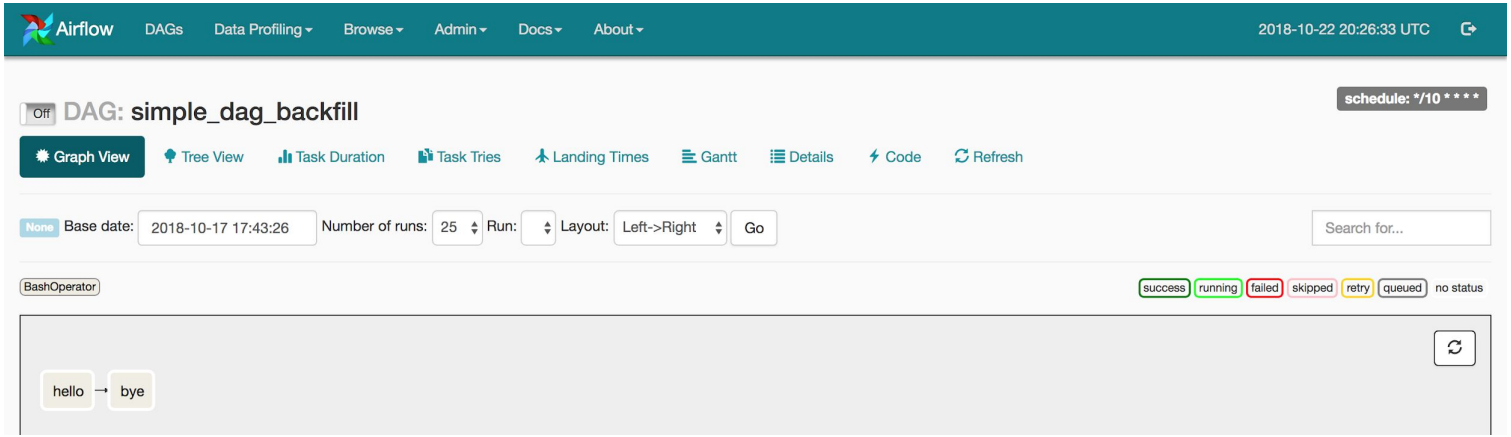
A Quick Play With Backfill and Catchup

A quick example

Instruction

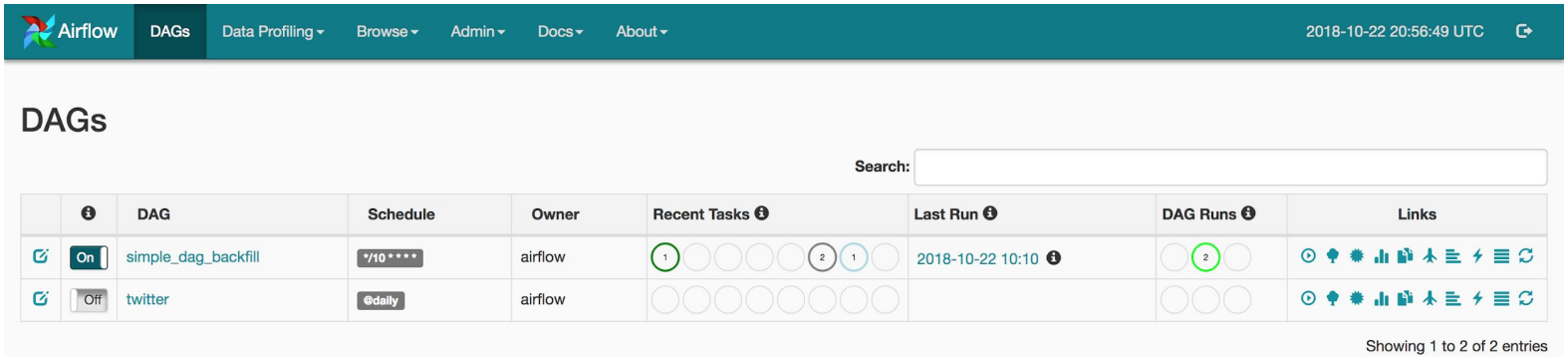
- `cd ~/airflow/dags`
- `wget http://www.marclamberti.com/airflow/simple_dag_backfill.py`
- `vim simple_dag_backfill.py`
- `a`
- Change the `date dt.datetime(2018, 10, 22, 10, 00, 00)` with your current date.
- Wait for the Scheduler to take into consideration your DAG (5 min max)
- Turn ON the scheduler toggle from Airflow UI

Airflow UI


Airflow
DAGs
Data Profiling
Browse
Admin
Docs
About
2018-10-22 20:26:33 UTC
Off
DAG: simple_dag_backfill
schedule: */10 * * * *
Graph View
Tree View
Task Duration
Task Tries
Landing Times
Gantt
Details
Code
Refresh
None
Base date:
2018-10-17 17:43:26
Number of runs:
25
Run:
Layout:
Left->Right
Go
Search for...
BashOperator
success
running
failed
skipped
retry
queued
no status
hello
bye

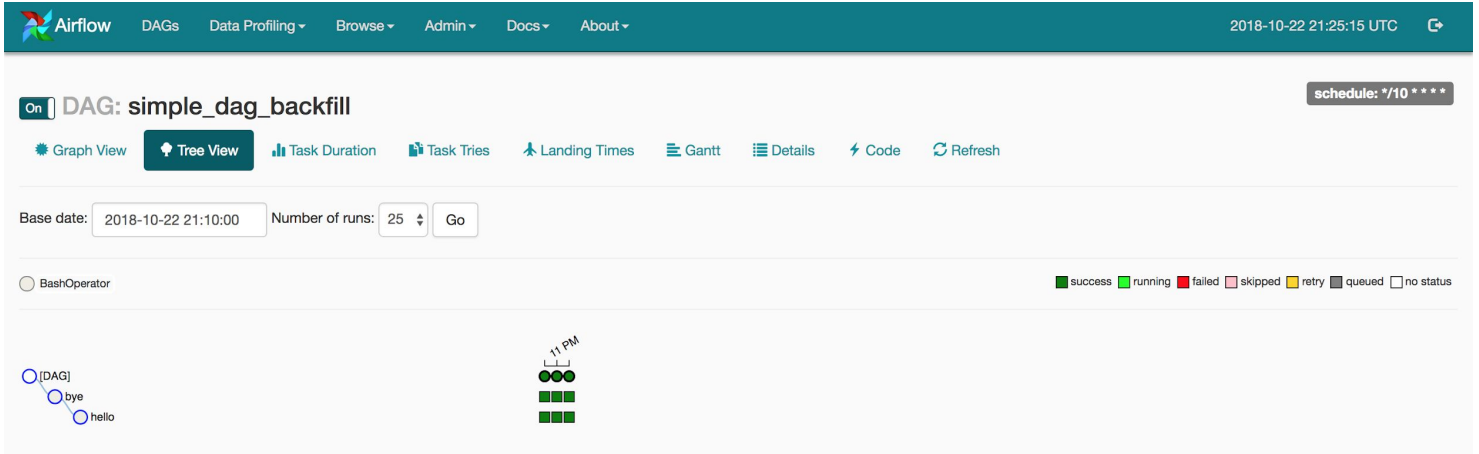
Airflow UI

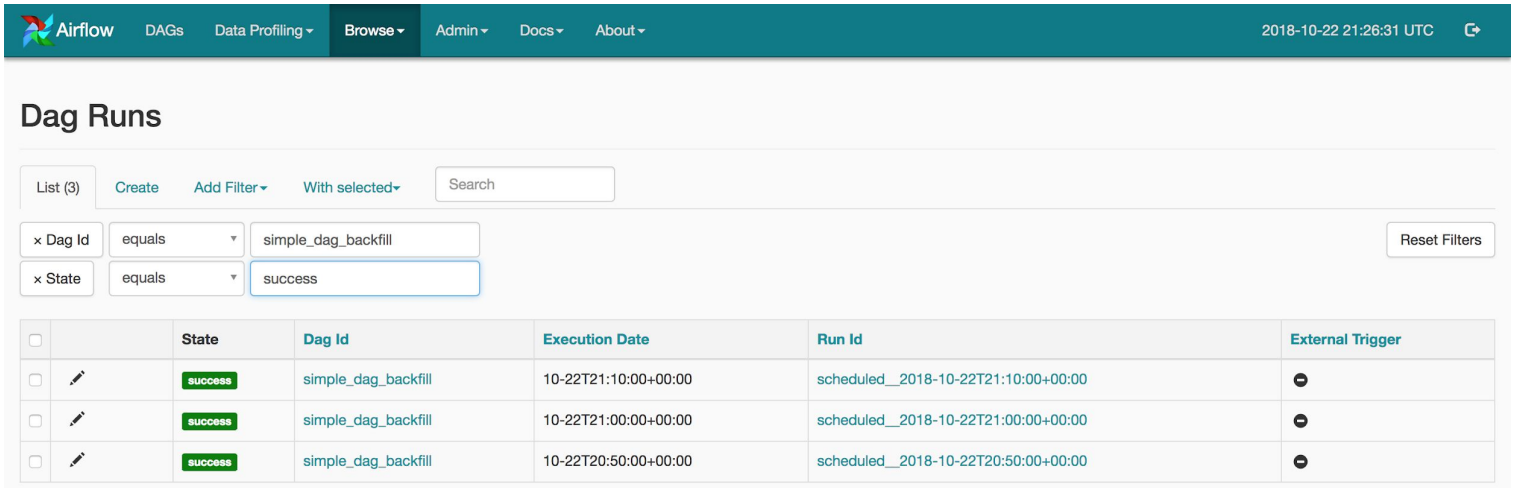
After having turned ON the toggle, simple_dag_backfill will run automatically the past DagRun until the current date.



Airflow UI

As you can see from the following screenshot, the past DagRuns have been all executed by the Scheduler until the current_time - schedule_interval.



Airflow UI

The last execution_date is equals to 21:10:00 but it has been actually done at 21:20:00. **Remember, the DAG is run ONCE the schedule_time is ELAPSED.**