

Exercise-6: Installation of Hadoop in Ubuntu System

Download and configure Hadoop

- Once hadoop user is created, it is switched to this user.

```
shruthimohan@EDITH:~$ su - hadoop
Password:
hadoop@EDITH:~$ |
```

- Start the SSH server manually in WSL.

1. sudo service ssh start
2. ssh localhost

```
hadoop@EDITH:~$ sudo service ssh start
[sudo] password for hadoop:
* Starting OpenBSD Secure Shell server sshd
hadoop@EDITH:~$ ssh localhost
Welcome to Ubuntu 20.04.6 LTS (GNU/Linux 5.10.60.1-microsoft-standard-WSL2 x86_64)

* Documentation:  https://help.ubuntu.com
* Management:    https://landscape.canonical.com
* Support:       https://ubuntu.com/pro

Last login: Tue Oct  1 11:01:07 2024 from 127.0.0.1
hadoop@EDITH:~$ |
```

- Start the Hadoop Cluster: start-all.sh


```
hadoop@EDITH:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [0.0.0.0]
Starting datanodes
Starting secondary namenodes [EDITH]
Starting resourcemanager
Starting nodemanagers
hadoop@EDITH:~$ |
```

- To confirm that all necessary Hadoop services are up and running, check their status using the 'jps' command

```
hadoop@EDITH:~$ jps
8228 SecondaryNameNode
8966 Jps
8441 ResourceManager
7993 DataNode
8569 NodeManager
7868 NameNode
```

- Access the Hadoop Namenode: <http://192.168.1.6:9870/>
- Access the Resource Manager: <http://192.168.1.6:8088/>

- If you're working on a single-node Hadoop setup (pseudo-distributed mode), you can use localhost: <http://localhost:8088>



All Applications

- Cluster
- About
- Nodes
- Node Labels
- Applications
- NEW
- NEW SAVING
- SUBMITTED
- ACCEPTED
- RUNNING
- FINISHED
- FAILED
- KILLED
- Scheduler
- Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Used Resources	Total Resources	Reserved Resources
0	0	0	0		<memory:0 B, vCores:0>	<memory:8 GB, vCores:8>	<memory:0 B, vCores:0>

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes
1	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Resource
Capacity Scheduler	[memory-mb (unit-M), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

Show 20 entries

ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU V-Cores	Allocated Memory MB	Allocated GPUs	Reserved CPU V-Cores	Reserved Memory MB
No data available in table																	

Showing 0 to 0 of 0 entries

- Stop Hadoop Services: stop-all.sh

```
hadoop@EDITH:~$ stop-all.sh
WARNING: Stopping all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: Use CTRL-C to abort.
Stopping namenodes on [0.0.0.0]
Stopping datanodes
Stopping secondary namenodes [EDITH]
Stopping nodemanagers
Stopping resourcemanager
```