Managing HDFS Directories

Now let us have a look at how to create directories and manage ownership.

- By default hdfs is superuser of HDFS
- `hadoop fs -mkdir` or `hdfs dfs -mkdir` – to create directories
- `hadoop fs -chown` or `hdfs dfs -chown` – to change ownership of files
- `chown` can also be used to change the group. We can change the group using `-chgrp` command as well. Make sure to run `-help` on chgrp and check the details.
- Here are the steps to create user space. Only users in HDFS group can take care of it.
  - Create directory with user id `itversity` under /user
  - Change ownership to the same name as the directory created earlier (/user/itversity)
  - You can validate permissions by using `hadoop fs -ls` or `hdfs dfs -ls` command on /user. Make sure to grep for the user name you are looking for.
- Let's go ahead and create user space in HDFS for `itversity`. I have to login as sudoer and run below commands.

```
sudo -u hdfs hdfs dfs -mkdir /user/itversity
sudo -u hdfs hdfs dfs -chown -R itversity:students /user/itversity
hdfs dfs -ls /user|grep itversity
```

- You should be able to create folders under your home directory.

```
%%sh

hdfs dfs -ls /user/${USER}
```

```
%%sh

hdfs dfs -mkdir /user/${USER}/retail_db
```

```
%%sh

hdfs dfs -ls /user/${USER}
```

- You can create the directory structure using `mkdir -p`. The existing folders will be ignored and non existing folders will be created.
  - Let us run `hdfs dfs -mkdir -p /user/${USER}/retail_db/orders/year=2020`.
  - As `/user/${USER}/retail_db` already exists, it will be ignored.
  - Both `/user/${USER}/retail_db/orders` as well as `/user/${USER}/retail_db/orders/year=2020` will be created.

```
%%sh

hdfs dfs -help mkdir
```

```
%%sh

hdfs dfs -ls -R /user/${USER}/retail_db
```

```
%%sh

hdfs dfs -mkdir -p /user/${USER}/retail_db/orders/year=2020
```

```
%%sh

hdfs dfs -ls -R /user/${USER}/retail_db
```

- We can delete non empty directory using `hdfs dfs -rm -R` and empty directory using `hdfs dfs -rmdir`. We will explore `hdfs dfs -rm` in detail later.

```
%%sh

hdfs dfs -help rmdir
```

```
%%sh

hdfs dfs -rmdir /user/${USER}/retail_db/orders/year=2020
```

```
%%sh

hdfs dfs -rm /user/${USER}/retail_db
```

```
%%sh

hdfs dfs -rmdir /user/${USER}/retail_db
```

```
%%sh

hdfs dfs -rm -R /user/${USER}/retail_db
```

```
%%sh

hdfs dfs -ls /user/${USER}
```

By Durga Gadiraju