Happy DAGs + Happy Teammates

How a little CICD can go a long way

Leah Cole

May 24, 2022
Hi! I'm Leah.

- Developer Relations Engineer
- Cloud Composer
- Airflow Committer
- Fun fact: I love yarn crafts and cats!

Leah Cole
Google Cloud Developer Relations Engineer

Millie
Frank

@leahecole
Have you ever...

... had a DAG that works locally but not when deployed?

... lost track of what you have and haven't tried to make your DAG work?

... seen the big red "Broken DAG" error?

... realized you forgot to deploy changes you made in your DAG to Airflow?
Have you ever...

... had a DAG that works locally but not when deployed?

... lost track of what you have and haven't tried to make your DAG work?

... seen the big red "Broken DAG" error?

... realized you forgot to deploy changes you made in your DAG to Airflow?
Potential fix: Multiple Airflow environments
Multiple Airflow Environments

Local ➔ Dev ➔ Prod

@leahecole
Have you ever...

... had a DAG that works locally but not when deployed?

... lost track of what you have and haven't tried to make your DAG work?

... seen the big red "Broken DAG" error?

... realized you forgot to deploy changes you made in your DAG to Airflow?
Potential fix:
Use version control
Use version control

Solo flow

● Make a commit with a detailed message
● Push to your branch

Team flow

● Make a commit with a detailed message
● Push to your branch
● Open a pull request
● Have a teammate review
● Merge in changes
Have you ever...

... had a DAG that works locally but not when deployed?

... lost track of what you have and haven't tried to make your DAG work?

... seen the big red "Broken DAG" error?

... realized you forgot to deploy changes you made in your DAG to Airflow?
Potential fix:
Test your DAGs
Test your DAGs

- Airflow CLI
- Unit tests
- Local tests
- Running in Dev environment

Resources:

- Testing DAGs Composer
- Community contributions to Airflow test guide
- Airflow testing best practices
- 2020 Airflow Summit Talk
- GoDataDriven blog post (2019)
- Astronomer testing Airflow guide
Have you ever...

... had a DAG that works locally but not when deployed?

... lost track of what you have and haven't tried to make your DAG work?

... seen the big red "Broken DAG" error?

... realized you forgot to deploy changes you made in your DAG to Airflow?
Potential fix: Automate
Automate
Use CI/CD to Keep DAGs in sync with Version Control

1. PR opened with a DAG change
2. GitHub
   - Presubmit Cloud Build
   - DAG file merged
   - PR Review
3. DAGs copied to Dev Composer Environment Cloud Build
4. Development Environment Cloud Composer
   - DAG success
5. Production Environment Cloud Composer
   - Manually promote DAG to production
6. GCP
   - Address DAG failures
   - DAG failure
   - DAG success

@leahecole
Multiple Airflow Environments

Use CI/CD to Keep DAGs in sync with Version Control

- PR opened with a DAG change
- GitHub
  - Presubmit
  - Cloud Build
  - DAG file merged
  - PR Review
- DAGs copied to Dev Composer Environment
- Cloud Build
- Development Environment
- Cloud Composer
- DAG success
- Manually promote DAG to production
- Production Environment
- Cloud Composer
- Address DAG failures
- DAG failure
- GCP
Version control

Use CI/CD to Keep DAGs in sync with Version Control

GitHub

- Presubmit Cloud Build
- PR Review
- DAG file merged

GitHub

- PR opened with a DAG change

Google Cloud

- Address DAG failures
- DAG failure
- Development Environment Cloud Composer
- DAG success
- Manually promote DAG to production

Production Environment Cloud Composer
Thank you.
Questions?

Additional Links

Airflow summit: airflowsummit.org
