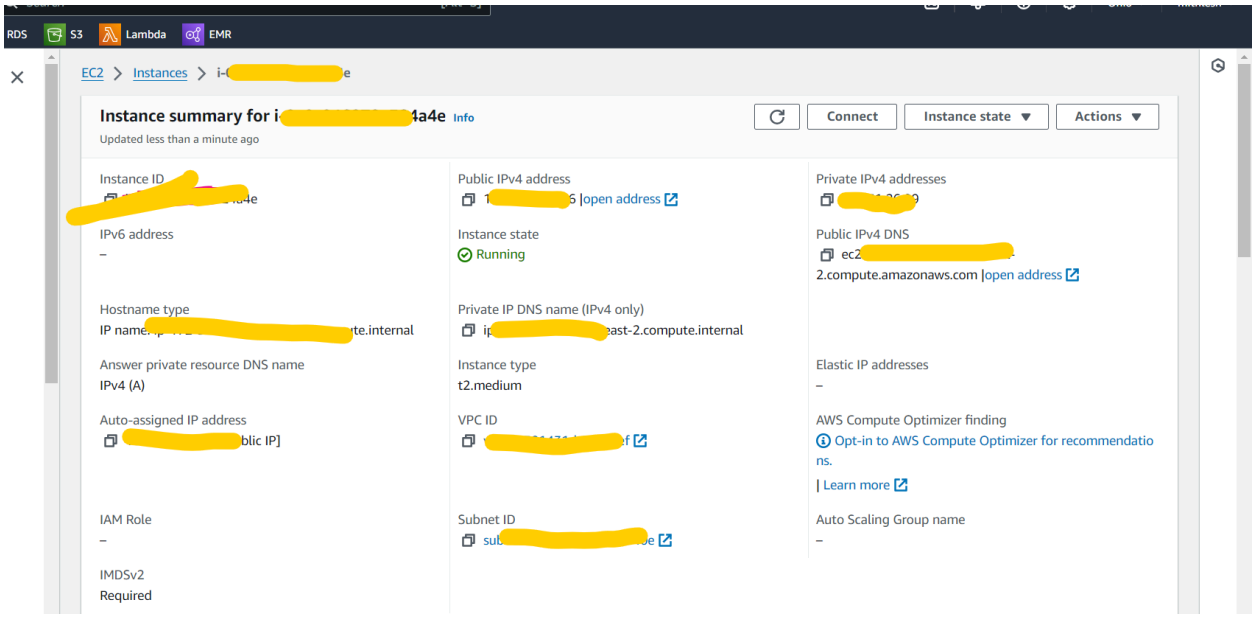
Airflow examples

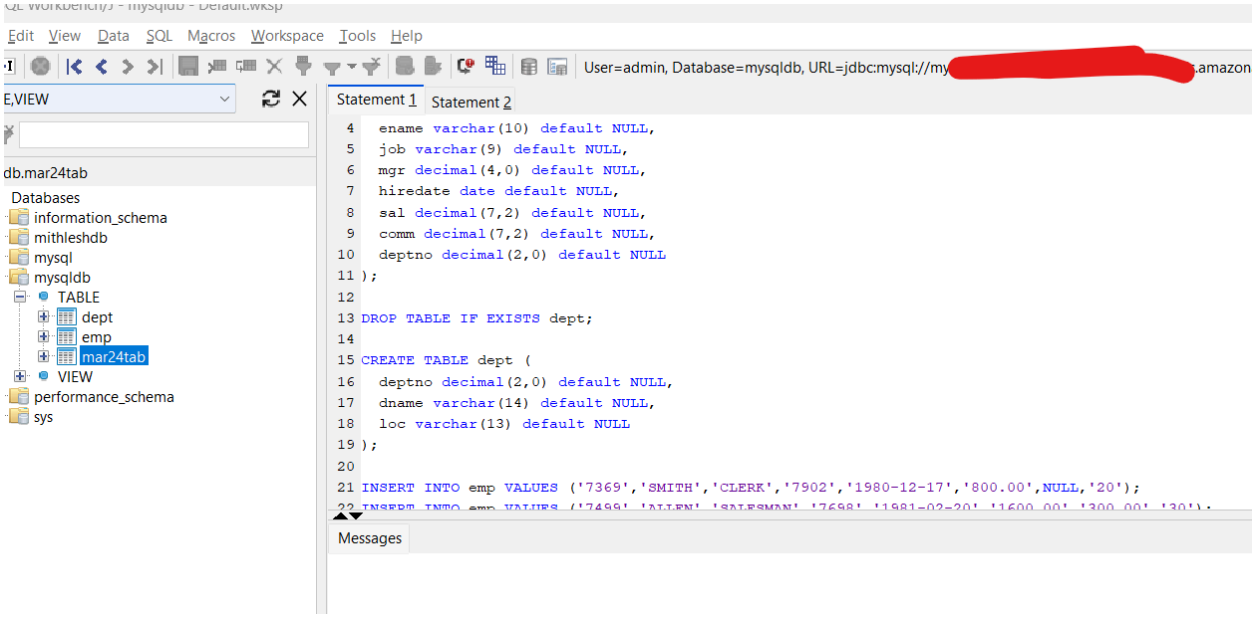
Create ec2 instance:



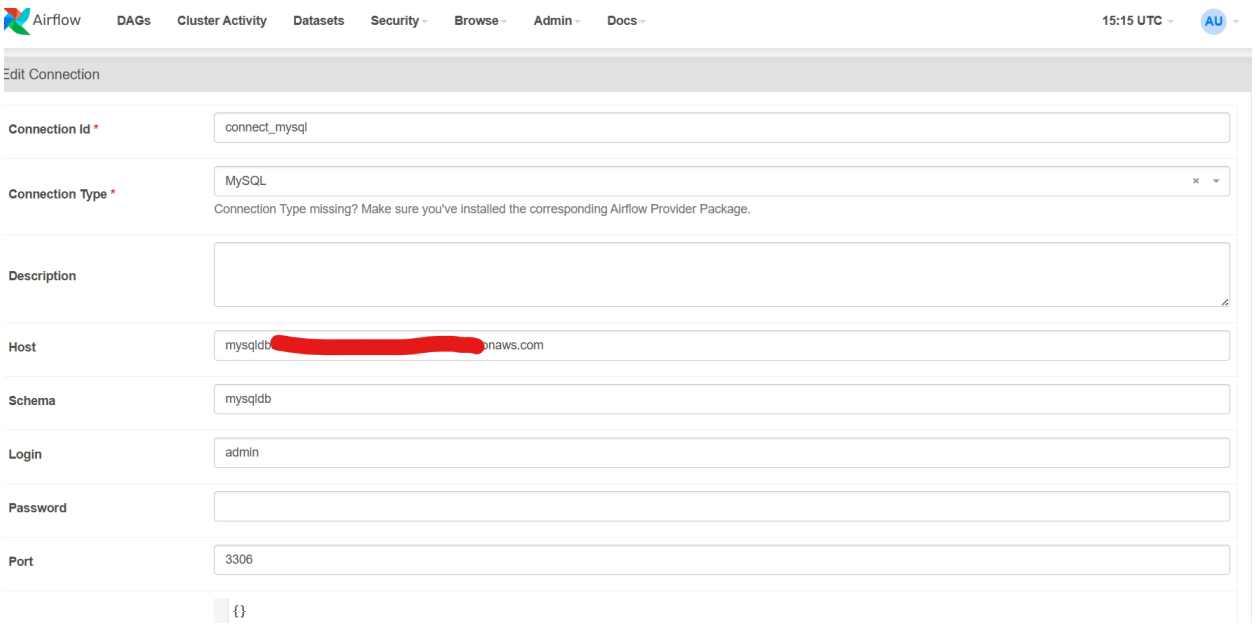
Update security groups inbound rules
8080 customip

Updated instance memory config to 10gb

Update connections as below:
This should match the RDS mysqldb connection as below:

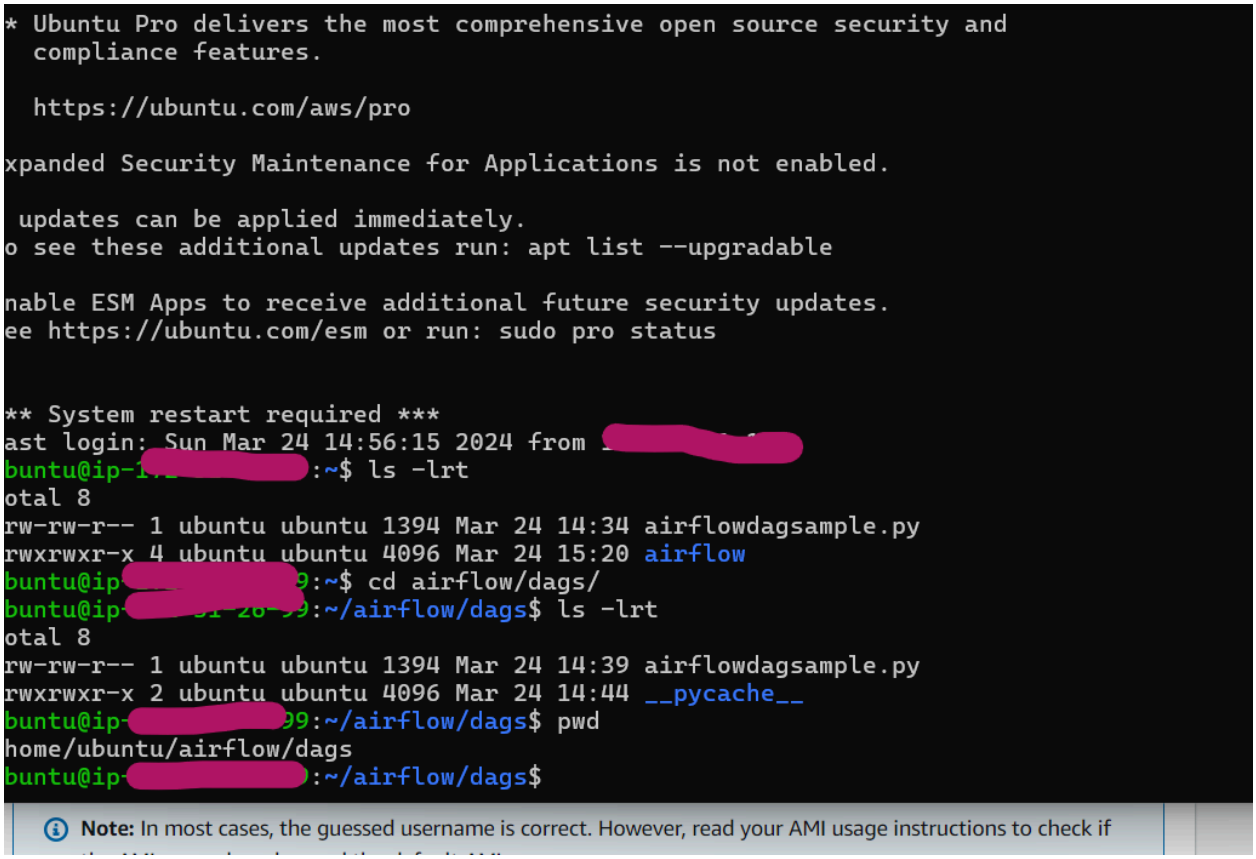


DAG connection:

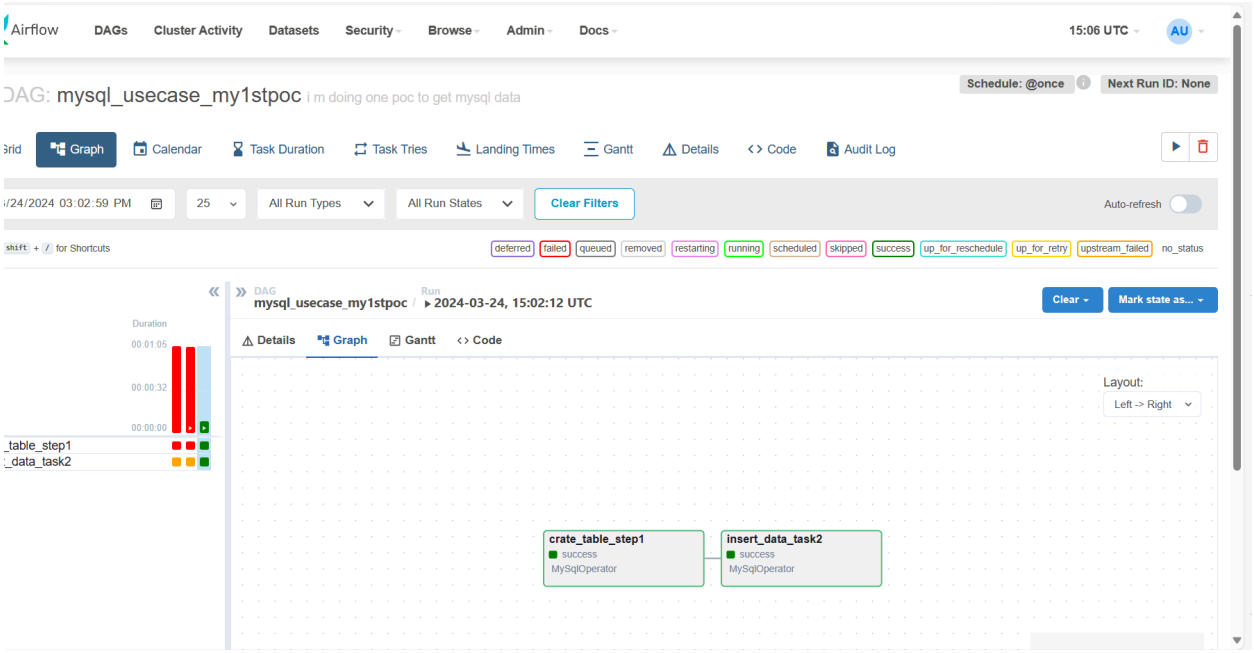


In EC2: terminal

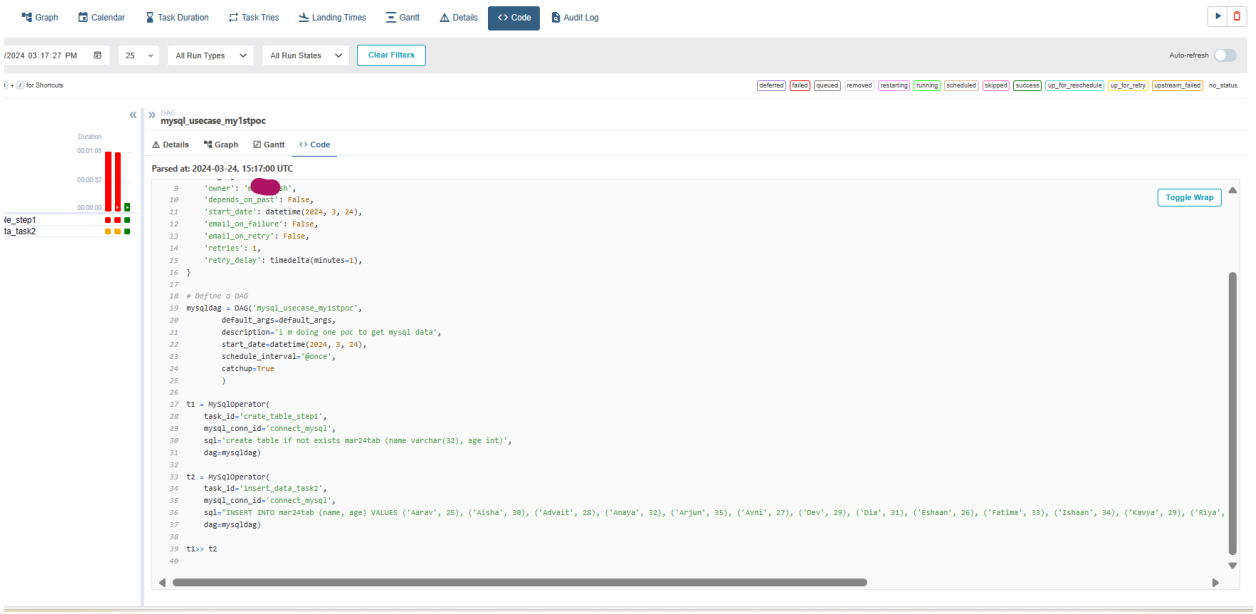
Create a dag folder under airflow



Check the DAG graph to identify process:


Airflow
DAGs
Cluster Activity
Datasets
Security
Browse
Admin
Docs
15:06 UTC
AU
DAG: mysql_usecase_my1stpoc i m doing one poc to get mysql data
Schedule: @once
Next Run ID: None
Graph
Calendar
Task Duration
Task Tries
Landing Times
Gantt
Details
Code
Audit Log
Clear Filters
Auto-refresh
Clear
Mark state as...
Layout:
Left -> Right
crate_table_step1
success
MySqlOperator
insert_data_task2
success
MySqlOperator

Code: DAG


Graph
Calendar
Task Duration
Task Tries
Landing Times
Gantt
Details
Code
Audit Log
Clear Filters
Auto-refresh
mysql_usecase_my1stpoc
Details
Graph
Gantt
Code
Parsed at: 2024-03-24, 15:17:00 UTC
Toggle Wrap
'depends_on_past': False,
'start_date': datetime(2024, 3, 24),
'email_on_failure': False,
'email_on_retry': False,
'retries': 1,
'retry_delay': timedelta(minutes=1),
}
# Define a DAG
mysqldag = DAG('mysql_usecase_my1stpoc',
default_args=default_args,
description='i m doing one poc to get mysql data',
start_date=datetime(2024, 3, 24),
schedule_interval='@once',
catchup=True
)
t1 = MySqlOperator(
task_id='crate_table_step1',
mysql_conn_id='connect_mysql',
sql='create table if not exists mar24tab (name varchar(32), age int)',
dag=mysqldag)
t2 = MySqlOperator(
task_id='insert_data_task2',
mysql_conn_id='connect_mysql',
sql="INSERT INTO mar24tab (name, age) VALUES ('Aarav', 25), ('Aisha', 30), ('Advait', 28), ('Anaya', 32), ('Arjun', 35), ('Avni', 27), ('Dev', 29), ('Dia', 31), ('Eshaan', 26), ('Fatima', 33), ('Ishaan', 34), ('Kavya', 29), ('Riya',
dag=mysqldag)
t1>> t2

Finally it should load data from into RDS mysqldb using airflow from DAG

1) Make sure update inbound rules to 8080 port
2) Update cfg file loadexamples to TRUE in ec2 terminal under airflow folder
3) During installation make sure
4) sudo apt install default-libmysqlclient-dev -y
   pip install apache-airflow-providers-ssh
   pip install apache-airflow-providers-mysql
   airflow standalone