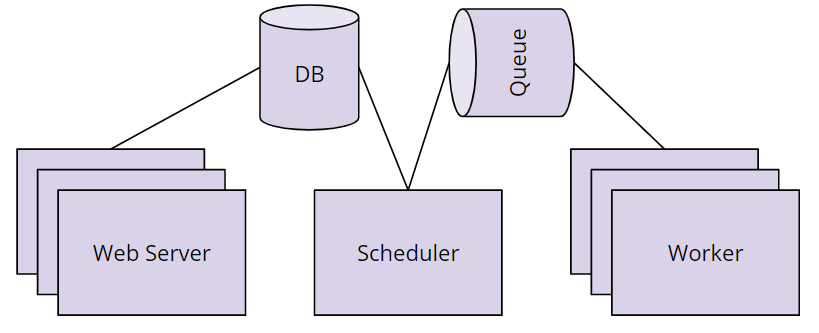
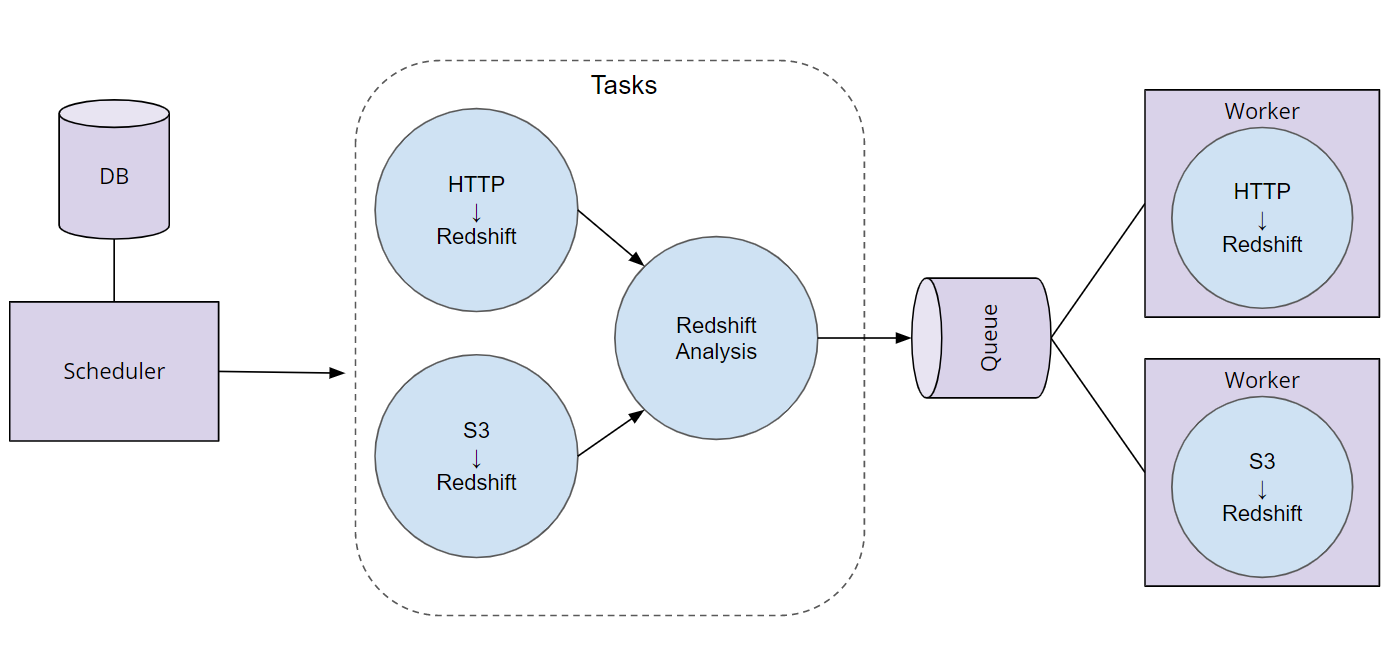
**Components of Airflow**

**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/1fa8af52-682d-440b-a391-16183cb0df08/lessons/af88b28f-ceb2-4239-8262-2a55a3b67bd5/concepts/f5fd1f7f-6eac-4551-a490-48414e93ad41)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/1fa8af52-682d-440b-a391-16183cb0df08/lessons/af88b28f-ceb2-4239-8262-2a55a3b67bd5/concepts/f5fd1f7f-6eac-4551-a490-48414e93ad41)**

* **Scheduler** orchestrates the execution of jobs on a trigger or schedule. The Scheduler chooses how to prioritize the running and execution of tasks within the system. You can learn more about the Scheduler from the official [**Apache Airflow documentation**](https://airflow.apache.org/scheduler.html).
* **Work Queue** is used by the scheduler in most Airflow installations to deliver tasks that need to be run to the **Workers**.
* **Worker** processes execute the operations defined in each DAG. In most Airflow installations, workers pull from the **work queue** when it is ready to process a task. When the worker completes the execution of the task, it will attempt to process more work from the **work queue** until there is no further work remaining. When work in the queue arrives, the worker will begin to process it.
* **Database** saves credentials, connections, history, and configuration. The database, often referred to as the *metadata database*, also stores the state of all tasks in the system. Airflow components interact with the database with the Python ORM, **[SQLAlchemy](https://www.sqlalchemy.org/" \t "_blank)**.
* **Web Interface** provides a control dashboard for users and maintainers. Throughout this course you will see how the web interface allows users to perform tasks such as stopping and starting DAGs, retrying failed tasks, configuring credentials, The web interface is built using the [**Flask web-development microframework**](http://flask.pocoo.org/).

**How Airflow Works**

**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/1fa8af52-682d-440b-a391-16183cb0df08/lessons/af88b28f-ceb2-4239-8262-2a55a3b67bd5/concepts/f5fd1f7f-6eac-4551-a490-48414e93ad41)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/1fa8af52-682d-440b-a391-16183cb0df08/lessons/af88b28f-ceb2-4239-8262-2a55a3b67bd5/concepts/f5fd1f7f-6eac-4551-a490-48414e93ad41)**

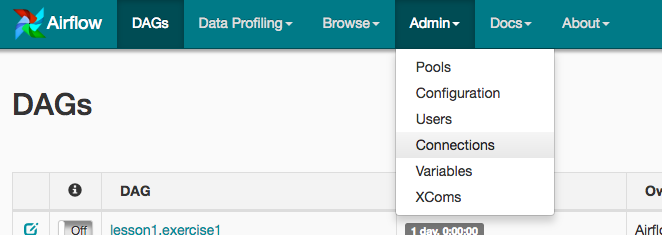
**Order of Operations For an Airflow DAG**

* The Airflow Scheduler starts DAGs based on time or external triggers.
* Once a DAG is started, the Scheduler looks at the steps within the DAG and determines which steps can run by looking at their dependencies.
* The Scheduler places runnable steps in the queue.
* Workers pick up those tasks and run them.
* Once the worker has finished running the step, the final status of the task is recorded and additional tasks are placed by the scheduler until all tasks are complete.
* Once all tasks have been completed, the DAG is complete.

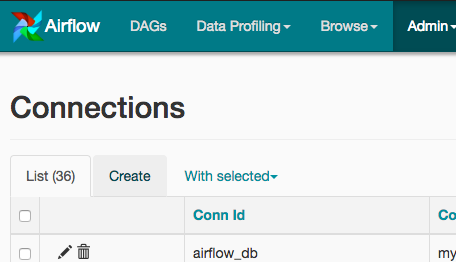
# Add Airflow Connections

Here, we'll use Airflow's UI to configure your AWS credentials and connection to Redshift.

1. To go to the Airflow UI:
   * You can use the Project Workspace here and click on the blue **Access Airflow** button in the bottom right.
   * If you'd prefer to run Airflow locally, open [**http://localhost:8080**](http://localhost:8080/) in Google Chrome (other browsers occasionally have issues rendering the Airflow UI).
2. Click on the **Admin** tab and select **Connections**.

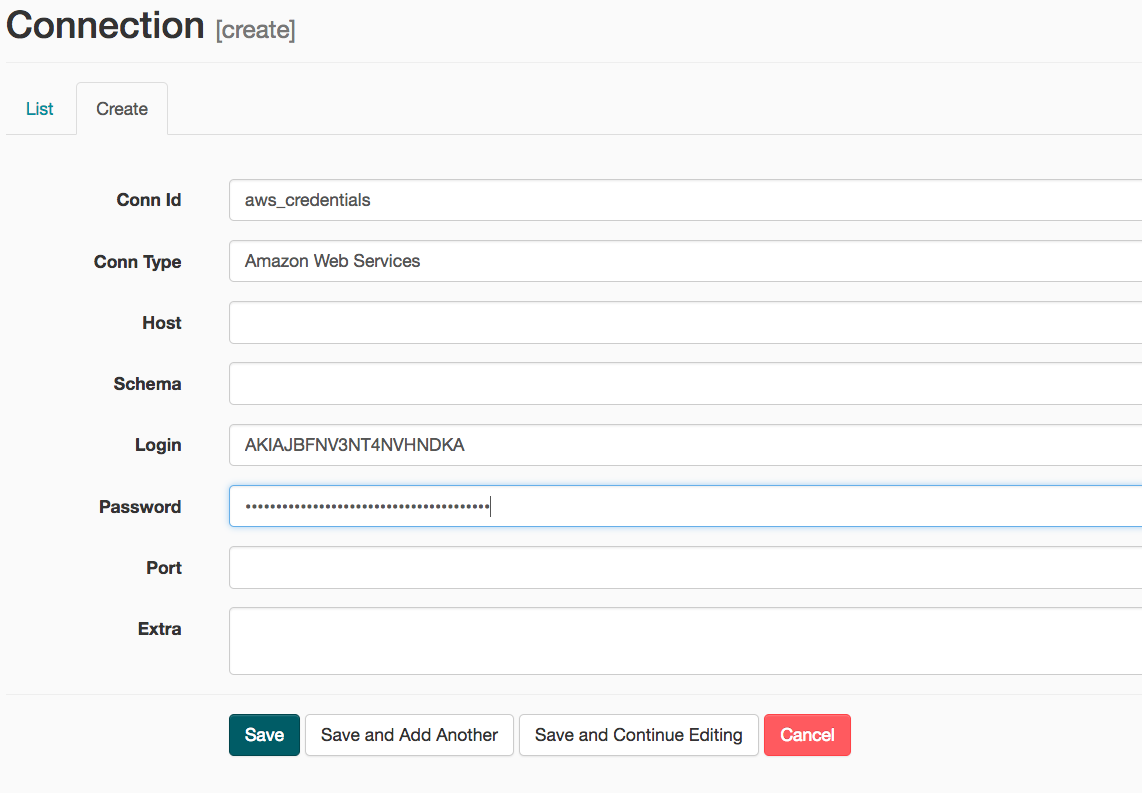
**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)**

1. Under **Connections**, select **Create**.

**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)**

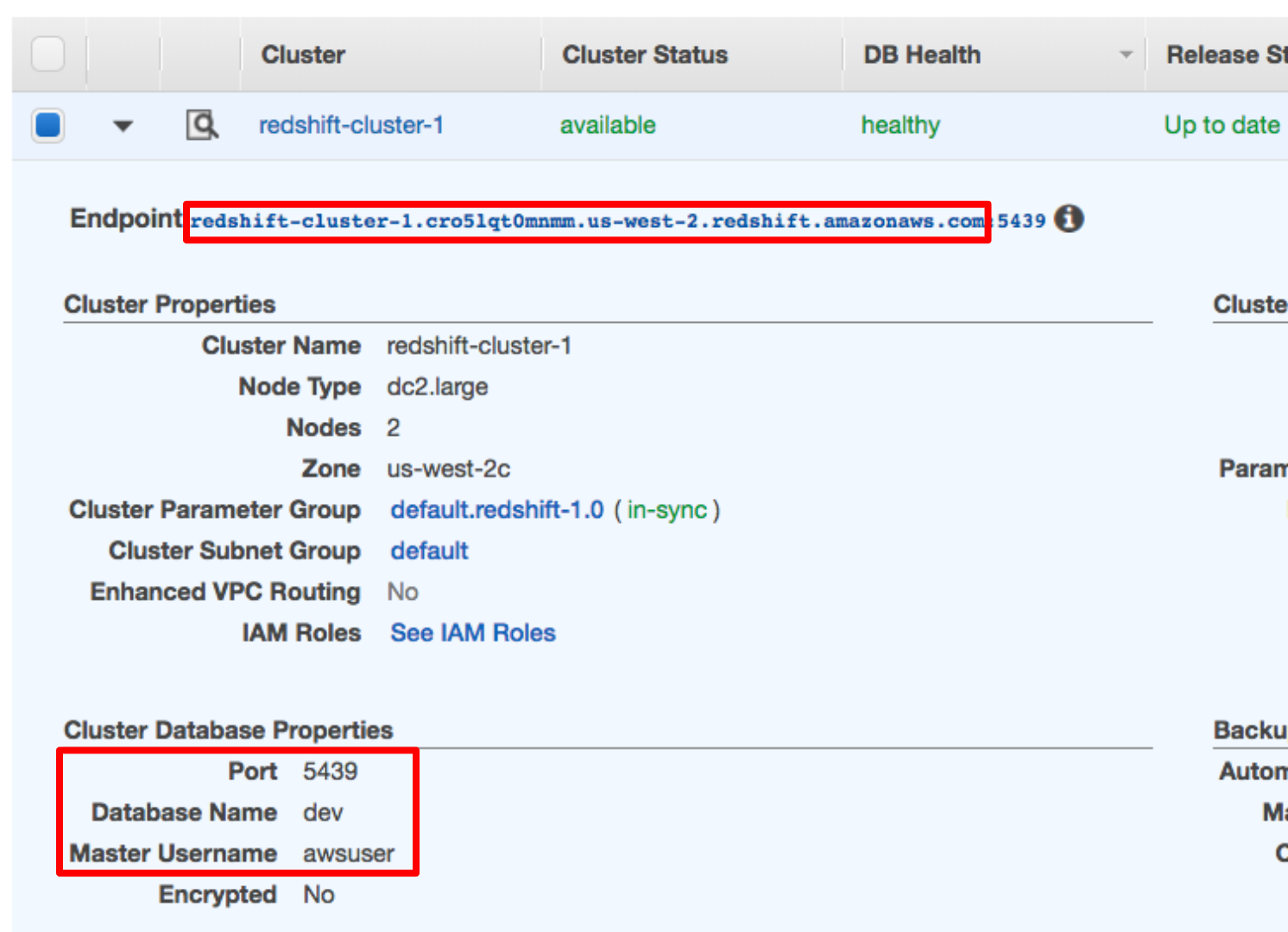
1. On the create connection page, enter the following values:
   * **Conn Id**: Enter aws\_credentials.
   * **Conn Type**: Enter Amazon Web Services.
   * **Login**: Enter your **Access key ID** from the IAM User credentials you downloaded earlier.
   * **Password**: Enter your **Secret access key** from the IAM User credentials you downloaded earlier.

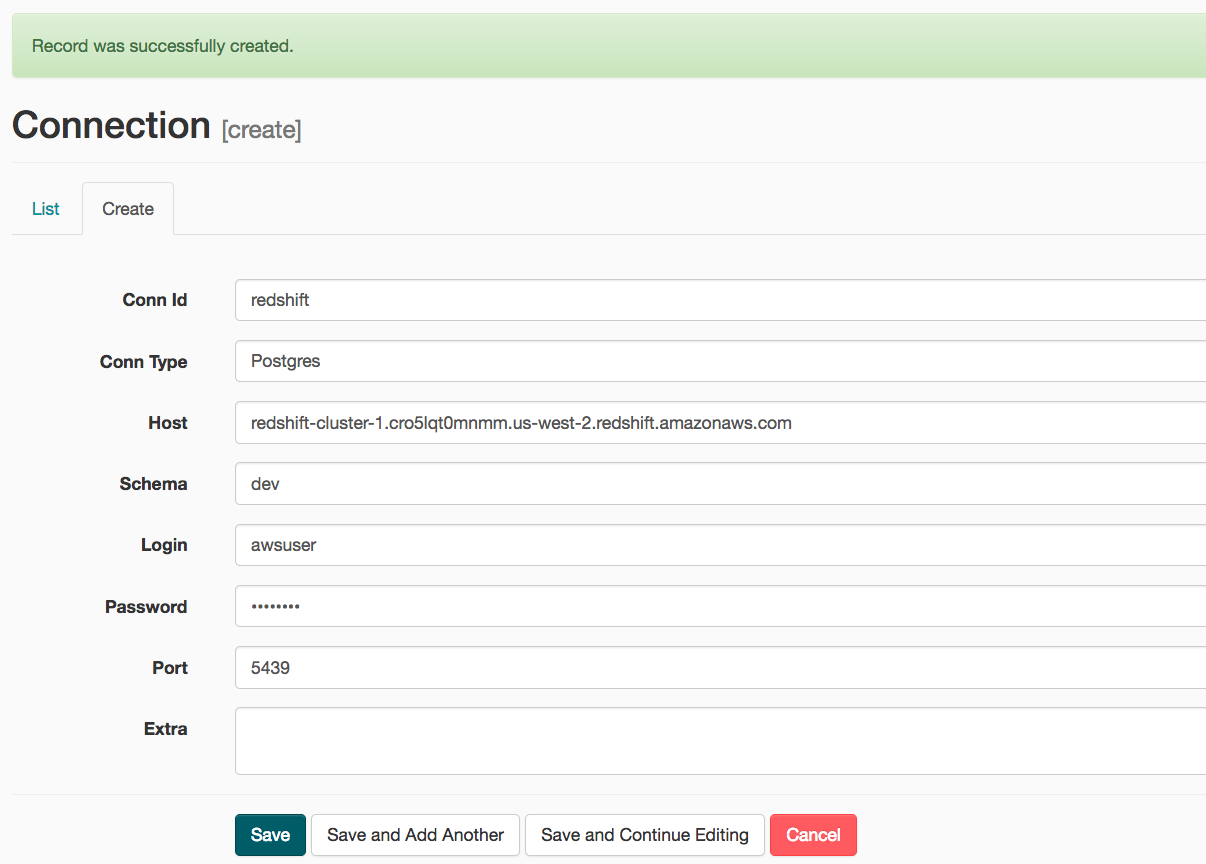
Once you've entered these values, select **Save and Add Another**.

**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)**

1. On the next create connection page, enter the following values:
   * **Conn Id**: Enter redshift.
   * **Conn Type**: Enter Postgres.
   * **Host**: Enter the endpoint of your Redshift cluster, excluding the port at the end. You can find this by selecting your cluster in the **Clusters** page of the Amazon Redshift console. See where this is located in the screenshot below. IMPORTANT: Make sure to **NOT** include the port at the end of the Redshift endpoint string.
   * **Schema**: Enter dev. This is the Redshift database you want to connect to.
   * **Login**: Enter awsuser.
   * **Password**: Enter the password you created when launching your Redshift cluster.
   * **Port**: Enter 5439.

Once you've entered these values, select **Save**.

**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)**

**[[](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)](https://classroom.udacity.com/nanodegrees/nd027/parts/45d1c3b1-d87b-4578-a6d0-7e86bb5fea6c/modules/2adf57ae-57cb-42f6-bd65-a2c383797ce3/lessons/4d1d5892-2cab-4456-8b1a-fb2b5fa1488d/concepts/f5c91ffd-60dc-4af3-bc86-44efce885834)**

Awesome! You're now all configured to run Airflow with Redshift.