



Orchestrating Data Pipelines with Prefect

Budapest Data Forum
Kevin Kho



About Me

- Open Source Community Engineer at Prefect
- Data Scientist for 4 Years
- Fugue Contributor
- Orlando Machine Learning Meetup

Agenda

01

Workflow Orchestration

Use Cases and the History of Prefect

03

Deployment

Work Queues and Agents

02

Simple Problem

Prefect Basics

04

The Future of Prefect

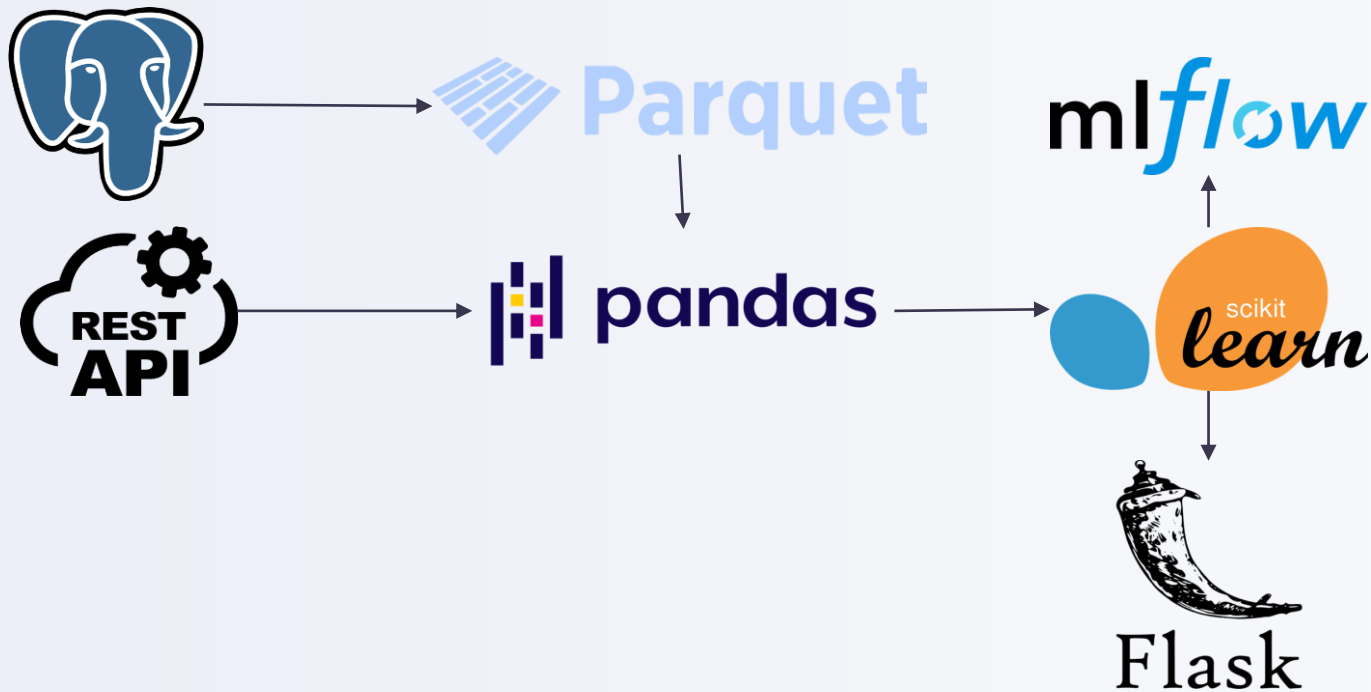
Prefect Orion: The Next Generation Workflow Engine



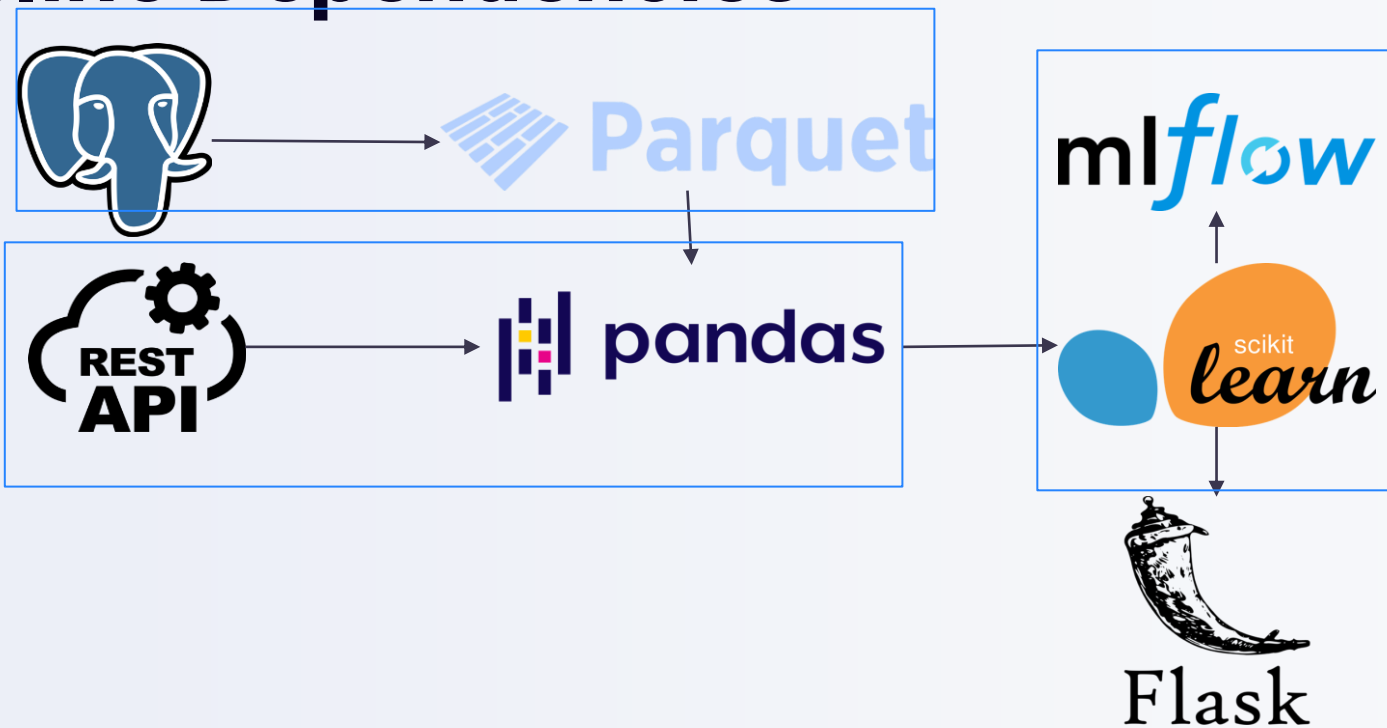
Workflow Orchestration

Use Cases and the History of Prefect

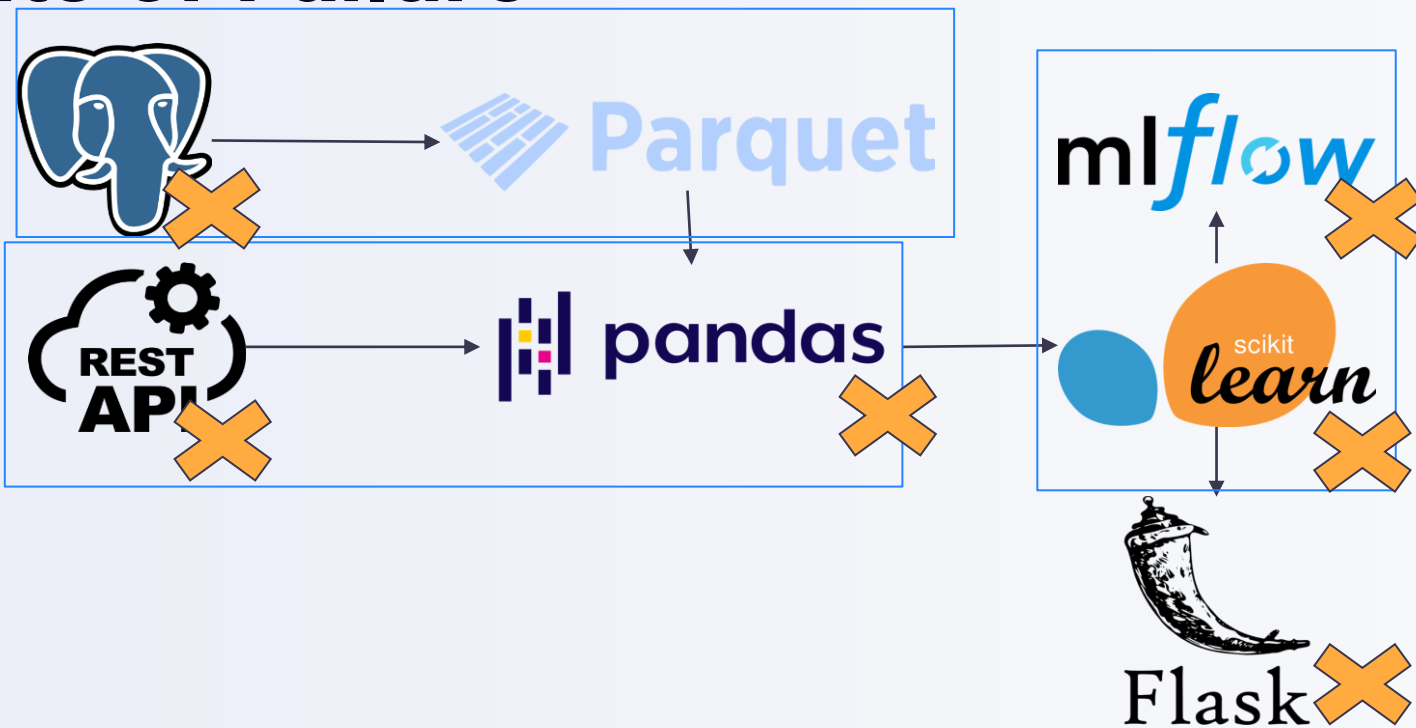
Machine Learning Pipeline



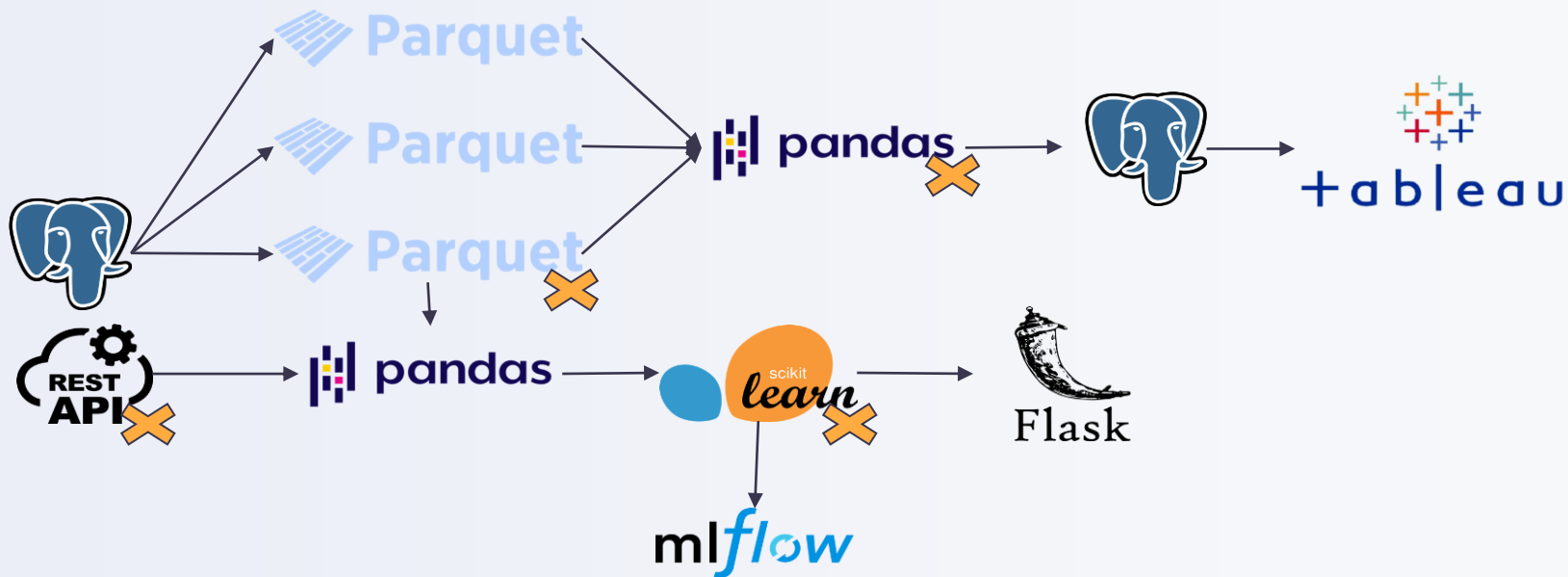
Pipeline Dependencies



Points of Failure



Interconnected Pipelines



Negative Engineering

90% of engineering time spent

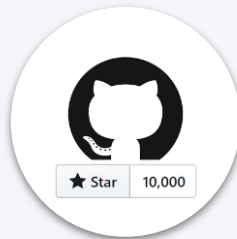
- Retries when APIs go down
- Malformed Data
- Notifications
- Observability into Failure
- Conditional Failure Logic
- Timeouts



Introducing Prefect

Eliminating Negative Engineering

- Open-source workflow orchestration
- Python-based
- Modern data stack
- Native Dask integration
- Very active community
- Prefect Cloud/Prefect Server
- Prefect Orion (Prefect 2.0)



8700+



15800+



Current State of Prefect



Prefect Core



Prefect Orion

- Embracing dynamic, DAG-free workflows
- An extraordinary developer experience
- Transparent and observable orchestration rules

```
from prefect import flow, task
from typing import List
import httpx

@task(retries=3)
def get_stars(repo: str):
    url = f"https://api.github.com/repos/{repo}"
    count = httpx.get(url).json()["stargazers_count"]
    print(f"{repo} has {count} stars!")

@flow(name="Github Stars")
def github_stars(repos: List[str]):
    for repo in repos:
        get_stars(repo)

# run the flow!
github_stars(["PrefectHQ/Prefect", "PrefectHQ/mitter-
design"])
```



Simple Problem

Seeing Prefect in Action



Future of Prefect

Redefining Workflow Orchestration

Resources

- Prefect Orion Docs
 - <https://orion-docs.prefect.io/>
- Announcing Orion
 - <https://www.prefect.io/blog/announcing-prefect-orion/>