of the exercise, you will verify some important settings prior to installing software. Although in these steps you will only verify settings on the `cmhost` machine, most of the settings below should be in place on all the machines that will be part of the cluster.

1. Test the repo URL that contains files to install Cloudera Manager. From **cmhost**, start a web browser and enter this URL: http://cmhost:8060/cm7.3.1/

---

2. Verify the Oracle JDK is installed and that `JAVA_HOME` is defined and referenced in the system `PATH`.

   In a terminal:

   ```
   $ java -version
   ```

   The message returned in the terminal should show the java version is 1.8.0_232.

   ```
   $ echo $JAVA_HOME
   ```

   The result should show that `/usr/java/default` is the `JAVA_HOME` location.

   ```
   $ env | grep PATH
   ```

   The `PATH` value returned includes a reference to `/usr/java/default/bin`.

---

3. Verify Python is installed. It is a requirement for Hue, which you will install later in the course.

```
$ rpm -q python-2.7*
```

The command should return the package name of the installed version of python 2.7.

---

4. Verify MySQL Server is installed and running on `cmhost`.

```
$ systemctl status mysqld
```

The results of the command should show that the `mysqld.service` service is `"active" (running)`.

Note: Note that in a true production deployment you would also move the old InnoDB log files to a backup location and update the `/etc/my.cnf` MySQL configurations to conform with requirements as documented here.

---

5. Confirm no blocking is being done by Security-Enhanced Linux (SELinux)

```
$ sestatus
```

Note: Although it is not a requirement to set SELinux to disabled or permissive, it is important that SELinux not block during installation.

---

6. Verify IPv6 is disabled. This is a CDP requirement.

```
$ ip addr show
```

Notice that there is an IPv4 (`inet`) address for the `eth0` network interface, however there is no `inet6` address.

---

7. Check firewall settings.

```
$ systemctl list-unit-files | grep firewalld
```

The results of the command should show that the `firewalld.service` service is `"disabled"`.

Note: Although it is not a requirement to disable the firewall, there are many ports that must not be blocked during and after installation.

---

8. Check Transparent Hugepage compaction is disabled.

```
$ cat /proc/meminfo | grep -i hugepages_total
```

A response of `HugePages_Total: 0` indicates the hugepage feature is turned off.

Note: The OS feature called transparent hugepages interacts poorly with Hadoop workloads and can seriously degrade performance. Cloudera recommends it be turned off.

---

9. Check the `vm.swappiness` Linux kernel setting.

```
$ cat /proc/sys/vm/swappiness
```

Note: Cloudera recommends that you set `vm.swappiness` to a value between 1 and 10, in order to reduce lengthy garbage collection pauses which can affect stability and performance.

---

10. Verify the service that ensures time consistency across machines is running.

```
$ ntpstat
```

The results of the command should show that the time is synchronized wth the ntp server.

---

11. Verify the MySQL JDBC Connector is installed. Sqoop (a part of CDP that you will install in this course) does not ship with a JDBC connector, but does require one.

```
$ ls -l /usr/share/java/mysql*
```

You should see that a symlink has been defined at `/usr/share/java/mysql-connector-java.jar` that points to a specific version of a MySQL connector JAR file that exists in the same directory.

Configure the External Cloudera Manager Database

As is typical in a production cluster, Cloudera Manager uses an external database system instead of the embedded PostgreSQL system.

12. Optional: Review the script you will run to create the required databases.

```
$ cat ~/training_materials/admin/scripts/mysql-setup.sql
```

This script creates databases and users for the Cloudera Manager database and databases required by other services on the cluster.

---

13. Run the script to create the databases.

```
$ mysql -u root -p < ~/training_materials/admin/scripts\
/mysql-setup.sql
```

When prompted for the password enter **training**

---

14. Confirm the databases were created.

```
$ mysql -u root -p -e "SHOW DATABASES"
```

When prompted for the password enter **training**

Confirm that there are 13 databases, include these that were created by the script: `amon`, `cmserver`, `hue`, `metastore`, `oozie`, `registry`, `rman`, and `streamsmsgmgr`. (The list will also include several previously created databases.)

---

15. Make your MySQL installation secure.

```
$ sudo /usr/bin/mysql_secure_installation
 [...]
 Enter current password for root (enter for none): training
 OK, successfully used password, moving on..
 [...]
 Change the root password? [Y/n] N
 Remove anonymous users? [Y/n] Y
 [...]
 Disallow root login remotely? [Y/n] Y
 [...]
 Remove test database and access to it [Y/n] Y
 [...]
 Reload privilege tables now? [Y/n] Y
 All done!
 [...]
 Thanks for using MySQL!
 Cleaning up . . .
```

---

16. Verify the Cloudera Manager local software repository.

Your instances contain a local yum repository of Cloudera Manager software to save download time in this course.

CentOS (and Red Hat) store software repository references in `/etc/yum.repos.d`.

- Issue the commands below to edit the yum repository settings.

```
$ sudo cp ~/config/cloudera-manager.repo /etc/yum.repos.d/
$ sudo nano /etc/yum.repos.d/cloudera-manager.repo
```

- Update the file with the following settings:

  [cloudera-manager]

  baseurl = http://cmhost:8060/cm7.3.1/

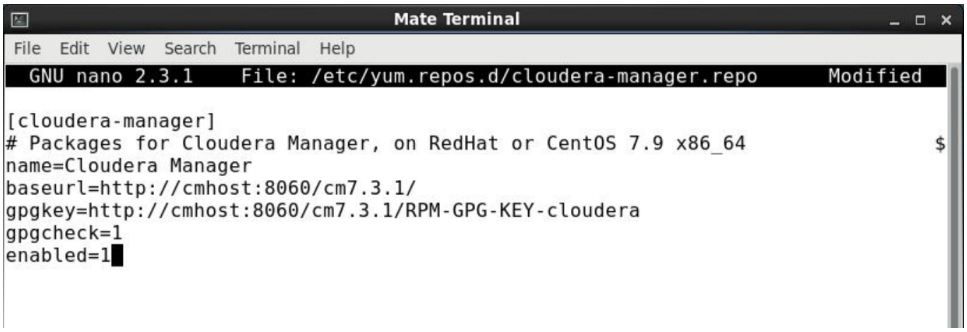
  enabled = 1

  gpgcheck = 1

  gpgkey = http://cmhost:8060/cm7.3.1/RPM-GPG-KEY-cloudera

  name = Cloudera Manager

Notice the base URL in the output of the last command. You will type this URL into Cloudera Manager later when you install Cloudera Manager agents.



Then save the file using `Ctrl + O`, and hit enter to accept the location. Exit with `Ctrl + X`.

Install Cloudera Manager Server

17. Install Cloudera Manager Server.

```
$ cd ~/software/cm7.3.1/RPMS/x86_64
$ sudo yum localinstall -y cloudera-manager-daemons-7* \
cloudera-manager-server-7*
```

Note: The `-y` option provides an answer of `yes` in response to an expected confirmation prompt.

Note: This command will take some time due to the size of the new CDP RPM packages.

---

18. Run the script to prepare the Cloudera Manager database.

```
$ sudo /opt/cloudera/cm/schema/scm_prepare_database.sh \
mysql cmserver cmserveruser password
```

After running the command above you should see the message, “All done, your SCM database is configured correctly!”

---

19. Start the Cloudera Manager Server.

```
$ sudo systemctl start cloudera-scm-server
```

---

20. Confirm that Cloudera Manager Server process started successfully.

```
$ ps -ef | grep cloudera-scm-server
```

The results of the `ps` command above show that Cloudera Manager Server is using the JDBC MySQL connector to connect to MySQL. It also shows logging configuration and other details.

Note: You should see multiple entries for the `cloudera-scm-server` process. If there is only one entry for the `cloudera-scm-server` process that displays, it means that the service did not start correctly. This can usually be corrected by rebooting `cmhost`. Use `sudo reboot` and enter password `training` when prompted. After rebooting, you will need to reconnect your browser session to `cmhost`.

---

21. Review the Cloudera Manager Server log file to see what took place.

The path to the log file is `/var/log/cloudera-scm-server/cloudera-scm-server.log`. Note that you must use `sudo` to access Cloudera Manager logs because of restricted permissions on the Cloudera Manager log file directories.

```
$ sudo tail -f /var/log/cloudera-scm-server/cloudera-scm-\
server.log
```

Tip: Bash tab-completion won't work for the filename in the above command, because the permissions on the directory do not allow the training user to view its contents until the `sudo` command actually runs.

Tip: Press the `Enter` key to scroll through the log file contents. Press the `Ctrl + c` to quit the `less` command file viewer.

This is the end of the exercise.