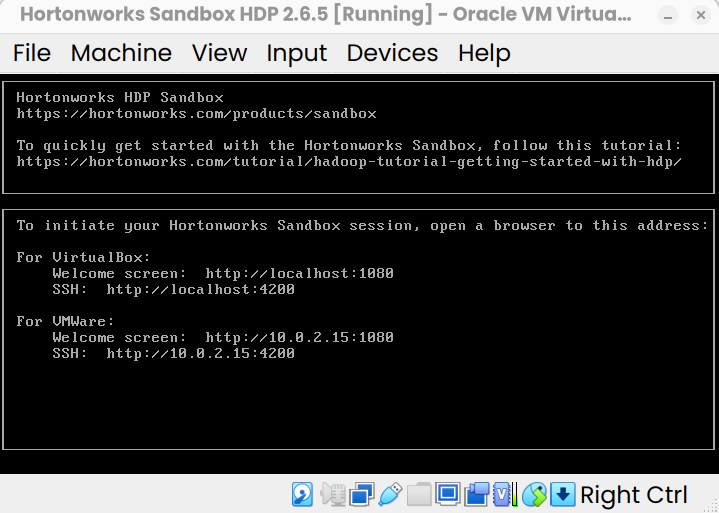
LAB 3

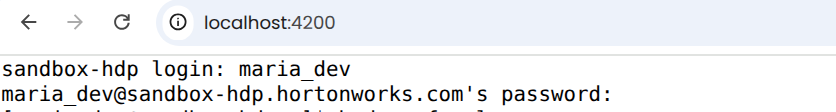
**Aim:** Use of HDFS file management commands

**Procedure:**

1. Start the HortonWorks it shows the IP address of machine.



1. Navigate to [http://localhost:](http://localhost:1080/)4200
2. Use the username and password: maria\_dev to login.



1. Execute the following commands one-by-one:

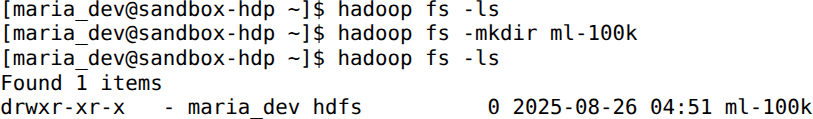
# hadoop fs -ls

The command **hadoop fs -ls** is used to **list files and directories stored in Hadoop Distributed File System (HDFS)** or other supported file systems (like local FS, S3, etc., depending on configuration). It also shows the metadata:

* + - File permissions
    - Replication factor
    - Owner & group
    - File size (in bytes)
    - Last modification date & time
    - Path

# hadoop fs -mkdir ml-100k

The **hadoop fs -mkdir** command is used to **create new directories in Hadoop Distributed File System (HDFS)** (or any other file system supported by Hadoop, like S3, local FS, etc., depending on your configuration)



We will give the command for creating a directory for a movielens dataset

* + **wget** [**http://media.sundog-soft.com/hadoop/ml-100k/u.data**](http://media.sundog-soft.com/hadoop/ml-100k/u.data)

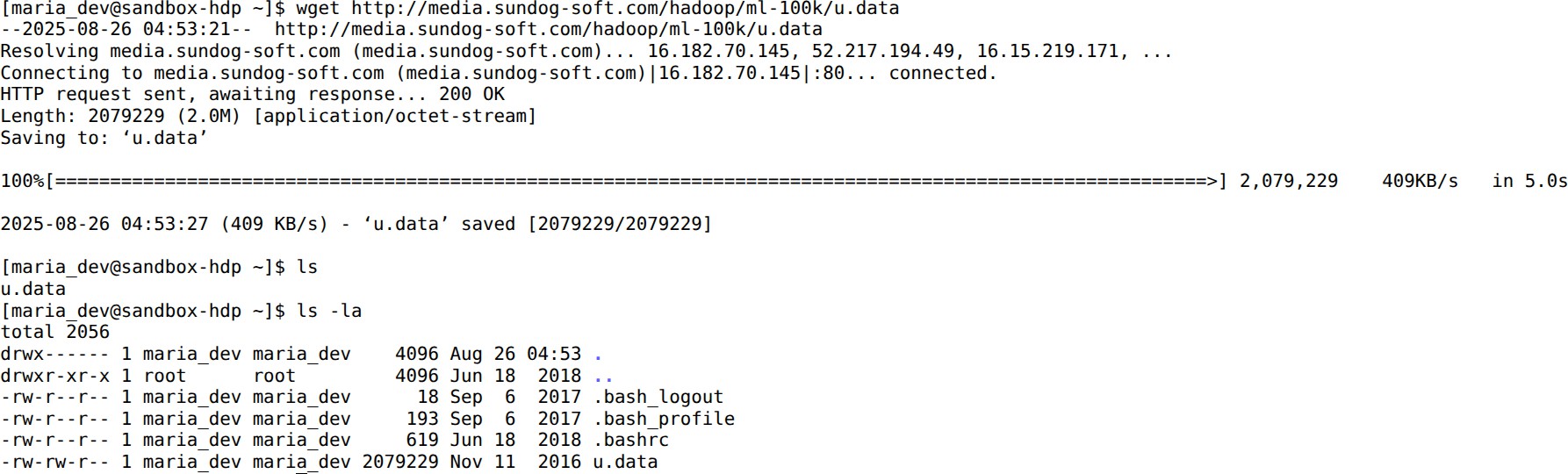
The above command is used to copy the data from web server to the Hadoop file system

# ls

Give the command ls to see whether the data is downloaded Once it is downloaded, you will see the name as u.data

# ls -la

In **Linux**, ls -la lists **all files including hidden ones** (those starting with .), with detailed information (long format).



# hadoop fs -copyFromLocal u.data ml-100k/u.data

The file will get copied from local file system to the Hadoop named as u.data

# hadoop fs -ls

The **hadoop fs -ls** command is used to **list files and directories in Hadoop Distributed File System (HDFS)** or in any other file system supported by Hadoop (like local FS, S3, etc., depending on configuration)

# hadoop fs -rm ml-100k/u.data

**📌 Purpose**

* + - To **remove (delete) files** from HDFS.
    - Works similar to Linux rm, but operates on HDFS.

# hadoop fs -rmdir ml-100k

The **hadoop fs -rmdir** command is used to **remove (delete) empty directories from HDFS**.

**📌 Purpose**

* + - To delete **empty directories** in Hadoop Distributed File System (HDFS).
    - It is similar to the Linux rmdir command.
    - ⚠️ Unlike -rm -r, it **cannot delete directories that contain files or subdirectories**.

# hadoop fs -ls

The commands checks where the directory is removed from the hadoop

