

CDP Developer Participant Environment Setup

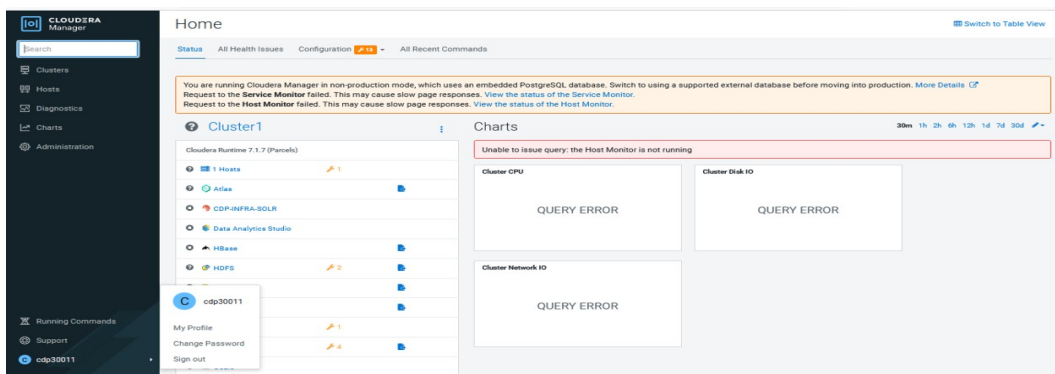
1. Software Stacks for CDP 3001 Certification Training:

1. Cloudera Platform : CDP Private Cloud on GCP
2. SDKs : Java, Scala, Python and Spark
3. IDEs : Jupyter Notebook, PyCharm Edu and IntelliJ Edu
4. Remote Terminal : WinSCP and Smart Putty
5. Utility : Notepad++, 7Zip
6. DevSecOps and MLOps : Jenkins, Git, GitHub and Splunk

For download links and version, refer “**CDP Developer Workstation for CDP 3001 Certification 1.0.docx**”

2. CDP Cloud Systems:

1. **CDP Private Cloud:** Installation on Cloud/Data Centre, private installation. CDP Private Cloud CDP Private Cloud Address: <http://35.223.46.173:7180> . It is installed on Ubuntu 20.04 LTS in GCP System with all Services.



2. **CDP Public Cloud:** No Installation, Cloudera SASS. CDP Public Cloud Data Services: URL: <https://www.cloudera.com/campaign/try-cdp-public-cloud.html>

Thank you for signing up for the
CDP Public Cloud trial

We will contact you shortly to provide access to CDP in your existing environment or to set up a trial.

In the meantime, sign up for free training on CDP via an [on-demand course](#).

Try CDP Public Cloud for free for 60-days using your own data and workloads

Experience:

- A powerful analytics platform that simplifies the management of hybrid and multi-cloud data
- Easy-to-use analytics services such as Data Warehouse and Machine Learning that scale to petabytes of data and thousands of users
- Centralized security and governance for pain-free portability across clouds

3. Software Stack for CDP-3001-Installations Steps:

1. Java 1.8 and Java 11 Installation:

- i. Download installers of JDK 1.8 and JDK 11
- ii. Run setup with default settings and click on finish
- iii. Post installation, configure JAVA_HOME in Environment Properties

- iv. Open Windows Command Prompt and type gpupdate, java -version and echo %JAVA_HOME%



```
Command Prompt
Microsoft Windows [Version 10.0.19042.2130]
(c) Microsoft Corporation. All rights reserved.

C:\Users\AmritChhetriB>gpupdate
Updating policy...

Computer Policy update has completed successfully.
User Policy update has completed successfully.

C:\Users\AmritChhetriB>java -version
java version "1.8.0_202"
Java(TM) SE Runtime Environment (build 1.8.0_202-b08)
Java HotSpot(TM) 64-Bit Server VM (build 25.202-b08, mixed mode)

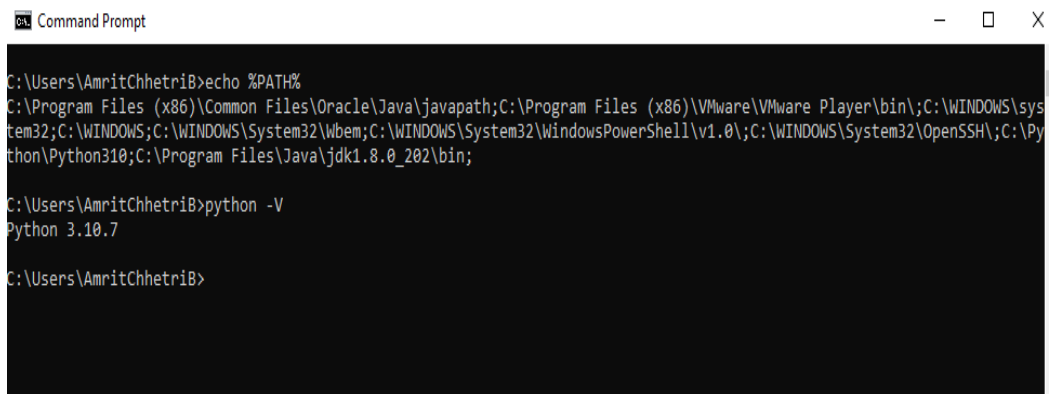
C:\Users\AmritChhetriB>echo %JAVA_HOME%
C:\Program Files\Java\jdk1.8.0_202

C:\Users\AmritChhetriB>
```

Follow similar steps for Java 11 for Jenkins and features comparisons. Site <https://www.oracle.com/in/java/technologies/javase/jdk11-archive-downloads.html>

2. Python 3.6 and Python 3.10:

- i. Get installer of python 3.10 and change installation directory to C:\Python\Python310
- ii. Post Installation, add python home folder C:\Python\Python310 in path
- iii. Type echo %PATH%, python -V



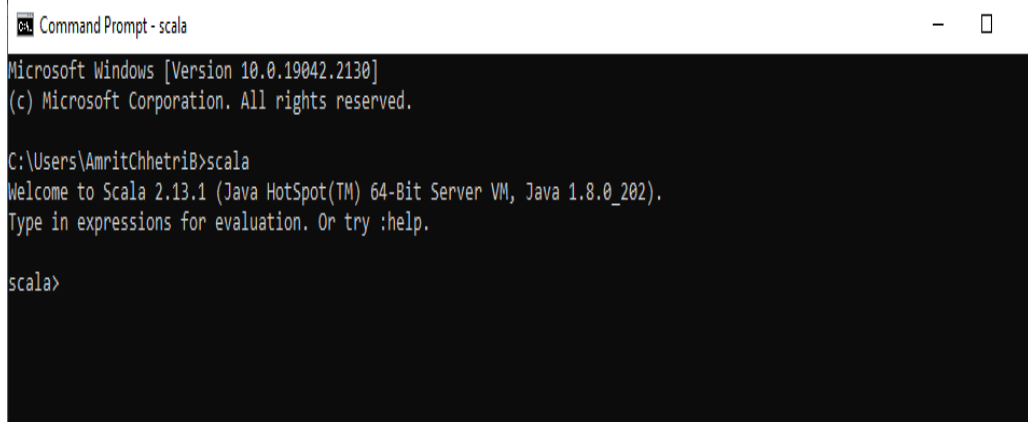
```
Command Prompt
C:\Users\AmritChhetriB>echo %PATH%
C:\Program Files (x86)\Common Files\Oracle\Java\javapath;C:\Program Files (x86)\VMware\VMware Player\bin\;C:\WINDOWS\system32;C:\WINDOWS;C:\WINDOWS\System32\Wbem;C:\WINDOWS\System32\WindowsPowerShell\v1.0\;C:\WINDOWS\System32\OpenSSH\;C:\Python\Python310;C:\Program Files\Java\jdk1.8.0_202\bin;

C:\Users\AmritChhetriB>python -V
Python 3.10.7

C:\Users\AmritChhetriB>
```

3. Scala Installation:

- i. Get installer of Scala and install
<https://downloads.lightbend.com/scala/2.13.1/scala-2.13.1.msi>
- ii. Post Installation, open Windows Prompt and type scala on Prompt



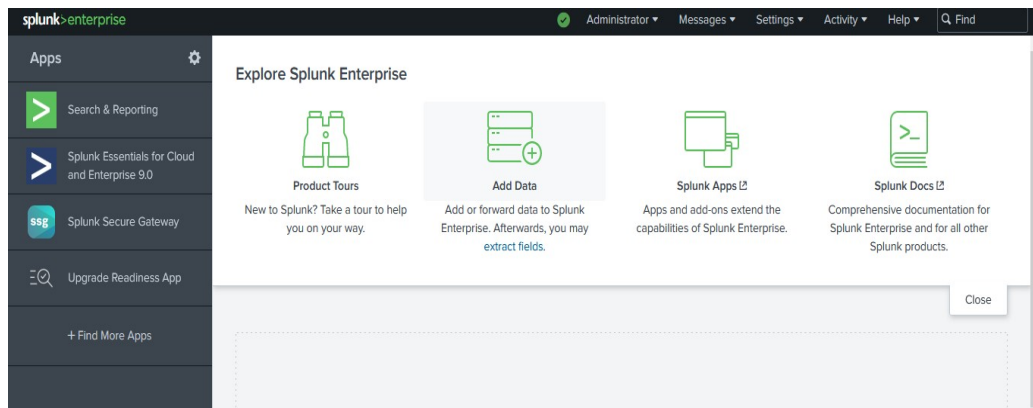
```
Command Prompt - scala
Microsoft Windows [Version 10.0.19042.2130]
(c) Microsoft Corporation. All rights reserved.

C:\Users\AmritChhetriB>scala
Welcome to Scala 2.13.1 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_202).
Type in expressions for evaluation. Or try :help.

scala>
```

4. Splunk Installation:

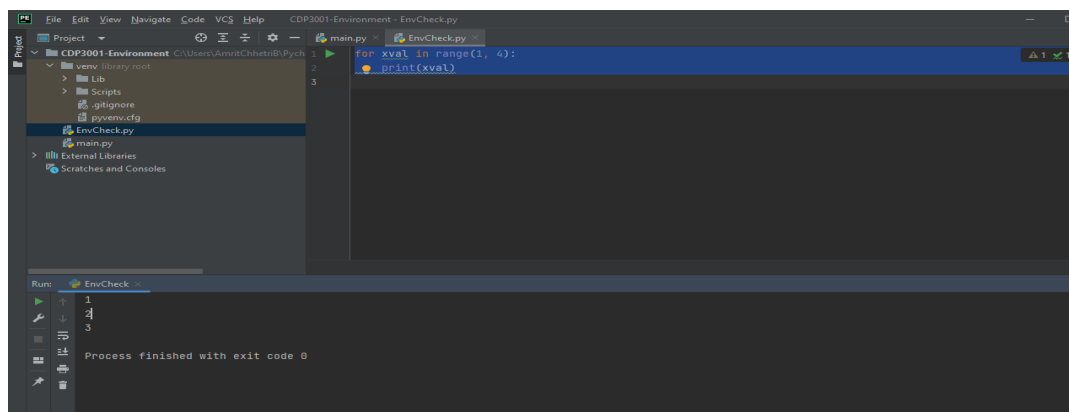
- Get installer of Splunk and install with default settings, whenever prompted create your user credentials - username/password: amritchhetrib/Qwerty123456
- Post Installation, access by typing <http://127.0.0.1:8000>



5. PyCharm Installation:

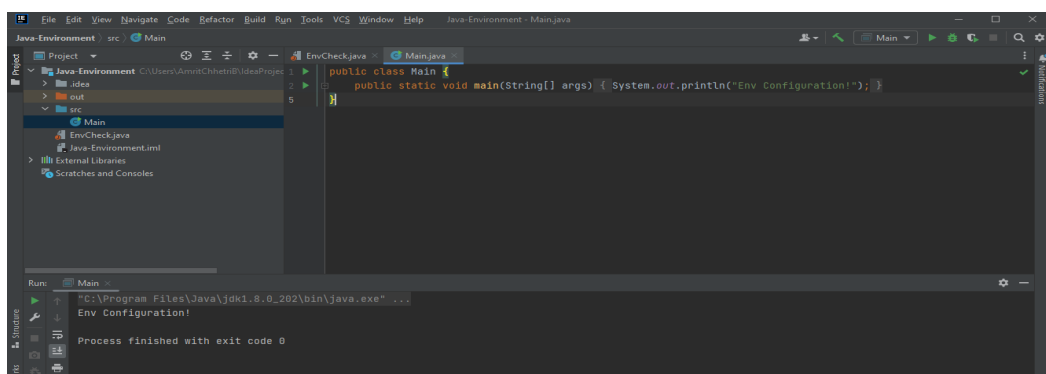
- Download/Get installer of PyCharm and install with default settings
- Post Installation, open and click on New ->Project->Python-> create file EnvCheck.py and write

```
for xval in range(1, 4):  
    print(xval)
```



6. IntelliJ Edu Installation:

- Download/Get installer of IntelliJ and install with default settings
- Post Installation, Post Installation, open and click on New -Project-Java Project



7. Git Installation:

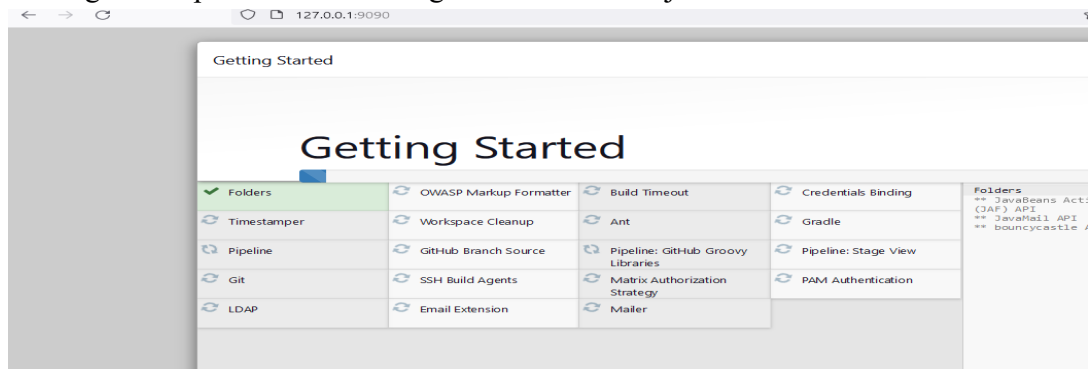
- Get installer of Git and install with default settings
- Post Installation, type git clone

```
--bare create a bare repository
--mirror create a mirror repository (implies bare)
--local clone from a local repository
--no-hardlinks don't use local hardlinks, always copy
--shallow clone a shallow repository
--recurse-submodules[=<pathspec>] initialize submodules in the clone
--recursive ... alias of --recurse-submodules
--jobs <n> number of submodules cloned in parallel
--template <template-directory> from which templates will be used
--reference <repo> reference repository
--reference-if-able <repo> reference repository
--dissociate use --reference only while cloning
--origin <name> use <name> instead of 'origin' to track upstream
--branch <branch> checkout <branch> instead of the remote's HEAD
--upload-pack <path> path to git-upload-pack on the remote
--depth <depth> create a shallow clone of that depth
--shallow-since <time> create a shallow clone since a specific time
--shallow-exclude <revision> deepen history of shallow clone, excluding rev
--single-branch clone only one branch, HEAD or --branch
--no-tags don't clone any tags, and make later fetches not to fetch tags
--shallow-submodules any cloned submodules will be shallow
--separate-git-dir <gitdir> separate git dir from working tree
--config <key=value> set config inside the new repository
--server-option <server-specific> option to transmit
--ipv4 use IPv4 addresses only
--ipv6 use IPv6 addresses only
--filter <args> object filtering
--also-filter-submodules apply partial clone filters to submodules
--remote-submodules any cloned submodules will use their remote-tracking b
--sparse initialize sparse-checkout file to include only files
--bundle-uri <uri> a URI for downloading bundles before fetching from ori
git remote
```

Create a folder CDP-3001-Certification in C or D and Examples
Open Windows Prompt, go to Examples and types
git -c http.sslVerify=false clone <https://github.com/spark-examples/spark-scala-xamples.git>

8. Jenkins Installation:

- Get installer of Jenkins and install with default settings
- Post Installation, access typing <http://ip:9090> , as given during installation
- Login with password in C:\ProgramData\Jenkins\jenkins\secrets\initialAdminPassword



4. Software Stack for CDP-3001- Check Steps:

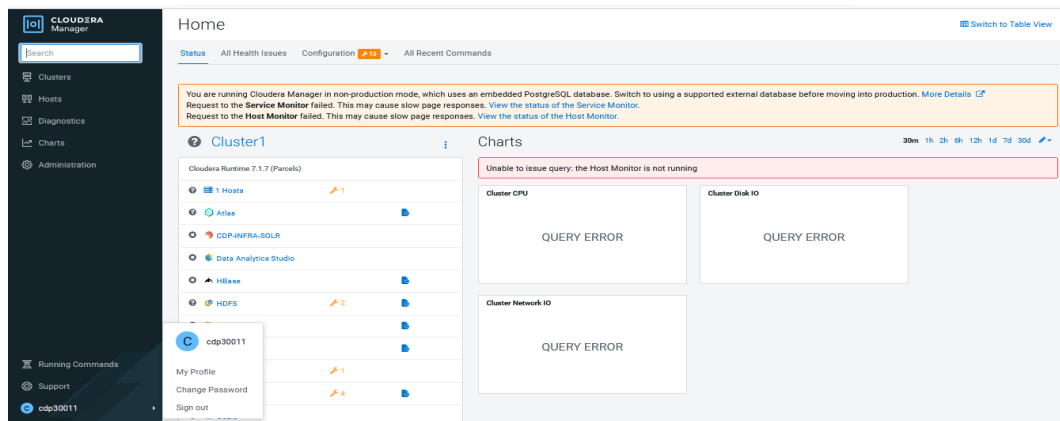
1. CDP Private Cloud(setup for Training):

Web Address: <http://35.223.46.173:7180>

Username: cdp30011

Password : Qwerty123456\$

Access Type: Private Cloud



2. CDP Public Cloud (SAAS) – SAAS Access :

3. HDFS Access: with SSH

IP Address: 35.223.46.173 Port: 22

Username: cdp30011

Password : Qwerty123456\$

Create directory to load file:

Leave Safe Mode: `sudo -u hdfs hdfs dfsadmin -safemode leave`

`sudo -u hdfs hadoop fs -mkdir /cdproot`

`sudo -u hdfs hadoop fs -mkdir /cdproot/spark`

`sudo -u hdfs hadoop fs -mkdir /cdproot/nifi`

`sudo -u hdfs hadoop fs -mkdir /cdproot/spark`

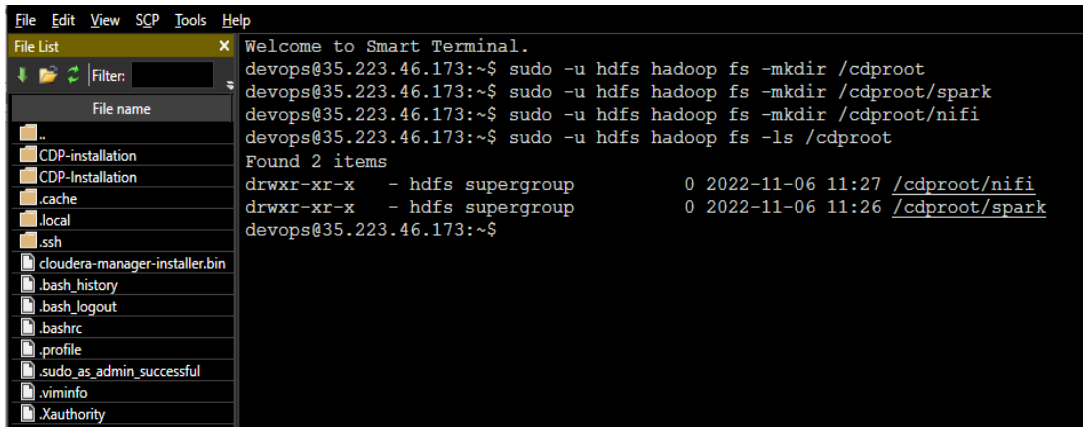
`sudo -u hdfs hadoop fs -ls /cdproot`

`sudo -u hdfs hadoop fs -ls /cdproot`

`sudo -u hdfs hadoop fs -put Data.txt /cdproot/data`

`sudo -u hdfs hadoop fs -cat /cdproot/data/Data.txt`

Deleting files: `sudo -u hdfs hadoop fs -rmdir /cdproot/spark`



The screenshot shows a Smart Terminal window with a menu bar (File, Edit, View, SCP, Tools, Help) and a file list on the left. The terminal output shows the following commands and results:

```
Welcome to Smart Terminal.
devops@35.223.46.173:~$ sudo -u hdfs hadoop fs -mkdir /cdproot
devops@35.223.46.173:~$ sudo -u hdfs hadoop fs -mkdir /cdproot/spark
devops@35.223.46.173:~$ sudo -u hdfs hadoop fs -mkdir /cdproot/nifi
devops@35.223.46.173:~$ sudo -u hdfs hadoop fs -ls /cdproot
Found 2 items
drwxr-xr-x - hdfs supergroup          0 2022-11-06 11:27 /cdproot/nifi
drwxr-xr-x - hdfs supergroup          0 2022-11-06 11:26 /cdproot/spark
devops@35.223.46.173:~$
```

The file list on the left includes: ., CDP-installation, CDP-Installation, .cache, .local, .ssh, cloudera-manager-installer.bin, .bash_history, .bash_logout, .bashrc, .profile, .sudo_as_admin_successful, viminfo, and .Xauthority.