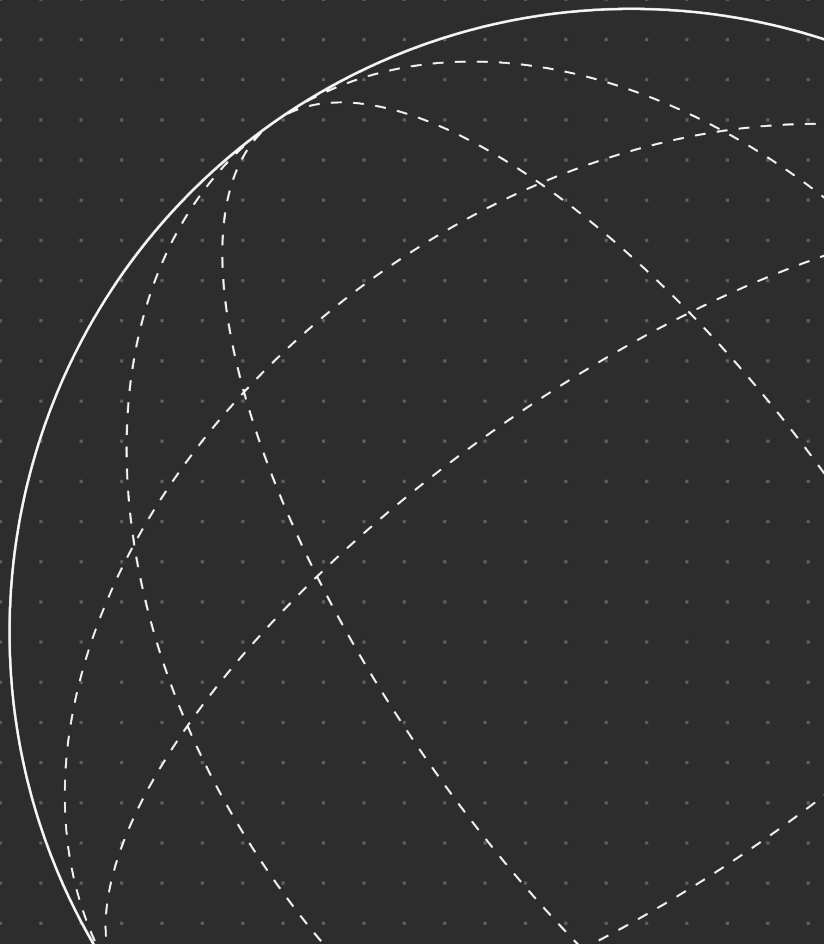


glassnode

# Glassnode Data Platform Journey

Festa OS 2024 - Girona, Octubre 2024



- State of Blockchain
- Que és Glassnode?
- Blockchain Ingestion
- Microservices Mesh
- Medaillon Architecture
- Següents passos



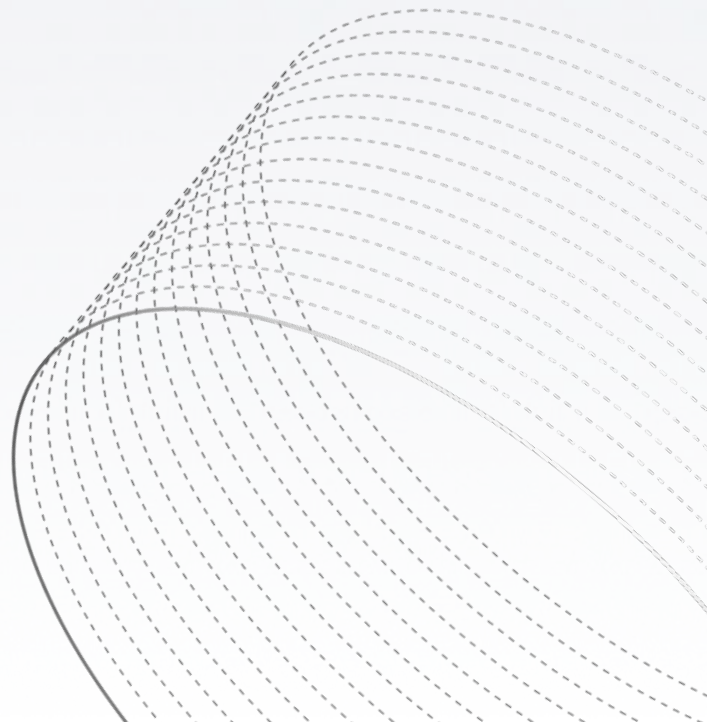
**Jordi Planadecursach**

VP Engineering @ Glassnode



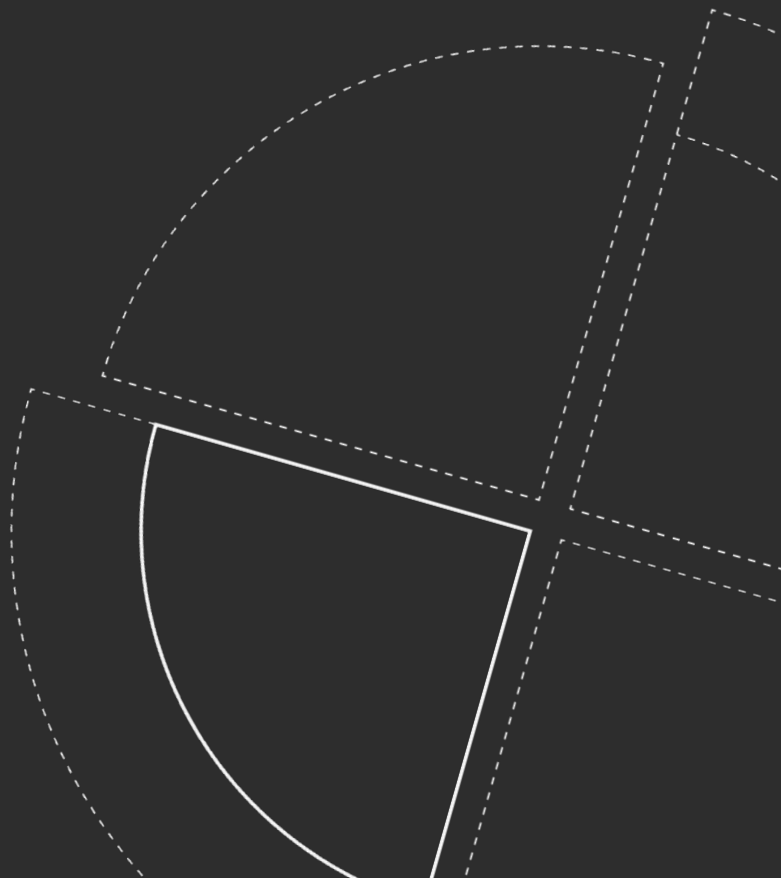
Adevinta

Sunweb















# State of Blockchain







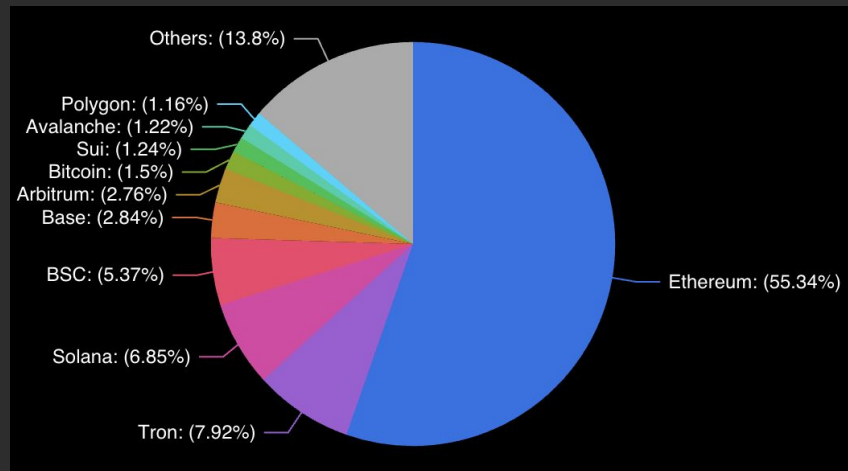
## State of Blockchain

#	Name	Price	1h %	24h %	7d %	Market Cap ⓘ	Volume(24h) ⓘ	Circulating Supply ⓘ
★ 1	 Bitcoin BTC	\$68,314.95	▼0.05%	▲0.49%	▲9.06%	\$1,350,566,701,627	\$30,806,009,949 450,829 BTC	19,769,709 BTC
★ 2	 Ethereum ETH	\$2,639.41	▼0.10%	▲0.48%	▲8.04%	\$317,753,959,351	\$14,746,432,465 5,580,743 ETH	120,388,141 ETH
★ 3	 Tether USDT	\$0.9999	▼0.00%	▲0.00%	▲0.02%	\$120,096,026,563	\$52,321,236,544 52,330,179,480 USDT	120,103,979,668 USDT
★ 4	 BNB BNB	\$597.38	▼0.05%	▲0.06%	▲3.56%	\$87,175,929,584	\$1,581,791,395 2,647,625 BNB	145,931,592 BNB
★ 5	 Solana SOL	\$154.47	▼0.18%	▲0.69%	▲6.11%	\$72,590,348,822	\$1,566,556,078 10,135,584 SOL	469,939,335 SOL
☆ 6	 USDC USDC	\$0.9998	▼0.01%	▼0.05%	▼0.02%	\$34,995,137,099	\$5,229,390,574 5,230,246,879 USDC	35,000,867,500 USDC
☆ 7	 XRP XRP	\$0.5471	▲0.14%	▼0.32%	▲1.67%	\$31,026,566,756	\$909,810,596 1,662,836,784 XRP	56,706,436,160 XRP
★ 8	 Dogecoin DOGE	\$0.1449	▲0.35%	▲7.82%	▲30.99%	\$21,214,580,088	\$2,149,899,273 14,839,148,435 DOGE	146,428,396,384 DOGE
☆ 9	 TRON TRX	\$0.1581	▼0.03%	▼0.63%	▼2.31%	\$13,682,465,574	\$232,085,669 1,467,594,205 TRX	86,521,099,153 TRX
☆ 10	 Toncoin TON	\$5.27	▲0.07%	▲1.11%	▲0.77%	\$13,389,895,256	\$143,805,188 27,271,993 TON	2,539,332,099 TON



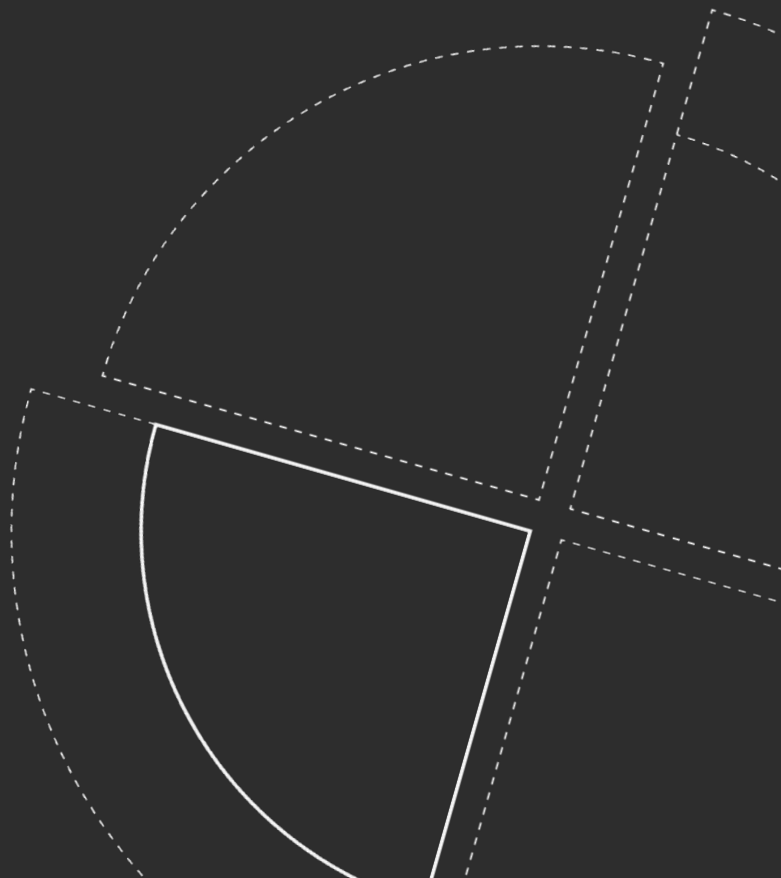
## State of Blockchain

Name	Protocols ↕	Addresses 🗳 ↕	1d Change ↕	7d Change ↕	1m Change ↕	TVL ↕
1  Ethereum	1192	430,121	+0.54%	+6.48%	+10.90%	\$48.647b
2  Tron	34	2.14m	-6.05%	-4.52%	-14.38%	\$6.961b
3  Solana	165	5.91m	+1.66%	+8.91%	+27.71%	\$6.021b
4  BSC	803	893,361	+1.01%	+4.64%	+9.32%	\$4.717b
5  Base	368	1.48m	+1.90%	+8.64%	+56.15%	\$2.494b
6  Arbitrum	714	447,098	+1.24%	+1.41%	-1.34%	\$2.429b
7  Bitcoin	37	765,780	+1.33%	+4.64%	+42.37%	\$1.32b
8  Sui	40		+0.77%	+7.07%	+48.79%	\$1.095b
9  Avalanche	409	29,894	+0.33%	+4.86%	+21.26%	\$1.07b
10  Polygon	579	453,440	+0.24%	+8.40%	+16.02%	\$1.017b
11  Scroll	104	123,239	-0.96%	+9.54%	+37.22%	\$978.22m
12  Hyperliquid	3		+0.19%	+6.53%	+7.81%	\$747.73m
13  Aptos	48		-0.30%	+18.30%	+68.44%	\$719.45m





# Què és Glassnode?





## Què és Glassnode?

Oferim intel·ligència sobre les blockchains en format de **Dades i Anàlisis** mitjançant una **Web App**. Distribuïm les dades a través dels nostres canals de difusió: API, Static Files, Data Shares.

**Quant traders**

**Risk Management**

**Discretionary Trading**

**Research & Content Creation**

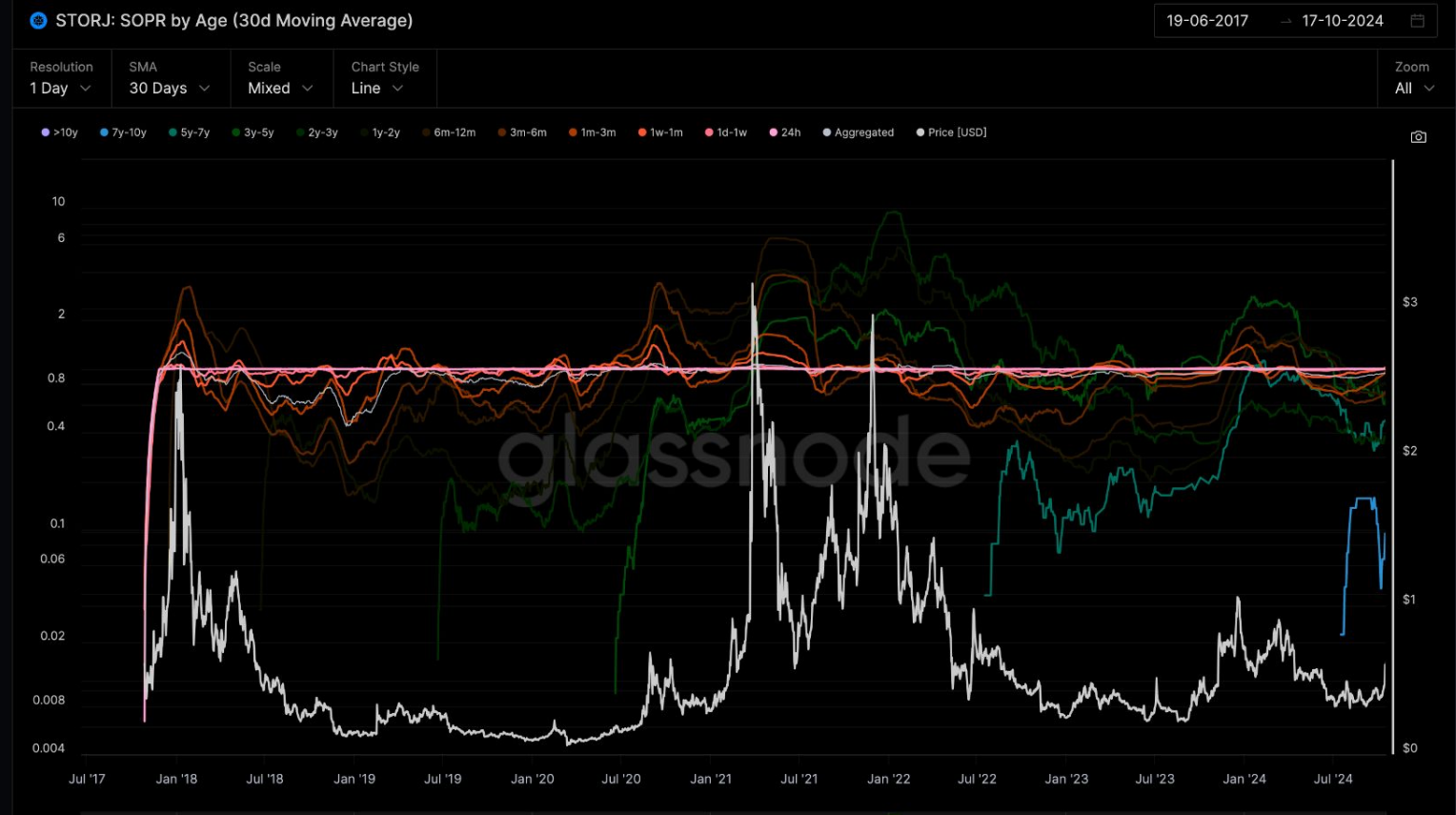


Filter Metrics

- All Categories
- Realized Profit by Age
- Realized Profit by LTH/STH
- Realized Profit by Wallet Size
- SOPR
- SOPR by Age
- SOPR by LTH/STH
- SOPR by Wallet Size
- Spent Volume
- Spent Volume in Loss by Age
- Spent Volume in Loss by LTH/STH
- Spent Volume in Loss by Wallet Size
- Spent Volume in Profit by Age
- Spent Volume in Profit by LTH/STH
- Spent Volume in Profit by Wallet Size
- Spent Volume by Age

Favorite

Compare





Què és Glassnode?

## Algunes dades

**X**

treballadors

**X**

engineers

**X**

data scientists

**X**

total registered users

**X**

B2B

**20**

exchanges

**650**

tokens

**800**

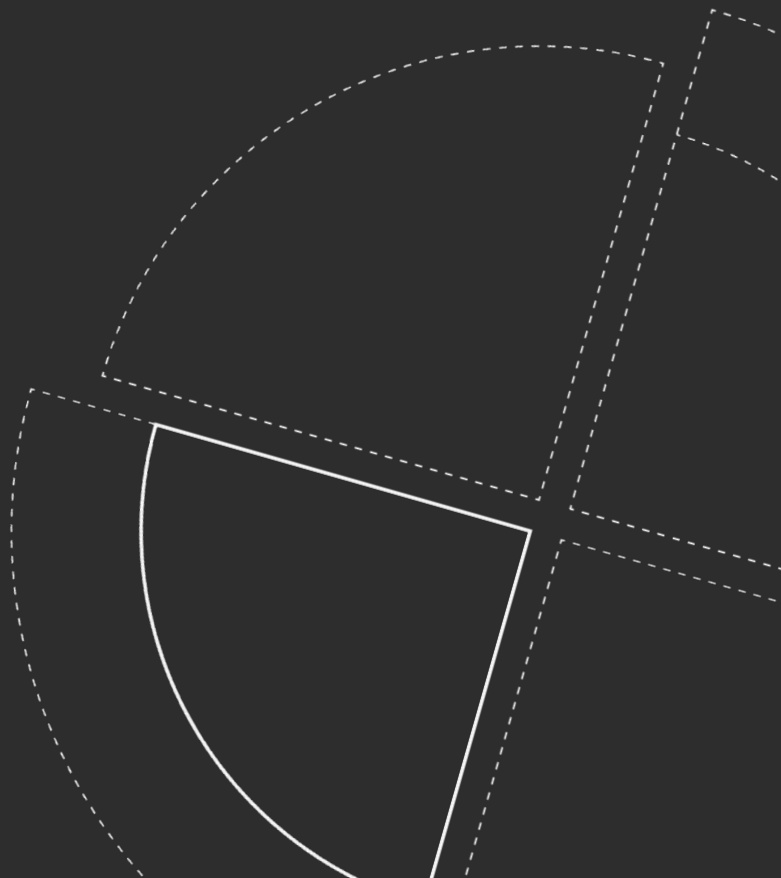
mètriques

**7**

chains



# Blockchain Ingestion



Blockchain	Total	Comprimit	Freqüència blocs	Blocs per dia	Generació diària
Bitcoin	7.6Tb	1.9Tb	10m	144	200Mb
Ethereum	39Tb	6.1Tb	12s	7200	700Mb
Solana					

Blockchain	Total	Comprimit	Freqüència blocs	Blocs per dia	Generació diària
Bitcoin	7.6Tb	1.9Tb	10m	144	200Mb
Ethereum	39Tb	6.1Tb	12s	7200	700Mb
Solana			400ms	216K	

Blockchain	Total	Comprimit	Freqüència blocs	Blocs per dia	Generació diària
Bitcoin	7.6Tb	1.9Tb	10m	144	200Mb
Ethereum	39Tb	6.1Tb	12s	7200	700Mb
Solana			400ms	216K	900Gb

Blockchain	Total	Comprimit	Freqüència blocs	Blocs per dia	Generació diària
Bitcoin	7.6Tb	1.9Tb	10m	144	200Mb
Ethereum	39Tb	6.1Tb	12s	7200	700Mb
Solana	1.000Tb	255Tb	400ms	216K	900Gb



**1Pb**

**1.000Tb**

**1.000.000 Gb**

**1.000.000.000 Mb**

**1.000.000.000.000 Kb**

**1.000.000.000.000.000 Bytes**





**1Pb**

**1.000Tb**

**1.000.000 Gb**

**1.000.000.000 Mb**

**1.000.000.000.000 Kb**

**1.000.000.000.000.000 Bytes**

**750M disquets**



**1Pb**

**1.000Tb**

**1.000.000 Gb**

**1.000.000.000 Mb**

**1.000.000.000.000 Kb**

**1.000.000.000.000.000 Bytes**

**750M disquets**

**1.5M CD**



**1Pb**

**1.000Tb**

**1.000.000 Gb**

**1.000.000.000 Mb**

**1.000.000.000.000 Kb**

**1.000.000.000.000.000 Bytes**

**750M disquets**

**1.5M CD**

**200K DVD**



**1Pb**

**1.000Tb**

**1.000.000 Gb**

**1.000.000.000 Mb**

**1.000.000.000.000 Kb**

**1.000.000.000.000.000 Bytes**

**750M disquets**

**1.5M CD**

**200K DVD**

**41K Blue rays**



**1Pb**

**1.000Tb**

**1.000.000 Gb**

**1.000.000.000 Mb**

**1.000.000.000.000 Kb**

**1.000.000.000.000.000 Bytes**

**750M disquets**

**1.5M CD**

**200K DVD**

**41K Blue rays**

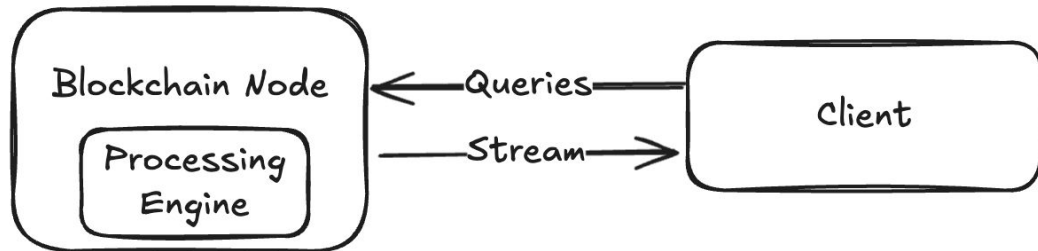
**14 m3 espai**

## Web3

&lt; 2022

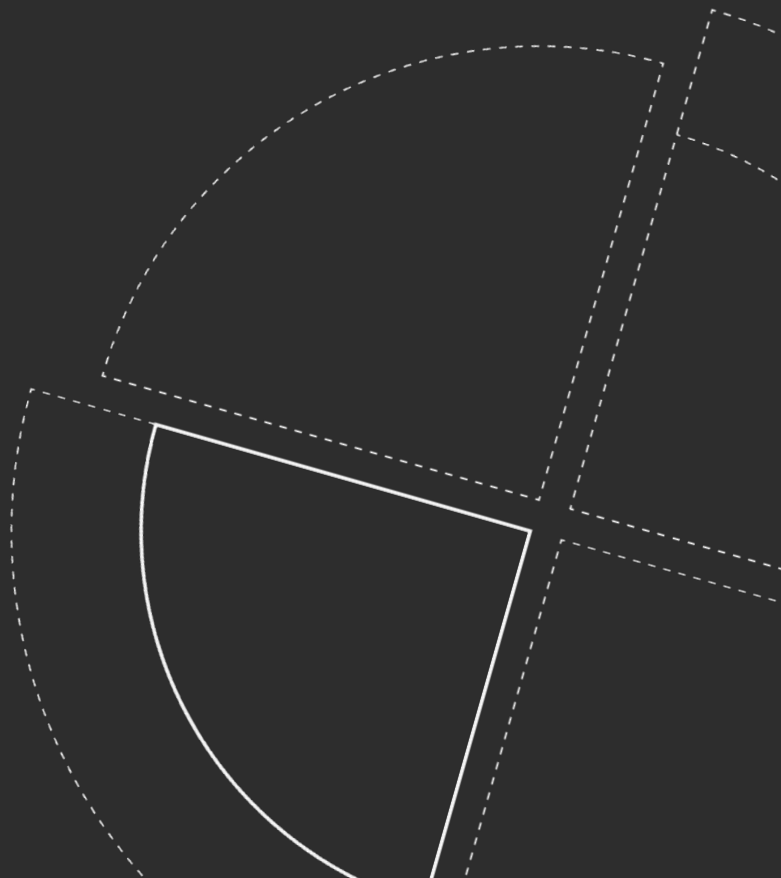


&gt;= 2022-23





# Microservices mesh



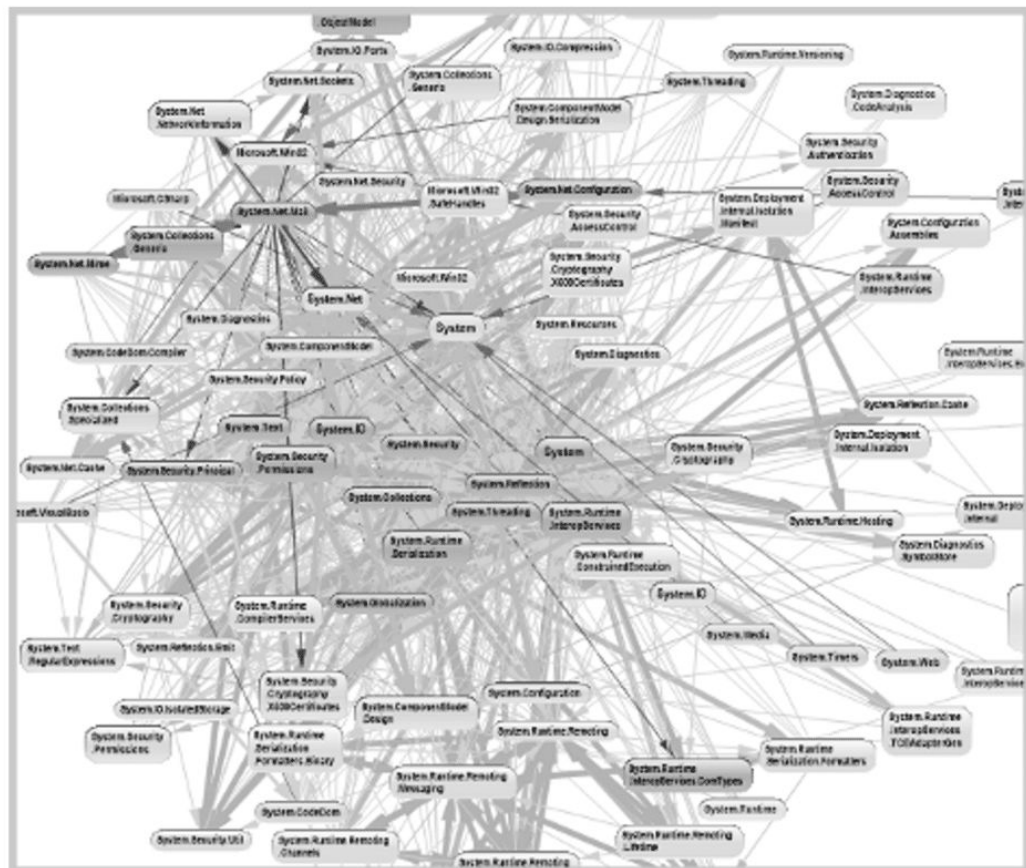
# Big ball of mud

**Brian Foote** and **Joseph Yoder**

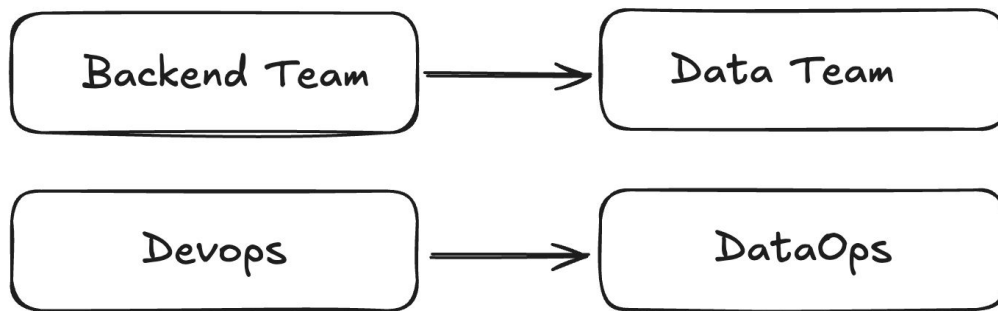
June 26, 1999

<http://laputan.org/mud/mud.html#BigBallOfMud>



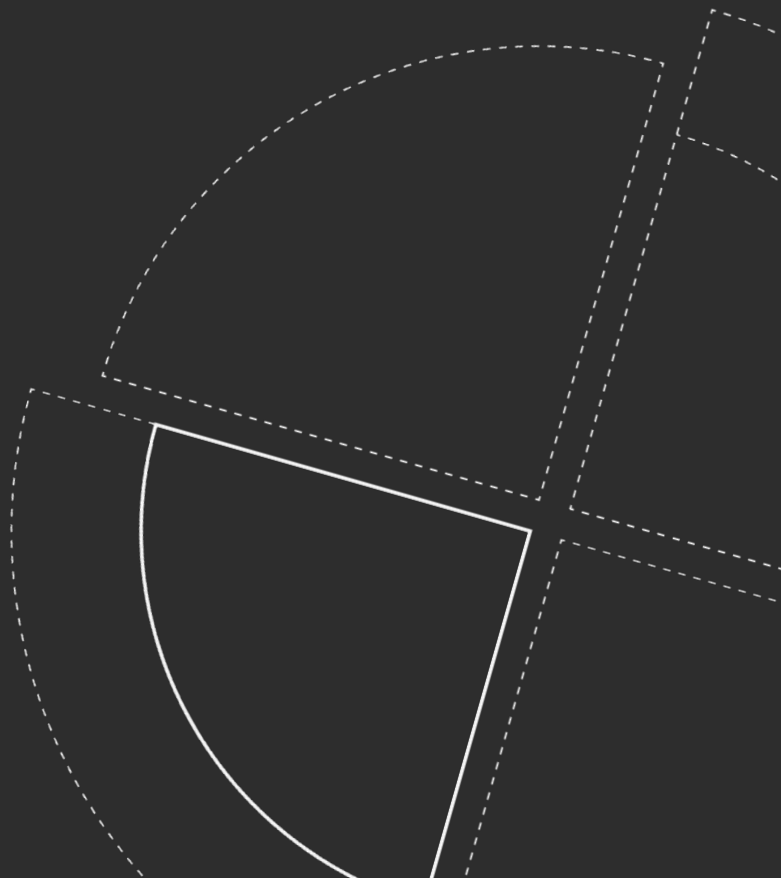


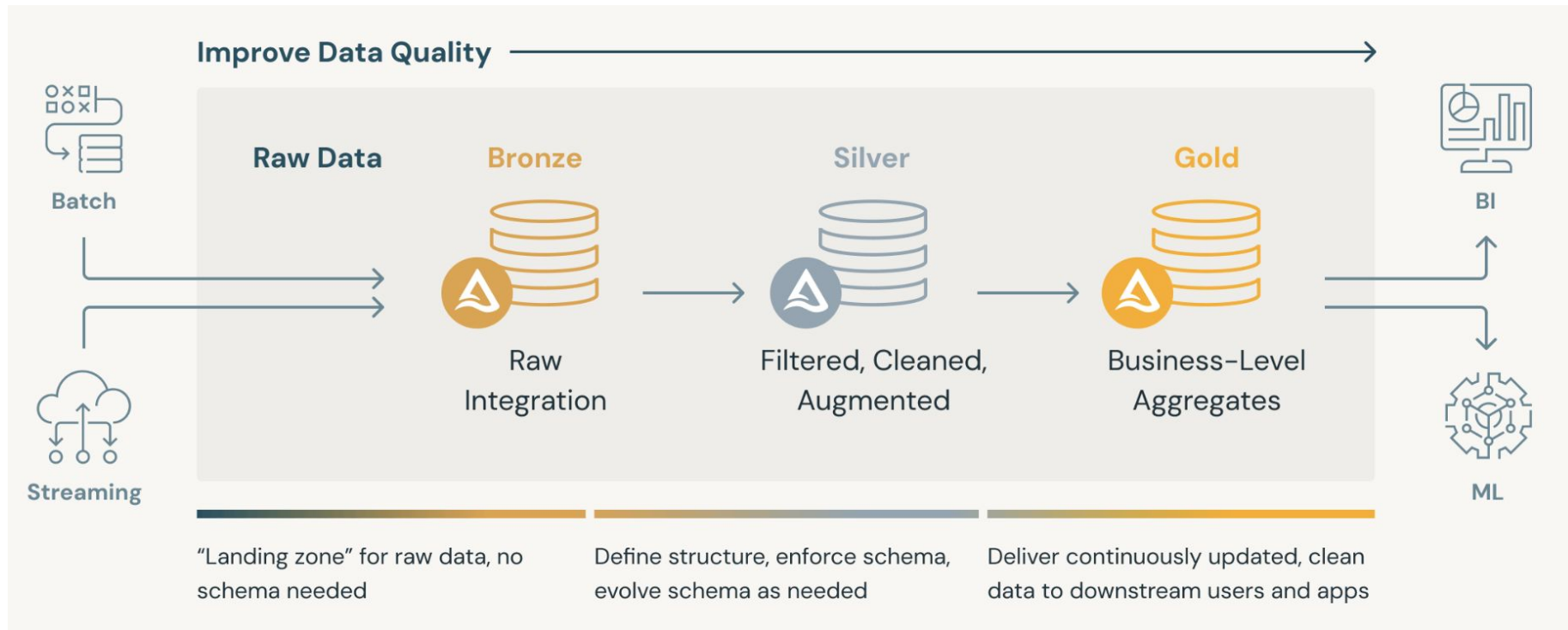


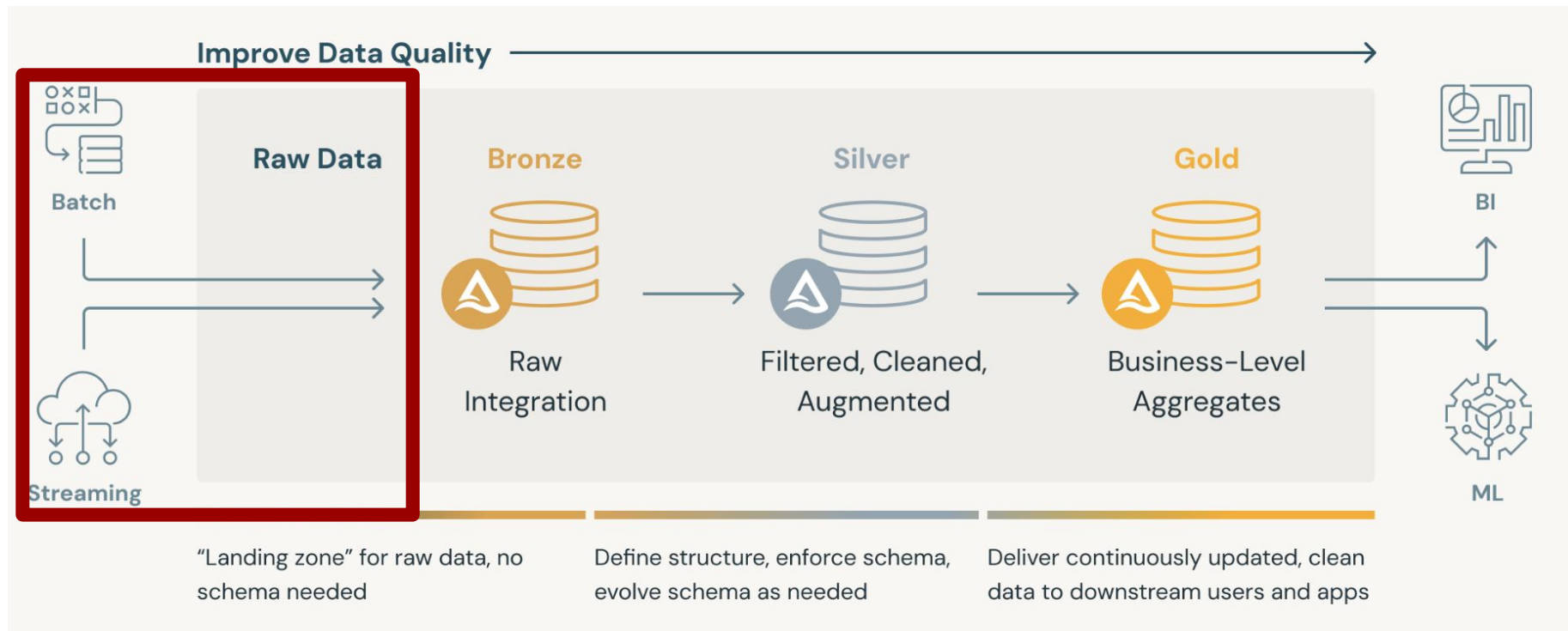




# Medaillon architecture







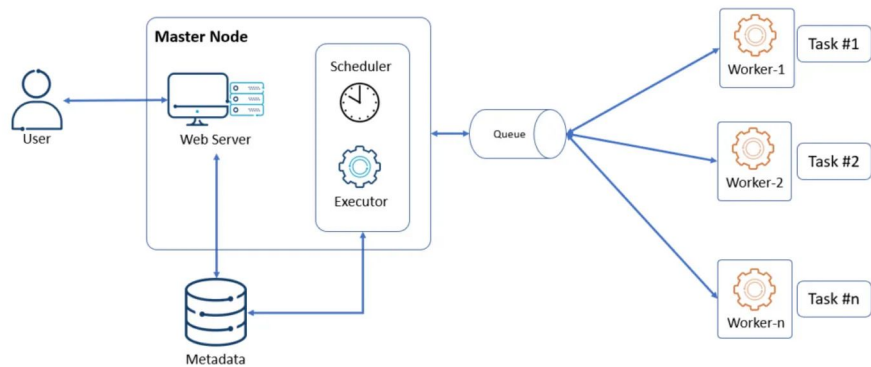


# Ingestion


- Serveis
  - WS Apis
  - Consumers
- Crons
  - Scrapers
  - RPC clients




# Orchestrator






 Airflow

DAGs Cluster Activity Datasets Security Browse Admin Docs

12:05 EDT (-04:00) 

 DAG: example\_bash\_operator


Schedule: 0 0 \* \* \* ⓘ Next Run ID: 2024-04-01, 20:00:00 ▶ ⌵

04/02/2024, 12:03:29 PM ⌵ All Run Types ▼ All Run States ▼ Clear Filters

Auto-refresh ☐ 25 ▼

Press **shift** + **/** for Shortcuts

deferred failed queued removed restarting running scheduled shutdown skipped success up\_for\_reschedule up\_for\_retry upstream\_failed no\_status



Task	runme_0	runme_1	runme_2	also_run_this	this_will_skip	run_after_loop	run_this_last
Duration	00:00:12	00:00:06	00:00:06	00:00:04	00:00:02	00:00:01	00:00:01

**example\_bash\_operator**

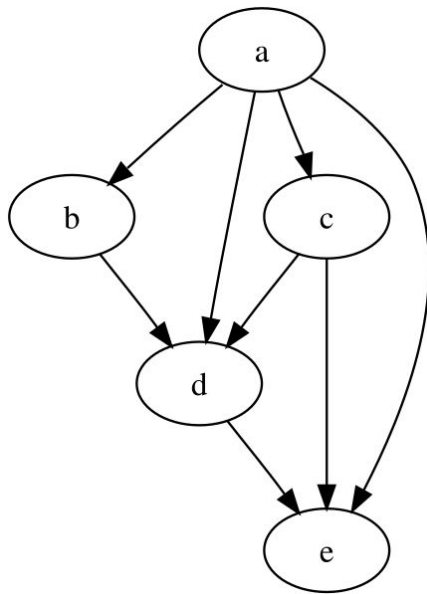
Details Graph Gantt **<> Code** Audit Log Run Duration Calendar

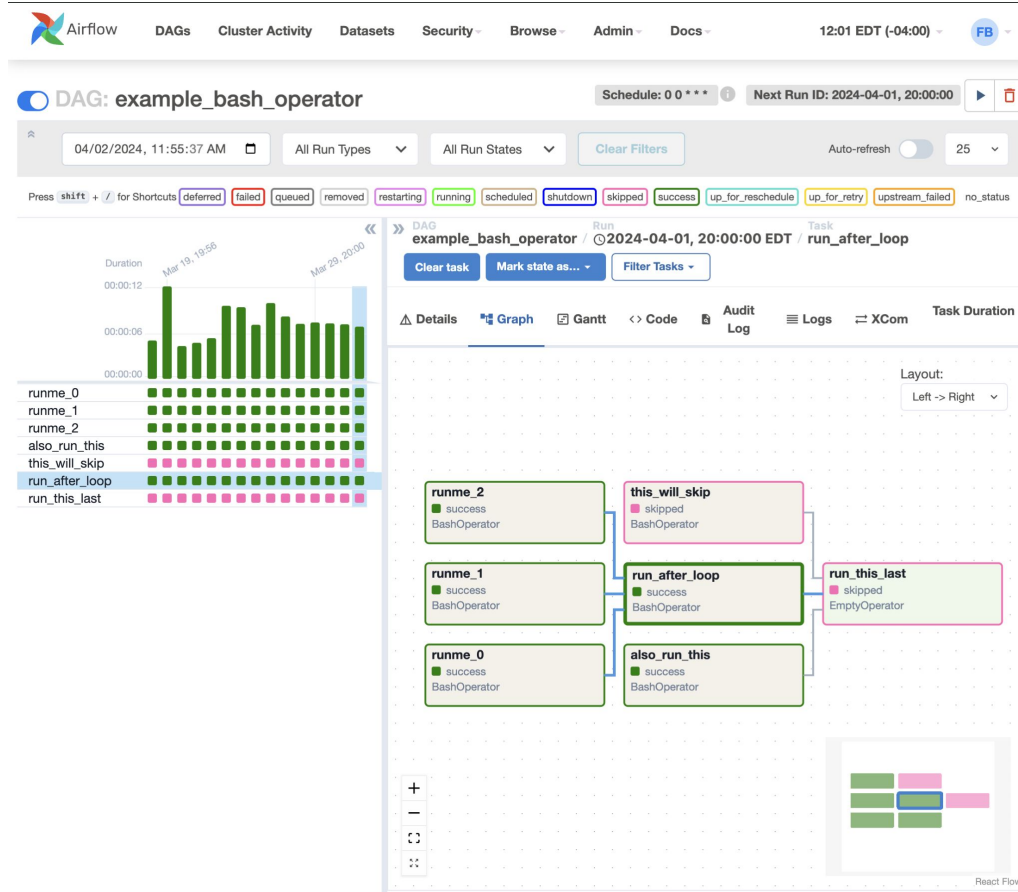
Parsed at: 2024-04-02, 12:03:04 EDT

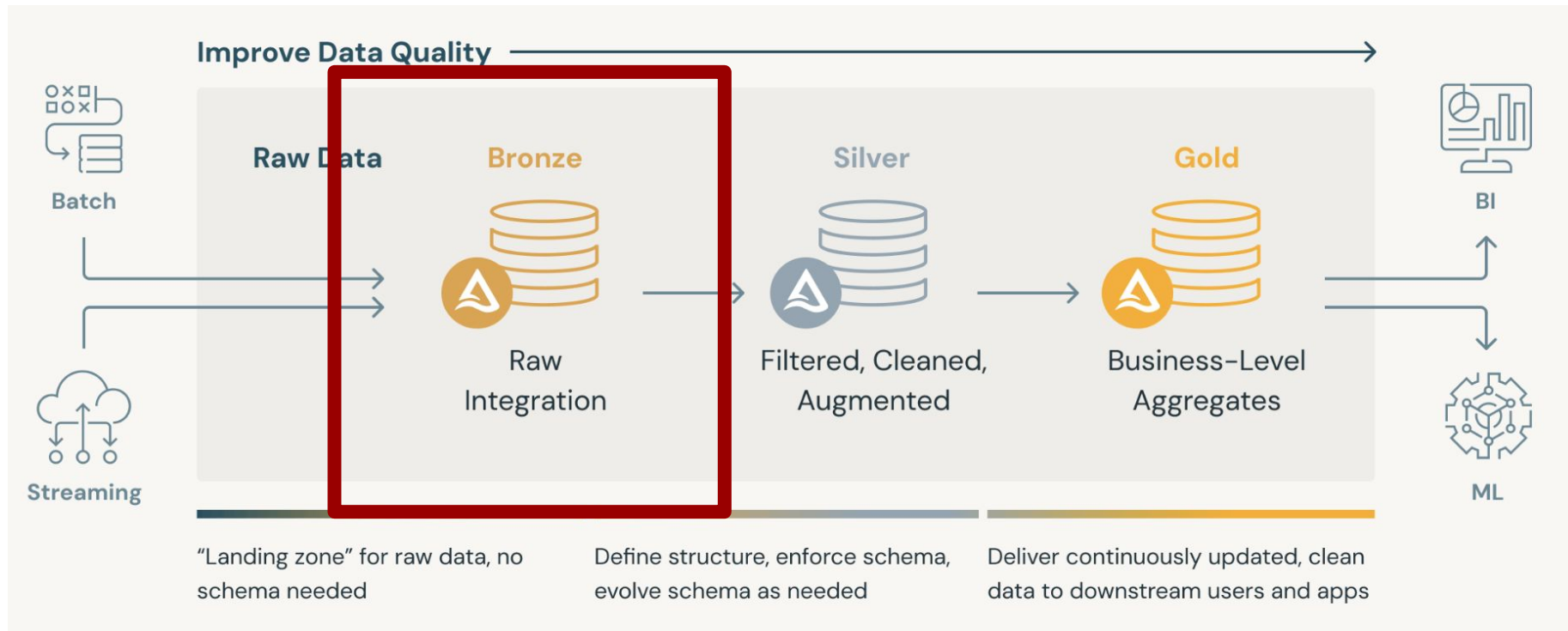
```
19
20 from __future__ import annotations
21
22 import datetime
23
24 import pendulum
25
26 from airflow.models.dag import DAG
27 from airflow.operators.bash import BashOperator
28 from airflow.operators.empty import EmptyOperator
29
30 with DAG(
31     dag_id="example_bash_operator",
32     schedule="0 0 * * *",
33     start_date=pendulum.datetime(2021, 1, 1, tz="UTC"),
34     catchup=False,
35     dagrun_timeout=datetime.timedelta(minutes=60),
36     tags=["example", "example2"],
37     params={"example_key": "example_value"},
38 ) as dag:
39     run_this_last = EmptyOperator(
40         task_id="run_this_last",
41     )
42
43     # [START howto_operator_bash]
44     run_this = BashOperator(
45         task_id="run_after_loop",
```

Toggle Wrap

## DAG: Directed acyclic graph







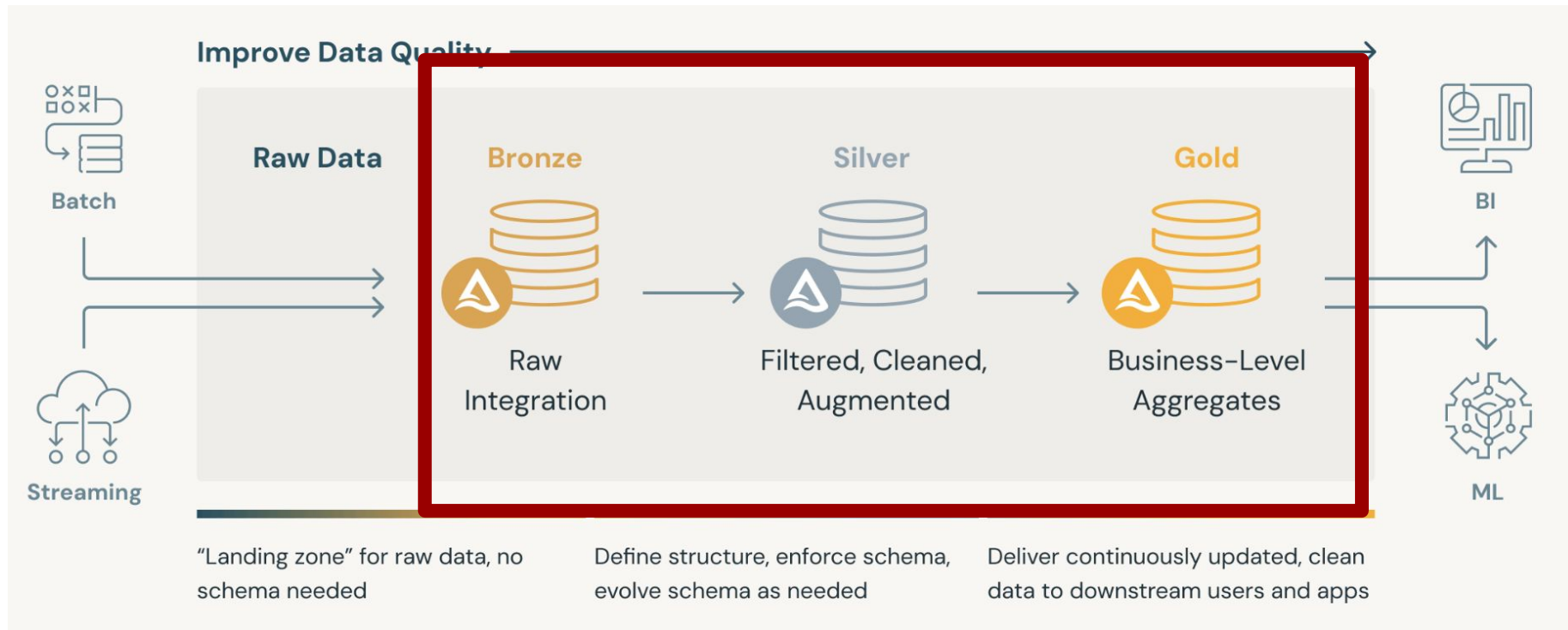


Google Cloud Storage

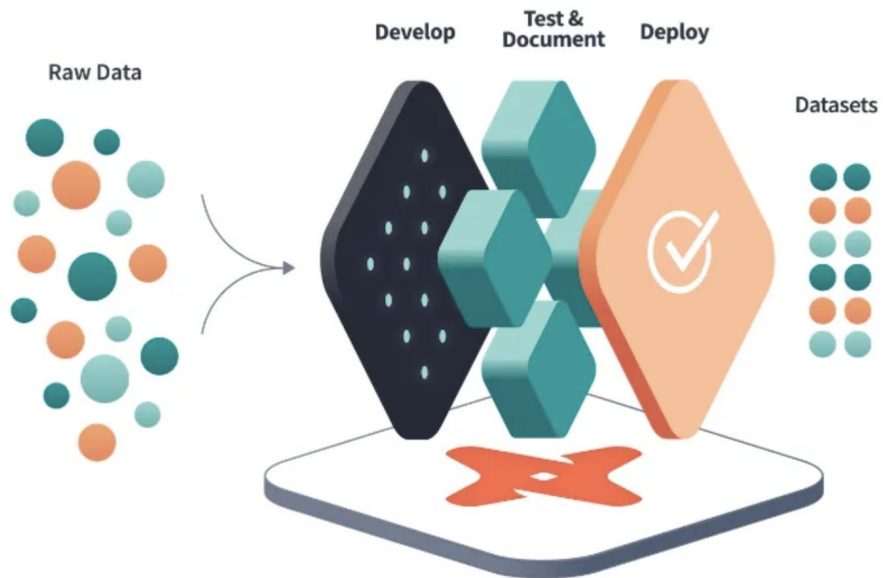


Parquet

- Format columnar (en lloc de per fila)
- Compensió i eficiència
- Lectura selectiva
- Metadades per columna
- Suporta canvis d'esquema



## Orchestrator



# What is DBT?

- Used to manage/execute data transformations in a Data Warehouse/Lake
  - We're using it with BigQuery
  - Batch processing
  - <https://www.getdbt.com/>
- ELT (Extract, Load, Transform) - DBT is the T
- DBT runs as a python process (cron job) in our GKE cluster, that sends sql jobs BigQuery and tracks results
- Acts as an orchestration layer on top of our data warehouse (plus testing, documentation)
- Benefits:
  - brings many software engineering best practices into the realm of analytics engineering
  - automates a lot of the repetitive, manual parts of data transformations
  - Auto document generation, automated testing for transformations and source freshness

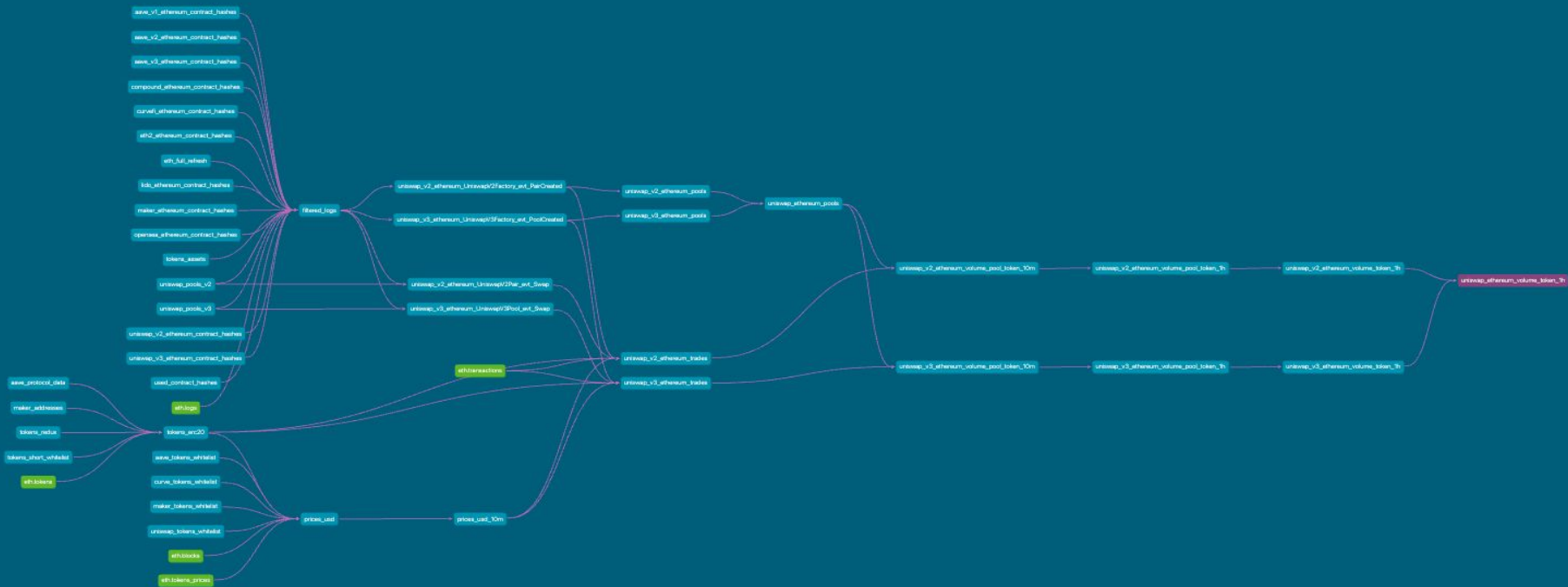


# The DBT src folder

```
> jobs
> logs
✓ src
  > macros
  ✓ models
    > subfolder
      btc_sources.yml
      metric_1d.sql
      model_1.sql
      model_2.sql
      schema.yml
    > seeds
      dbt_project.yml
      profiles.yml
      README.md
    .gitignore
    .gitlab-ci.yml
    README.md
> External Libraries
≡ Scratches and Consoles
```

```
2
3 {% set partitions_to_replace = [
4     'DATE_TRUNC(CURRENT_DATE(), WEEK)',
5     'DATE_TRUNC(DATE(TIMESTAMP_SUB(CURRENT_TIMESTAMP(), INTERVAL ' ~ buffer_
6 ] %}
7
8 {{ config(
9     partition_by={
10         "field": "week",
11         "data_type": "date",
12         "granularity": "day"
13 },
14     materialized='incremental',
15     incremental_strategy='insert_overwrite',
16     partitions=partitions_to_replace,
17     on_schema_change='fail'
18 ) }}
19
20 SELECT
21     TIMESTAMP_TRUNC(block_timestamp, DAY) AS timestamp,
22     DATE_TRUNC(DATE(block_timestamp), WEEK) AS week,
23     COUNT(1) AS count
24 FROM {{ ref('model_2') }}
25 {% if is_incremental() %}
26 WHERE block_week IN ({{ partitions_to_replace | join(',') }})
27 AND {{ complete_intervals('TIMESTAMP_TRUNC(block_timestamp, DAY)') }}
28 {% endif %}
29 GROUP BY 1, 2
30
```

**Model:** uniswap\_ethereum\_volume\_token\_1h

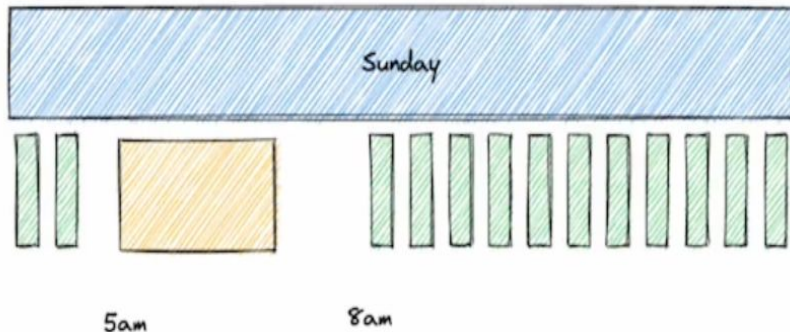


# Full Refresh Job vs. Incremental DBT Job

## Full Refresh:

Drops any changed models and recomputes from scratch.

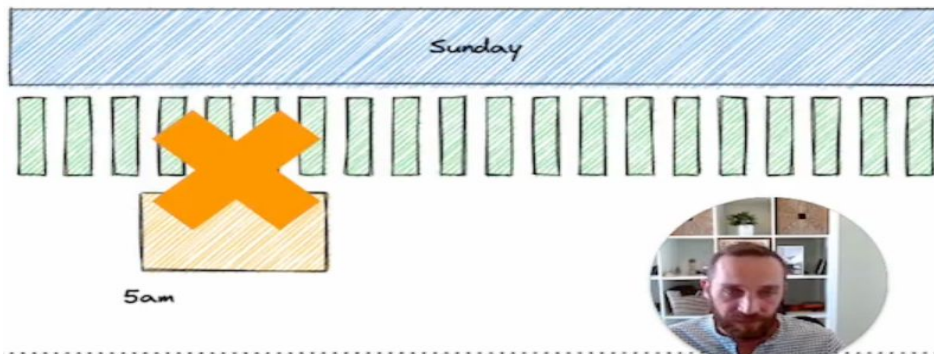
dbt run *--select state:modified --state old\_target*  
<more params>

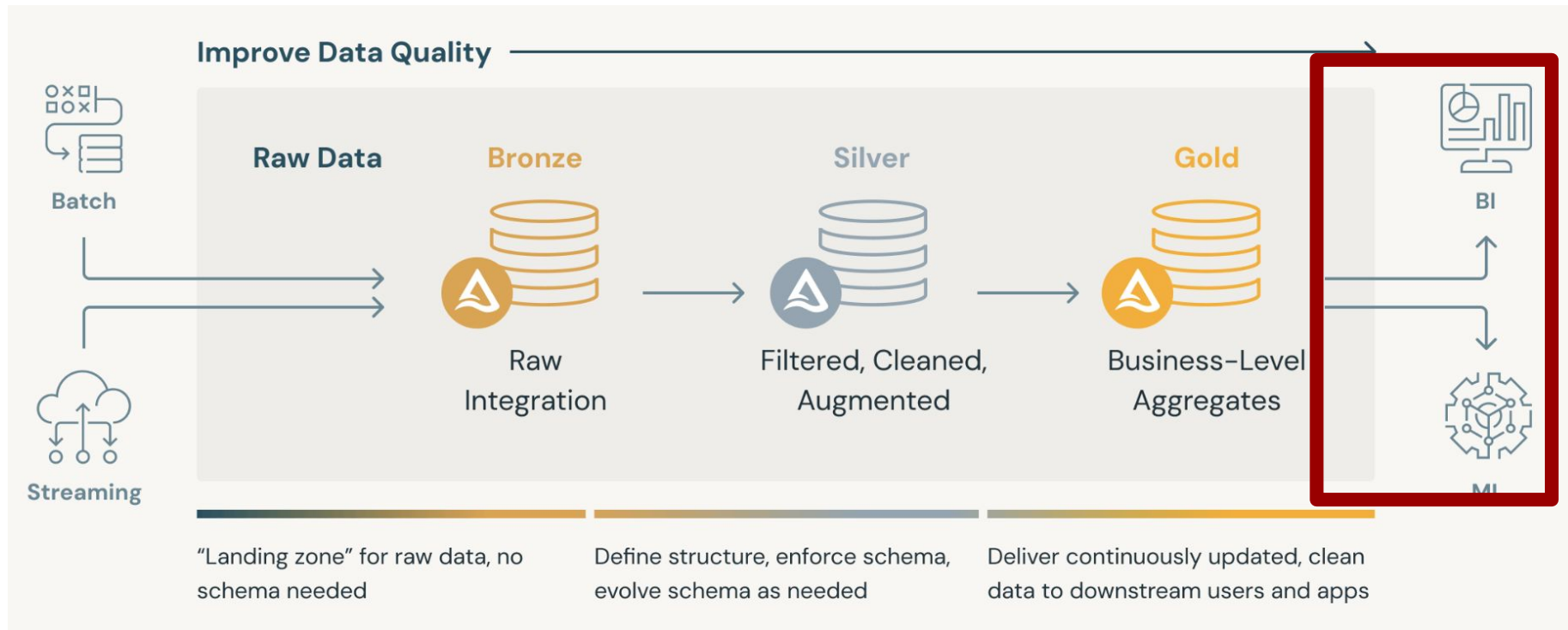


## Incremental:

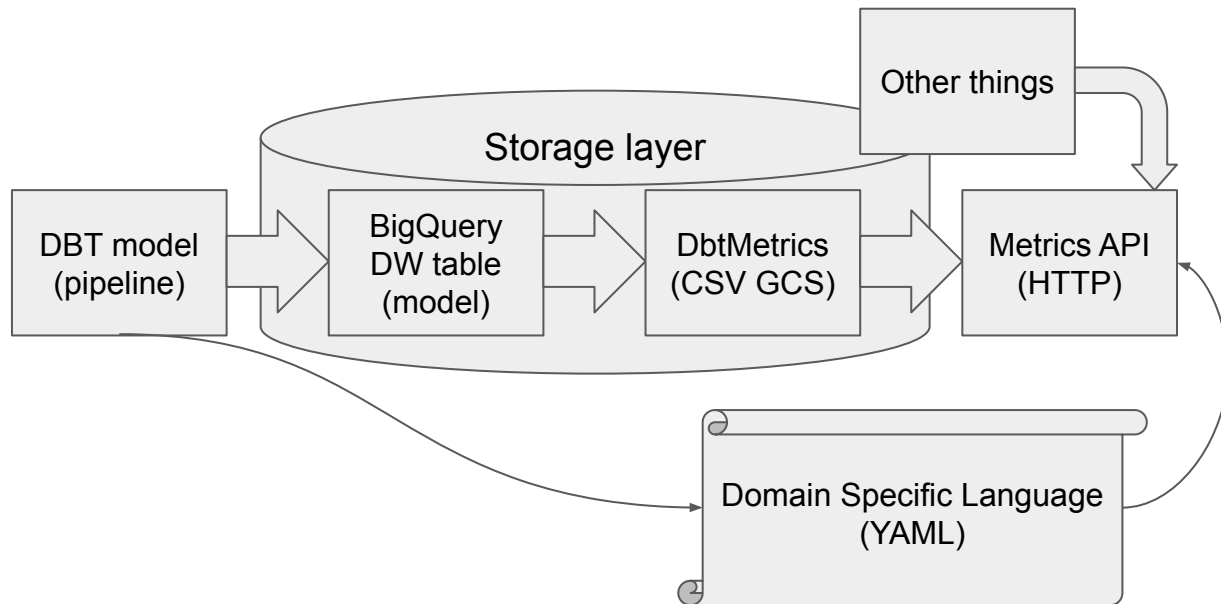
Updates the model with new data since the last run, more efficient for

incremental job  
full refresh



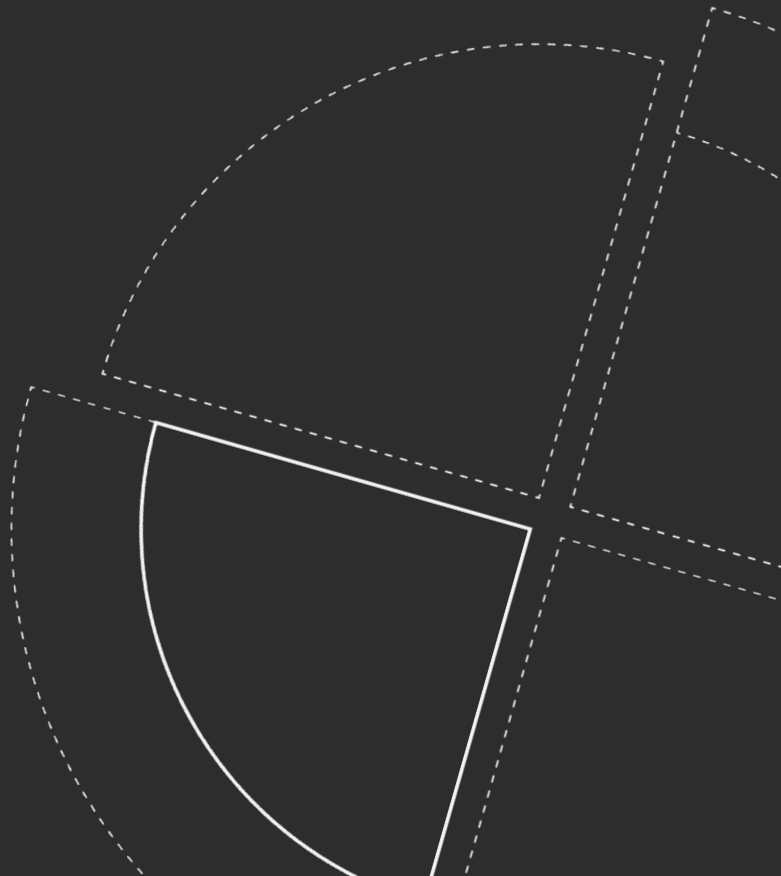


## Metric DSL





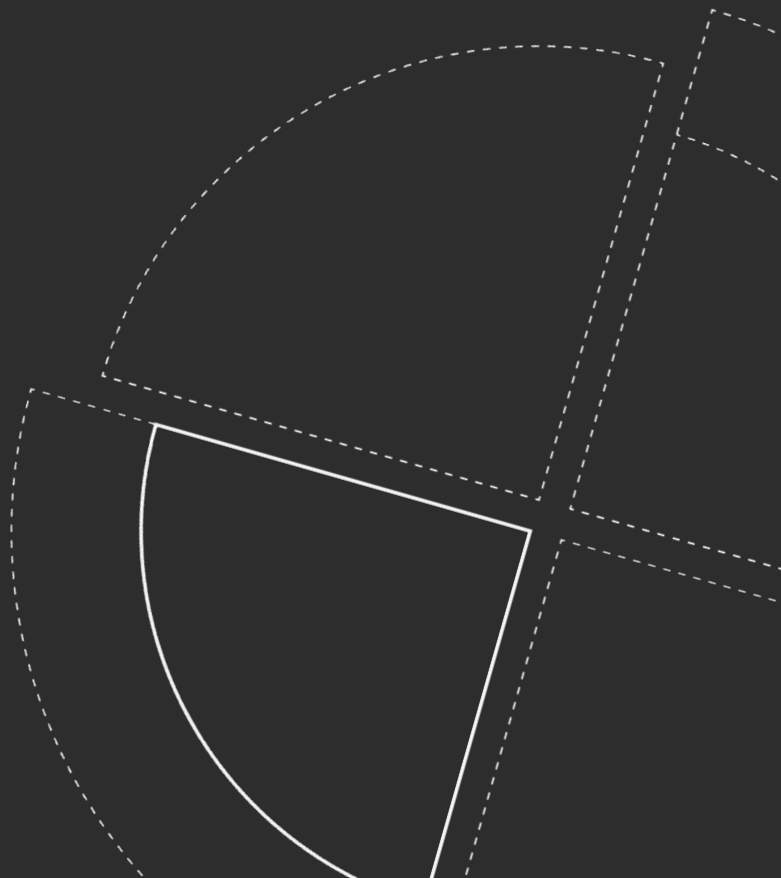
# Next steps



- Cost scalability
- Multi cloud
  - Snowflake
- Split DBT
  - 1m latency
- Data Lake -> Delta House
  - Data Catalogs



Gràcies 🎉







# Q&A

