

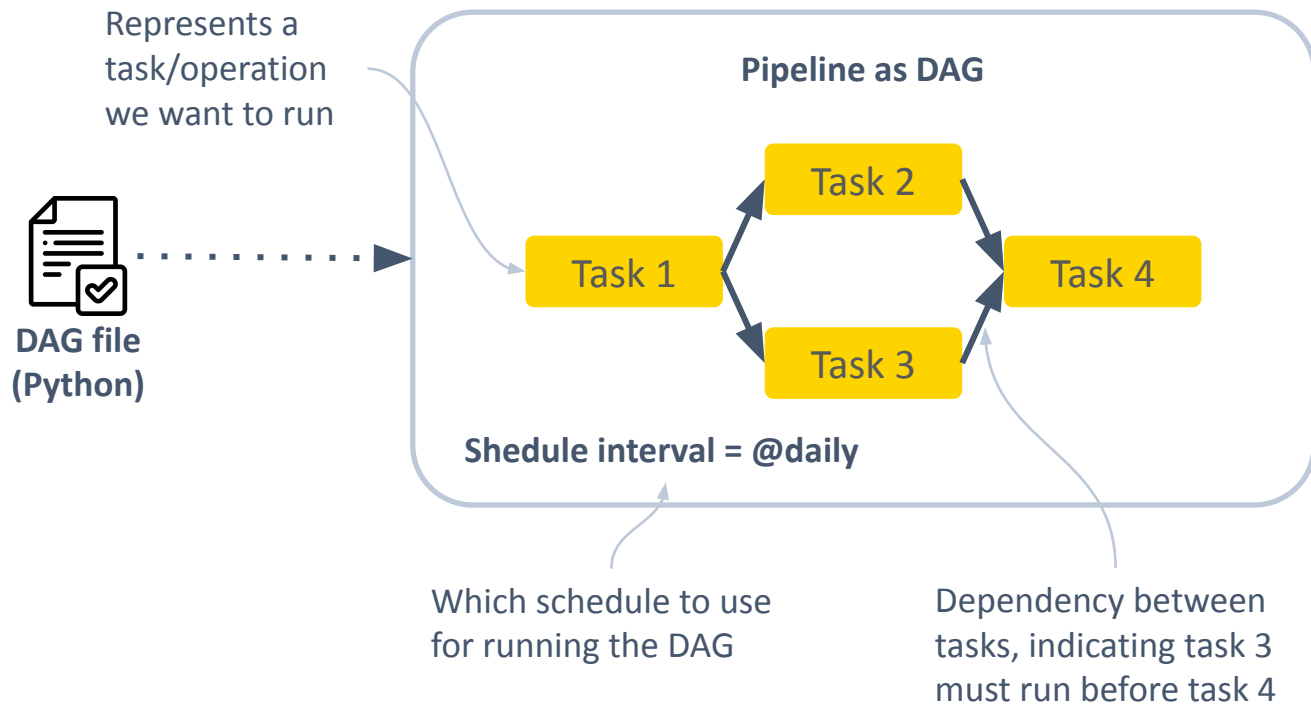
Дата-инженер

DAG

Ася Гайламазян



Механизм DAG



Создание DAG 4-мя способами

Стандартный конструктор,
передающий DAG
операторам

Менеджер контекста with

DAG декоратор @dag

TaskFlow API

Стандартный конструктор

```
my_dag = DAG("my_dag_name",  
start_date=pendulum.datetime(2023, 1, 1, tz="UTC"),  
schedule="@daily", catchup=False)  
  
op = EmptyOperator(task_id="task", dag=my_dag)
```



Менеджер контекста with

```
with DAG(
    "my_dag_name", start_date=pendulum.datetime(2023, 1, 1, tz="UTC"),
    schedule="@daily", catchup=False
) as dag:
    op = EmptyOperator(task_id="task")
```



DAG-декоратор @dag

```
@dag(start_date=pendulum.datetime(2023, 1, 1, tz="UTC"),
     schedule="@daily", catchup=False)
def generate_dag():
    op = EmptyOperator(task_id="task")
    ...
dag = generate_dag()
```



TaskFlow API. Декораторы @task

```
@dag(schedule=None,start_date=pendulum.datetime(2023, 1, 1, tz="UTC"),catchup=False,tags=["example"],)
def example_dag_decorator(email: str = "example@example.com"):
    get_ip = GetRequestOperator(task_id="get_ip", url="http://httpbin.org/get")

    @task(multiple_outputs=True)
    def prepare_email(raw_json: dict[str, Any]) -> dict[str, str]:
        external_ip = raw_json["origin"]

        return {"subject": f"Server connected from {external_ip}","body": f"Seems like today your server executing Airflow is
connected from IP {external_ip}<br>"}

    email_info = prepare_email(get_ip.output)

    EmailOperator(
        task_id="send_email", to=email, subject=email_info["subject"], html_content=email_info["body"])

example_dag = example_dag_decorator()
```

TaskFlow API

```
from airflow.decorators import task
from airflow.operators.email import EmailOperator

@task
def get_ip():
    return my_ip_service.get_main_ip()

@task
def compose_email(external_ip):
    return {'subject':f'Server connected from {external_ip}',
            'body': f'Your server executing Airflow is connected from the external
IP {external_ip}<br>'}

email_info = compose_email(get_ip())
EmailOperator(task_id='send_email',to='example@example.com',subject=email_info
['subject'],html_content=email_info['body'])
```



Параметры запуска DAG

```
my_dag = DAG("my_dag_name",  
start_date=pendulum.datetime(2023, 1, 1, tz="UTC"),  
schedule="@daily", catchup=False)
```

[Документация по параметрам DAG](#)

