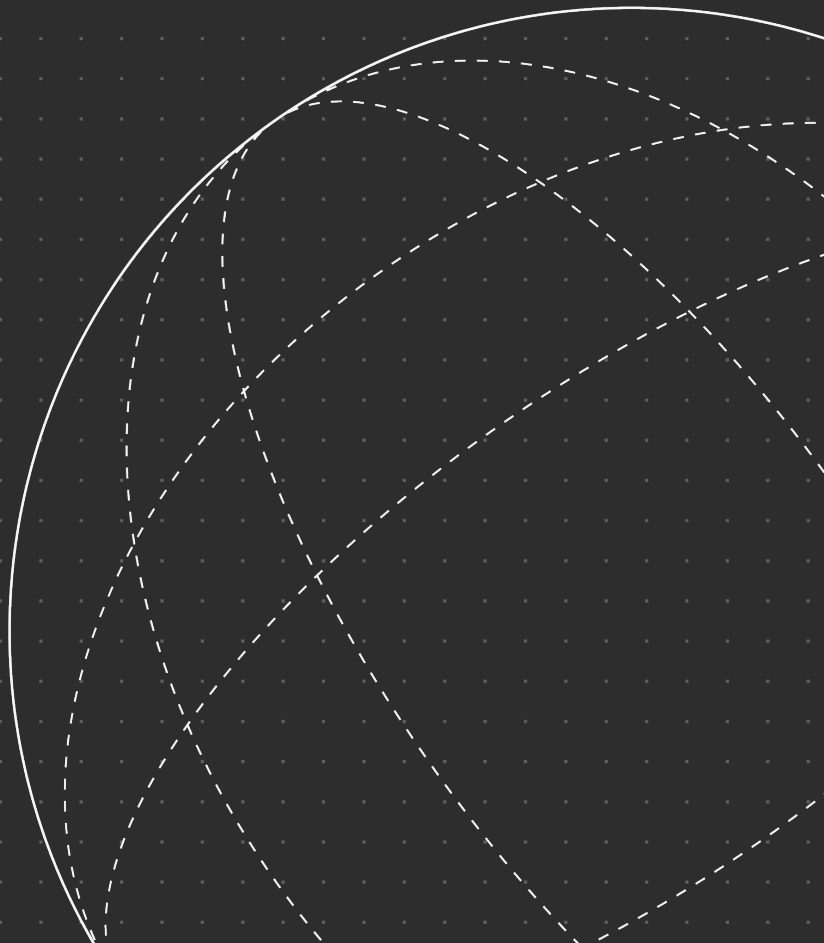
glassnode

Glassnode Data Platform Journey

Festa OS 2024 - Girona, Octubre 2024



Agenda

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

Jordi Planadecursach

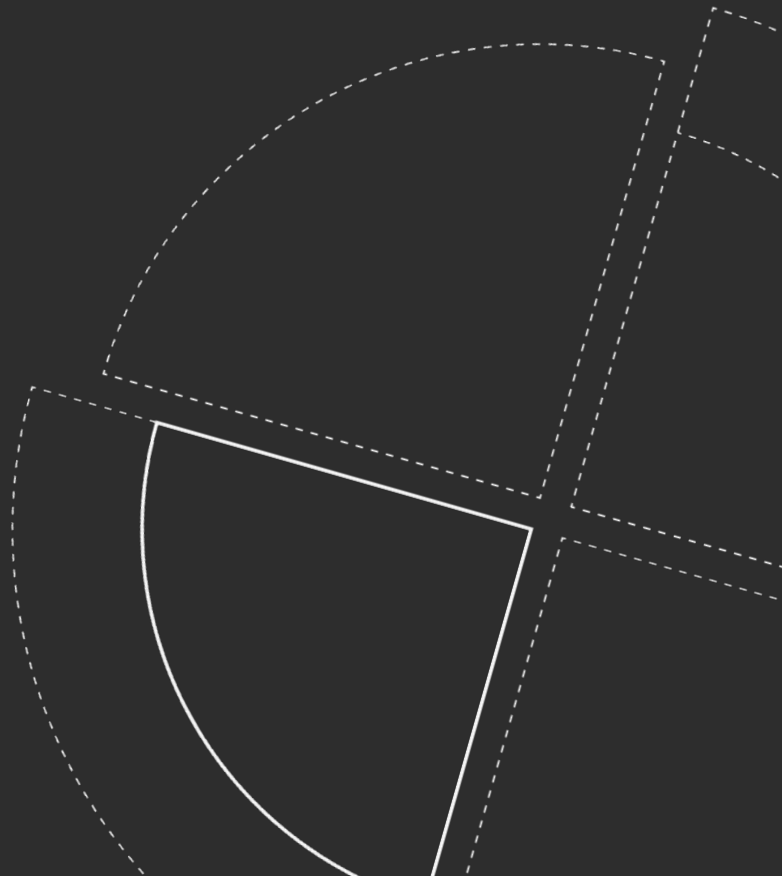
VP Engineering @ Glassnode

planadecu

Adevinta

Sunweb

State of Blockchain

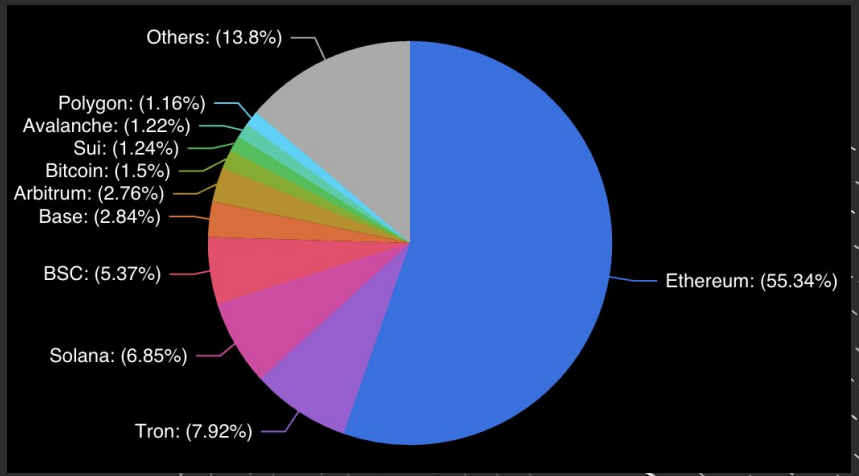


State of Blockchain

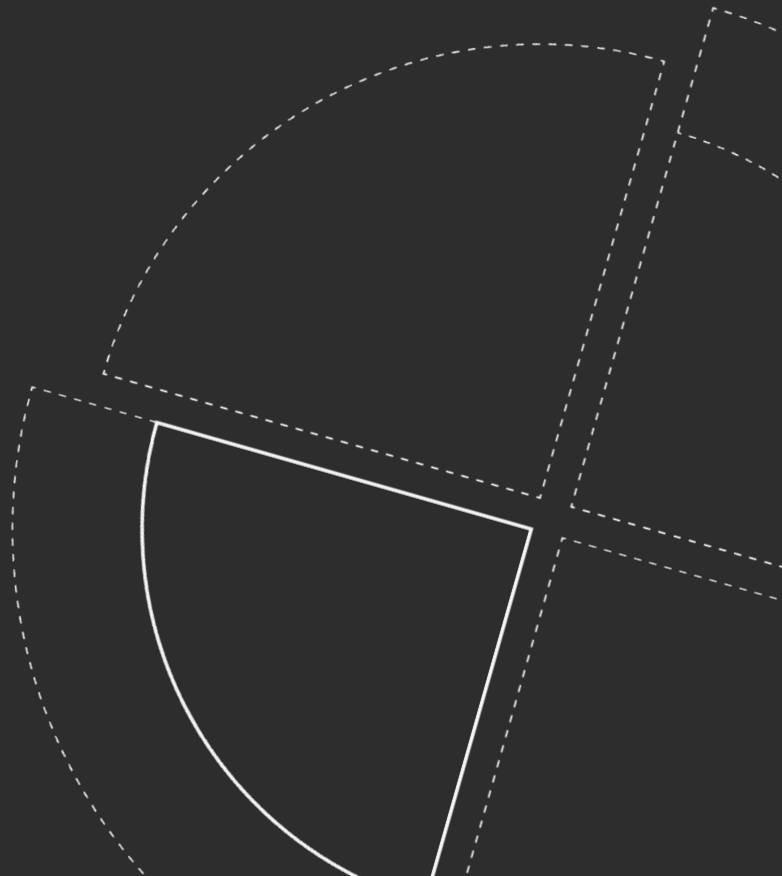
	#	Name	Price	1h %	24h %	7d %	Market Cap	Volume(24h)	Circulating Supply
★	1	Bitcoin BTC	$68,314.95	0.05%	0.49%	9.06%	$1,350,566,701,627	$30,806,009,949 450,829 BTC	19,769,709 BTC
★	2	Ethereum ETH	$2,639.41	0.10%	0.48%	8.04%	$317,753,959,351	$14,746,432,465 5,580,743 ETH	120,388,141 ETH
★	3	Tether USDT	$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

T3 Realized Profit by Age
T3 Realized Profit by LTH/STH
T3 Realized Profit by Wallet Size

SOPR

T3 SOPR by Age
T3 SOPR by LTH/STH
T3 SOPR by Wallet Size

Spent Volume

T3 Spent Volume in Loss by Age
T3 Spent Volume in Loss by LTH/STH
T3 Spent Volume in Loss by Wallet Size
T3 Spent Volume in Profit by Age
T3 Spent Volume in Profit by LTH/STH
T3 Spent Volume in Profit by Wallet Size
T3 Spent Volume by Age

Favorite | Compare

STORJ: SOPR by Age (30d Moving Average)

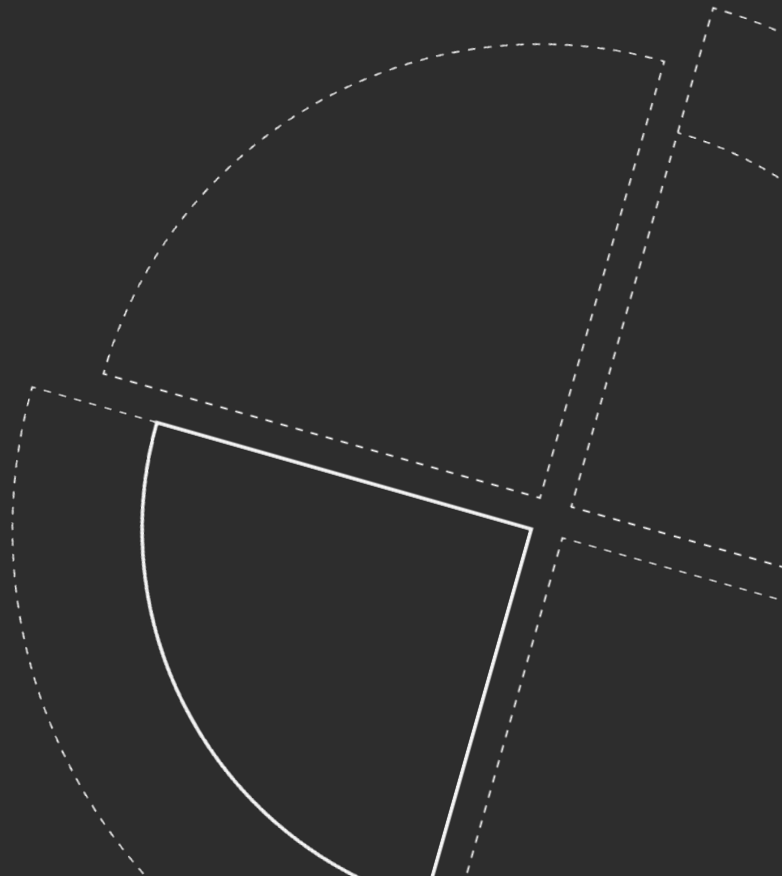
19-06-2017 → 17-10-2024

Resolution: 1 Day | SMA: 30 Days | Scale: Mixed | Chart Style: Line | Zoom: All

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	Freqüencia blocs	Blocs per dia	Generació diaria
Bitcoin	7.6Tb	1.9Tb	10m	144	200Mb
Ethereum	39Tb	6.1Tb	12s	7200	700Mb
Solana					

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

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

Blockchain	Total	Comprimit	Freqüencia blocs	Blocs per dia	Generació diaria
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

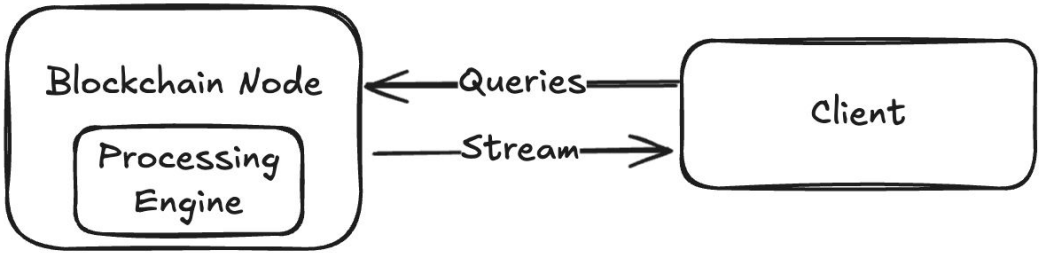
41K Blue rays

14 m3 espai

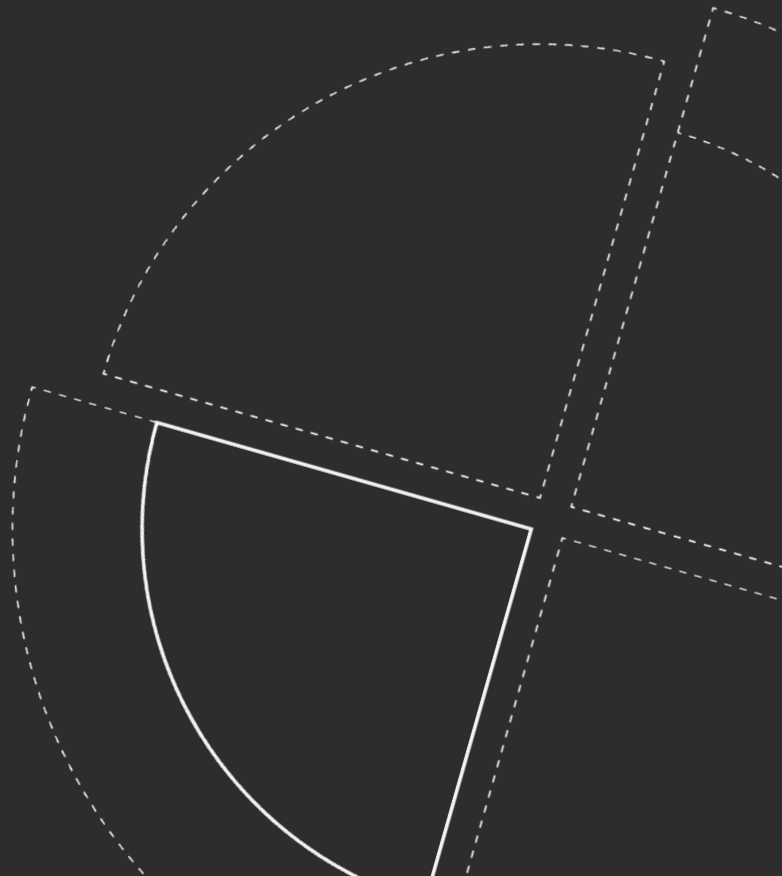
Web3

< 2022

>= 2022-23



Microservices mesh

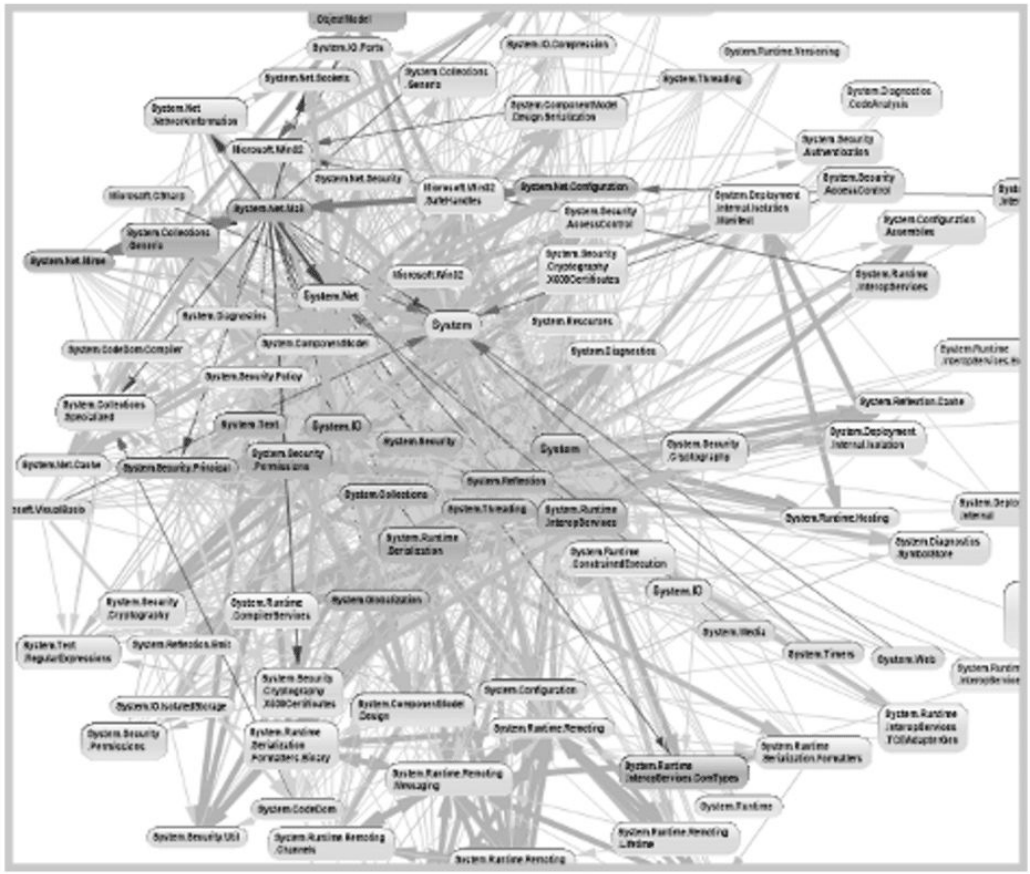


Big ball of mud

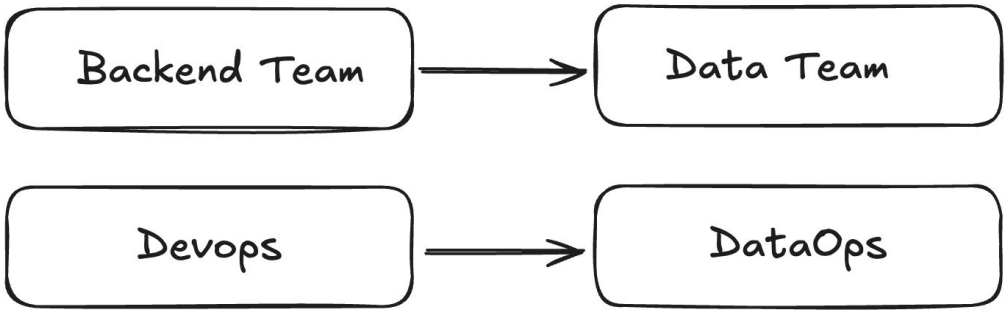
Brian Foote and **Joseph Yoder**
June 26, 1999

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

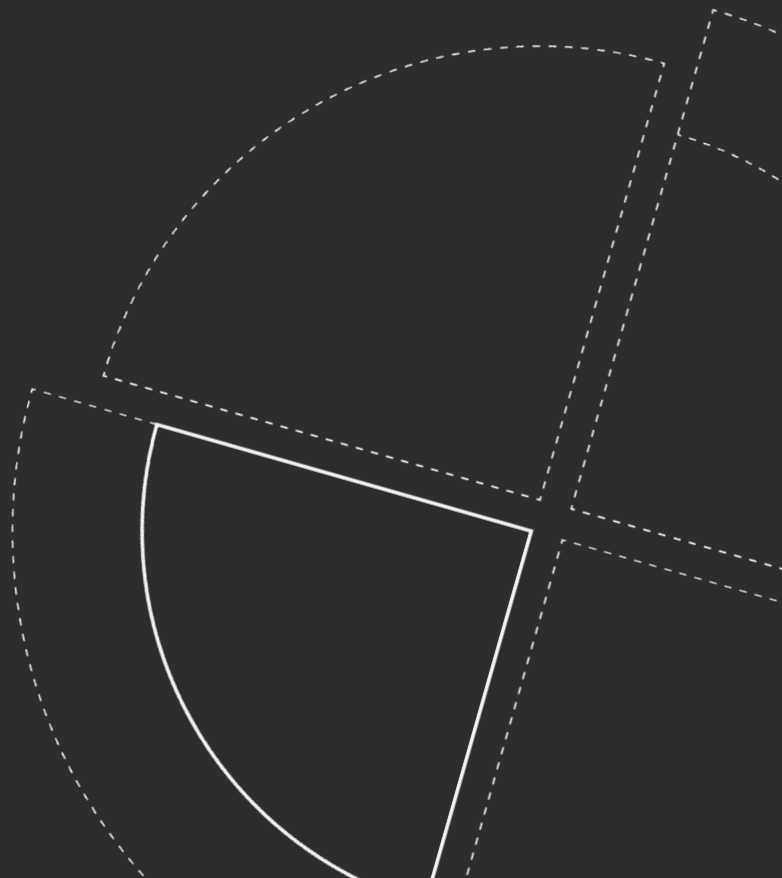
Microservices mesh



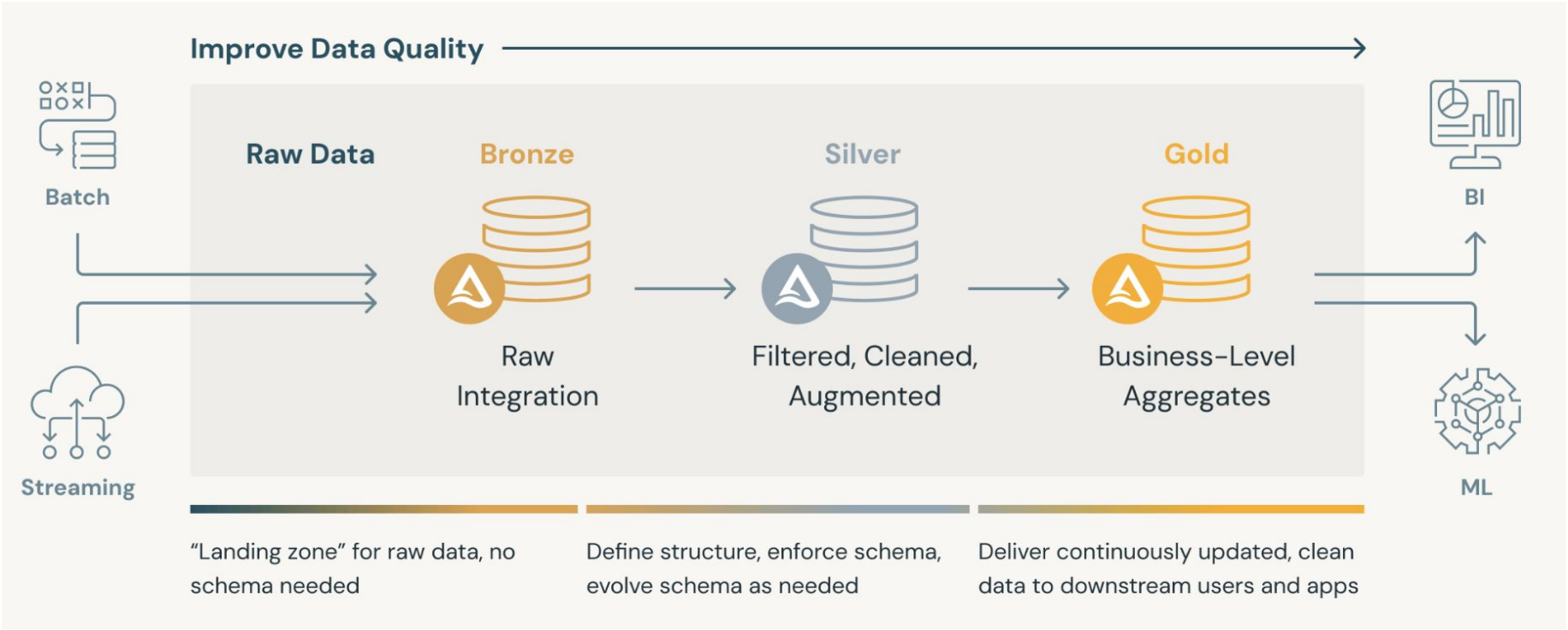
Microservices mesh


Backend Team
Data Team
Devops
DataOps

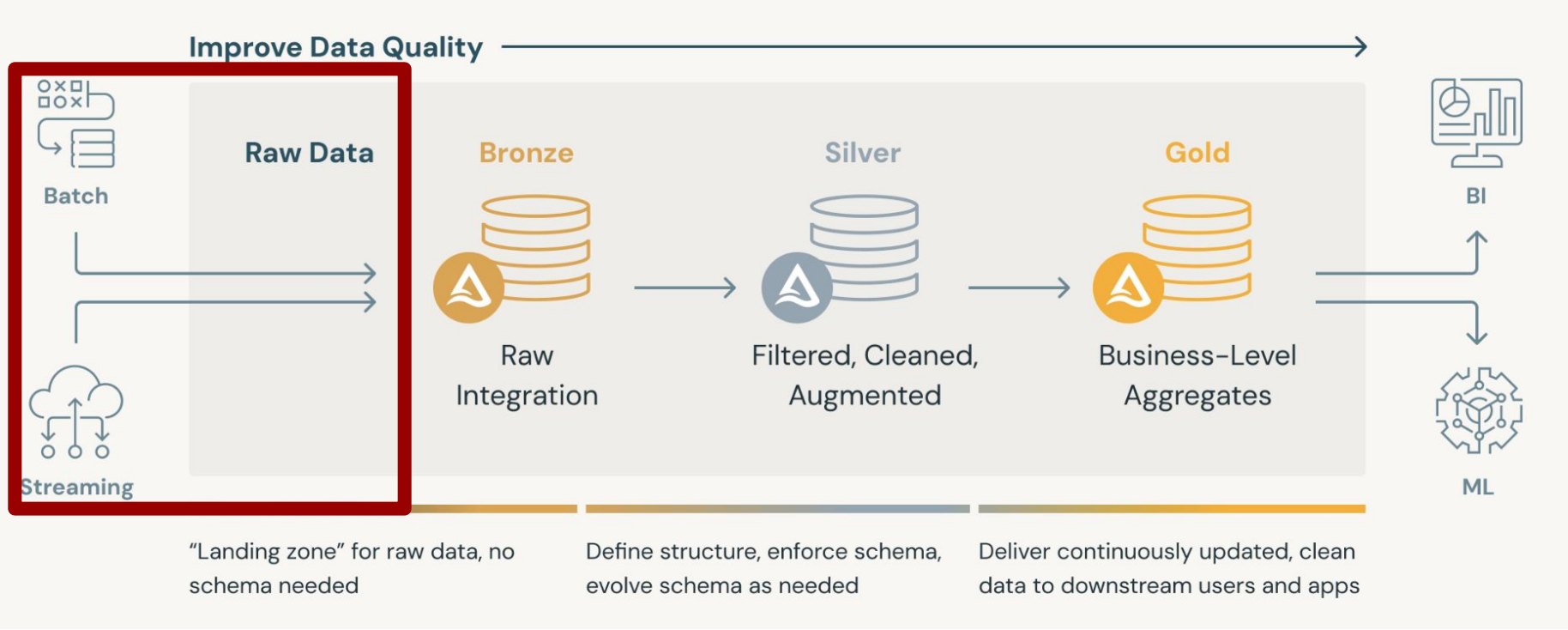
Medaillon architecture



Medaillon architecture



Medaillon architecture: Ingestion



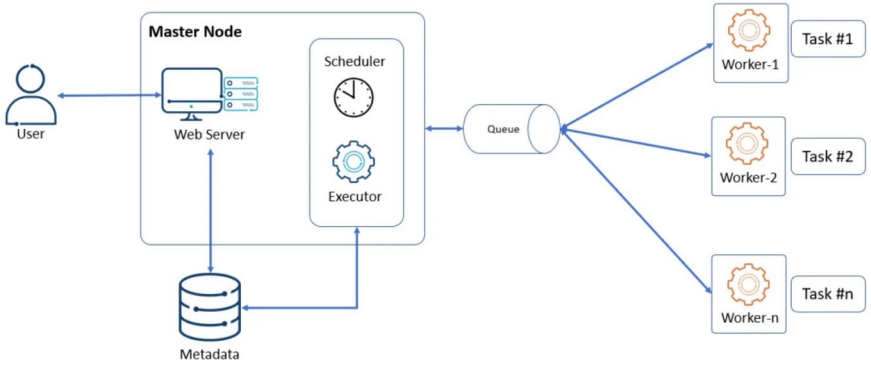
Ingestion

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

Airbyte

Fivetran

Orchestrator

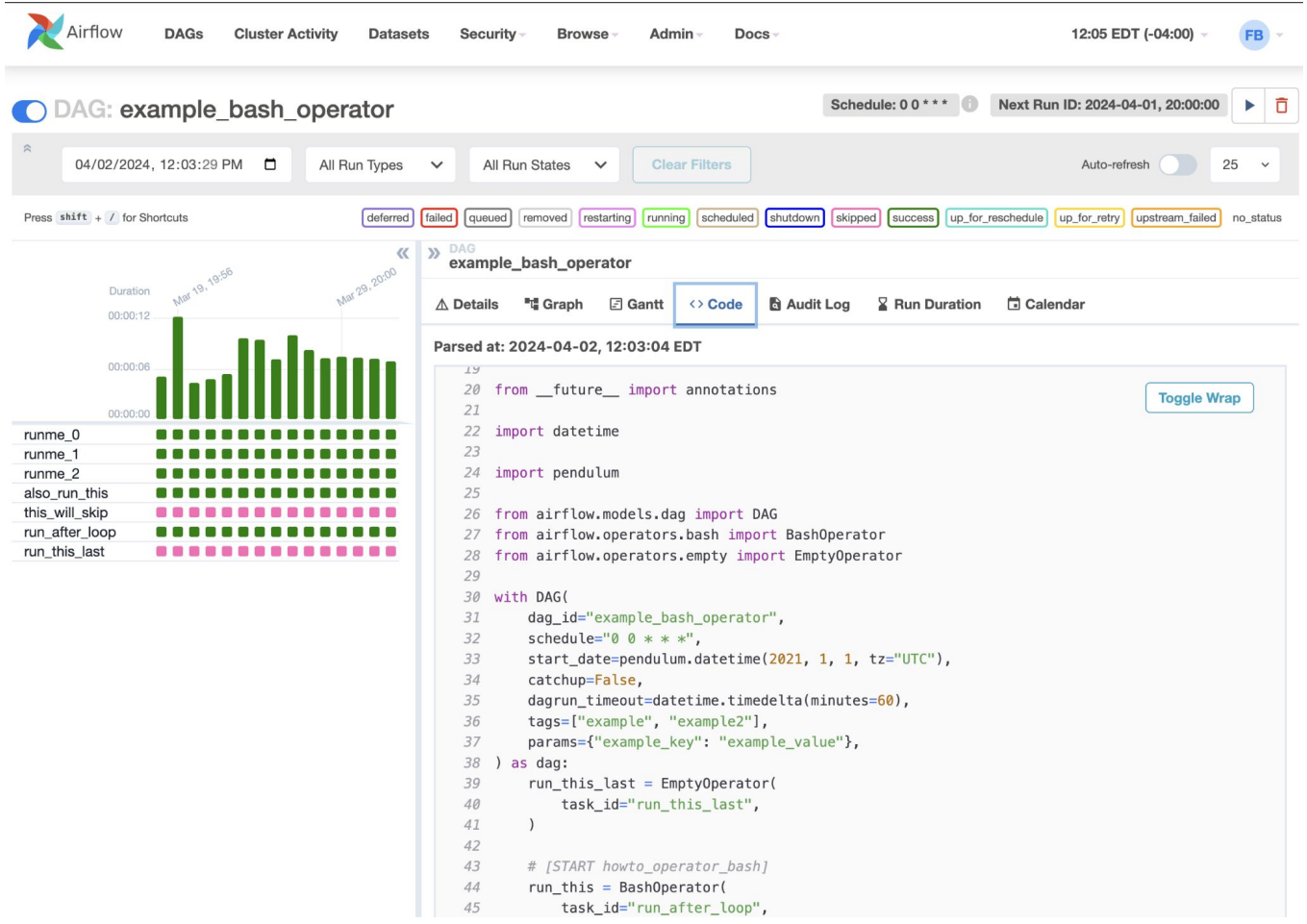


Apache Airflow

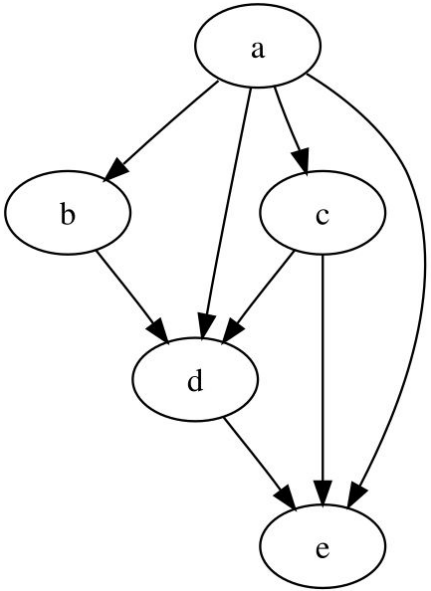
dagster

argo

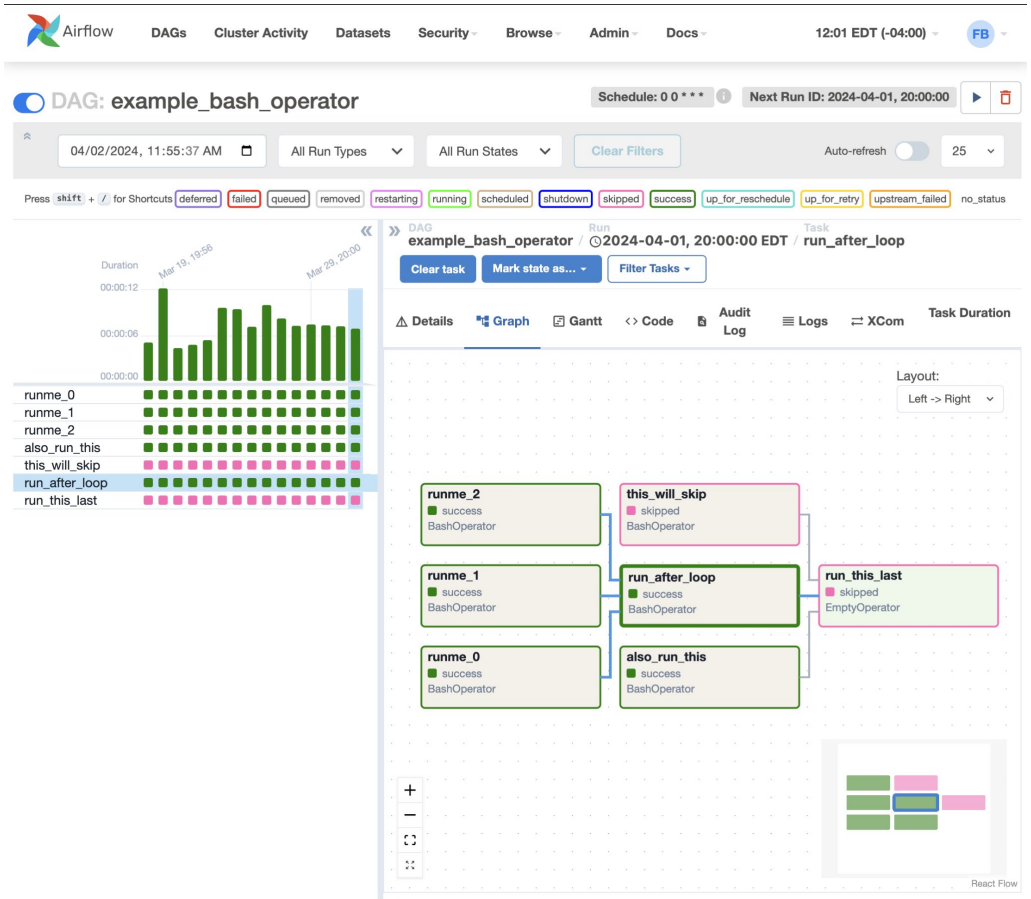
Medaillon architecture: Orchestrator



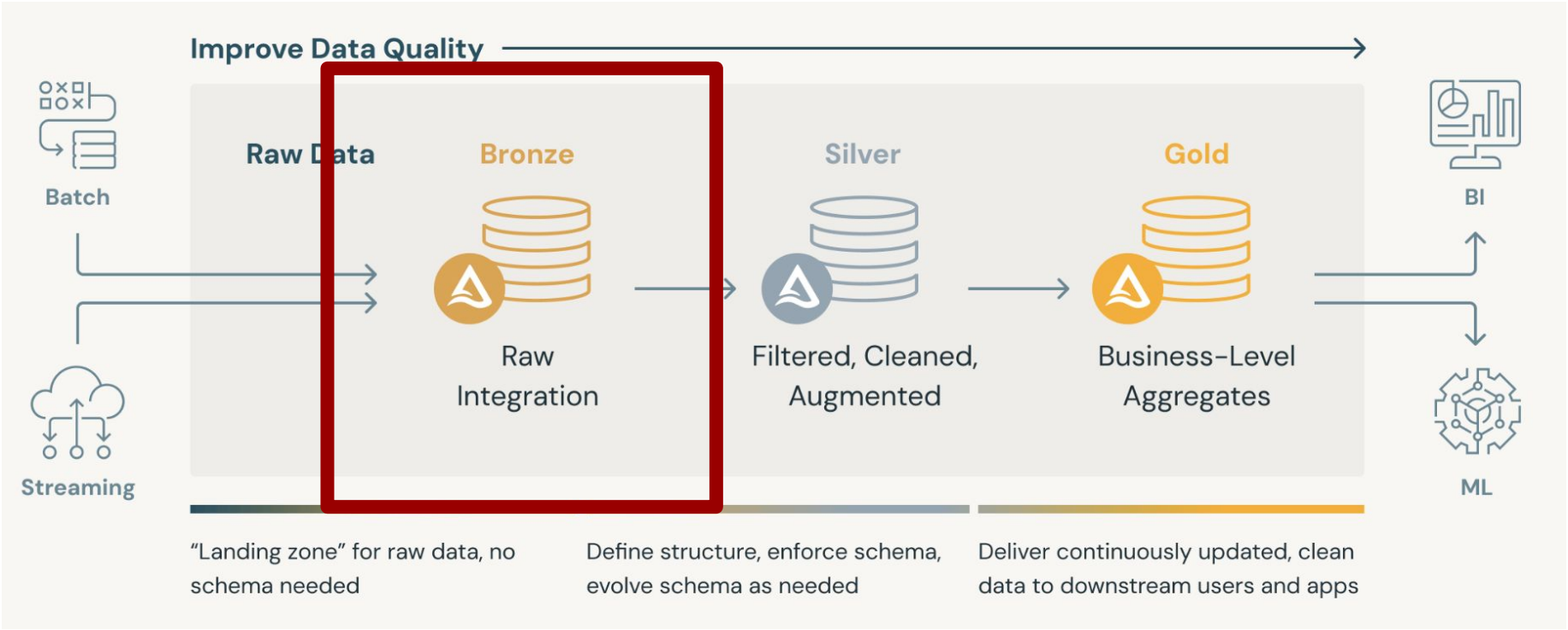
DAG: Directed acyclic graph



Medaillon architecture: Orchestrator



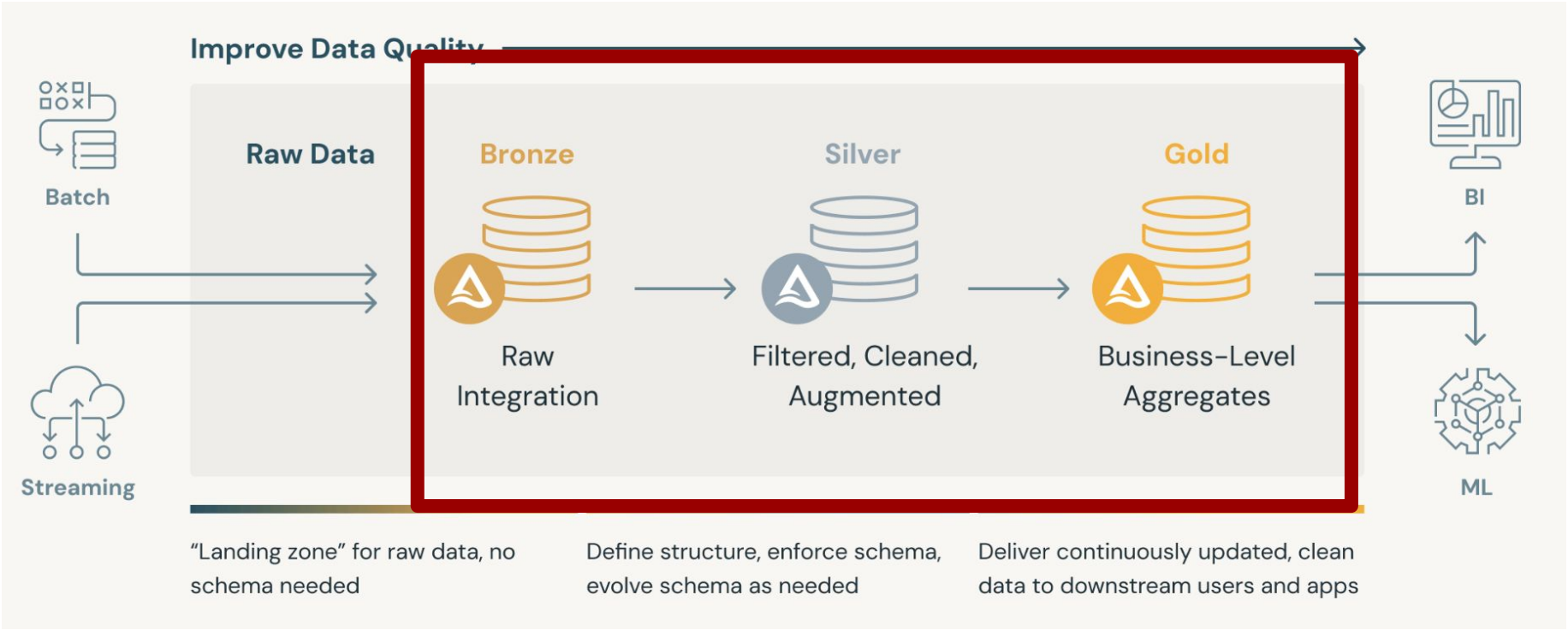
Medaillon architecture



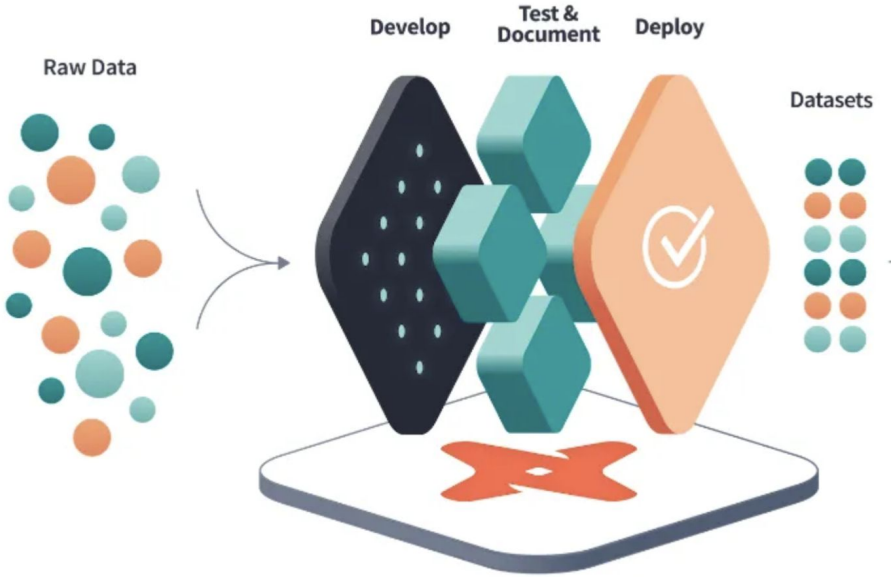
Medaillon architecture: Bronze layer

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

Medaillon architecture



Orchestrator

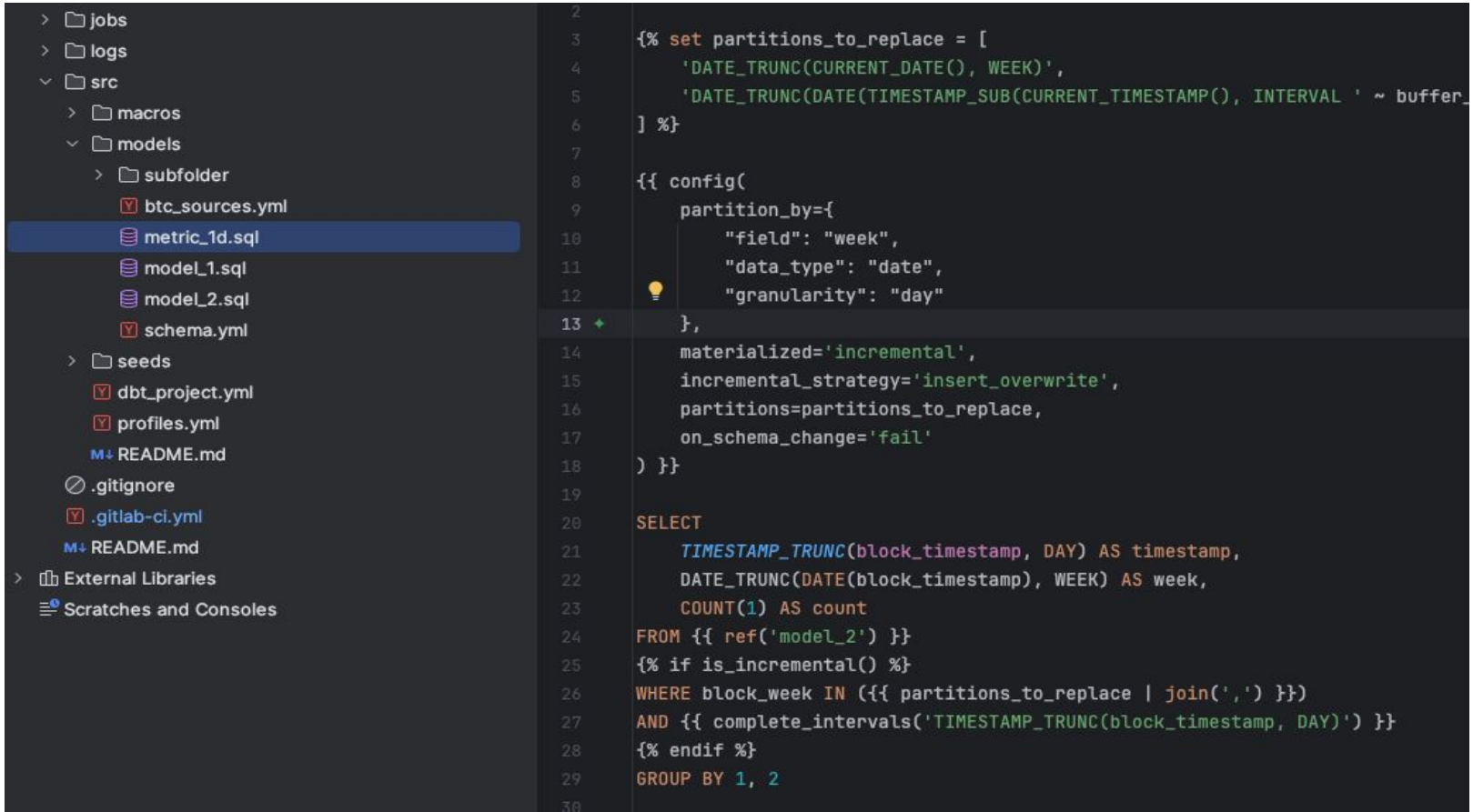


dbt

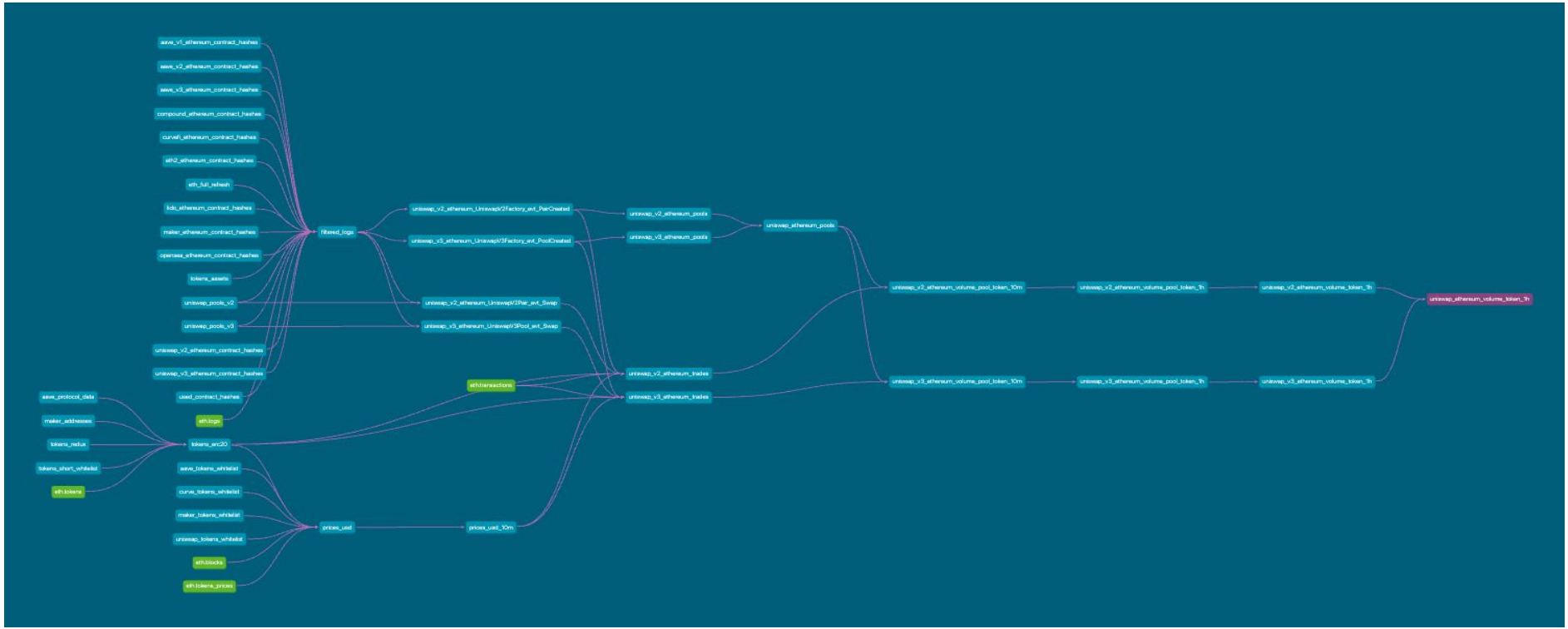
What is DBT?

- 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



Model: uniswap_ethereum_volume_token_1h

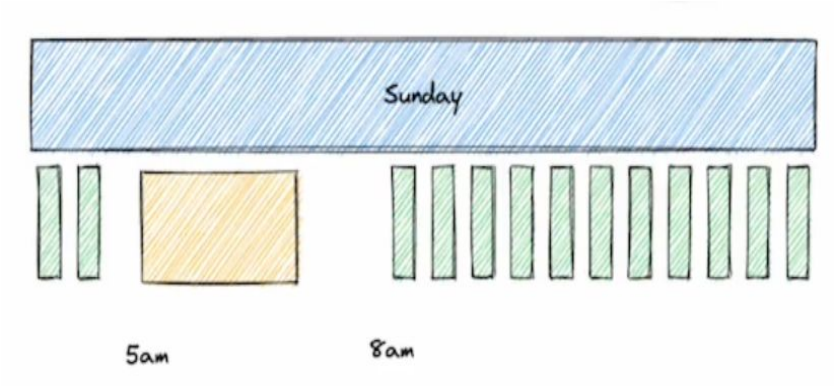


Full Refresh Job vs. Incremental DBT Job

Full Refresh:

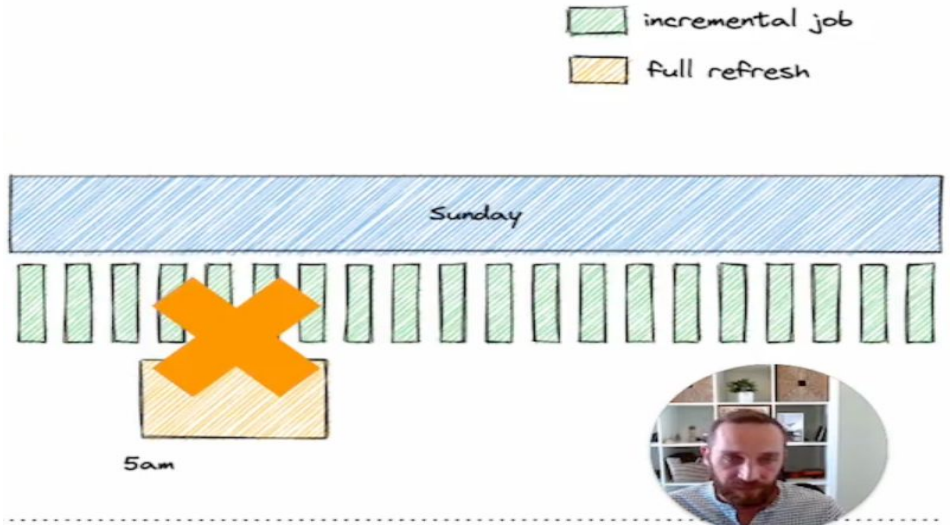
Drops any changed models and recomputes from scratch.

dbt run *--select state:modified --state old_target*
<more params>

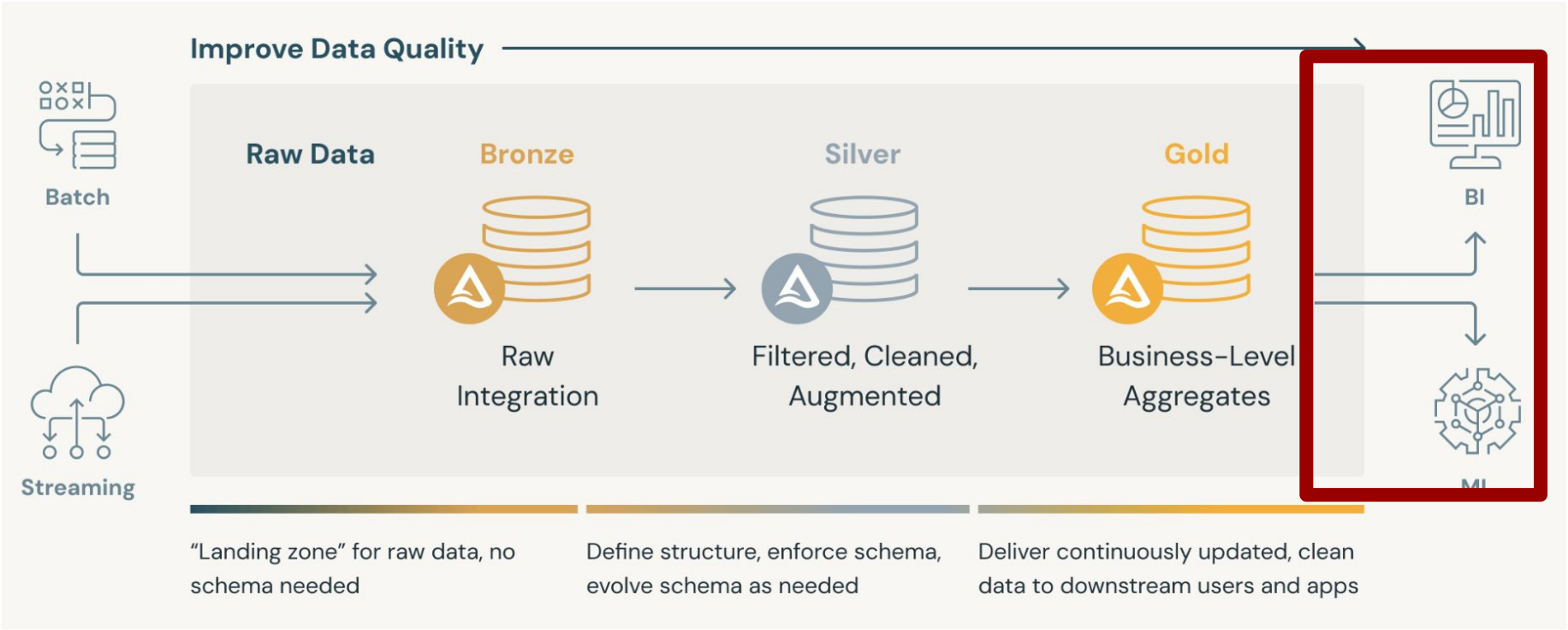


Incremental:

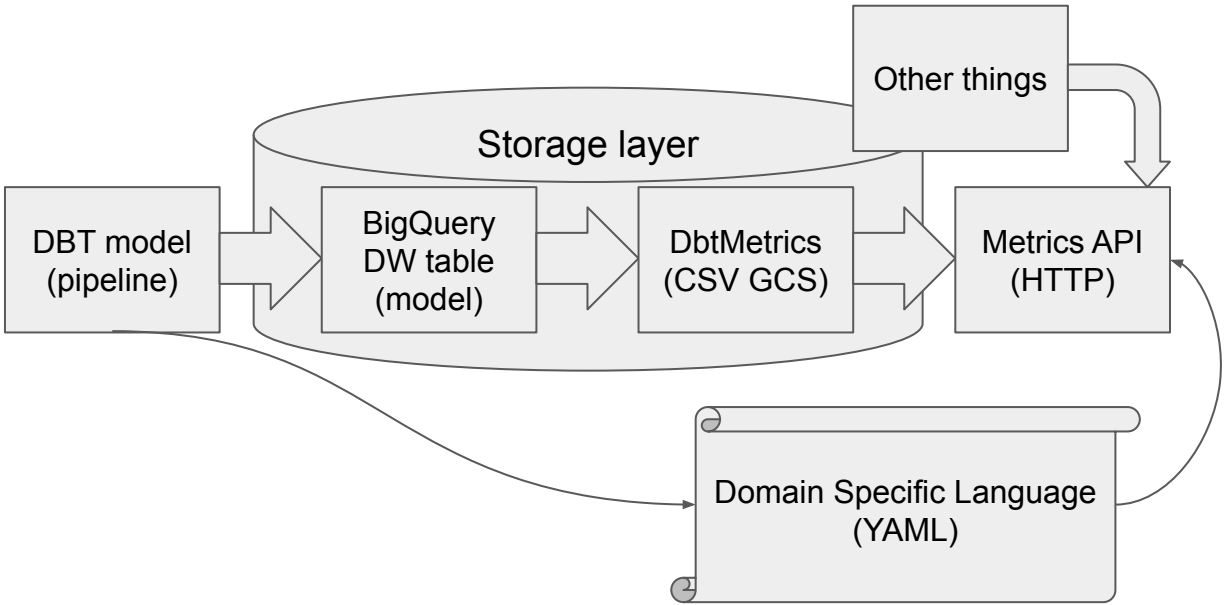
Updates the model with new data since the last run, more efficient for



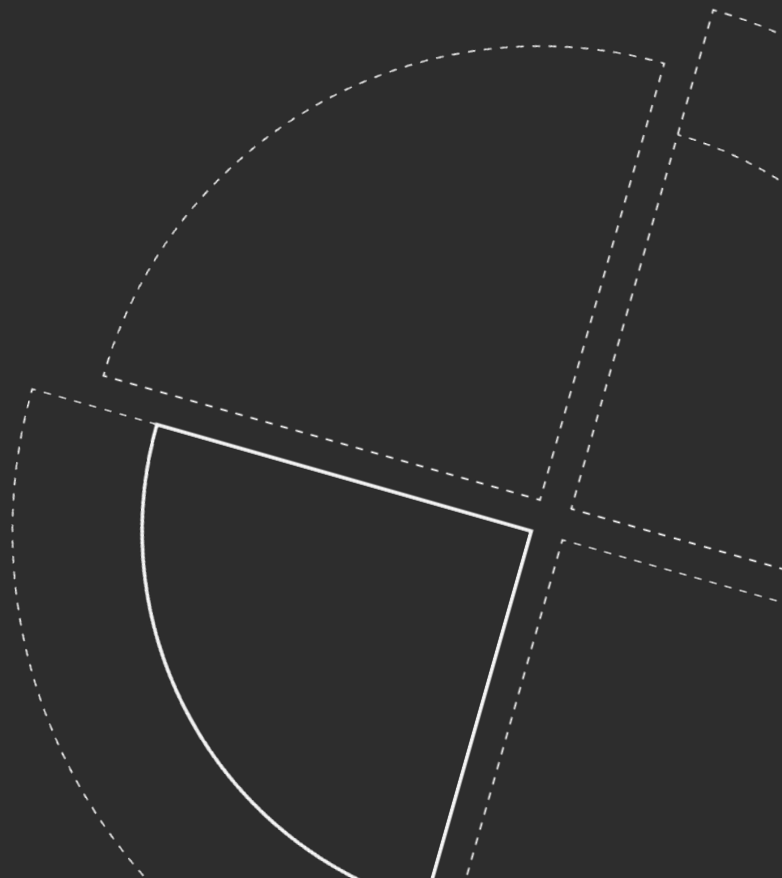
Medaillon architecture



Metric DSL


Other things
Storage layer
DBT model
(pipeline)
BigQuery
DW table
(model)
DbtMetrics
(CSV GCS)
Metrics API
(HTTP)
Domain Specific Language
(YAML)

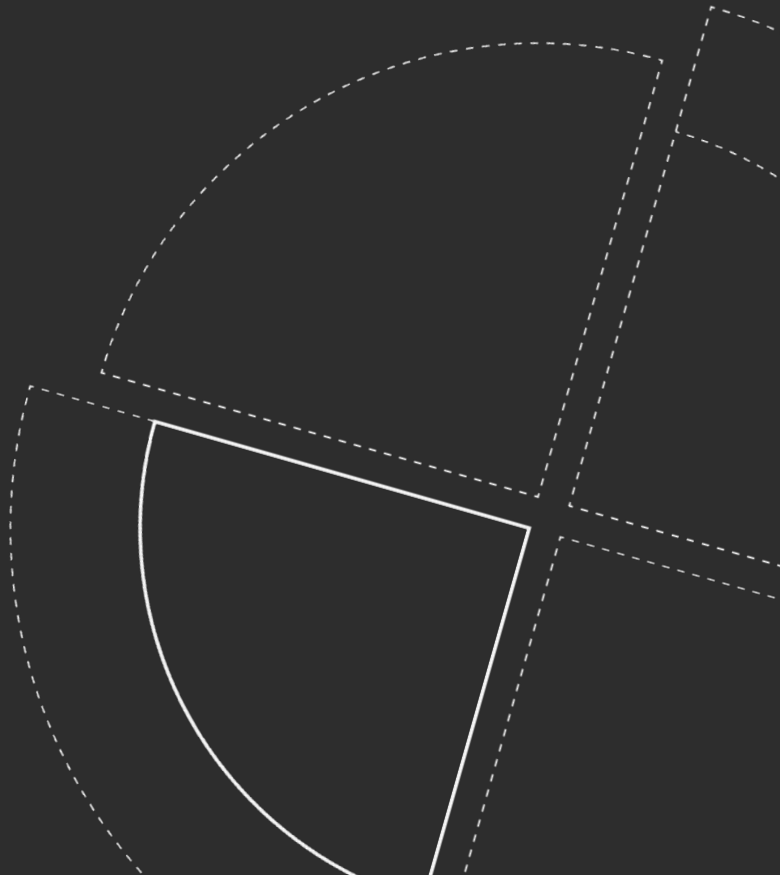
Next steps



Medaillon architecture: Next steps

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

Gràcies 🎉



Q&A