Orchestrating ELT with Fivetran and Airflow
Agenda

Introduction ................................................................. 1

Case Study ........................................................................ 3

What is Fivetran ............................................................. 7

Fivetran Airflow Provider ............................................... 10

Demo .............................................................................. 14

Trying it out! .................................................................... 16
Fivetran + Airflow Case Study
Fivetran, Airflow, and Data Warehousing costs

- Fivetran pricing *independent* of data sync frequency
- Some data warehouses’ pricing is *dependent* of data sync frequency
- Created a dynamic schedule for database replication to warehouse with Airflow
Fivetran, Airflow, and Data Warehousing costs

- Fivetran pricing *independent* of data sync frequency
- Some data warehouses’ pricing is *dependent* of data sync frequency
- Created a dynamic schedule for database replication to warehouse with Airflow
- 20% reduction in data warehouse costs!
What makes Fivetran different?

- Automatic Data Updates (DML)
- Automatic Schema Migrations (DDL)
- Automated Recovery from Failure (Idempotent)
- Micro-batched architecture
- Extensible with the Fivetran Airflow Provider
Fivetran + Airflow - Top Uses Cases

Syncing Syncs
- Programmatically prioritize syncs
- Inherent benefits of automated data ingestion

Triggering Transformations
- Transformations across sources
- Transform too late - latency problems and SLA issues
- Transform too early - missing data and integrity issues

DataOps Management
- Manage data tasks before Fivetran and after dbt
- Integrate projects across data teams
Fivetran Airflow Provider
The Astronomer Registry
A deep library of standard patterns for tight integration between Airflow and almost any tool in the modern data stack
Including...

Fivetran
FivetranHook

- Abstracts away logic of Fivetran and Airflow interactions
- Leveraged by other components of provider for execution
- For **advanced** users and use cases only
FivetranOperator & FivetranSensor

- The FivetranOperator starts a Fivetran data sync
- The FivetranSensor monitors the execution of a Fivetran data sync
- All operators interfacing with Fivetran API
Fivetran + Airflow

Linkedin Ads → Airflow

Twitter Ads → Airflow

Data Warehouse
- Target Database
- ad_reporting package
- Modeled Data

Compute Engine
For more...
https://5tran.co/AirflowSummit
Thank You
nick.acosta@fivetran.com