



# How Dremio provides you fast and easy data access while saving you money

Presented by Alex Merced



# Alex Merced

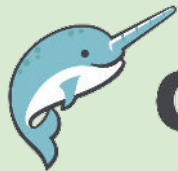
## Developer Advocate, Dremio

Alex Merced is a developer advocate for Dremio, a developer, and a seasoned instructor with a rich professional background. Having worked with companies like GenEd Systems, Crossfield Digital, CampusGuard, and General Assembly.

Alex is a co-author of the O'Reilly Book "Apache Iceberg: The Definitive Guide." With a deep understanding of the subject matter, Alex has shared his insights as a speaker at events including Data Day Texas, OSA Con, P99Conf and Data Council.

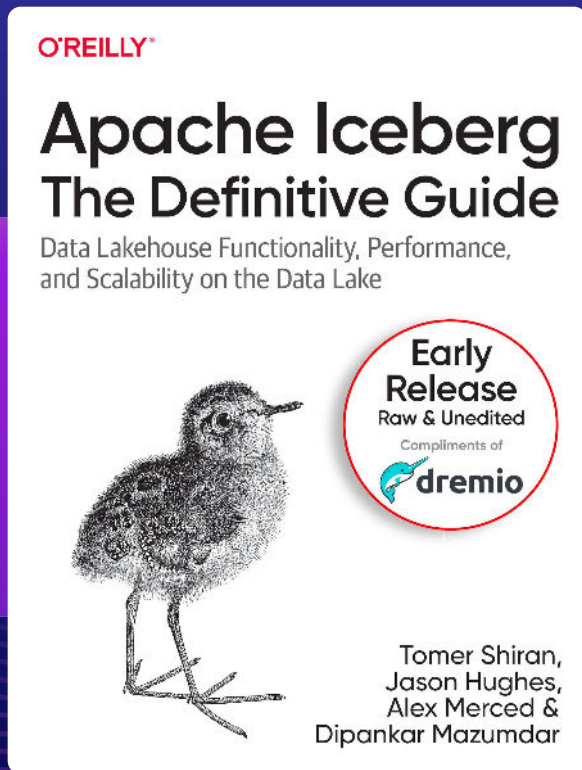
Driven by a profound passion for technology, Alex has been instrumental in disseminating his knowledge through various platforms. His tech content can be found in blogs, videos, and his podcasts, Datanation and Web Dev 101.

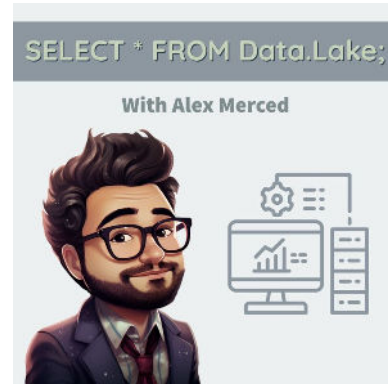
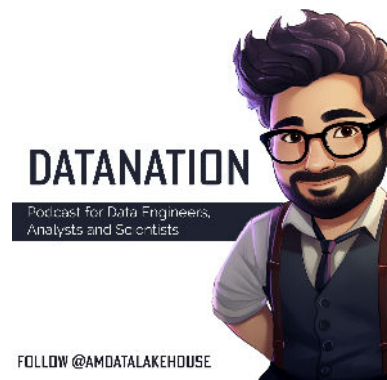
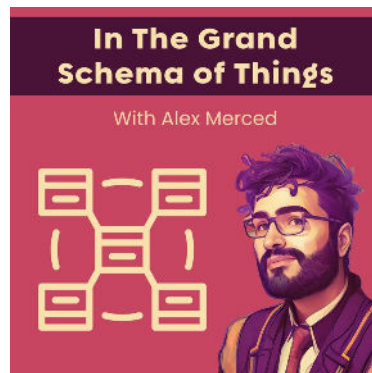
Moreover, Alex Merced has made contributions to the JavaScript and Python communities by developing a range of libraries. Notable examples include SencilloDB, CoquitoJS, and dremio-simple-query, among others.



# dremio

# Apache Iceberg: The Definitive Guide





Subscribe on Spotify/iTunes

# Did you know



Saved 60%  
on their  
storage  
costs



Saw 10x-30x  
performance  
improvement  
and reduce  
consultant  
spend.



+10% efficient,  
30x  
performance  
increase, lower  
compute and  
maintenance  
costs.

This isn't even factoring value from increased productivity and data democratization

What did they have in common?



## Cloud Infrastructure Cost Drivers

- Data Storage  
(more copies, more costs)
- Compute  
(more pipelines, more costs)
- Network Fees  
(Bandwidth, File Requests, Egress)

## On-Prem Infrastructure

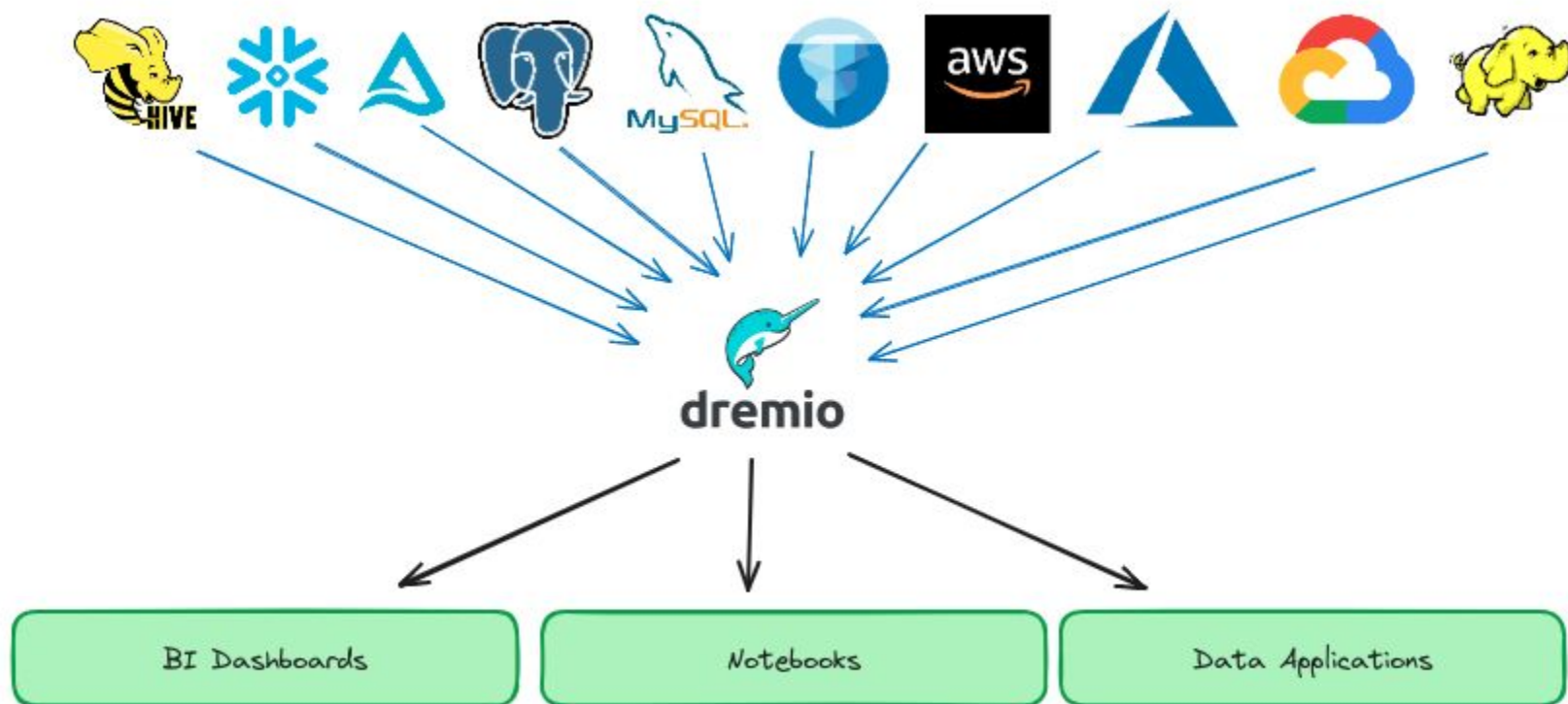
- Licensing Fees
- Costs of Additional Nodes  
(Coupled Storage and Compute)



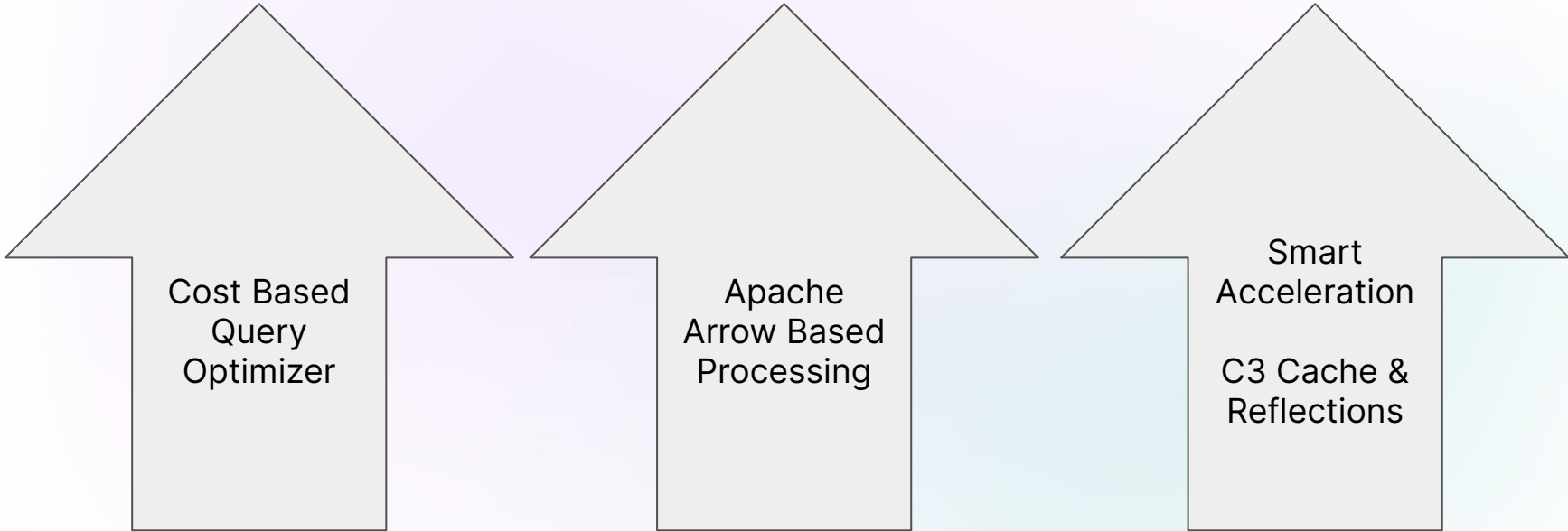
## How do you solve for these costs

- Federate sources to avoid need for data movement
- Better performance
- Smarter Acceleration

# Federate Sources



# Better Performance

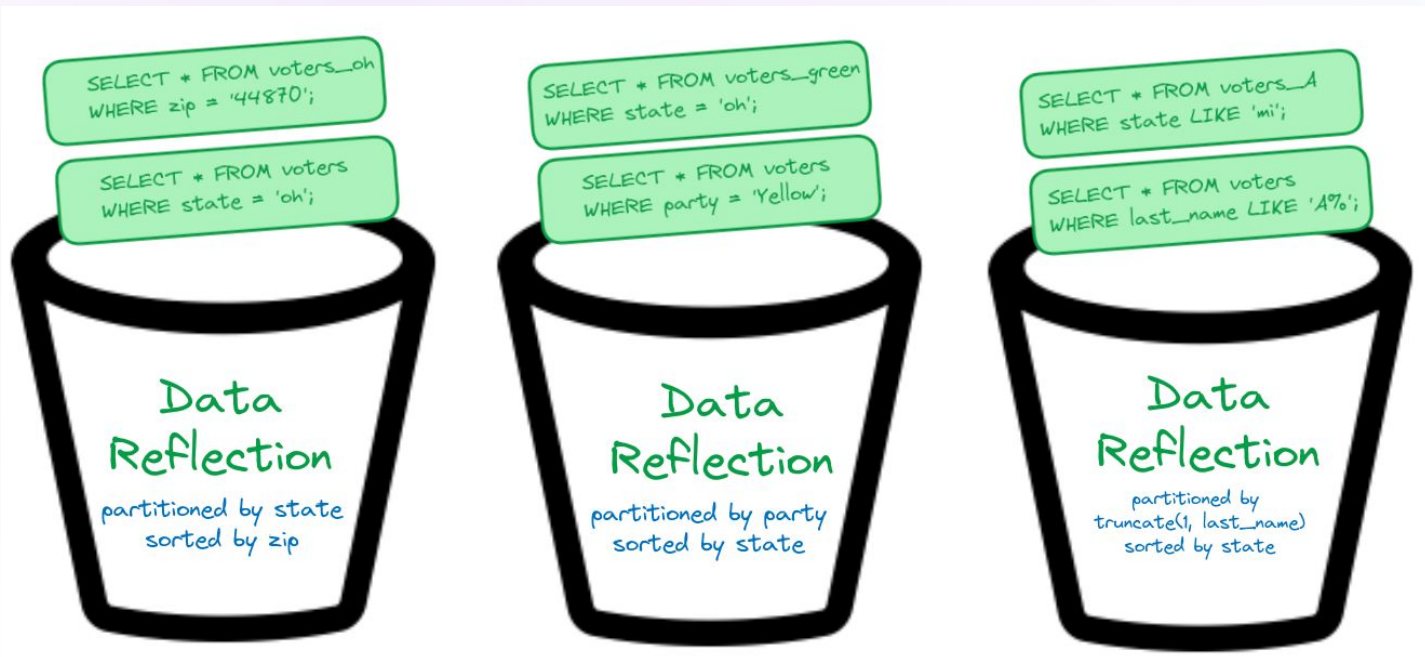
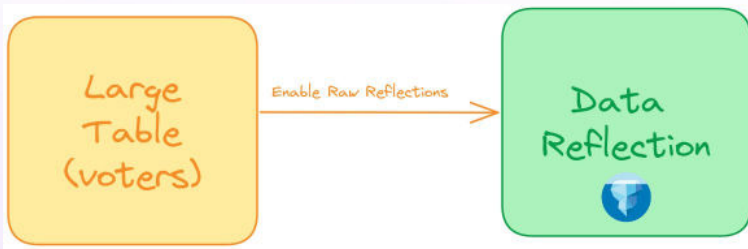


Cost Based  
Query  
Optimizer

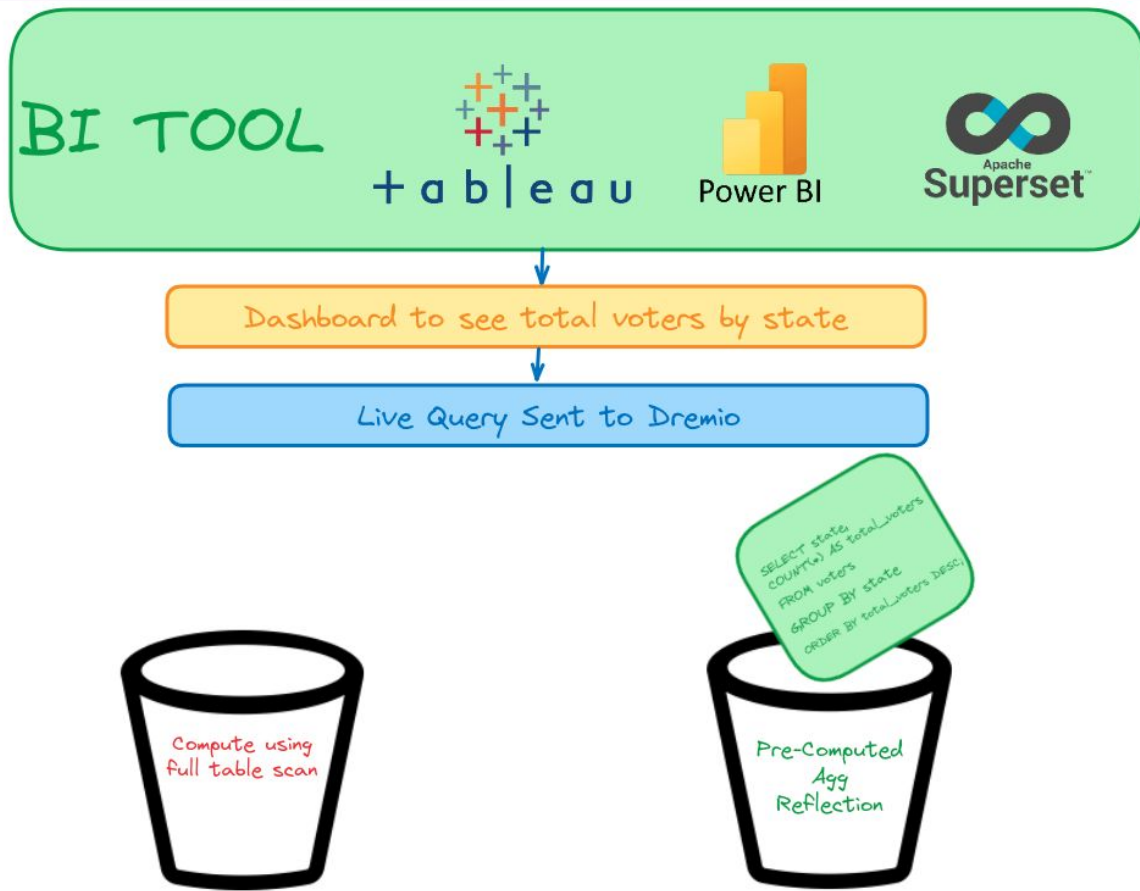
Apache  
Arrow Based  
Processing

Smart  
Acceleration  
  
C3 Cache &  
Reflections

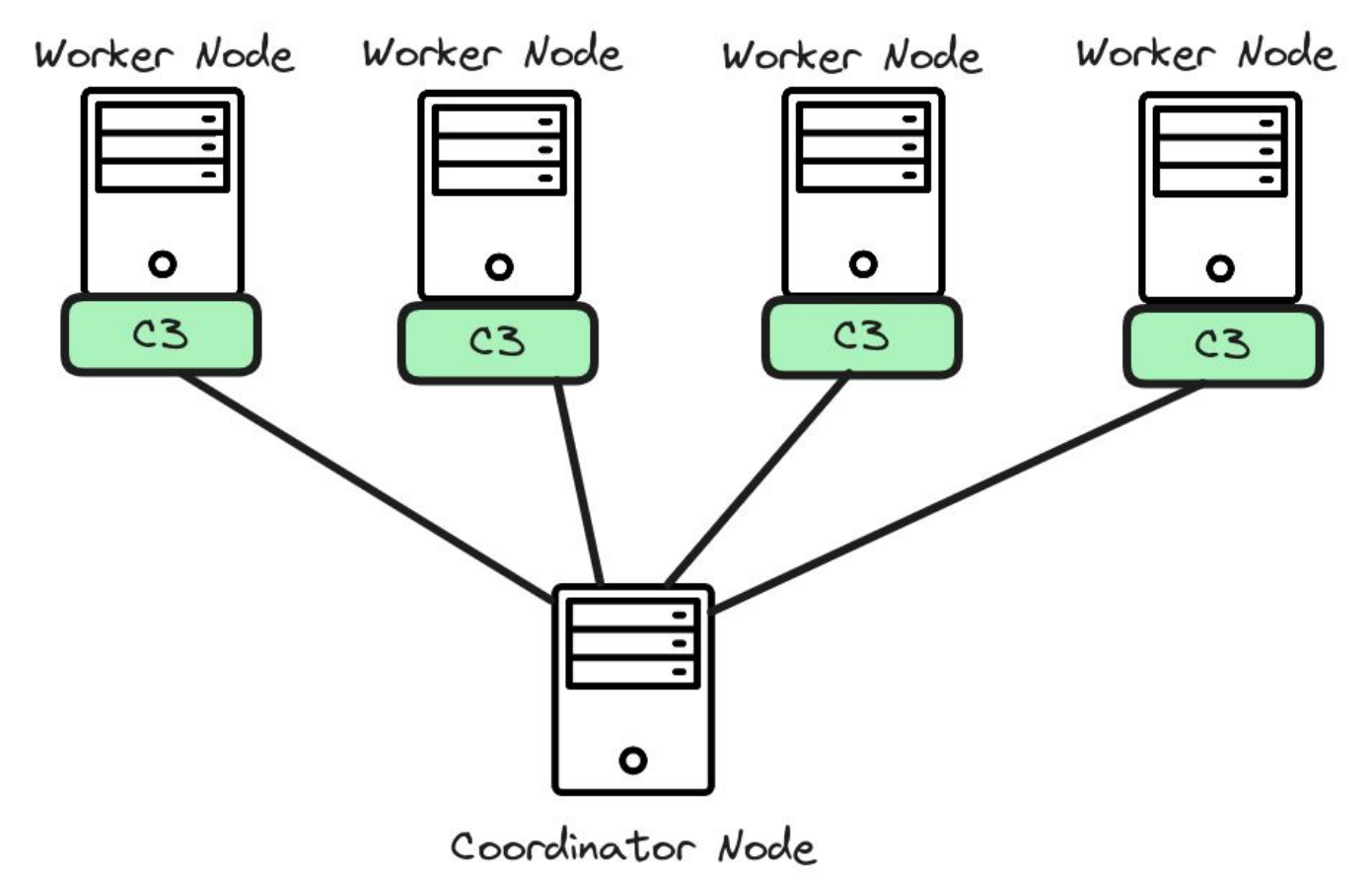
# Raw Reflections



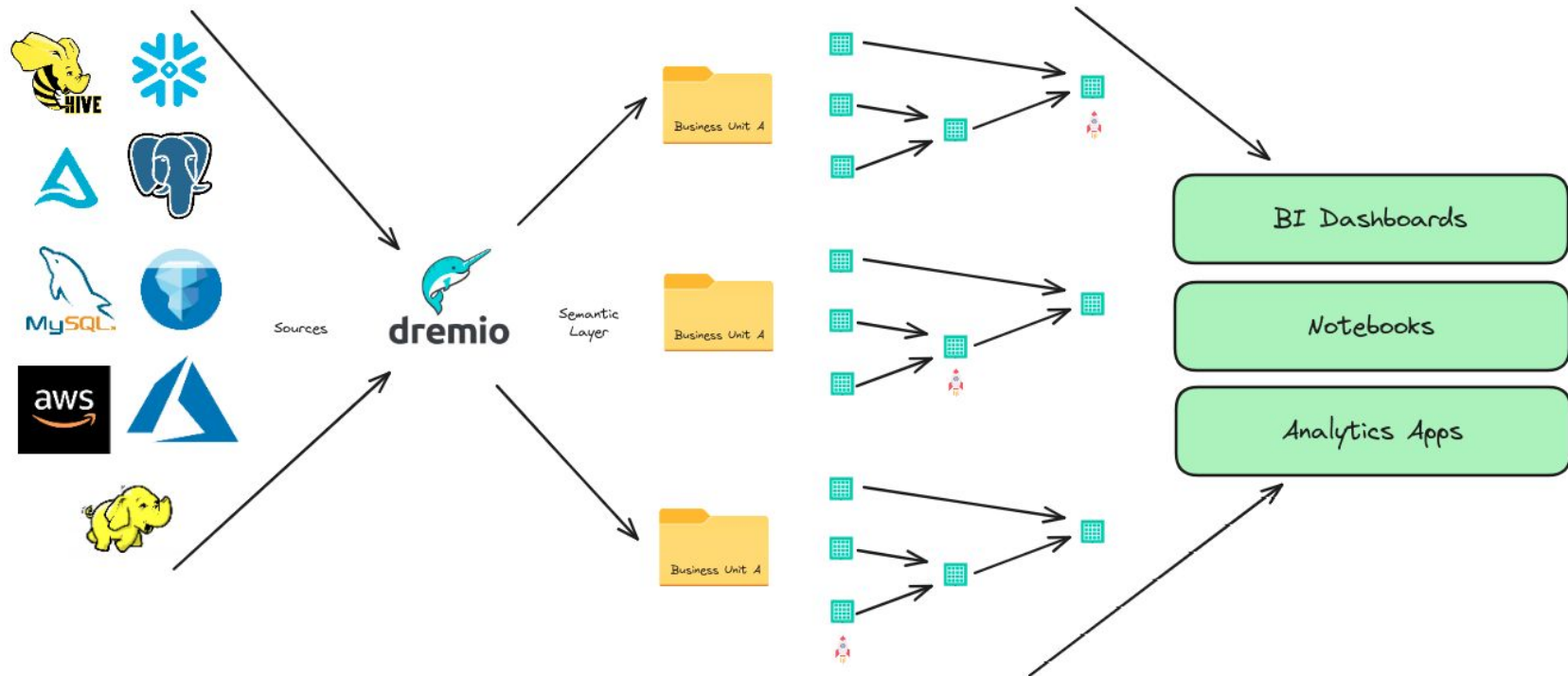
# Aggregate Reflections



# C3 - Columnar Cloud Cache



# When you put it all together



# Alex Merced

## follow @amdatalakehouse

