

airflow-gcp-examples

Repository with examples and smoke tests for the GCP Airflow operators and hooks.

Setup

This Google Cloud Examples does assume you will have a standard Airflow setup up and running. This tutorial does work perfectly locally as in a production setup because the only requirement is that you have a service key, that we'll explain next. But first a quick rundown of what you need:

- Running Airflow (as of this writing you need Airflow master branch!!!)
- Create a service account (Cloud Console)
- Setup a Google Cloud Connection in Airflow
- Setup variables that the DAG's will need
- Copy the DAG's to your dags folder

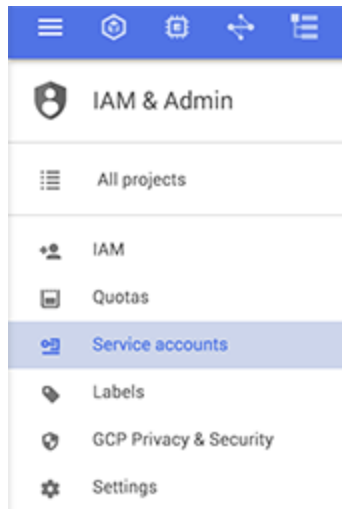
Airflow setup

- Checkout master of Airflow
- pip install google-api-python-client
- python setup.py install

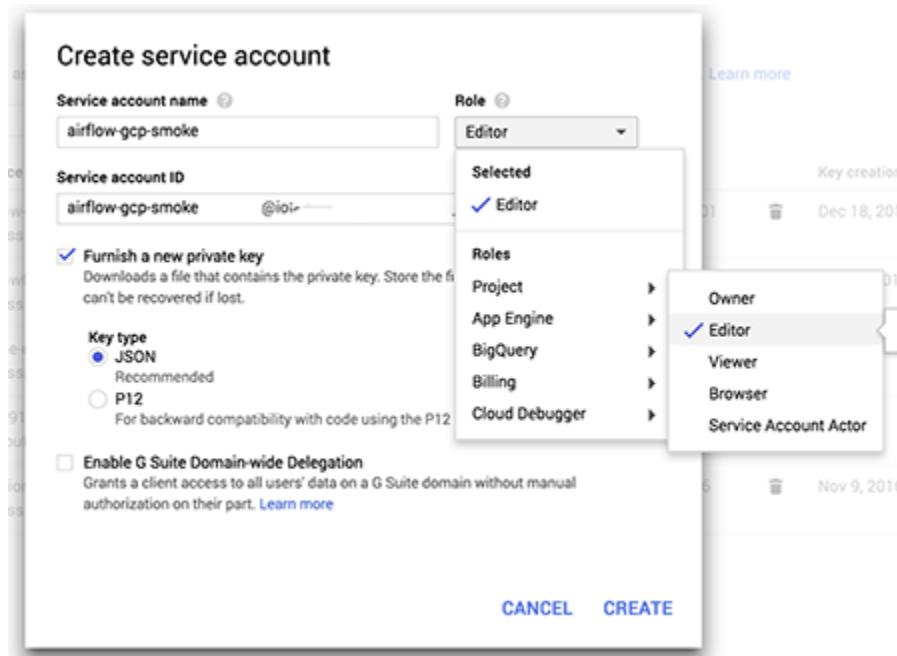
Make sure you're running the LocalExecutor and have a decent database setup.

Google Cloud Service Key

Next create a service account where you want your smoke tests and examples to run in. Go to the console:



And create a service key. Provision a JSON private key and give it Editor's rights



Airflow Connection

In Airflow you need to define the `gcp_smoke` named connection to your project:

Connection [edit]

List Create Edit

Conn Id gcp_smoke

Conn Type Google Cloud Platform

Project Id `project-id-123456789012`

Keyfile Path `/Users/sjebobblair/airflow/airflow-gcp-examples/airflow-gcp-smoke.json`

Scopes (comma separated) `https://www.googleapis.com/auth/cloud-platform`

Save Save and Add Another Save and Continue Editing Cancel

Supply the path to the downloaded private key, supply the *project_id* and define the minimum scope of <https://www.googleapis.com/auth/cloud-platform>





Airflow Variables

You need to setup variables that are used in the examples. You can tweak them to suite your environment.

Variables

Choose File No file chosen Import Variables

List (2) Create Add Filter With selected Search

		Key	Val
<input type="checkbox"/>	 	gcp_smoke_dataset	airflow
<input type="checkbox"/>	 	gcp_smoke_tempset	airflow_temp

variable	example value	note
gc_project	my-project	Project where the examples will run in
gcq_dataset	airflow	BigQuery dataset for examples

gcq_tempset	airflow_temp	BiqQuery dataset with 1 day retentions
gcs_bucket	airflow-gcp-smoke	Storage bucket
gcs_root	data	Storage root path (required, no start and end with slash)