Using MLflow in a Collaborative Scenario



Paweł Kordek

Software Engineer – Data and Machine Learning

@pawel_kordek



Overview



The need for a collaborative ML environment

Closer look at the current setup

Individual components

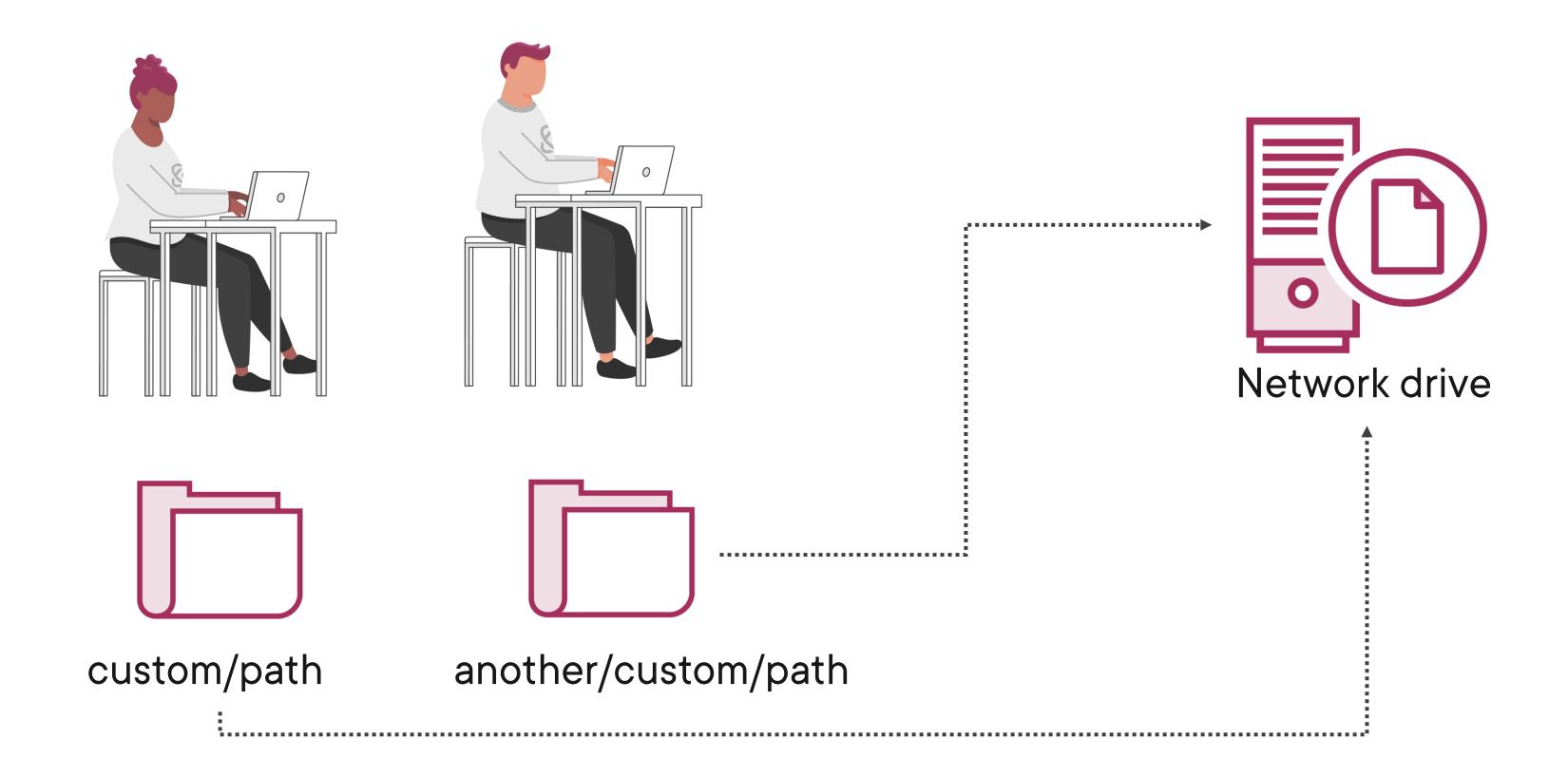
Complete MLflow deployment







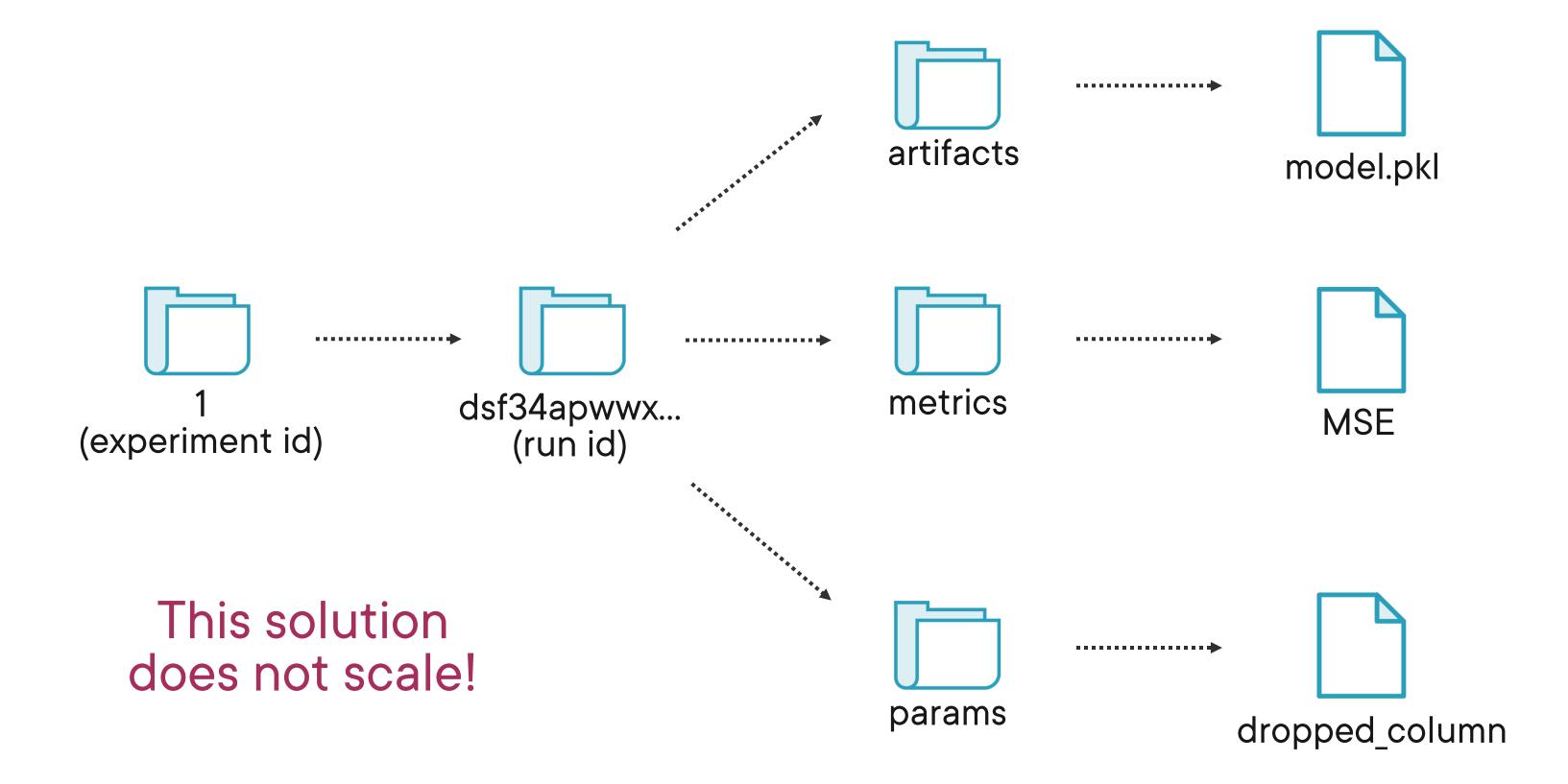




Changing MLflow Storage Location



Inside 'mlruns'

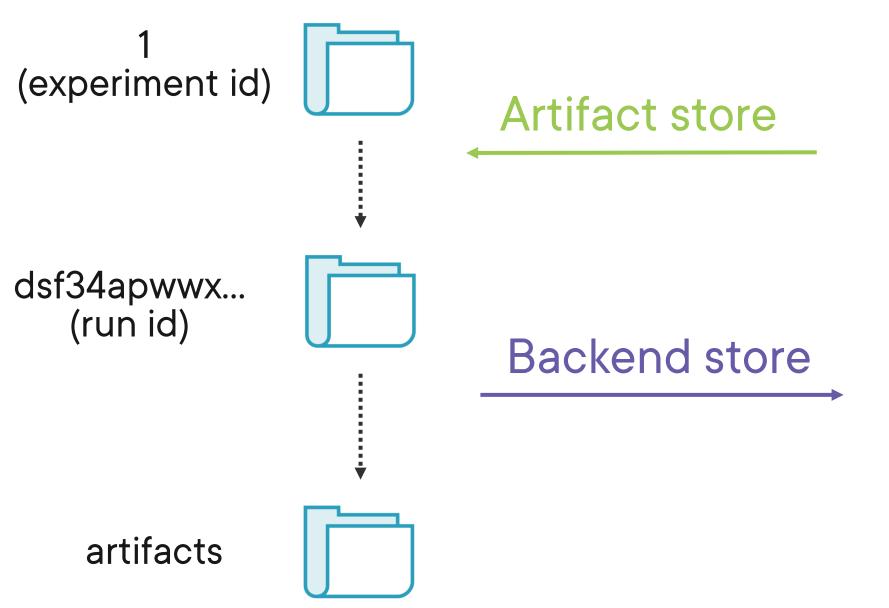


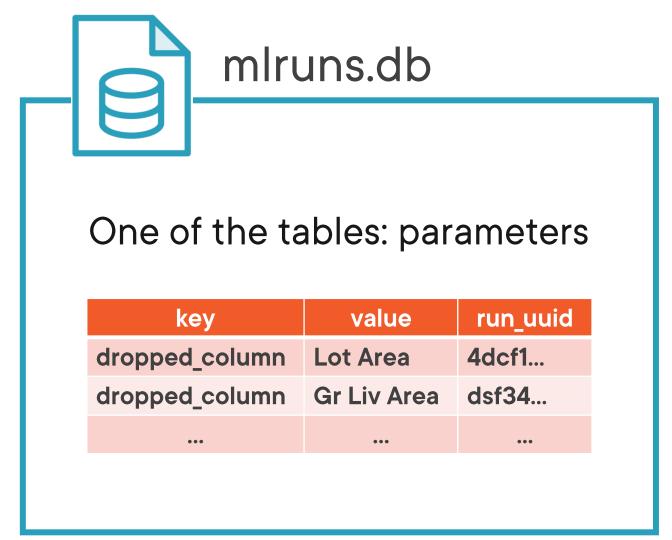


Using SQLite

Artifacts: Same as before

Everything else: Single SQLite file





Using SQLite Backend Store

```
in the Python code
mlflow.set_tracking_uri(
    "sqlite:///custom/location/mlruns.db"
mlflow ui
--backend-store-uri sqlite:///custom/location/mlruns.db
                                                                starting the UI
```





Artifact store cannot be customized in this setup.

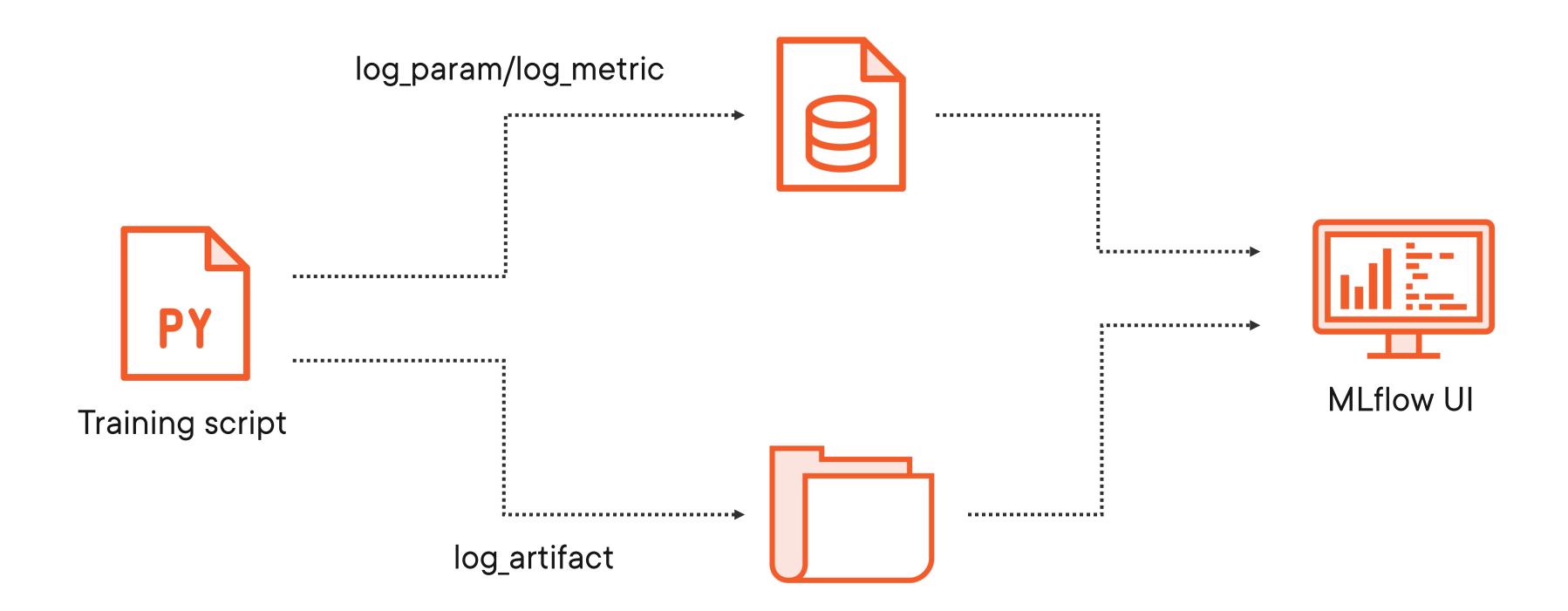
Up Next: MLflow Server



MLflow Server



Without MLflow Server



Using MLflow Server

```
mlflow server --host 0.0.0.0

--backend-store-uri

sqlite:///data/mlflow/mlruns.db

--default-artifact-root

file:///data/mlflow/artifacts

mlflow.set_tracking_uri("localhost:5000")

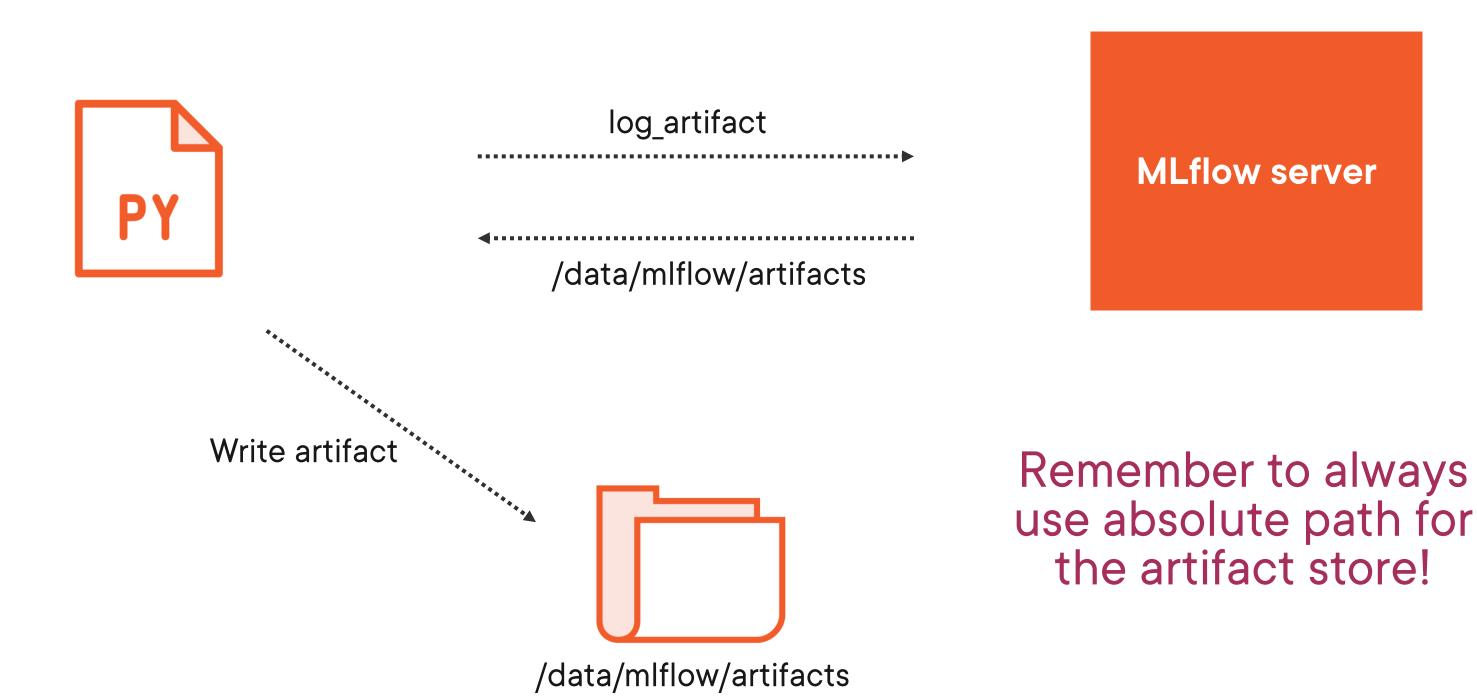
connecting to the server server in the Python script
```



Backend Store with MLflow Server



Artifact Store with MLflow Server





Demo



MLflow server

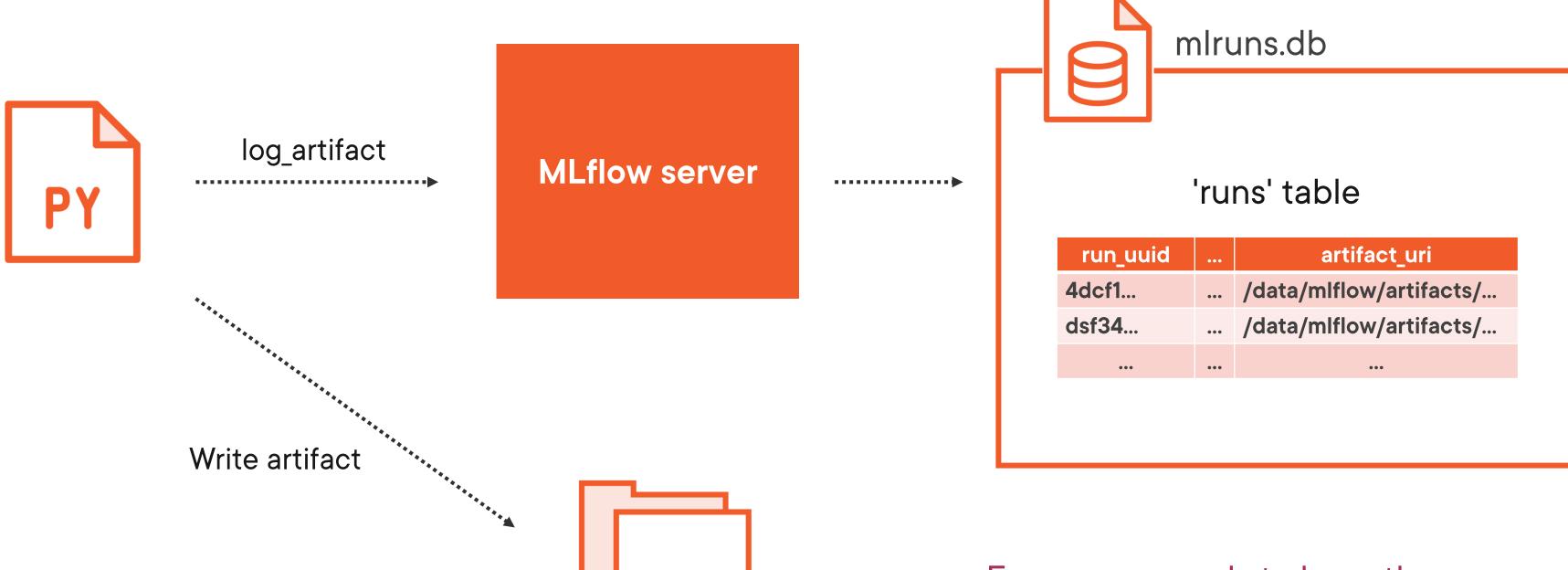


Complete MLflow Deployment

Everyone needs to run their own MLflow server.

/data/mlflow/artifacts/...

SQLite can be slow.



Every user needs to have the artifacts folder mounted on the same path.



Some Artifact Store Alternatives



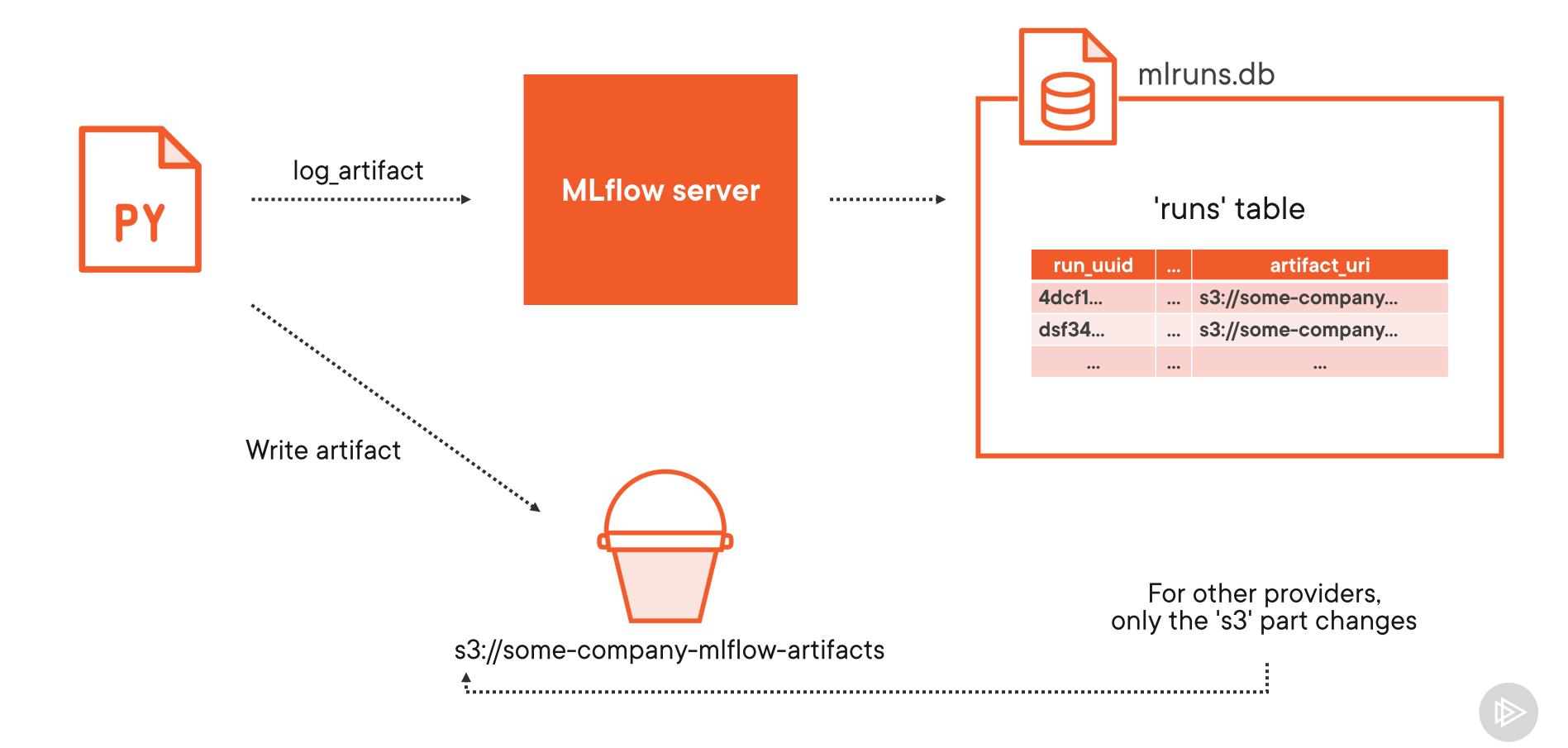
Some Artifact Store Alternatives

Amazon S3

Azure Blob Storage

Google Cloud
Storage





```
mlflow server --host 0.0.0.0
    --backend-store-uri
    sqlite:///data/mlflow/mlruns.db
    --default-artifact-root
    s3://some-company-mlflow-artifacts
```

Cloud Storage for Artifacts

Amazon S3

Other Artifact Store Options



Remote Server



mlflow.set_tracking_uri("mlflow.companyname.net"





Summary



Shortcomings of the default setup

Using SQLite

Distinction between artifact store and backend store

MLflow server

Architecture of a complete MLflow deployment

