

IMPORTANT COMMANDS TO KNOW

Hadoop

To start Hadoop, we need to HDFS and the resource manager yarn with the following commands

- start-dfs.sh
- start-yarn.sh

To stop Hadoop

- stop-dfs.sh
- stop-yarn.sh

HBase

To start HBase, we need to make sure Hadoop is running. Check with command “jps” if Hadoop is not on, then start Hadoop before starting HBase since Hbase databases are stored in HDFS

- start-yarn.sh
- start-dfs.sh
- start-hbase.sh

Check hdfs /hbase folder in Hadoop repertory

- hdfs dfs -ls /hbase

Hive

To start Hive, we need to make sure Hadoop is running. Check with command “jps” if Hadoop is not on, then start Hadoop before starting Hive since Hive metastore is in HDFS

- start-yarn.sh
- start-dfs.sh
- hive --service metastore

Open a new terminal window to launch Hive when service metastore is active and type

- hive

Kafka

First start ZOOKEEPER

- zookeeper-server-start.sh \$KAFKA_HOME/config/zookeeper.properties

OR

- zookeeper-server-start.sh -daemon \$KAFKA_HOME/config/zookeeper.properties (daemon mode is preferred)

Then start KAFKA once ZOOKEEPER is running

- kafka-server-start.sh \$KAFKA_HOME/config/server.properties

OR

- kafka-server-start.sh -daemon \$KAFKA_HOME/config/server.properties (daemon mode is preferred)

To create a topic called “test” in Kafka, type the following commands

- kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic test

To stop KAFKA, first stop ZOOKEEPER

- zookeeper-server-stop.sh
- kafka-server-stop.sh

If you have to kill Kafka in case you cannot stop it the regular way, do “jps” to get the port number and type the following command

- kill -9 -<Port Number>

By default, Zookeeper runs on port 2181 and the default port for consumer group 9092 in the configuration setting

Create a producer called “kafkaTest”

- kafka-console-producer.sh --broker-list localhost:9099 --topic kafkaTest

To check the topics in Zookeeper

- kafka-topics.sh --list --zookeeper localhost:2181

Describe the topic to check that the topic called “kafkaTest” is actually created

- kafka-topics.sh --describe --zookeeper localhost:2181 --topic kafkaTest

Create a consumer

- kafka-console-consumer.sh --bootstrap-server localhost:9099 --topic kafkaTest

Initialize a consumer that reads from previous messages

- kafka-console-consumer.sh --bootstrap-server localhost:9099 --topic kafkaTest --from-beginning

HDFS

Most of the basic Linux commands work in HDFS in the following syntax

- hdfs dfs -<Linux command>

Check the following website for more details

<https://hadoop.apache.org/docs/r2.4.1/hadoop-project-dist/hadoop-common/FileSystemShell.html>

Linux Commands

Access directory
cd directory

List all files in a long listing (detailed) format
ls -al

Display the present working directory
pwd

Create a directory
mkdir directory

Remove (delete) file
rm file

Remove the directory and its contents recursively
rm -r directory

Force removal of file without prompting for confirmation
rm -f file

Forcefully remove directory recursively
rm -rf directory

Copy file1 to file2
cp file1 file2

Copy source_directory recursively to destination. If destination exists, copy source_directory into destination, otherwise create destination with the contents of source_directory.
cp -r source_directory destination

Rename or move file1 to file2. If file2 is an existing directory, move file1 into directory file2
mv file1 file2

Create symbolic link to linkname
ln -s /path/to/file linkname

Create an empty file or update the access and modification times of file.
touch file

View the contents of file
cat file

Browse through a text file

less file

Display the first 10 lines of file
head file

Display the last 10 lines of file
tail file

Display the last 10 lines of file and "follow" the file as it grows.
tail -f file

Extract a gzip compressed tar file.
tar xzf archive.tar.gz

Please check the links below for more commands

<https://www.linuxtrainingacademy.com/linux-commands-cheat-sheet/>

<https://www.loggly.com/wp-content/uploads/2015/05/Linux-Cheat-Sheet-Sponsored-By-Loggly.pdf>