**1. Server Access**

* **Linux Terminal Login:**

css

ssh PMO@10.10.11.242

Password: PMO@1234

* **MySQL Login (Workbench or Terminal):**

**10.10.11.242 port 3306**

* + Database: RME\_DH
  + Database: RME\_TEST
    - User: gamal
    - Password: password123Y$
  + Newly Created User:
    - Username: Omar2
    - Password: Omar\_54321

**2. Basic Linux Commands**

* List directory: ls
* Read text file: cat <filename>
* Edit text file: nano <filename>
* Remove files or directories:

bash

sudo rm -r <filename/directory>

**3. File Transfer (Install Filezilla Instead)**

1. **Upload Script (Windows to Linux):**

bash

pscp -pw PMO@1234 D:\test.csv PMO@10.10.11.242:/home/PMO/airflow/dags

1. **Download Script (Linux to Windows):**

bash

pscp -r -pw PMO@1234 PMO@10.10.11.242:/home/PMO/airflow/dags D:

1. **Remove Cache Files:**

bash

sudo rm -r \_\_pycache\_\_

**4. Airflow Management**

* **Access Airflow:**
  + URL: http://10.10.11.242:8090/home
  + User: mohamedgamal
  + Password: gamal123
  + Newly Added User:
    - Username: Omar2
    - Password: Omar\_54321
* **Restart Airflow:**

bash

cd airflow

source venv/bin/activate

airflow db init

airflow scheduler -D

airflow webserver -D --port 8090

change port to 8082 (-D as a deamon )

airflow webserver -D --port 8082

restart

airflow variables set MYSQL\_HOST "10.10.11.242"

airflow variables set MYSQL\_USER "omar2"

airflow variables set MYSQL\_PASSWORD "Omar\_54321"

airflow variables set MYSQL\_DATABASE "RME\_TEST"

stop and start

 **Find the PID:**

Bash

ps aux | grep "airflow webserver"

ps aux | grep "airflow scheduler"

This will show you the processes related to the webserver and scheduler, including their PIDs.

 **Kill the processes:**

Bash

kill -9 <webserver\_PID>

kill -9 <scheduler\_PID>

or

pkill -f "airflow webserver"

pkill -f "airflow scheduler"

* **Kill Existing Webserver Process:**

bash

ps aux | grep "airflow webserver"

kill -9 <PID>

* **Enable API in Airflow Config:** Update airflow.cfg:

makefile

[api]

enable\_experimental\_api = True

auth\_backend = airflow.api.auth.backend.basic\_auth

* **Test Airflow API (PowerShell):**

powershell

curl -Method GET http://10.10.11.242:8080/api/v1/dags -Credential (New-Object System.Management.Automation.PSCredential ("omar2", (ConvertTo-SecureString "Omar\_54321" -AsPlainText -Force)))

**5. Spark Management**

* **Restart Spark:**

bash

cd /opt/spark/sbin

sudo bash start-master.sh -i 10.10.11.242 --webui-port 8080

sudo bash start-worker.sh spark://10.10.11.242:7077

sudo jps

* **Spark Ports:**
  + Master: 7077
  + UI: 8081

**6. MySQL Management**

* **Enter MySQL from Linux Terminal:**

bash

mysql -u omar2 -p

Enter password: Omar\_54321

* **Check MySQL Storage:**

bash

CopyEdit

df -h /var/lib/mysql

**7. Python & Pip**

* **Python Version:**

bash

python3.10 -m pip --version

* **Install Packages with Pip:**

bash

CopyEdit

pip install apache-airflow --trusted-host pypi.org --trusted-host files.pythonhosted.org

**8. Miscellaneous**

* **Generate SSH Key:**

bash

ssh-keygen -t ed25519

cat $env:USERPROFILE\.ssh\id\_ed25519.pub

* **Monitor Running Processes:**

bash

htop

* **OracleDB Connection:**

bash

CopyEdit

cd ~/Downloads/oracle/instantclient\_21\_16

export LD\_LIBRARY\_PATH=~/Downloads/oracle/instantclient\_21\_16:$LD\_LIBRARY\_PATH

export PATH=~/Downloads/oracle/instantclient\_21\_16:$PATH

sqlplus RMEDWH/DWHRME@'//10.0.11.59:1521/RMEDB'

**9. Useful Links**

* [Crontab Guru](https://crontab.guru/) - Schedule cron jobs.
* [Airflow Email Notifications Guide](https://chatgpt.com/share/67445981-0518-8007-937d-422cc1b02d97)
* [Run Airflow on Docker](https://youtu.be/aTaytcxy2Ck?feature=shared)
* [Apache Spark Installation](https://youtu.be/GxJTxz1EQwY?feature=shared)