Beyond Single Machine

Paweł Kordek
SOFTWARE ENGINEER

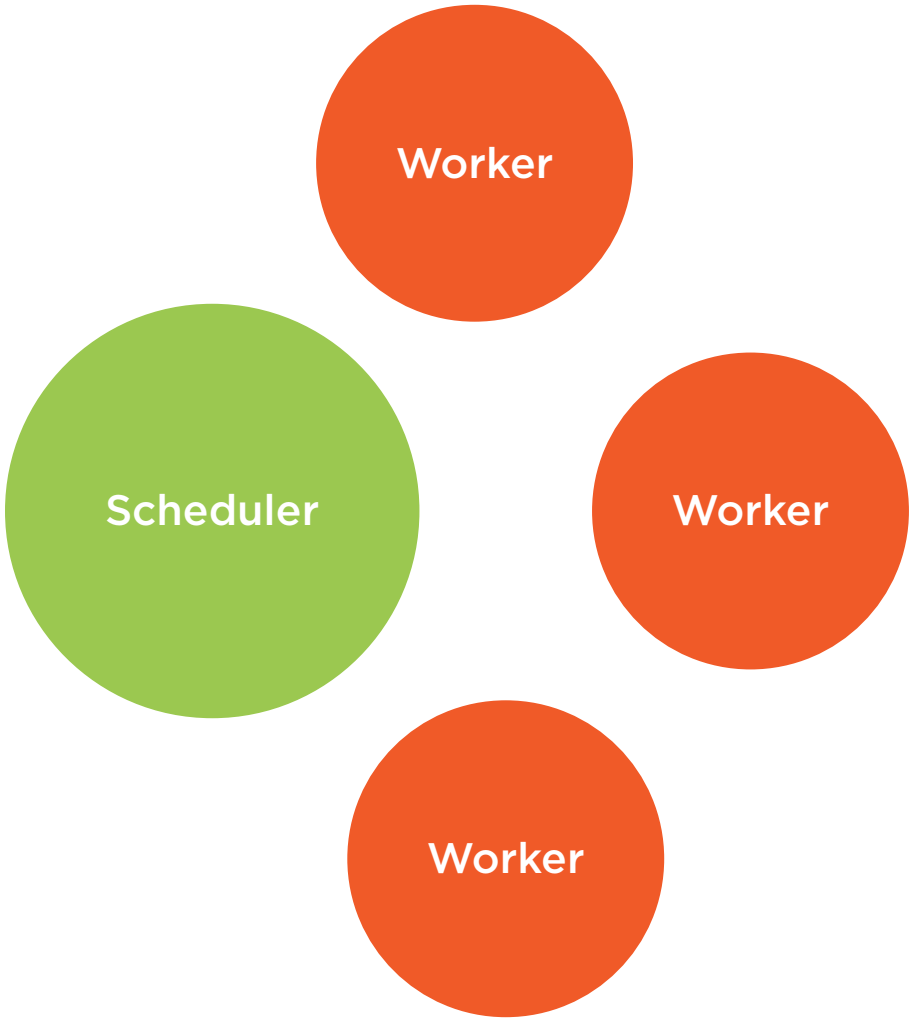
@pawel_kordek https://kordek.github.io

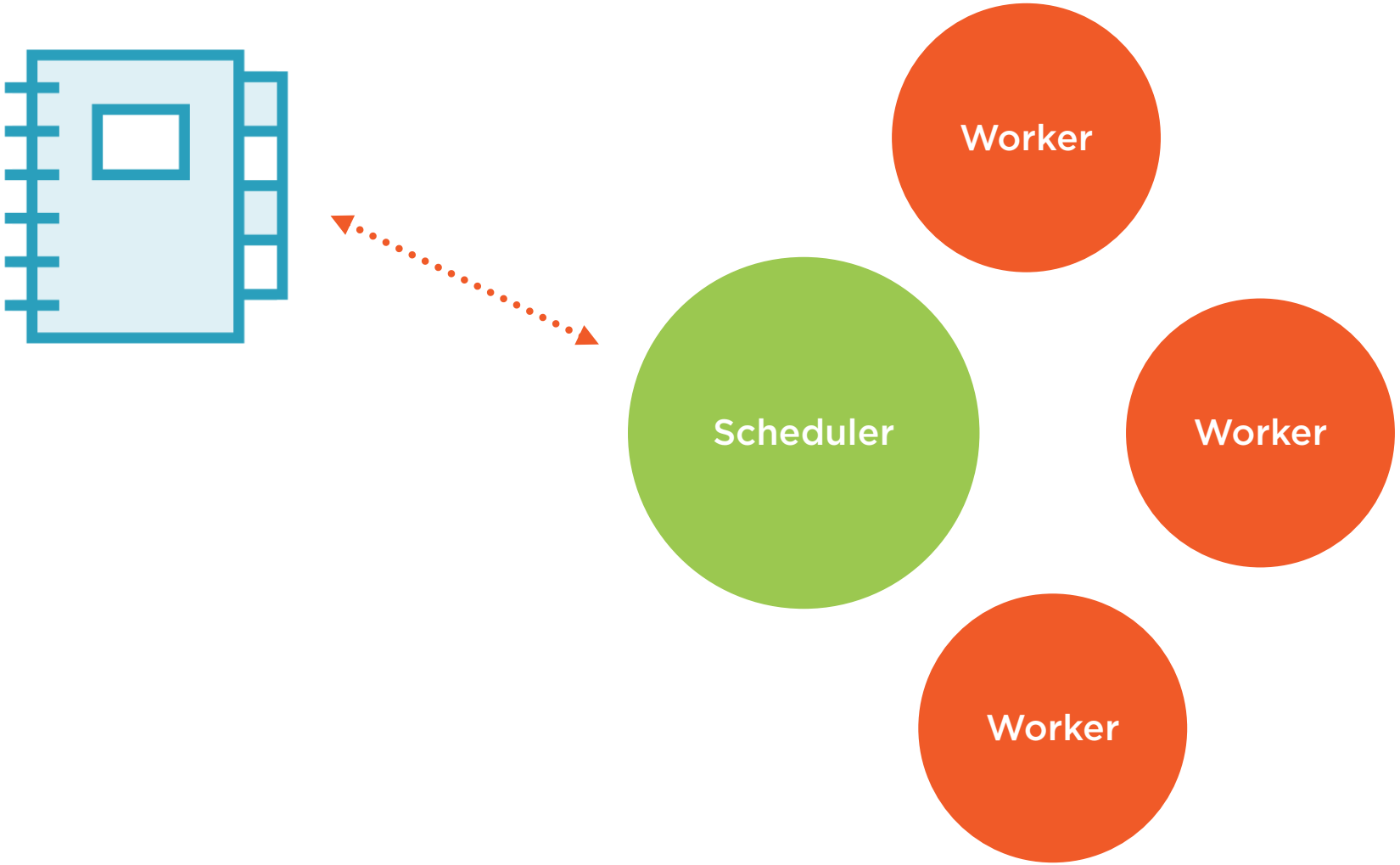
Distributed Dask

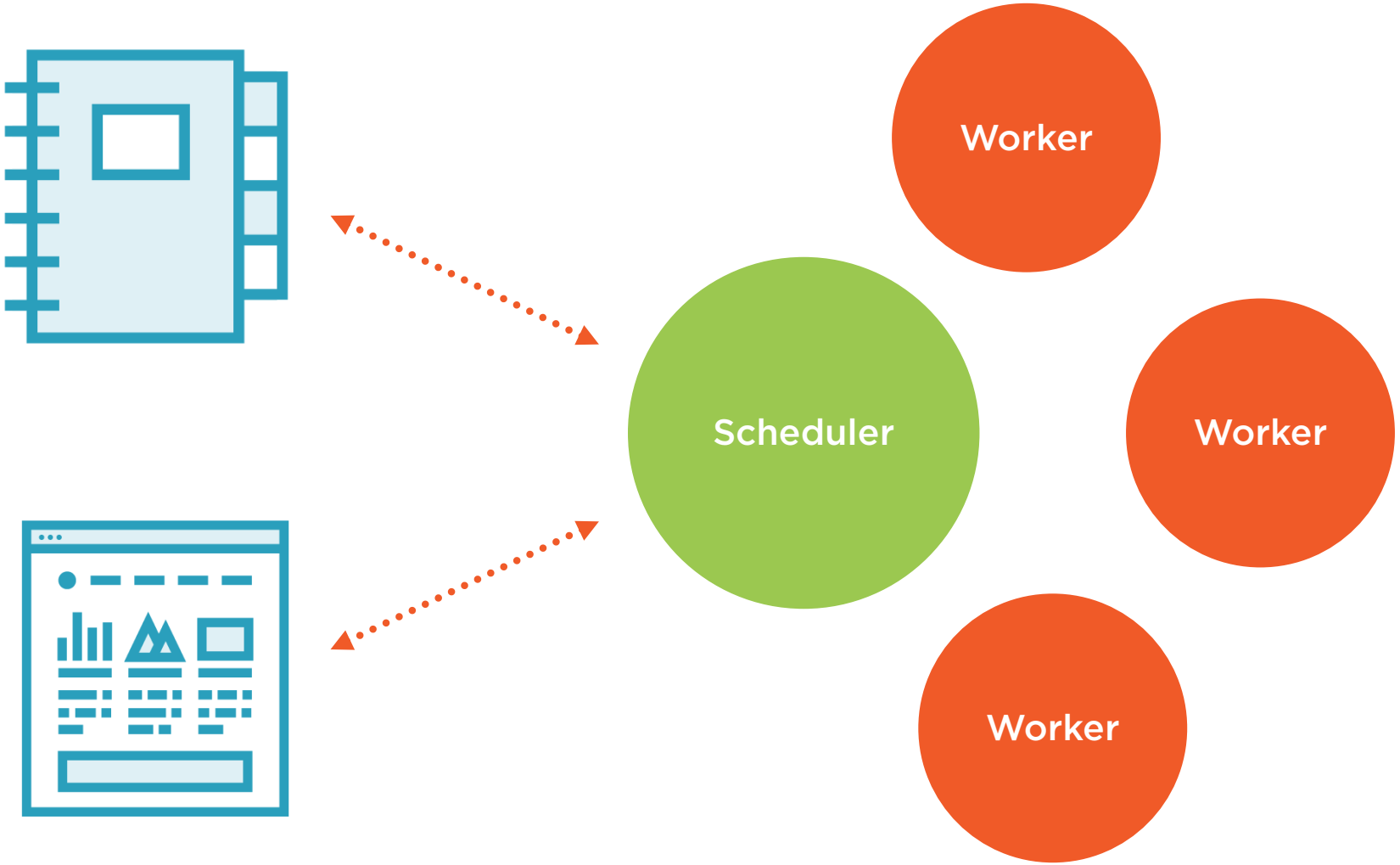
Architecture

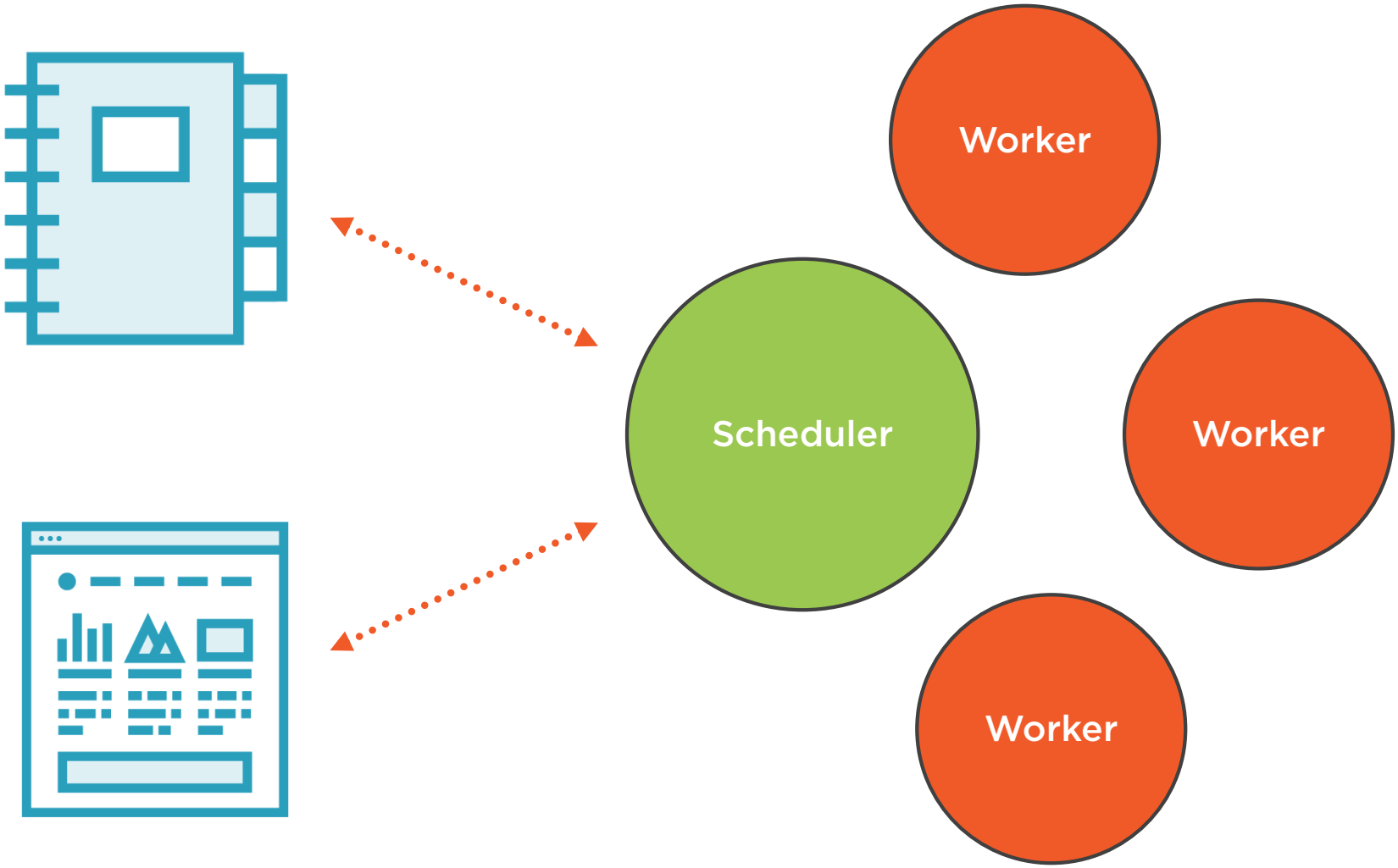
Docker environment

Test


Worker
Scheduler
Worker
Worker


Worker
Scheduler
Worker
Worker

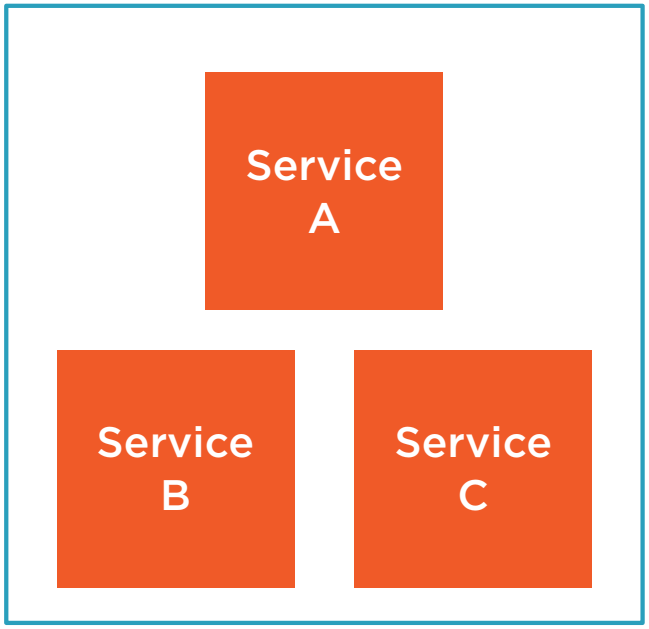

Worker
Scheduler
Worker
Worker


Worker
Scheduler
Worker
Worker

Docker Containers

'Pretend' they are separate machines.

Docker Compose



Orchestrates multiple containers.

Running the Services

Summary

Understanding Dask

Bags

Dashboard

Numpy and Pandas

Multiple Machines

Key Outcome

Confidence in writing Python data applications that scale.

Thank You!