Apache Airflow Basics

By Dhandapani Yedappalli Krishnamurthi Sep 11, 2025

1. What is Airflow?

- Apache Airflow is an open-source platform for orchestrating workflows.
- It helps you schedule, monitor, and manage data pipelines.
- Workflows are written as Python code.

👉 Think of Airflow as a task scheduler + pipeline manager.

---

2. Key Concepts

1. DAG (Directed Acyclic Graph)
    - A workflow in Airflow.
    - Directed → tasks have order.
    - Acyclic → no loops allowed.
2. Task
    - A unit of work (e.g., run SQL, execute Python script, call API).
3. Operator
    - A template for a task. Examples:
        - PythonOperator → run Python function
        - BashOperator → run shell commands
        - PostgresOperator → run SQL on Postgres
4. Task Instance
    - A specific execution of a task in a DAG run.
5. Scheduler
    - Decides when tasks should run.
6. Executor
    - Decides how tasks run (e.g., Local, Celery, Kubernetes).
7. Web UI

- User interface to view DAGs, runs, and logs.

---

3. How Airflow Works

1. You define a DAG in Python.
2. Scheduler reads DAG and schedules tasks.
3. Executor runs tasks.
4. Metadata is stored in Airflow database.
5. You monitor via Web UI.

```
from airflow import DAG
from airflow.operators.bash import BashOperator
from datetime import datetime

# Define DAG
with DAG(
dag_id="example_dag",
start_date=datetime(2025, 9, 1),
schedule_interval="@daily", # run daily
catchup=False
) as dag:

# Task 1
task1 = BashOperator(
task_id="print_date",
bash_command="date"
)

# Task 2
task2 = BashOperator(
task_id="say_hello",
bash_command="echo 'Hello, Airflow!'"
)

# Set dependency
task1 >> task2
```

5. Airflow Components

- airflow scheduler → schedules DAGs
- airflow webserver → UI for monitoring
- airflow worker → runs tasks
- airflow database → stores metadata

6. ---

6. Why Use Airflow?

✅ Automates workflows

✅ Scalable (can run thousands of tasks)

✅ Supports retries & error handling

✅ Integrates with databases, cloud services, big data tools

✅ Rich monitoring (logs, UI, alerts)