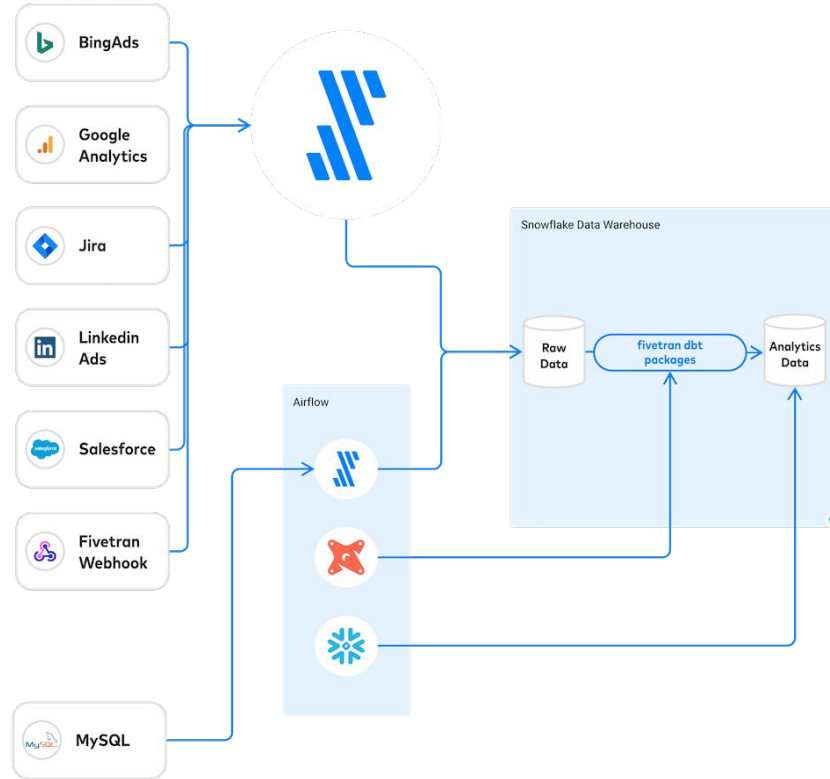

Orchestrating ELT with Fivetran and Airflow

Agenda

Introduction	1
Case Study	3
What is Fivetran	7
Fivetran Airflow Provider	10
Demo	14
Trying it out!	16

Fivetran + Airflow Case Study

Architecture



Fivetran, Airflow, and Data Warehousing costs

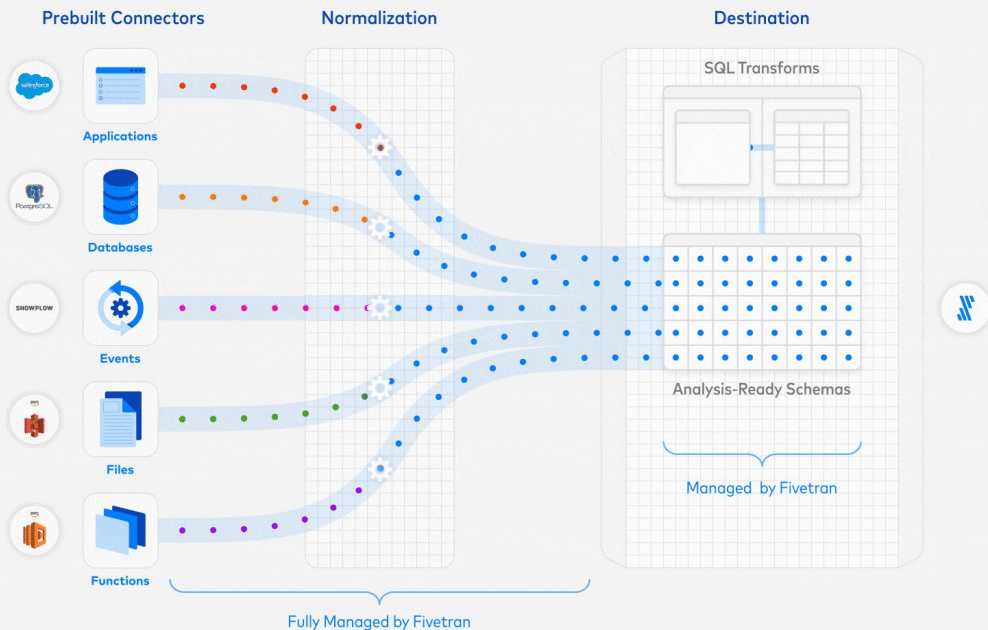
- Fivetran pricing ***independent*** of data sync frequency
- Some data warehouses' pricing is ***dependent*** of data sync frequency
- Created a dynamic schedule for database replication to warehouse with Airflow

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!

Fivetran

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

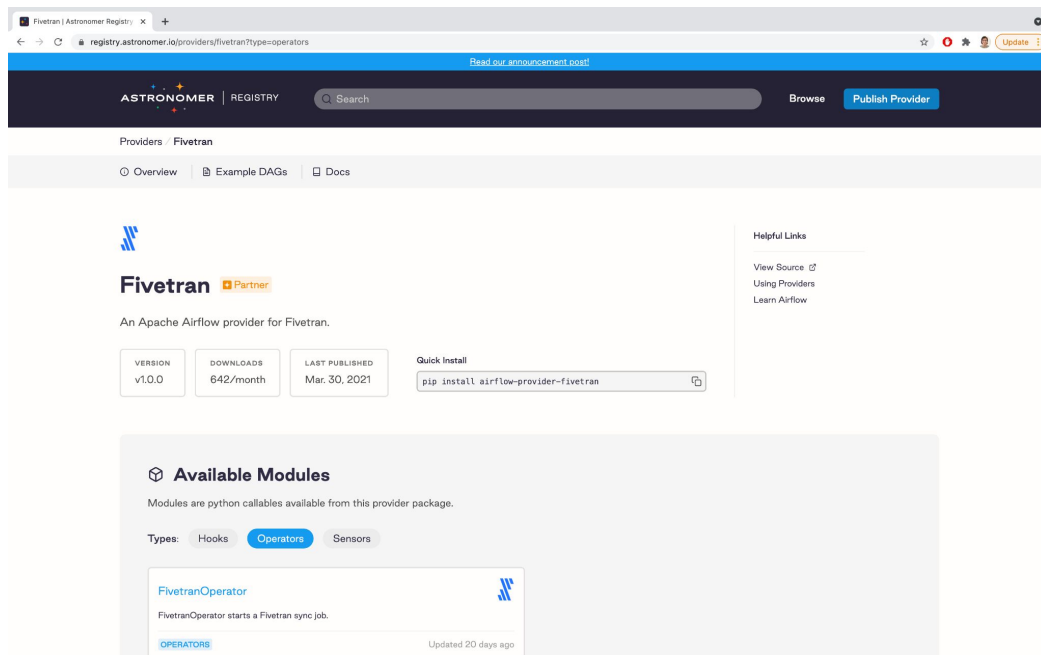
Fivetran Airflow Provider

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



The screenshot shows the Astronomer Registry page for the Fivetran provider. The page is titled "Fivetran" and is marked as a "Partner". It describes the provider as "An Apache Airflow provider for Fivetran." The page includes a "Quick Install" section with the command `pip install airflow-provider-fivetran`. Below this, there is an "Available Modules" section with tabs for "Hooks", "Operators", and "Sensors". The "Operators" tab is selected, showing the "FivetranOperator" module, which is described as "FivetranOperator starts a Fivetran sync job." and was updated 20 days ago.

Providers Fivetran

Overview Example DAGs Docs

Fivetran Partner

An Apache Airflow provider for Fivetran.

VERSION	DOWNLOADS	LAST PUBLISHED	Quick Install
v1.0.0	642/month	Mar. 30, 2021	<code>pip install airflow-provider-fivetran</code>

Available Modules

Modules are python callables available from this provider package.

Types: Hooks **Operators** Sensors

FivetranOperator

FivetranOperator starts a Fivetran sync job.

OPERATORS Updated 20 days ago

Helpful Links

- View Source
- Using Providers
- Learn Airflow

FivetranHook

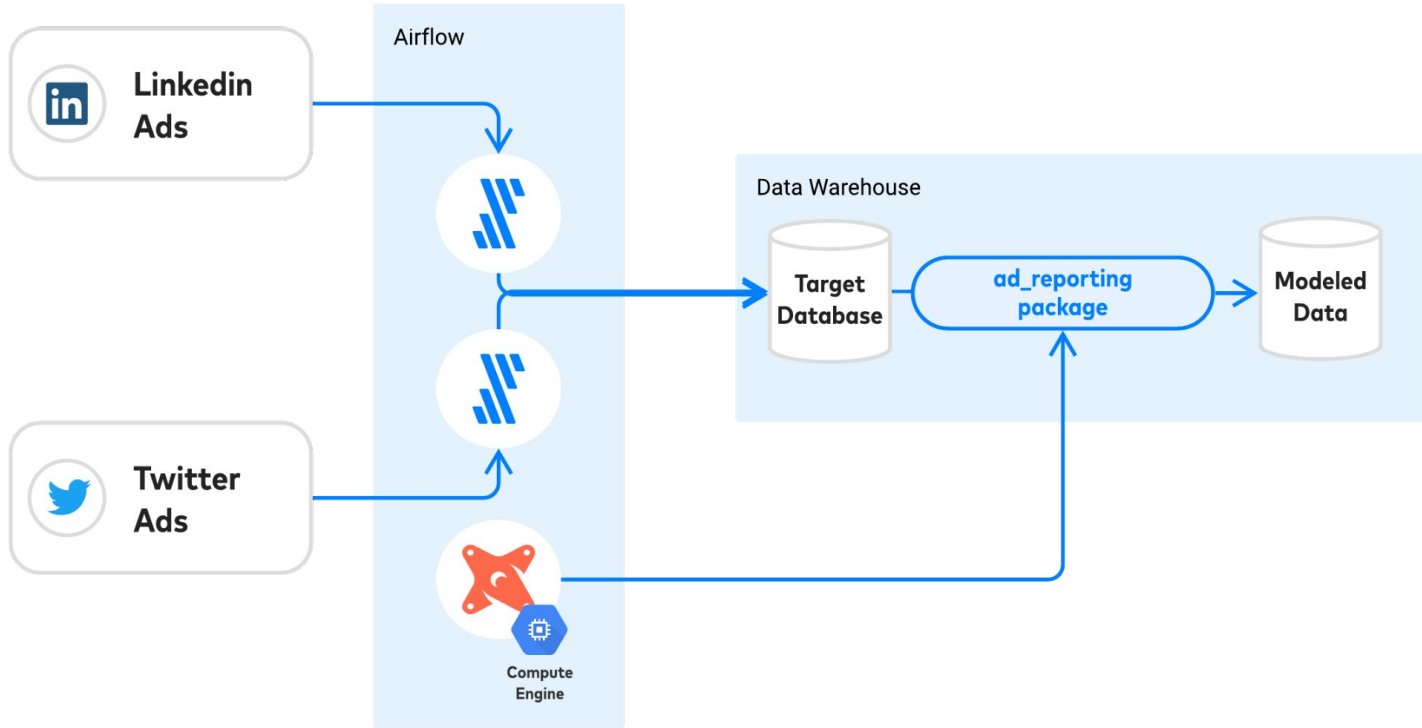
- Abstracts away logic of Fivetran and Airflow interactions
- Leveraged by other components of provider for execution
- For ***advanced*** users and use cases only

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

Demo

Fivetran + Airflow



For more...

<https://5tran.co/AirflowSummit>

Thank You

nick.acosta@fivetran.com