

Install AirFlow

1.Install Apache Airflow on Ubuntu on AWS

Airflow is one of the most popular workflow management solution, it author, schedule and monitor workflows.

Requirements

- Python 2.7
- pip
- Ubuntu 16

2.Install Python, pip, Airflow and dependencies

First we need to install `python` and the python package management tool `pip`.

By default , for ubuntu python package are already been installed.

To install dependencies follow below steps.

```
#####
```

```
sudo apt-get install python-setuptools
```

`pip` is a Python package management tool. We'll be using this for installing packages required in Airflow.

To install `pip`:

```
sudo apt-get install python-pip
```

Note: It is recommended to use latest `pip` version. For upgrading `pip` version, use `pip upgrade` command given below.

```
sudo pip install --upgrade pip
```

3. Installing PostgreSQL for Airflow

Airflow is shipped with a `sqlite` database backend. This database system will not be able to run data pipeline on webUI. We would require to have more powerful database system like PostgreSQL, it is an open source database management system, that comes with robust feature set, data integrity and extensibility. We will install PostgreSQL and configure it to use with Airflow.

```
sudo apt-get install postgresql postgresql-contrib
```

create a database for airflow and grant access to a sudo user. Lets access to `psql`, a command line tool for Postgres.

```
sudo -u postgres psql
```

```
root@ip-172-31-35-91:~# sudo -u postgres psql
could not change directory to "/root": Permission denied
psql (9.5.19)
Type "help" for help.

postgres=#
```

After logging in successfully, we will get psql prompt (postgres=#). We will create a new user and provide privileges to it.

```
CREATE ROLE ubuntu;
CREATE DATABASE airflow;
GRANT ALL PRIVILEGES ON database airflow to ubuntu;
ALTER ROLE ubuntu SUPERUSER;
ALTER ROLE ubuntu CREATEDB;
GRANT ALL PRIVILEGES ON ALL TABLES IN SCHEMA public to ubuntu;
```

Now connect to airflow database and get connection information.

```
postgres=# \c airflow
```

```
postgres=# \c airflow
You are now connected to database "airflow" as user "postgres".
airflow=#
```

After successful connection, prompt will be changed to airflow=#. We will verify this by fetching connection info

```
airflow=# \conninfo
```

```
airflow=# \conninfo
You are connected to database "airflow" as user "postgres" via socket in "/var/run/postgresql" at port "5432".
airflow=#
```

We'll change settings in pg_hba.conf file for required configuration as per Airflow. You can run command SHOW hba_file;

```
vi /etc/postgresql/9.5/main/pg_hba.conf
```

change ipv4 address to 0.0.0.0/0

And restart PostgreSQL to load changes.

```
sudo service postgresql restart
```

4. Install Airflow

As PostgreSQL is already installed and configured. Next, We will install Airflow and configure it.

Set AIRFLOW_HOME environment variable to ~/airflow.

```
export AIRFLOW_HOME=~/airflow
```

First install the following dependencies:

- `sudo apt-get install libmysqlclient-dev` (dependency for airflow[mysql] package)

```
root@ip-172-31-32-11:~# apt-get install libmysqlclient-dev
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  libmysqlclient20 mysql-common libibsig-dev
The following NEW packages will be installed:
  libmysqlclient-dev libmysqlclient20 mysql-common libibsig-dev
0 upgraded, 4 newly installed, 0 to remove and 3 not upgraded.
Need to get 2,154 kB of archives.
After this operation, 12.0 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://us-east-1-ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 mysql-common all 5.7.27-0ubuntu0.16.04.1 [14.7
Get:2 http://us-east-1-ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libmysqlclient20 amd64 5.7.27-0ubuntu0.16.04.1
Get:3 http://us-east-1-ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libibsig-dev amd64 1:1.2.8.dfsg-2ubuntu4.1 [148
Get:4 http://us-east-1-ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libmysqlclient-dev amd64 5.7.27-0ubuntu0.16.04
Fetched 2,154 kB in 0s (29.3 MB/s)
Selecting previously unselected package mysql-common.
(Reading database ... 120541 files and directories currently installed.)
Preparing to unpack .../mysql-common_5.7.27-0ubuntu0.16.04.1_all.deb ...
Unpacking mysql-common (5.7.27-0ubuntu0.16.04.1) ...
Selecting previously unselected package libmysqlclient20:amd64.
Preparing to unpack .../libmysqlclient20_5.7.27-0ubuntu0.16.04.1_amd64.deb ...
Unpacking libmysqlclient20:amd64 (5.7.27-0ubuntu0.16.04.1) ...
Selecting previously unselected package libibsig-dev:amd64.
Preparing to unpack .../libibsig-dev_1:1.2.8.dfsg-2ubuntu4.1_amd64.deb ...
Unpacking libibsig-dev:amd64 (1:1.2.8.dfsg-2ubuntu4.1) ...
Selecting previously unselected package libmysqlclient-dev.
Preparing to unpack .../libmysqlclient-dev_5.7.27-0ubuntu0.16.04.1_amd64.deb ...
Unpacking libmysqlclient-dev (5.7.27-0ubuntu0.16.04.1) ...
Processing triggers for libc-bin (2.23-0ubuntu1) ...
Processing triggers for man-db (2.7.5-1) ...
Setting up mysql-common (5.7.27-0ubuntu0.16.04.1) ...
update-alternatives: using /etc/mysql/my.cnf fallback to provide /etc/mysql/my.cnf (my.cnf) in auto mode
Setting up libmysqlclient20:amd64 (5.7.27-0ubuntu0.16.04.1) ...
Setting up libibsig-dev:amd64 (1:1.2.8.dfsg-2ubuntu4.1) ...
Setting up libmysqlclient-dev (5.7.27-0ubuntu0.16.04.1) ...
Processing triggers for libc-bin (2.23-0ubuntu1) ...
```

- `sudo apt-get install libssl-dev` (dependency for airflow[cryptography] package)

```

root@ip-172-31-35-91:~$ apt-get install libssl-dev
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  libssl-doc
The following NEW packages will be installed:
  libssl-dev libssl-doc
0 upgraded, 2 newly installed, 0 to remove and 3 not upgraded.
Need to get 2,421 kB of archives.
After this operation, 10.1 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libssl-dev amd64 1.0.2g-lubuntu4.15 [1,344 kB]
Get:2 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libssl-doc all 1.0.2g-lubuntu4.15 [1,077 kB]
Fetched 2,421 kB in 0s (23.5 MB/s)
Selecting previously unselected package libssl-dev:amd64.
(Reading database ... 120716 files and directories currently installed.)
Preparing to unpack .../libssl-dev_1.0.2g-lubuntu4.15_amd64.deb ...
Unpacking libssl-dev:amd64 (1.0.2g-lubuntu4.15) ...
Selecting previously unselected package libssl-doc.
Preparing to unpack .../libssl-doc_1.0.2g-lubuntu4.15_all.deb ...
Unpacking libssl-doc (1.0.2g-lubuntu4.15) ...
Processing triggers for man-db (2.7.5-1) ...
Setting up libssl-dev:amd64 (1.0.2g-lubuntu4.15) ...
Setting up libssl-doc (1.0.2g-lubuntu4.15) ...
root@ip-172-31-35-91:~$

```

- `sudo apt-get install libkrb5-dev` (dependency for airflow[kerbero] package)

```

root@ip-172-31-35-91:~$ apt-get install libssl-dev
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  libssl-doc
The following NEW packages will be installed:
  libssl-dev libssl-doc
0 upgraded, 2 newly installed, 0 to remove and 3 not upgraded.
Need to get 2,421 kB of archives.
After this operation, 10.1 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libssl-dev amd64 1.0.2g-lubuntu4.15 [1,344 kB]
Get:2 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libssl-doc all 1.0.2g-lubuntu4.15 [1,077 kB]
Fetched 2,421 kB in 0s (23.5 MB/s)
Selecting previously unselected package libssl-dev:amd64.
(Reading database ... 120716 files and directories currently installed.)
Preparing to unpack .../libssl-dev_1.0.2g-lubuntu4.15_amd64.deb ...
Unpacking libssl-dev:amd64 (1.0.2g-lubuntu4.15) ...
Selecting previously unselected package libssl-doc.
Preparing to unpack .../libssl-doc_1.0.2g-lubuntu4.15_all.deb ...
Unpacking libssl-doc (1.0.2g-lubuntu4.15) ...
Processing triggers for man-db (2.7.5-1) ...
Setting up libssl-dev:amd64 (1.0.2g-lubuntu4.15) ...
Setting up libssl-doc (1.0.2g-lubuntu4.15) ...
root@ip-172-31-35-91:~$ apt-get install libkrb5-dev
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  comerr-dev krb5-multidev libkrb5-dev libkrb5-doc libkrb5-kdc
Suggested packages:
  doc-base krb5-doc krb5-user
The following NEW packages will be installed:
  comerr-dev krb5-multidev libkrb5-dev libkrb5-doc libkrb5-kdc
0 upgraded, 5 newly installed, 0 to remove and 3 not upgraded.
Need to get 343 kB of archives.

```

- `sudo apt-get install libsasl2-dev` (dependency for airflow[hive] package):

```

Processing triggers for libc-bin (2.23-0ubuntu1) ...
root@ip-172-31-35-91:~$ apt-get install libsasl2-dev
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following NEW packages will be installed:
  libsasl2-dev
0 upgraded, 1 newly installed, 0 to remove and 3 not upgraded.
Need to get 254 kB of archives.
After this operation, 831 kB of additional disk space will be used.
Get:1 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 libsasl2-dev amd64 2.1.26.dfsg1-14ubuntu0.1 [254 kB]
Fetched 254 kB in 0s (2,266 kB/s)
Selecting previously unselected package libsasl2-dev.
(Reading database ... 122659 files and directories currently installed.)
Preparing to unpack .../libsasl2-dev_2.1.26.dfsg1-14ubuntu0.1_amd64.deb ...
Unpacking libsasl2-dev (2.1.26.dfsg1-14ubuntu0.1) ...
Processing triggers for man-db (2.7.5-1) ...
Setting up libsasl2-dev (2.1.26.dfsg1-14ubuntu0.1) ...
root@ip-172-31-35-91:~$

```

After installing these dependencies, we can install airflow and its packages. (You can modify these packages depending on need. Celery and RabbitMQ are needed to use the Web-based GUI)

```
pip install apache-airflow
```

```
root@ip-172-31-35-91:~# pip install apache-airflow
Collecting apache-airflow
  Downloading https://files.pythonhosted.org/packages/74/cb/3a4bcfdb75897607c0c4d9ec17f3c3d6eb49739302dfa59bd0b47c20a67/apache-airflow-1.10.6-py2.py3-none-any.whl (12.6MB)
100% |#####| 4.5MB 281kB/s
Collecting tzlocal<2.0.0,>=1.4 (from apache-airflow)
  Downloading https://files.pythonhosted.org/packages/cb/89/e3687d3ed99bc882793f82634e9824e62499fdcdc4b1ae39e211c5b05017/tzlocal-1.5.1.tar.gz (12.6MB)
Collecting Jinja2<2.11.0,>=2.10.1 (from apache-airflow)
  Downloading https://files.pythonhosted.org/packages/65/e0/eb35e762802015cab1ccce04e8a277b03fd8e53da3ec3106882ec42558b/Jinja2-2.10.3-py2.py3-none-any.whl (12.6MB)
100% |#####| 133kB 8.7MB/s
Collecting markdown<3.0,>=2.5.2 (from apache-airflow)
  Downloading https://files.pythonhosted.org/packages/6d/7d/488b90f470b96531a3f5f88cf12a93332f543dbab13c423a5e7ce96a0493/Markdown-2.6.11-py2.py3-none-any.whl (12.6MB)
100% |#####| 81kB 11.2MB/s
Collecting pygments<3.0,>=2.0.1 (from apache-airflow)
  Downloading https://files.pythonhosted.org/packages/5c/73/1dfa428150e3ccb0fa3e68db406e5be48698f2a979ccbcc795f28f44048/Pygments-2.4.2-py2.py3-none-any.whl (12.6MB)
100% |#####| 890kB 1.5MB/s
Collecting pandas<1.0.0,>=0.17.1 (from apache-airflow)
  Downloading https://files.pythonhosted.org/packages/b7/93/b544dd08092b457d88e10fc1e0989d9397fd32ca936fdcfbb2584178dd2b/pandas-0.25.3.tar.gz (12.6MB)
100% |#####| 12.6MB 99kB/s
Complete output from command python setup.py egg_info:
```

for other subpackages like celery, async, crypto, rabbitmq etc., you can check [apache airflow installation page](#).

After successfully installing airflow, we will initialise Airflow's database

```
airflow initdb
```

```
root@ip-172-31-35-91:~# airflow initdb
DEPRECATION: Python 2.7 will reach the end of its life on January 1st, 2020. Airflow 1.10 will be the last release series to support Python 2
DB: sqlite:///root/airflow/airflow.db
[2019-11-04 00:11:35.110] (0.0.0.0) INFO - Creating tables
INFO [alembic.runtime.migration] Context impl SQLiteImpl.
INFO [alembic.runtime.migration] Will assume non-transactional DDL.
INFO [alembic.runtime.migration] Running upgrade 0ba246ebd0c1 -> eba246ebd0c1, current schema
INFO [alembic.runtime.migration] Running upgrade eba246ebd0c1 -> 1507a7289a2f, create is_encrypted
/usr/local/lib/python2.7/dist-packages/alembic/dialects/sqlite.py:199: UserWarning: Skipping unsupported ALTER for creation of implicit constraint
"Skipping unsupported ALTER for "
INFO [alembic.runtime.migration] Running upgrade 1507a7289a2f -> 13eb51ff1627, maintain history for compatibility with earlier migrations
INFO [alembic.runtime.migration] Running upgrade 13eb51ff1627 -> 338e90f94d61, More logging into task_instance
INFO [alembic.runtime.migration] Running upgrade 338e90f94d61 -> 52d714495f0, job id indices
INFO [alembic.runtime.migration] Running upgrade 52d714495f0 -> 952898887f04, Adding extra to Log
INFO [alembic.runtime.migration] Running upgrade 952898887f04 -> 1b36cecf0b76e, add dagrun
INFO [alembic.runtime.migration] Running upgrade 1b36cecf0b76e -> 2e541a1dcfcd, task_duration
INFO [alembic.runtime.migration] Running upgrade 2e541a1dcfcd -> 40e67319e3a9, dagrun_config
INFO [alembic.runtime.migration] Running upgrade 40e67319e3a9 -> 341833c1c74b, add password column to user
INFO [alembic.runtime.migration] Running upgrade 341833c1c74b -> 4446e08589, dagrun_start_end
INFO [alembic.runtime.migration] Running upgrade 4446e08589 -> 1bc73705a13e, Add notification_sent column to sla_miss
INFO [alembic.runtime.migration] Running upgrade 1bc73705a13e -> 3ba6a76cf094, Add a column to track the encryption state of the 'extra' field in connection
INFO [alembic.runtime.migration] Running upgrade 3ba6a76cf094 -> 1948acfc0943, add is_encrypted column to variable table
INFO [alembic.runtime.migration] Running upgrade 1948acfc0943 -> 2e22a8bdf20, rename user table
INFO [alembic.runtime.migration] Running upgrade 2e22a8bdf20 -> 211e954da130, add TI state index
INFO [alembic.runtime.migration] Running upgrade 211e954da130 -> 44de9808f6d9, add task fails journal table
INFO [alembic.runtime.migration] Running upgrade 44de9808f6d9 -> 22ca12b8f418, add dag_state table
INFO [alembic.runtime.migration] Running upgrade 22ca12b8f418 -> 4a02fa1234f1, Add fractional seconds to mysql tables
INFO [alembic.runtime.migration] Running upgrade 4a02fa1234f1 -> 8504031e801b, add dag task indices
INFO [alembic.runtime.migration] Running upgrade 8504031e801b -> 3e1d17757c7a, add pid field to TaskInstance
INFO [alembic.runtime.migration] Running upgrade 3e1d17757c7a -> 127828f23fa7, Add dag_id/state index on dag_run table
INFO [alembic.runtime.migration] Running upgrade 127828f23fa7 -> c01e6423dc07, add max_retry column to task instance
INFO [alembic.runtime.migration] Running upgrade c01e6423dc07 -> bdaa743a6d14, Make snow value column a large binary
INFO [alembic.runtime.migration] Running upgrade bdaa743a6d14 -> 9476548f1dff, add ti job_id index
INFO [alembic.runtime.migration] Running upgrade 9476548f1dff -> d2ae31099d61, Increase text size for MySQL (not relevant for other DBs' text types)
INFO [alembic.runtime.migration] Running upgrade d2ae31099d61 -> 6e2a74e0fc9f, Add time zone awareness
INFO [alembic.runtime.migration] Running upgrade 6e2a74e0fc9f -> 33ae117a1f4, Kubernetes_resource_checkpointing
INFO [alembic.runtime.migration] Running upgrade 33ae117a1f4 -> 270ca30d7c24, Kubernetes_resource_checkpointing
INFO [alembic.runtime.migration] Running upgrade 270ca30d7c24 -> 86770d121500, add Kubernetes scheduler uniqueness
```

To set up the first-time configs. An `airflow.cfg` file is generated in the airflow home directory.

```
#####
home directory location

/root/airflow

#####
```

```
root@ip-172-31-35-91:~# cd /root/airflow
root@ip-172-31-35-91:~/airflow# ll
total 124
drwxr-xr-x 3 root root 4096 Nov  6 05:11 ./
drwx----- 8 root root 4096 Nov  6 05:16 ../
-rw-r--r-- 1 root root 31200 Nov  6 05:11 airflow.cfg
-rw-r--r-- 1 root root 72704 Nov  6 05:11 airflow.db
drwxr-xr-x 3 root root 4096 Nov  6 05:11 logs/
-rw-r--r-- 1 root root 2501 Nov  6 05:11 unittests.cfg
root@ip-172-31-35-91:~/airflow#
```

We should open it with a text editor, and change some configurations in the [core] section:

- For the executor, we should use CeleryExecutor instead of SequentialExecutor if we want to run the pipeline in the webUI:

```
# The executor class that airflow should use. Choices include
# SequentialExecutor, LocalExecutor, CeleryExecutor, DaskExecutor, KubernetesExecutor
#executor = SequentialExecutor
executor = CeleryExecutor
```

- For the backend DB connection, we should pass along the connection info of the postgresql database airflow we just created:

```
#####

sql_alchemy_conn =
postgresql+psycopg2://airflow:airflow@ip-172-31-35-91:5432/airflow

load_examples = false
```

After doing all these setting just save your configuration and exit.

For Loading new configurations, we should run

airflow initdb

5.Installing Rabbitmq

Rabbitmq is a message broker, that required to rerun airflow dags with celery. Rabbitmq can be installed with following command.

```
sudo apt-get install rabbitmq-server
```

```

$ curl -fsSL https://bit.ly/rabbitmq-debian | sudo apt-get install rabbitmq-server
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  erlang-asn1 erlang-base erlang-crypto erlang-diameter erlang-edoc erlang-eldap erlang-erl-smp erlang-eunit erlang-inets erlang-ldap erlang-mnesia erlang-os-mon erlang-parsetools erlang-percept erlang-public-key erlang-runtime-tools erlang-sasl erlang-sh erlang-ssl erlang-tls erlang-tools erlang-wctool erlang-xmerl libhidn0 libiodbc1 libncurses6
Suggested packages:
  erlang erlang-manpages erlang-doc xsltpdf grok erlang-to-java erlang-observer libmyxos0 odbc-postgresql tdsodbc unixodbc-bin libxslt-erlang
The following NEW packages will be installed:
  erlang-asn1 erlang-base erlang-crypto erlang-diameter erlang-edoc erlang-eldap erlang-erl-smp erlang-eunit erlang-inets erlang-ldap erlang-mnesia erlang-os-mon erlang-parsetools erlang-percept erlang-public-key erlang-runtime-tools erlang-sasl erlang-sh erlang-ssl erlang-tls erlang-tools erlang-wctool erlang-xmerl libhidn0 libiodbc1 libncurses6
0 upgraded, 30 newly installed, 0 to remove and 3 not upgraded.
Need to get 23.6 MB of archives.
After this operation, 41.3 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-base amd64 1:18.3~dfsg-lubuntu3.1 [7,002 KB]
Get:2 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-asn1 amd64 1:18.3~dfsg-lubuntu3.1 [709 KB]
Get:3 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-mnesia amd64 1:18.3~dfsg-lubuntu3.1 [464 KB]
Get:4 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-runtime-tools amd64 1:18.3~dfsg-lubuntu3.1 [144 KB]
Get:5 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-crypto amd64 1:18.3~dfsg-lubuntu3.1 [106 KB]
Get:6 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-public-key amd64 1:18.3~dfsg-lubuntu3.1 [534 KB]
Get:7 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-ssl amd64 1:18.3~dfsg-lubuntu3.1 [688 KB]
Get:8 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-inets amd64 1:18.3~dfsg-lubuntu3.1 [174 KB]
Get:9 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-corba amd64 1:18.3~dfsg-lubuntu3.1 [2,233 KB]
Get:10 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-syntax-tools amd64 1:18.3~dfsg-lubuntu3.1 [522 KB]
Get:11 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-diameter amd64 1:18.3~dfsg-lubuntu3.1 [642 KB]
Get:12 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-xmerl amd64 1:18.3~dfsg-lubuntu3.1 [958 KB]
Get:13 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-edoc amd64 1:18.3~dfsg-lubuntu3.1 [282 KB]
Get:14 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-eldap amd64 1:18.3~dfsg-lubuntu3.1 [39,9 KB]
Get:15 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-erl-smp amd64 1:18.3~dfsg-lubuntu3.1 [112 KB]
Get:16 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-eunit amd64 1:18.3~dfsg-lubuntu3.1 [134 KB]
Get:17 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 erlang-io amd64 1:18.3~dfsg-lubuntu3.1 [816 KB]
Get:18 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial/main amd64 libhidn0 amd64 2.4.6-0.1 [28.3 KB]
Get:19 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial/main amd64 libiodbc1 amd64 2.3.1-4.1 [180 KB]

```

change configuration `NODE_IP_ADDRESS=0.0.0.0` in configuration file located at **`/etc/rabbitmq/rabbitmq-env.conf`**

```
# Defaults to rabbit. This can be useful if you want to run more than one node
# per machine - RABBITMQ_NODENAME should be unique per erlang-node-and-machine
# combination. See the clustering on a single machine guide for details:
# http://www.rabbitmq.com/clustering.html#single-machine
#NODENAME=rabbit

# By default RabbitMQ will bind to all interfaces, on IPv4 and IPv6 if
# available. Set this if you only want to bind to one network interface or#
# address family.
#NODE_IP_ADDRESS=127.0.0.1
NODE_IP_ADDRESS=0.0.0.0
# Defaults to 5672.
#NODE_PORT=5672

~
```

Then start rabbitmq service

```
sudo service rabbitmq-server start
```

6.Installing Celery

Celery is a python api for rabbitmq. We can install celery using pip

```
sudo pip install celery
```

7.Starting Airflow

All the required installation and configuration is done. We will create a dags folder in airflow home directory .i.e; at /home/ubuntu/airflow/location

```
mkdir -p /home/ubuntu/airflow/dags/
```

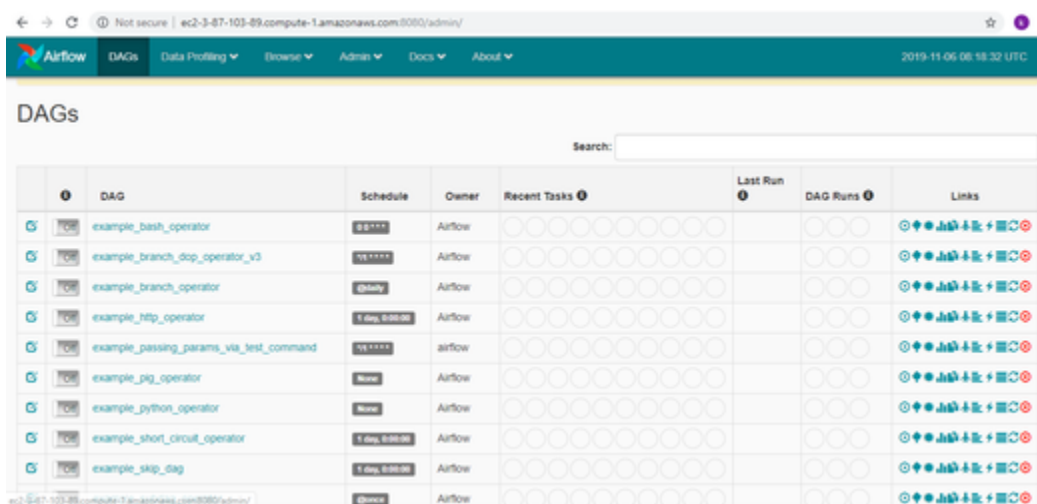
and then we'll start all airflow services to up airflow webUI

```
airflow webserver
airflow scheduler
airflow worker
```

If you want to up airflow continuously up, you should run these command with -D flag like
airflow webserver -D, this will run airflow as a Daemon in background. You required to do it for all the services, If you want to keep these services continuously up.

- Once webserver is up once with ip address with default port 8080
<http://ec2-3-87-103-89.compute-1.amazonaws.com:8080/admin/>

Dashboard for airflow:



The screenshot shows the Airflow web interface at the URL <http://ec2-3-87-103-89.compute-1.amazonaws.com:8080/admin/>. The interface has a teal header with the Airflow logo and navigation links: DAGs, Data Profiling, Browse, Admin, Docs, and About. The main content area is titled "DAGs" and includes a search bar. Below the search bar is a table listing various DAGs. Each row in the table contains a DAG icon, the DAG name, its schedule, owner, recent tasks, last run, DAG runs, and a set of links for further actions.

	DAG	Schedule	Owner	Recent Tasks	Last Run	DAG Runs	Links
	example_bash_operator	@@* * * *	Airflow				
	example_branch_dag_operator_v3	@@* * * *	Airflow				
	example_branch_operator	@daily	Airflow				
	example_http_operator	1 day, @00:00	Airflow				
	example_passing_params_via_test_command	@@* * * *	airflow				
	example_pig_operator	None	Airflow				
	example_python_operator	None	Airflow				
	example_short_circuit_operator	1 day, @00:00	Airflow				
	example_skip_dag	1 day, @00:00	Airflow				
	example_trigger_dag	@once	Airflow				