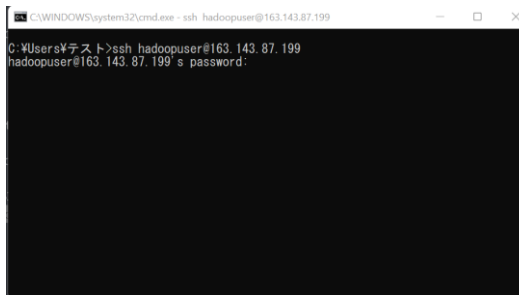
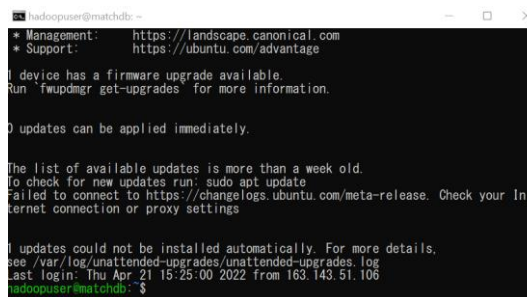


Steps to Access Hadoop and Spark Server machines

1. Set up the VPN by following the instructions available at <https://web-int.u-aizu.ac.jp/labs/istc/ipc/service/ains-vpn/vpn-e.html>
2. Connect to VPN if you are accessing from outside of the university.
3. Open terminal in your Linux/Mac machine. Windows users are requested to open command prompt.
4. Type “ssh [hadoopuser@163.143.87.199](https://web-int.u-aizu.ac.jp/labs/istc/ipc/service/ains-vpn/vpn-e.html)” and press Enter button



5. Type the password “hadoopmaster” and press Enter button.
6. Successful login into the remote server will show the below screen.



7. Create a directory with your studentID by executing the following command:

```
hdfs dfs -mkdir /user/studentID
```

```
hadoopuser@matchdb: ~  
hadoopuser@matchdb:~$ hdfs dfs -mkdir /user/ml111112345
```

8. Use the above created folder for storing and deleting your files.
9. Copy a file from the local machine into Hadoop file system by executing the following command:

```
hdfs dfs -put transactional_T10I4D100K.csv /user/d8222110/
```

```
hadoopuser@matchdb: ~  
hadoopuser@matchdb:~$ hdfs dfs -put transactional_T10I4D100K.csv /user/d8222110/  
2022-04-21 15:45:38,995 WARN util.NativeCodeLoader: Unable to load native-hadoop library f  
or your platform. . using builtin-java classes where applicable  
2022-04-21 15:45:39,998 INFO sasl.SaslDataTransferClient: SASL encryption trust check: loc  
alHostTrusted = false, remoteHostTrusted = false  
hadoopuser@matchdb:~$
```

10. Practice Hadoop commands by referring to the cheat code provided at

<https://images.linuxide.com/hadoop-hdfs-commands-cheatsheet.pdf>