GNARLY Data_Waves

PRESENTED BY dremio

EPISODE 17

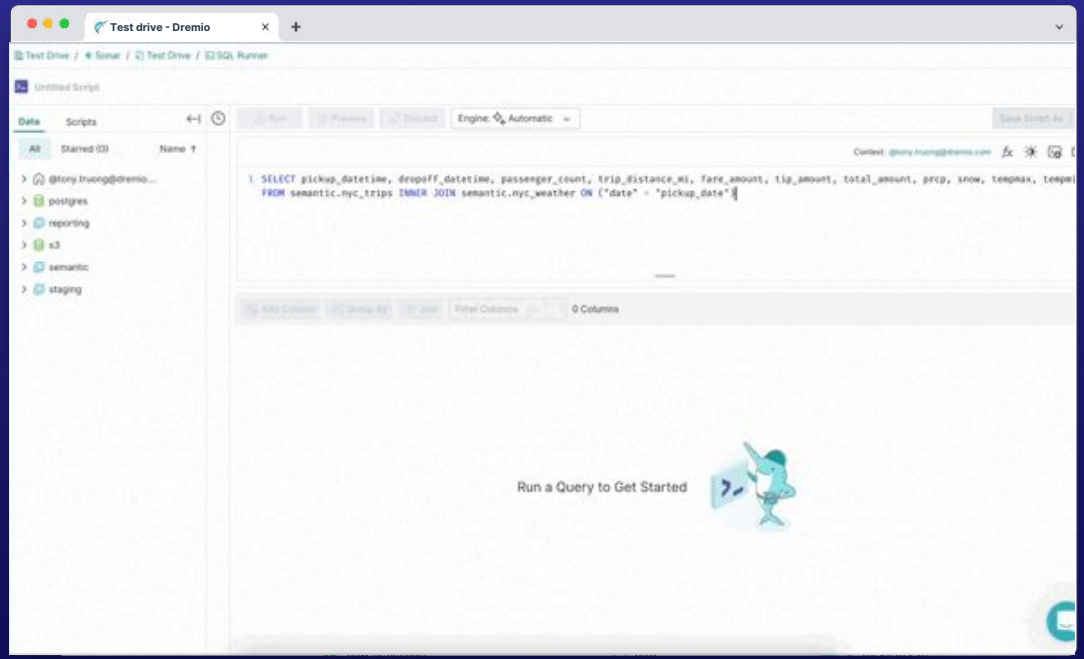
Unified Access for Your Data Mesh

Self-Service Data with Dremio's Semantic Layer

Experience the data lakehouse with Dremio Test Drive

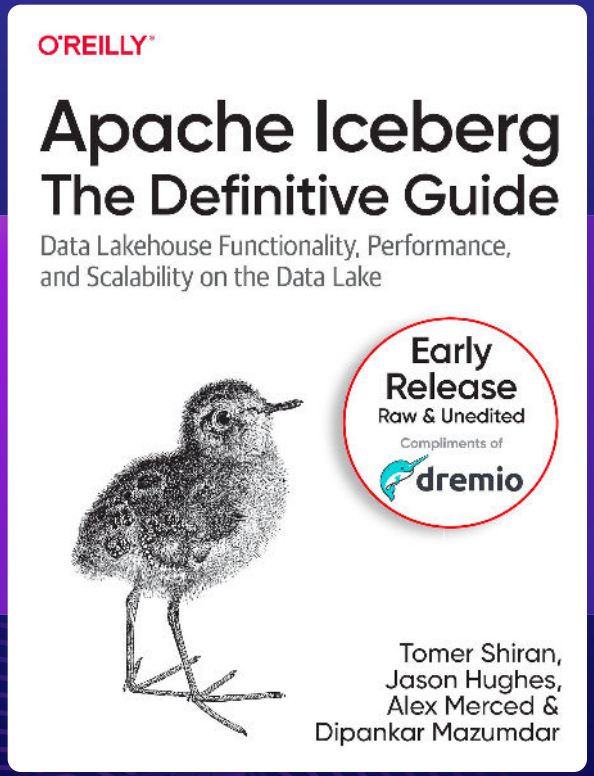
- Sub-second query on 1 billion rows of data joining Amazon S3 with a Postgres database
- Connect to Tableau or Power BI and build a dashboard with this dataset
- Everything hosted by Dremio - 100% free for you

Start Test Drive



GNARLY Data_Waves

Apache Iceberg: The Definitive Guide



Upcoming shows

Register now

EPISODE 16

Automatic Data Optimization with Dremio Arctic

May 9th, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 17

Unified Access for Your Data Mesh: Self-Service Data with Dremio's Semantic Layer

May 16, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 18

Easily Migrate Hadoop Workloads to AWS with Dremio

May 23, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 19

How MSK Accelerates Cancer Research with Dremio's Data Lakehouse

May 31, 2023 at 8AM PST | 11AM EST | 4PM GMT

dremio | aws | +ableau

AWS Dev Day: Chicago

Experience Dremio Cloud and Tableau on Amazon S3

May 17th, 10am CST-12pm | Lunch to follow

dremio | aws

AWS Summit Toronto

June 14th, 2023
Metro Toronto Convention Centre

dremio | aws | +ableau

AWS Dev Day: New York

Experience Dremio Cloud

July 25th, 10:00 am to 1:00 pm
Endeavor 1 | Courtyard - 461 W 34th St.
New York, NY 10001

dremio | aws

AWS Summit New York

July 26th, 2023
Javits Center

BIG DATA & AI WORLD | BARC

dremio

Coalesce by dbt

Oct 16-20, 2023
Hilton Bayfront San Diego

BIGDATA & AI by corp
PARIS

TIME TO ACCELERATE

September 25 & 26, 2023

Paris Convention Center

Unified Access for Your Data Mesh

Self-Service Data with Dremio's Semantic Layer

Alex Merced

Developer Advocate, Dremio

GNARLY Data_Waves

Data Utopia

- Data is consistent
- Data is fresh
- Data is correct
- Data is discoverable
- Data is accessible by the right people

Data Reality

- Data is inconsistent (so many copies being used with opaque lineage)
- Data is stale (too much data for a centralized team to make SLAs)
- Data is correct (too much data to monitor quality within SLAs)
- Data is discoverable (data in many places with little documentation)
- Data is accessible by the right people (due to many sources access rules are opaque and distributed)

Solution

Data Mesh

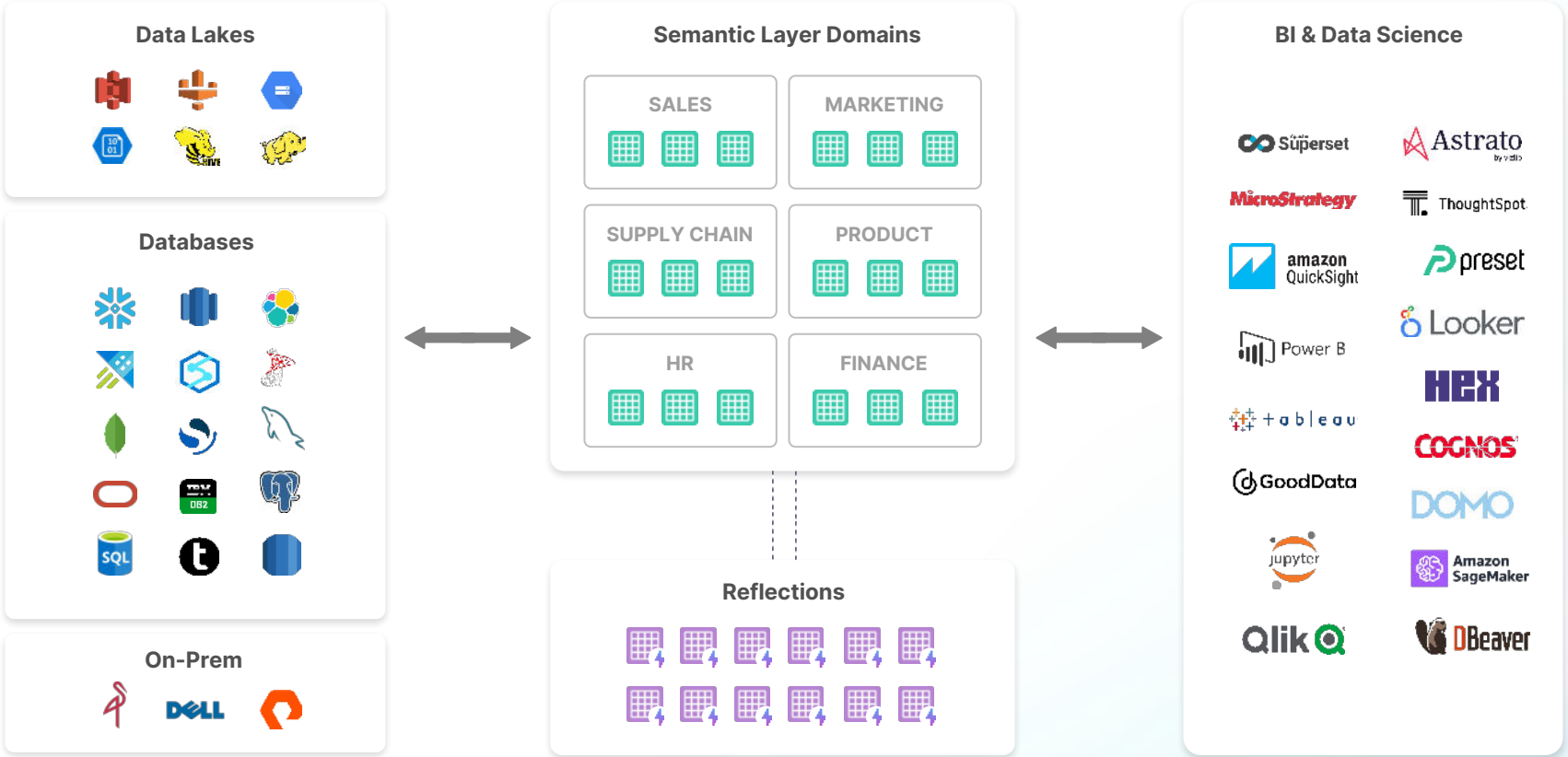
a. An architecture that relies on decentralized creation of data into data products using shared governance and standards.
    i. Meet SLAs as different teams focus on different parts of the data making freshness and correctness easier to achieve

Solution

Dremio's Semantic Layer

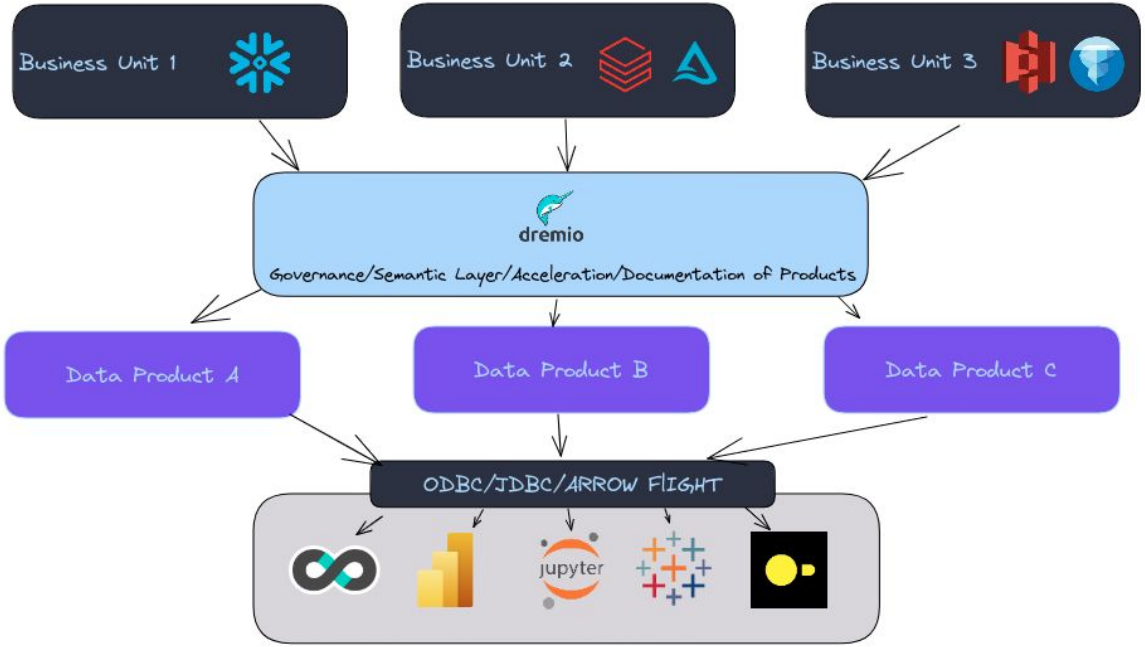
a. A layer for organizing, documenting, and governing your data across different sources with visibility into your data's linease, further enhanced by Dremio's no copy architecture.
    i. One central place for all end users to discover data across all your data products
    ii. One central place to manage access all your data by decentralized teams
    iii. Documentation and Lineage provide context to end users to better use data

Governed, Domain-based Semantic Layer



Confidential - Do Not Share or Distribute

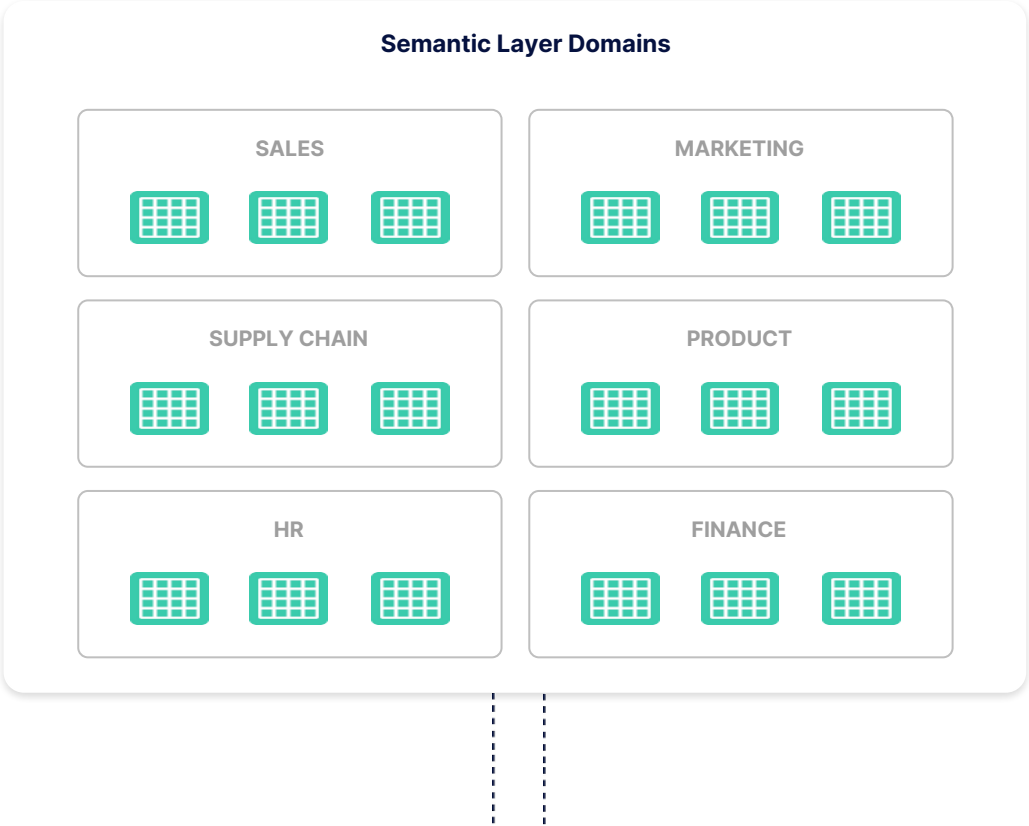
Understanding Data Mesh



Four Principles of Data Mesh

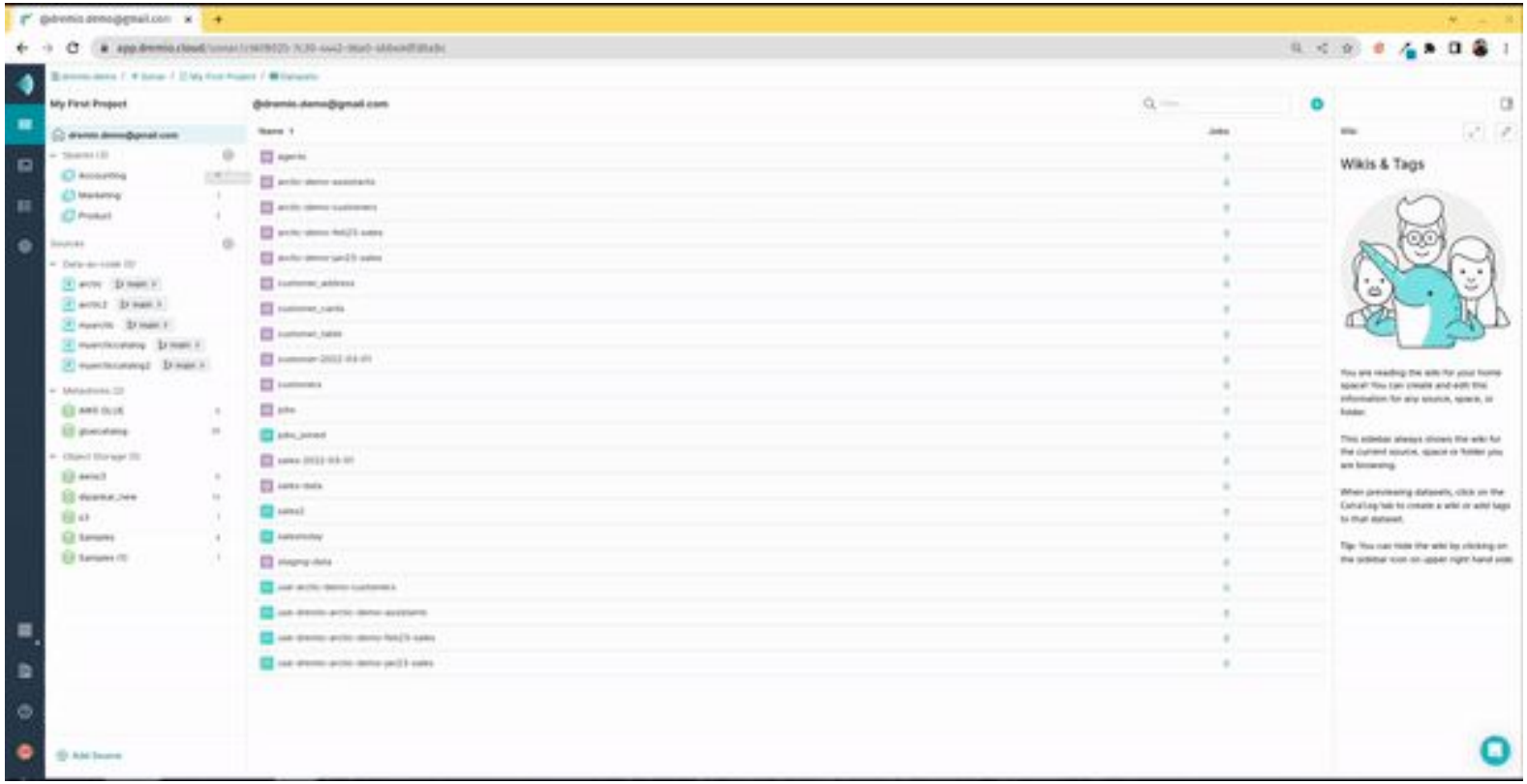
- Domain Oriented Decentralization
- Data as Product
- Self-Service Infrastructure
- Computational Governance

Step 1 - Determine Domains



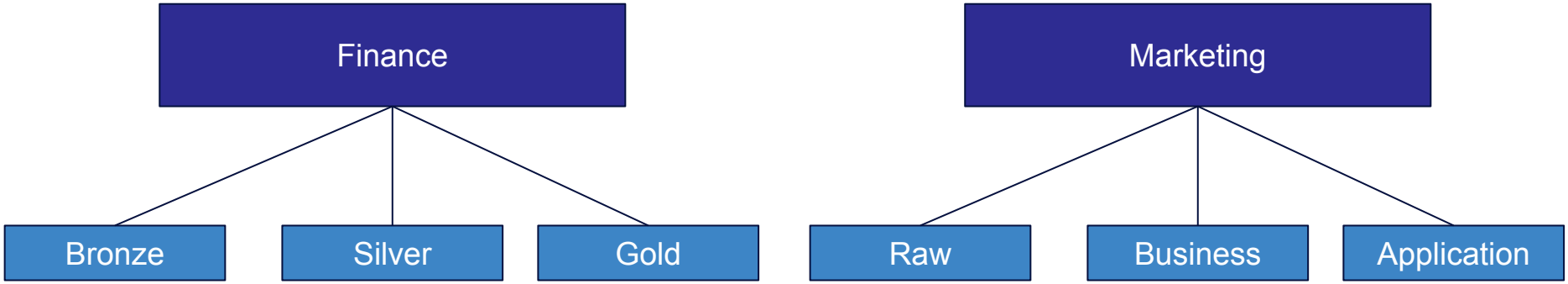
GNARLY Data_Waves
dremio

Step 2 - Create a Space for Each Domain



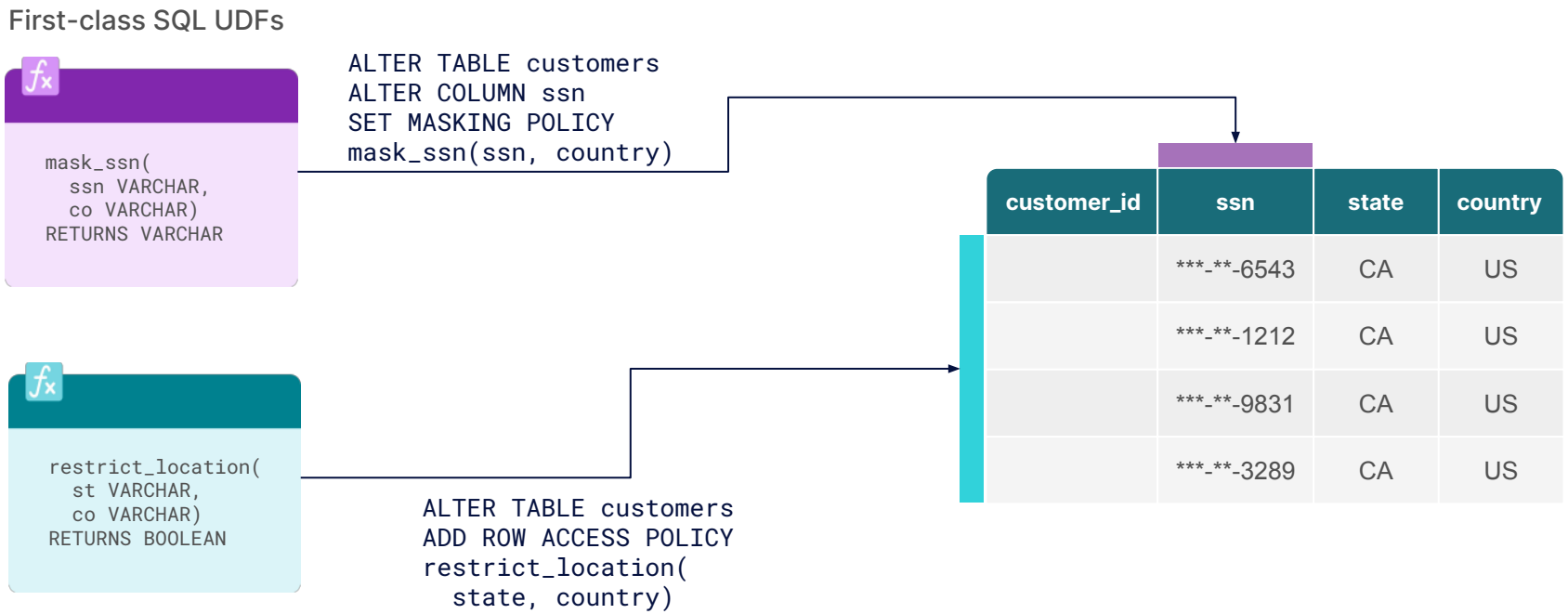
GNARLY
Data_Waves
dremio

Step 3 - Standardize Semantics

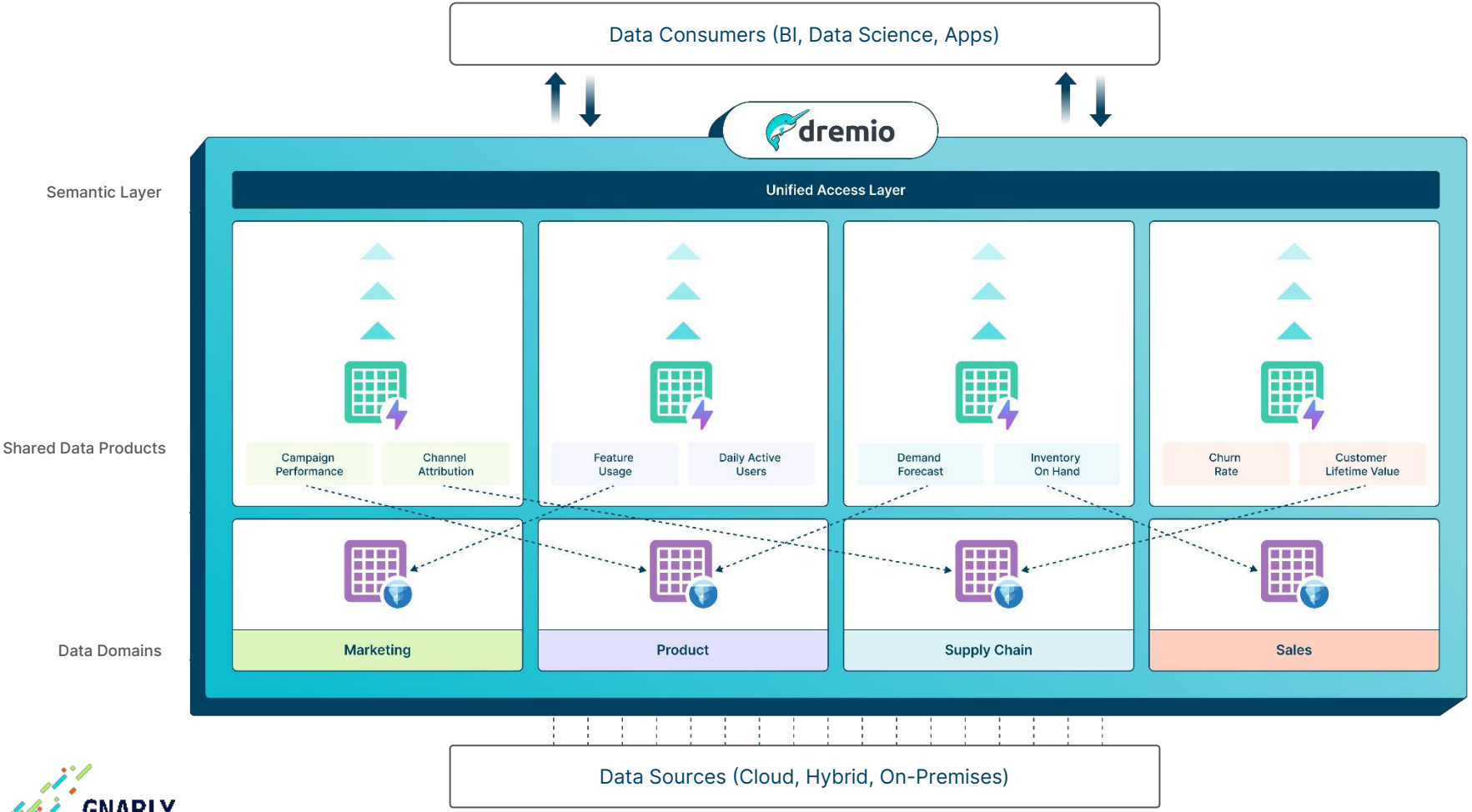


GNARLY
Data_Waves

Step 4 - Give end users access to production data



GNARLY Data Waves
dremio



GNARLY Data_Waves