Power your data orchestration with Apache Airflow® 3

ASTRONOMER

Agenda

- Background - 30 min
  - Airflow 3 feature overview
- Hands-on exercises - 45 min
  - Run Airflow 3 locally
  - Explore the new UI
  - Write a DAG with assets
  - Use backfills and DAG versioning
- Q&A - 15 min

Prerequisites

To complete the hands-on exercises you will need:

- A GitHub account
- The Astro CLI
  - bit.ly/astro-cli

Airflow is the open source standard for Workflow Management

30M+

Monthly downloads

1500+

Building blocks

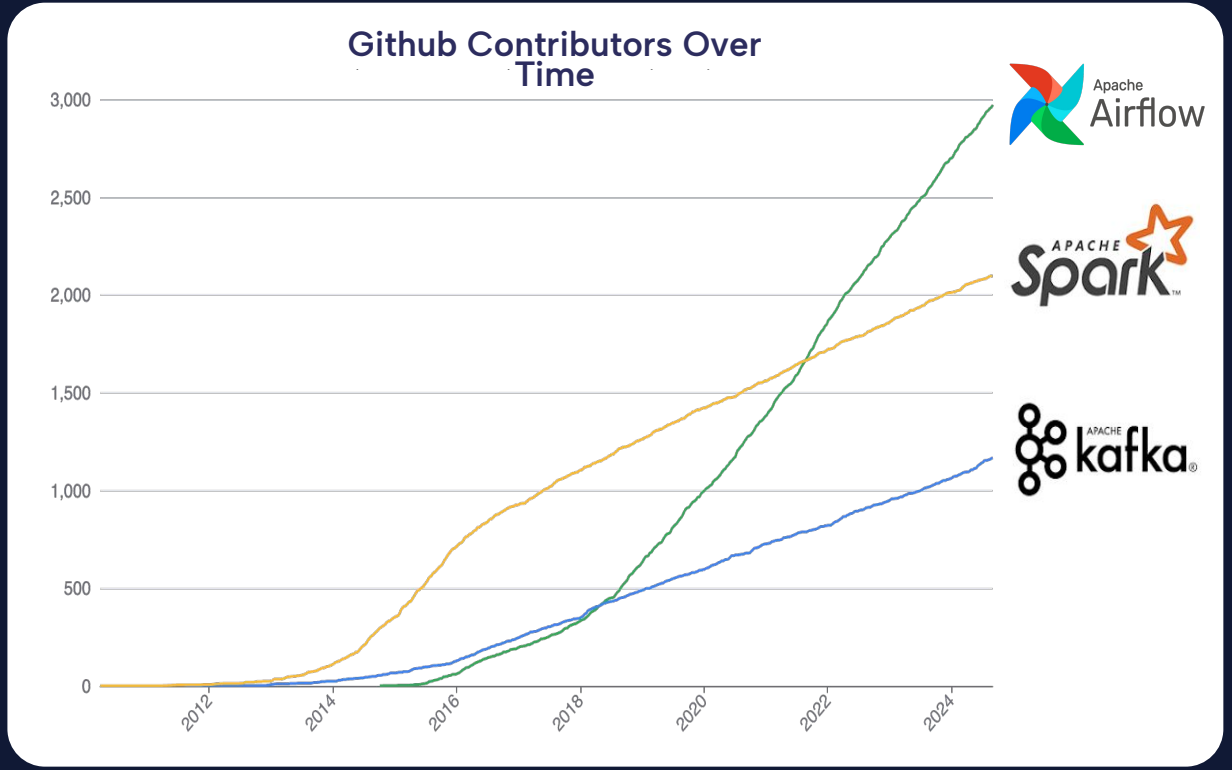
57k+

Slack community

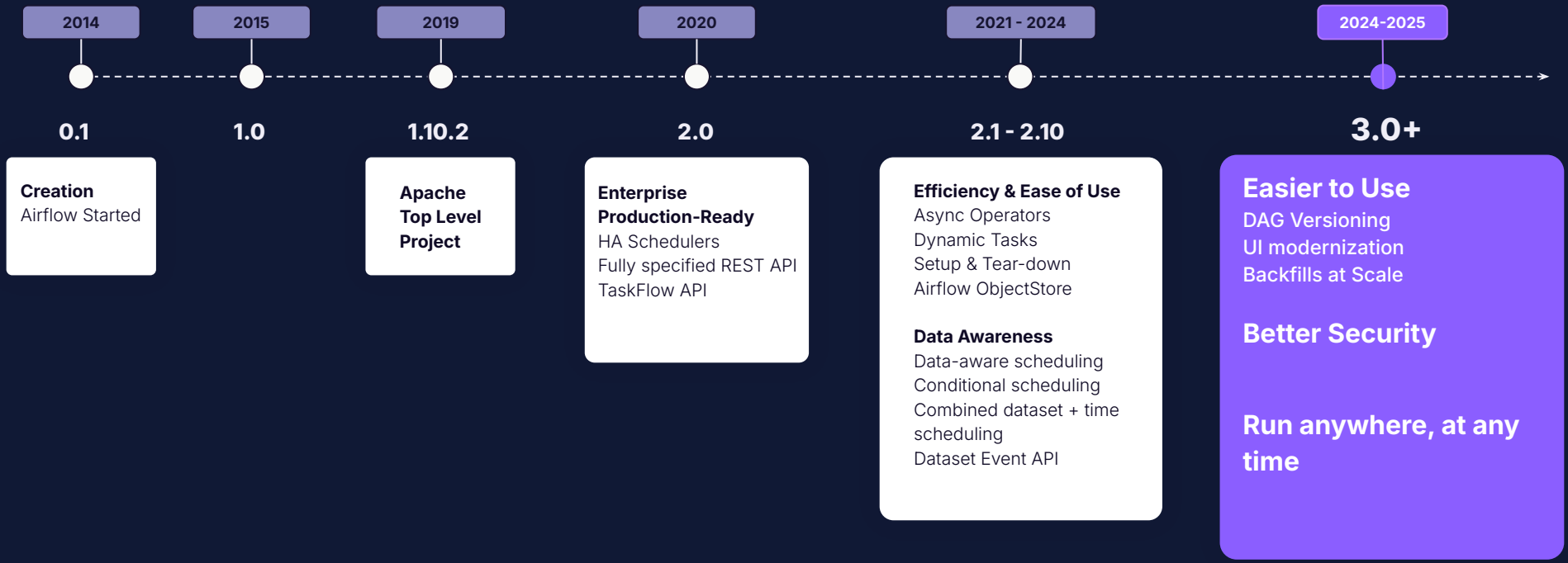
3k+

Contributors

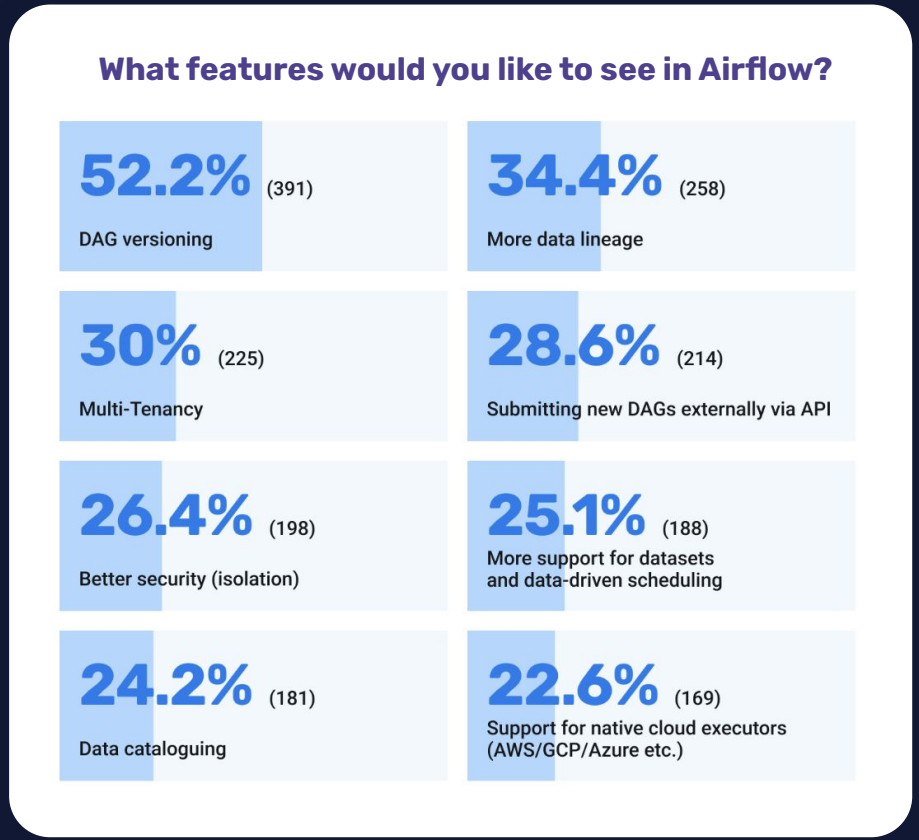
Airflow contributor growth



Key Airflow milestones



Airflow survey results



DAG Versioning always #1 on the wish list!

More:

- Security
- Execution / invocation
- Datasets
- Observability / Governance

Airflow 3.0 is the biggest release since 2020

Three key themes

Easier to use

Stronger Security

Run [tasks] anywhere, at any time

Airflow 3.0 is the biggest release since 2020

Three key themes

Easier to use

Stronger Security

Run [tasks] anywhere, at any time

Improve the UX by making Airflow easier to use

DAG Versioning
View historical versions of DAGs and all their run information

UI Modernization
React-based UI with easier navigation

Backfills at Scale
Run backfills easily from the UI, API, or CLI

Improve the UX by making Airflow easier to use

DAG Versioning
View historical versions of DAGs and all their run information

UI Modernization
React-based UI with easier navigation

Backfills at Scale
Run backfills easily from the UI, API, or CLI

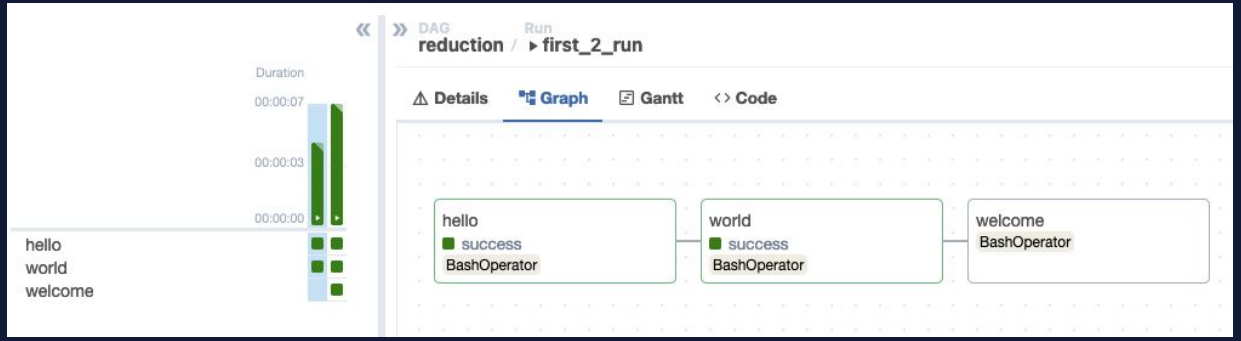
DAG versioning in Airflow 2.X

❌ Airflow assumes the most recent DAG code applies to all past runs

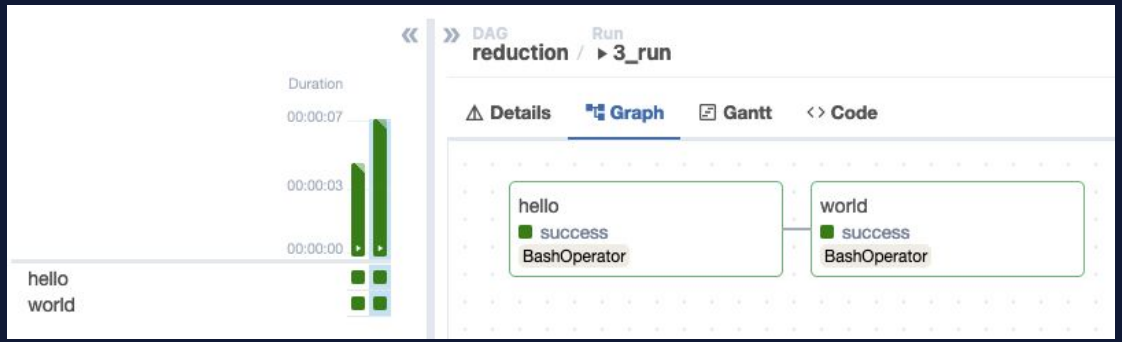
❌ Code updates during a run can cause tasks to execute different code versions

❌ DAG changes overwrite history, making past executions hard to trace

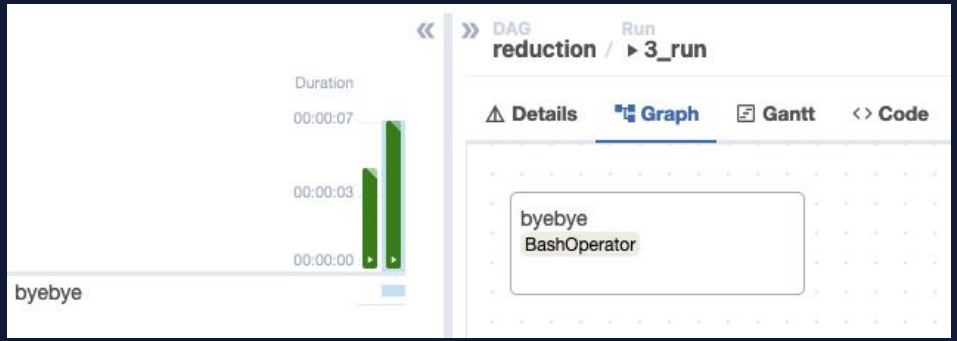
Went from 2 to 3 tasks



Reduce back to 2 tasks

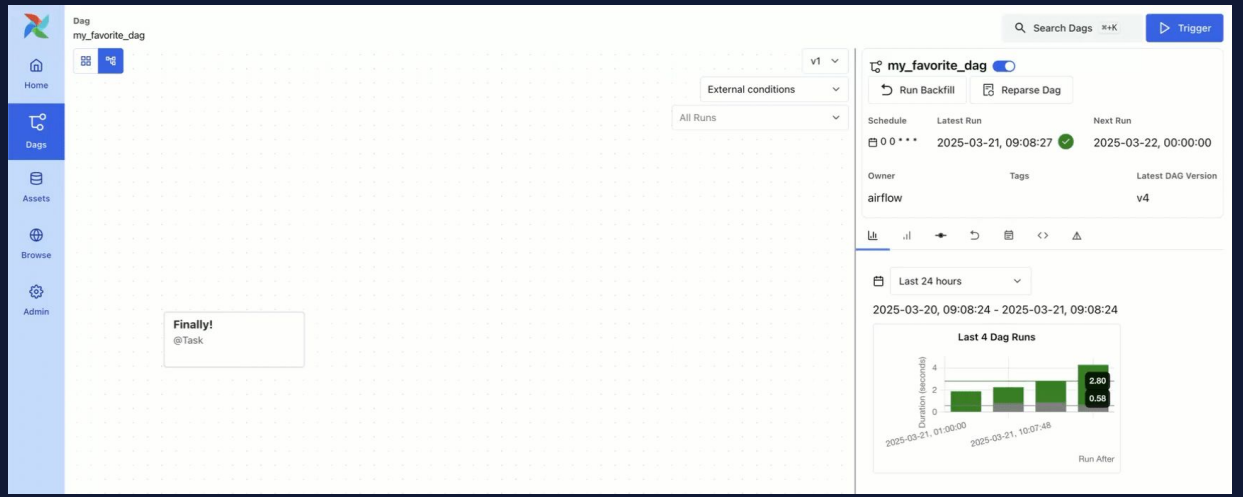


Completely different tasks



DAG versioning in Airflow 3.X

The #1 requested community feature is now delivered!



Improve the UX by making Airflow easier to use

DAG Versioning
View historical versions of DAGs and all their run information

UI Modernization
React-based UI with easier navigation

Backfills at Scale
Run backfills easily from the UI, API, or CLI

Backfills in Airflow 2.X

➡️ Airflow backfills can only run via the CLI

❌ Cannot trigger backfill via API/UI

➡️ CLI backfill acts as a second scheduler

❌ Scheduling logic for backfill and normal runs differs

➡️ If the CLI process dies, the backfill job dies

❌ Cannot observe progress, track status, or cancel

Backfills in Airflow 3.0

- ✅ Trigger backfills from UI, API, or CLI
- ✅ API support to trigger async backfills
- ✅ UI support to view progress and track status
- ✅ Backfills coordinated via the scheduler

Airflow 3.0 is the biggest release since 2020

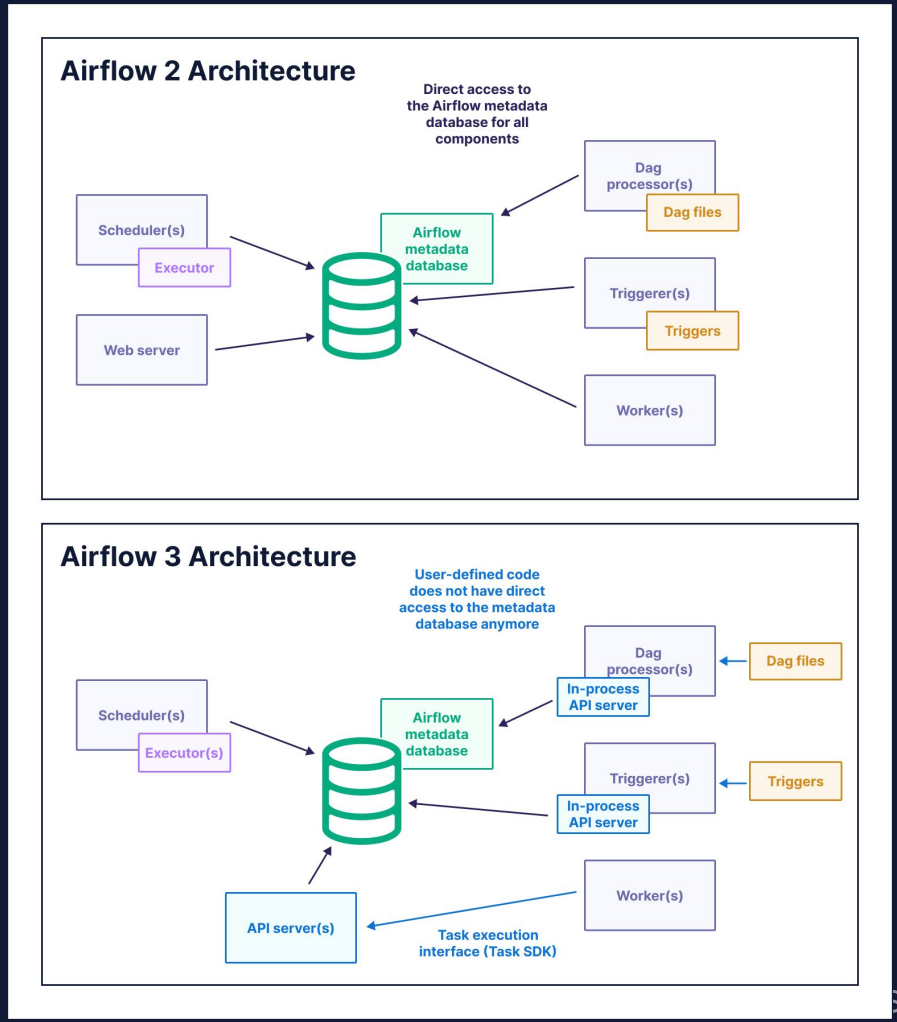
Three key themes

Easier to use

Stronger Security

Run [tasks] anywhere, at any time

Airflow 3.0 introduces an architectural change for task isolation



Airflow 3.0 is the biggest release since 2020

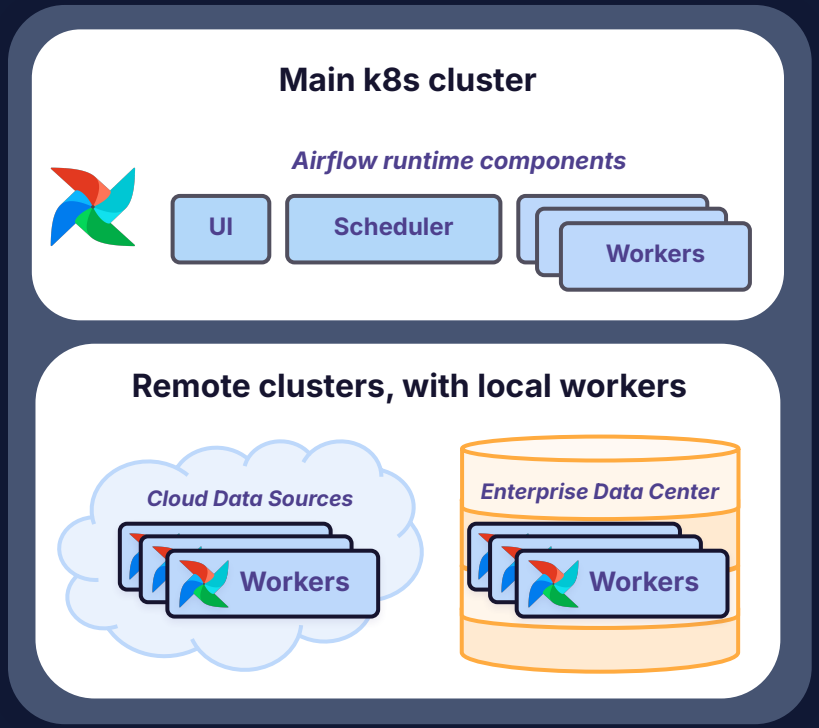
Three key themes

Easier to use

Stronger Security

Run [tasks] anywhere, at any time

Run anywhere builds on the foundation of task isolation in Airflow 3.0



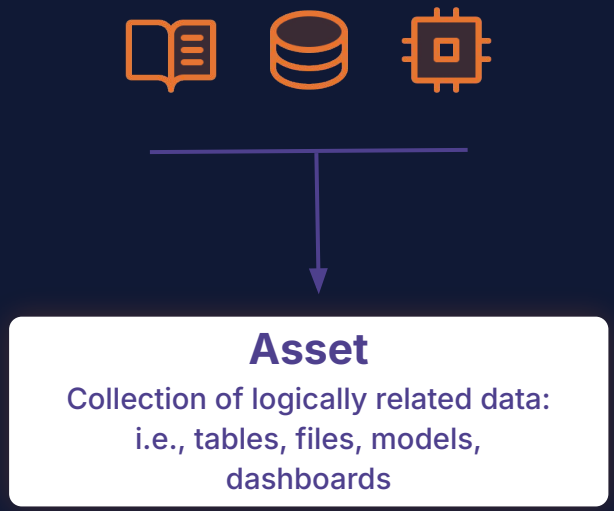
Remote / Edge execution: run tasks on workers in remote clusters

Benefits

- Deployment flexibility with workers on public, hybrid, private cloud, on-prem, edge, GPU clouds
- Higher resilience and scalability
- Improved security isolation
- Easier upgrades, fewer dependencies
- Better meet data locality mandates

Run at any time, Airflow 3.0 supports Scheduled (Batch), Event-Driven, and Adhoc Execution

- **Event-driven scheduling:** Run DAGs in a push or pull architecture based on updates to an external system. Built on top of data assets. SQS supported
- **Inference Execution:** Simultaneous execution of the same DAG.



What is an Asset ?

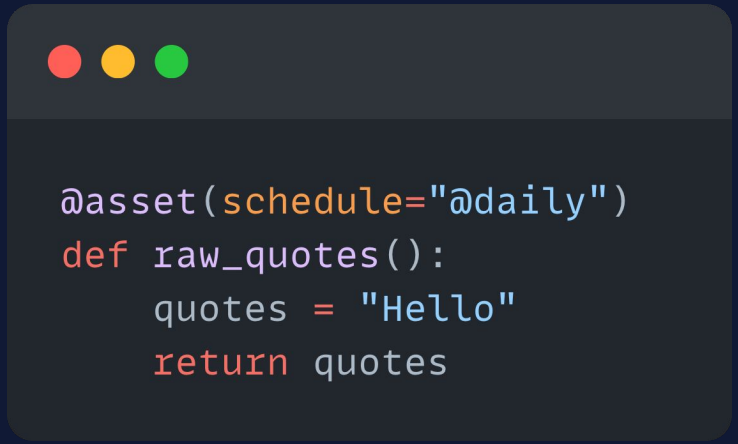
- An Asset is a **collection of logically related data**
  - The next evolution of datasets
- Assets can have a name, a URI, a group, and extra information. They are mainly identified by their name in the UI right now.
- An **asset event** is an event that is attached to an asset and created whenever that asset is updated
- Assets can have a **Watcher** for event-driven scheduling (only SQS right now)

How do you define/interact with an Asset?

- **Outlets**: Specify which assets a task *produces or updates*. They appear in the DAG graph and Assets tab UI as soon as the DAG is parsed—regardless of task execution.
- **Inlets**: Indicate which assets a task *depends on or reads from*. This gives the task access to all asset events for that asset
  - Note this DOES NOT affect the DAG's schedule

- **@asset**: a decorator that can be used to create one DAG with one task that produces an asset, by directly declaring the desired asset in Python with less boilerplate code
  - "Asset-oriented" approach to DAG writing, vs "task-oriented"

The Basics



```python
@asset(schedule="@daily")
def raw_quotes():
    quotes = "Hello"
    return quotes
```

Generating an Asset Event

There are several ways to update an asset:

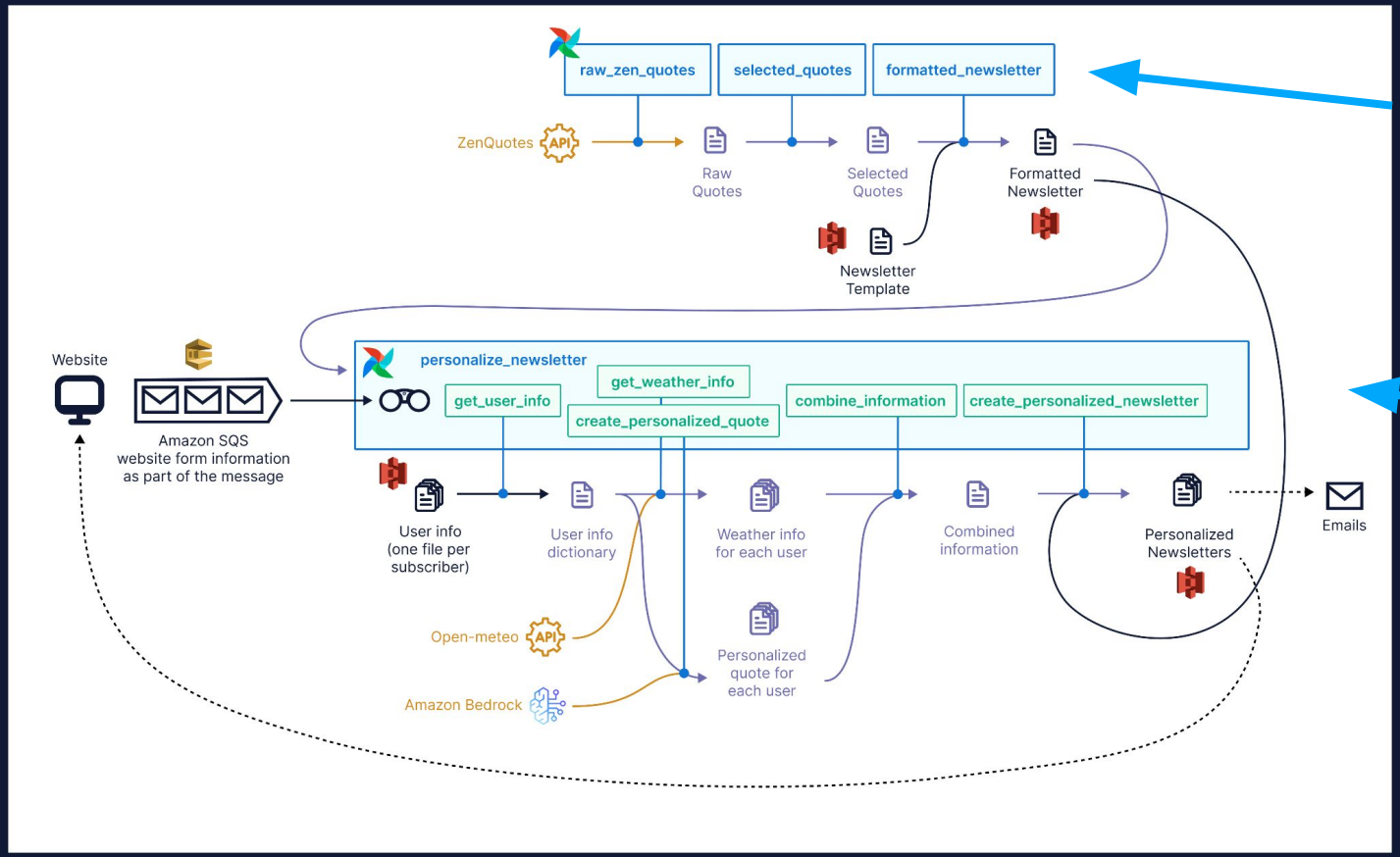
- A DAG defined using `@asset` completes successfully. Under the hood, `@asset` creates a DAG with one task which produces the asset.
- A task with an outlet parameter that references the asset completes successfully.
- A `POST` request to the assets endpoint of the Airflow REST API.
- An `AssetWatcher` that listens for a `TriggerEvent` caused by a message in a message queue. See event-driven scheduling for more information.
- A manual update in the Airflow UI by using the Create Asset Event button on the asset graph. Two options here:
    - Materialize: runs the full DAG that contains the task that produces the asset event
    - Manual: creates a new asset event without running any task. Useful for testing

Hands-on exercises

Instructions

- You will need:
  - A GitHub account
  - The Astro CLI - bit.ly/astro-cli
    - (If you can’t install the CLI, you can use GH codespaces. If using codespaces, choose 4 cores instead of the default 2 when setting it up.)
- Fork the repo here: **bit.ly/af3-workshop**
  - **Make sure you uncheck “fork the main branch only”**
  - Instructions are in the Readme
  - Pick the **airflow-3-0** branch

What you will build



ETL pipeline with Assets (3 separate DAGs)

GenAI pipeline (1 DAG)

Start Your 14-Day Free Trial of Astro

Run Airflow in 5 minutes

A simple platform and a suite of developer tools empower
your team to quickly develop, test, and deploy pipelines.

Manage data pipelines like a pro

Extend your capabilities beyond Airflow to easily and reliably build data pipelines, manage workflows, and optimize operations.

Scale data operations to success

Enterprise-grade security, 24/7 support from in-house experts and top committers, and end-to-end observability allow you to grow with confidence.

ASTRONOMER

Q&A

ASTRONOMER

AIRFLOW SUMMIT

Register now to get your

Early Bird Discount

Before June 30th

OCT 7-9, 2025, SEATTLE, USA

airflowsummit.org/