

Real Time Intelligence Experience in Microsoft Fabric

Nicola Paro

The logo for Data Saturdays, featuring the word "DATA" in a large, white, sans-serif font. To the left of the "D" is a blue icon consisting of three horizontal bars of increasing length, resembling a stylized "D" or a data bar chart. Below "DATA" is the word "SATURDAYS" in a smaller, white, sans-serif font.

DATA
SATURDAYS



Sponsors





TMPPRO-V1

€ 100

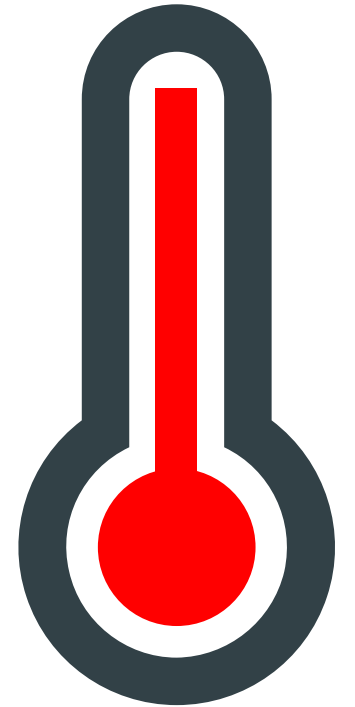


TMPPRO-V2

€ 1000

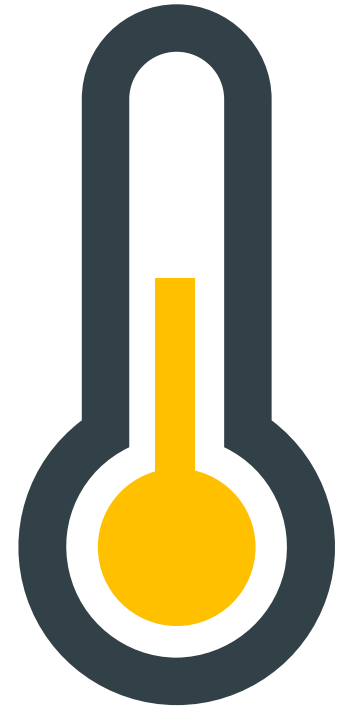
Hot - primo livello di monitoraggio

- Variabili «onchange» di cui tracciare il cambiamento
 - Stati, Allarmi, Settings, Counters
- Cardinalità di valori limitata
- Immediatamente notificati, oppure aggregati
- Inviati nell'ordine dei minuti
- Json+MQTT



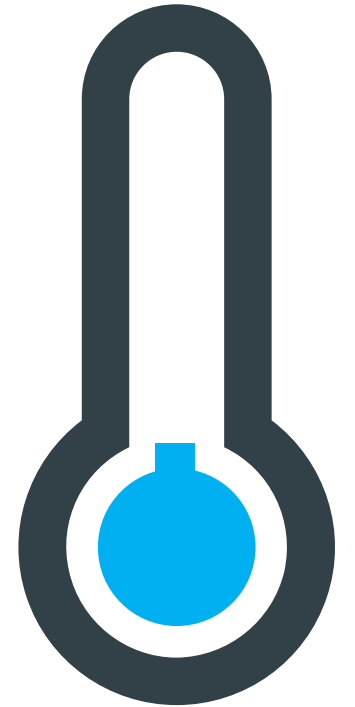
Warm: il secondo livello di monitoraggio

- Variabili di performance, con valori continui (decimali)
 - Se sono in allarme, o voglio fare una ottimizzazioni, allora guardo lo storico, che arriva su con calma e regolarità
- Possono essere anche campionati nell'ordine del secondo, ma sono bufferizzati ed inviati a pacchetti
- CSV, Parquet



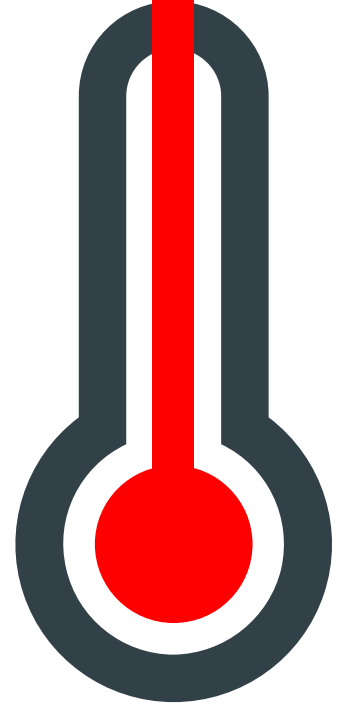
Cold: non è monitoraggio

- Dati relativi al contesto produttivo
 - Asset, produzione, mes, erp, per decodificare i dati che arrivano dagli asset
- Non è OT
- SQL, Data Integration, ETL



Hot Streaming: il terzo livello di monitoraggio

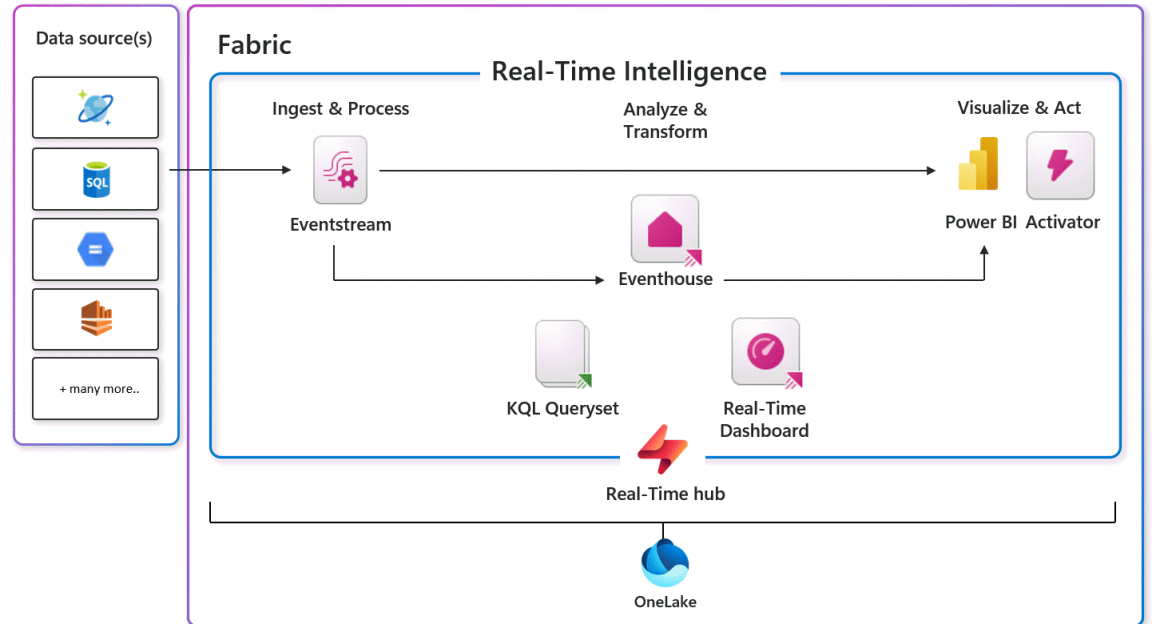
- Se serve il qui e ora, si fa lo streaming dei dati che normalmente arrivano come warm
- Lo si fa solamente «on demand» on/off, per colmare il gap dell'attesa dei dati warm.
- Json+MQTT



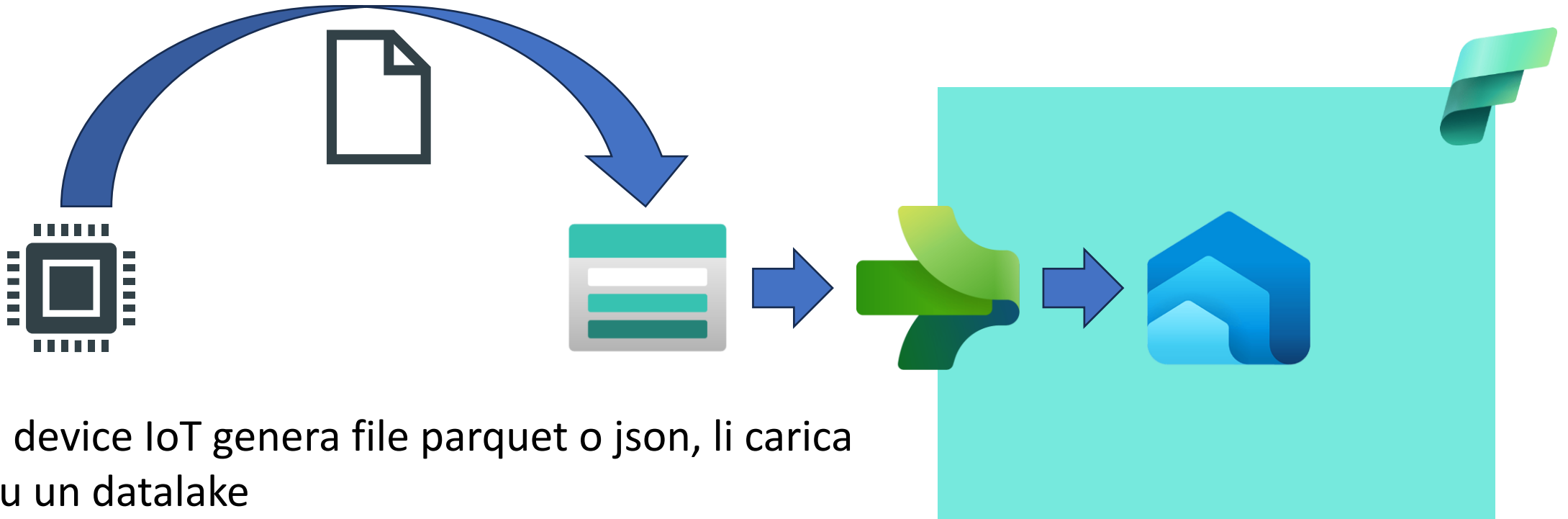
È uno dei workload/experience di Fabric

Si focalizza sul pattern e sui servizi che hanno a che fare con il concetto di «Evento»:

- Eventstream
- EventHouse
- Data Activator

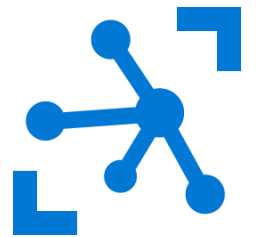


All'IoT piace LakeHouse e le shortcuts



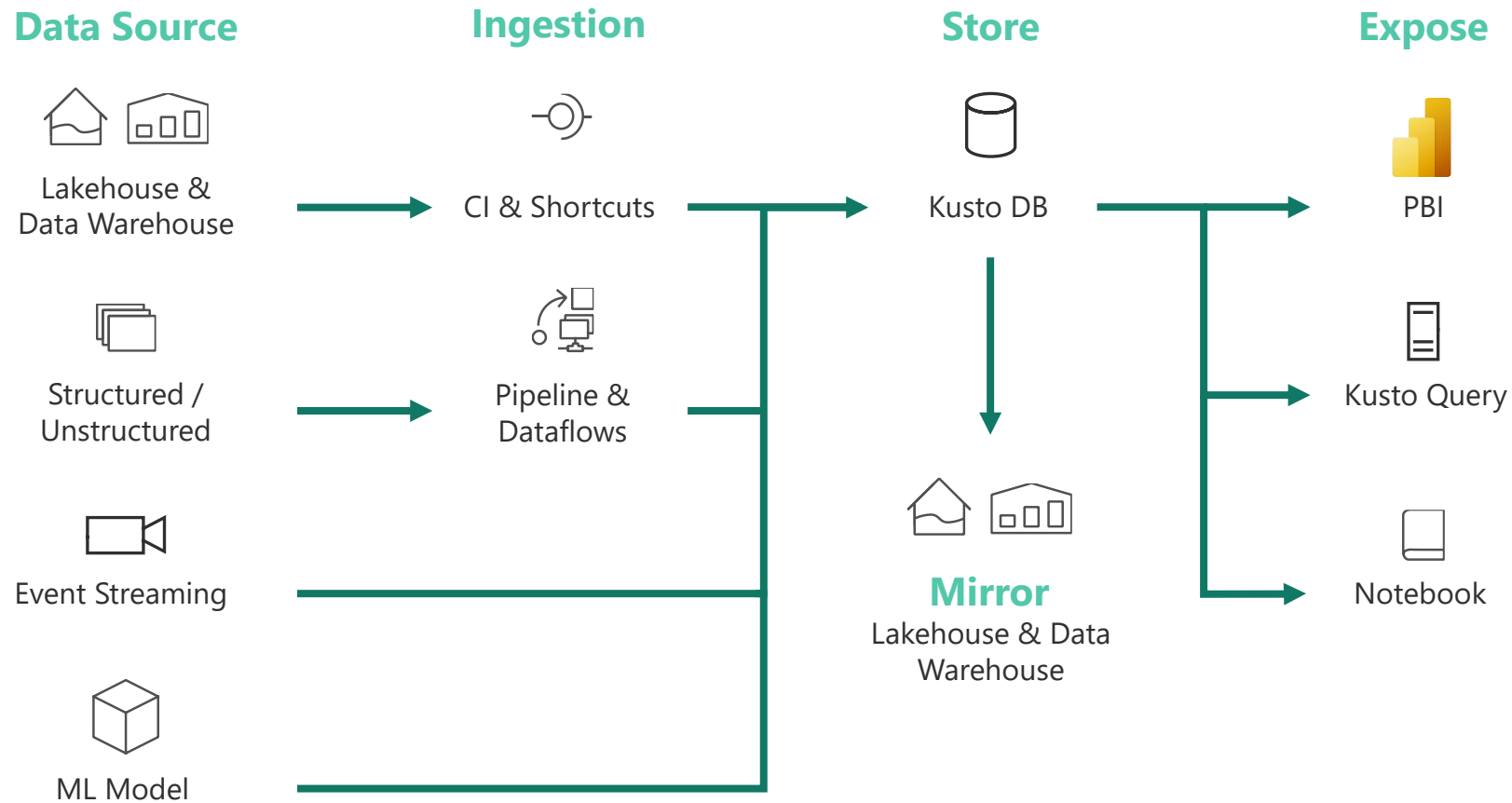
- Il device IoT genera file parquet o json, li carica su un datalake
- Posso caricare i dati in LakeHouse tramite shortcut o tramite una pipeline di Data Factory

- Azure Event Hub è un servizio di messaggistica in tempo reale completamente gestito di Microsoft Azure. È progettato per l'ingestion di grandi quantità di dati di eventi da varie fonti e la loro elaborazione in tempo reale.
- IoT Hub ha un endpoint compatibile EventHub
 - Gestisce anche le identità dei devices
- è una soluzione gestita che offre un'integrazione facile con altri servizi di Azure e una gestione semplificata



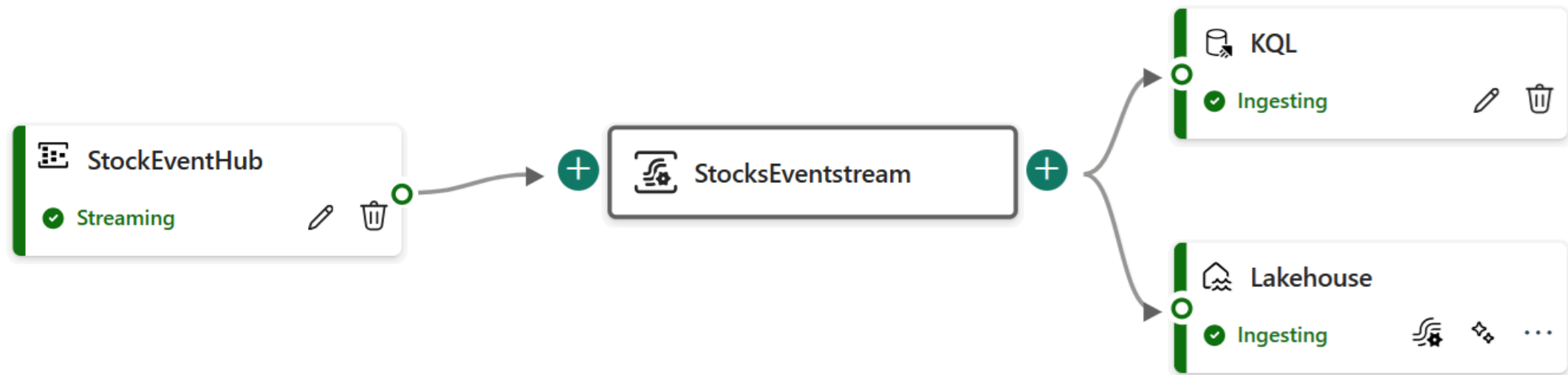
- Anche le applicazioni possono generare eventi
- Un log altro non è che la traccia della sequenza di eventi
- Fino a prova contraria anche un server può essere visto come un dispositivo IoT
- Uno storage può generare eventi ed in generale tutti i db con dei feed Change Data Capture sono buone sorgenti per Real Time Analytics

Common Real-time Analytics Patterns



Eventstreams

- Punto centralizzato dove raccogliere, trasformare e instradare gli eventi.
- Connettori per recuperare dati degli eventi da diverse fonti, come app personalizzate per inviare a Eventstreams, Azure Event Hub o Samples
- Esperienza Low Code nella definizione del flusso



Event House

L'esperienza KQL in Fabric

Fabric Workspace

KQL Database

Table

Schema

No Primary
Key

No Unique
Keys

No Foreign
Keys

Columnstore
Indexes

Data Sharding
(Extents)

Nome	Cognome	Data di Nascita	Sesso
Giovanni	Rossi	15/03/1987	M
Sonia	Bianchi	22/07/1995	F
Marco	Esposito	10/11/1980	M
Laura	Romano	05/09/1990	F
Luca	Russo	20/04/1975	M
Chiara	Colombo	18/06/1988	F
Matteo	Mioretti	30/12/1978	M
Alessia	Ferrari	08/02/1992	F
Federico	Conti	25/10/1983	M
Martina	Marini	12/07/1998	F

Giòwànni Ròssì Nşòğiã
 Biãñçhì GÑásçò Éş
 řòşîţşò NŁăũsă Rộặặặ
 GLũçả Rựşşộ
 NChĩảsă Cộặặặ
 GÑấţţệộ Nộsêţţĩ NÁlêşş
 iả Gêssásî GGêđêşîçộ
 Cộặặặ NÑấţţĩặ Nấşĩặ
 G

[illegible]

Quando ha senso usare un KQL Database?

Sliding
Window of
data

Tante letture

Tanti Insert /
Append

Poche Delete

NESSUN
Update

Kusto Query Language

Una query Kusto è una richiesta in sola lettura per il processamento dei dati e la produzione di risultati.

La richiesta è effettuata tramite testo, utilizzando un modello di data-flow che è semplice da leggere, scrivere ed automatizzabile.

Le query Kusto sono composte di una o più istruzioni.



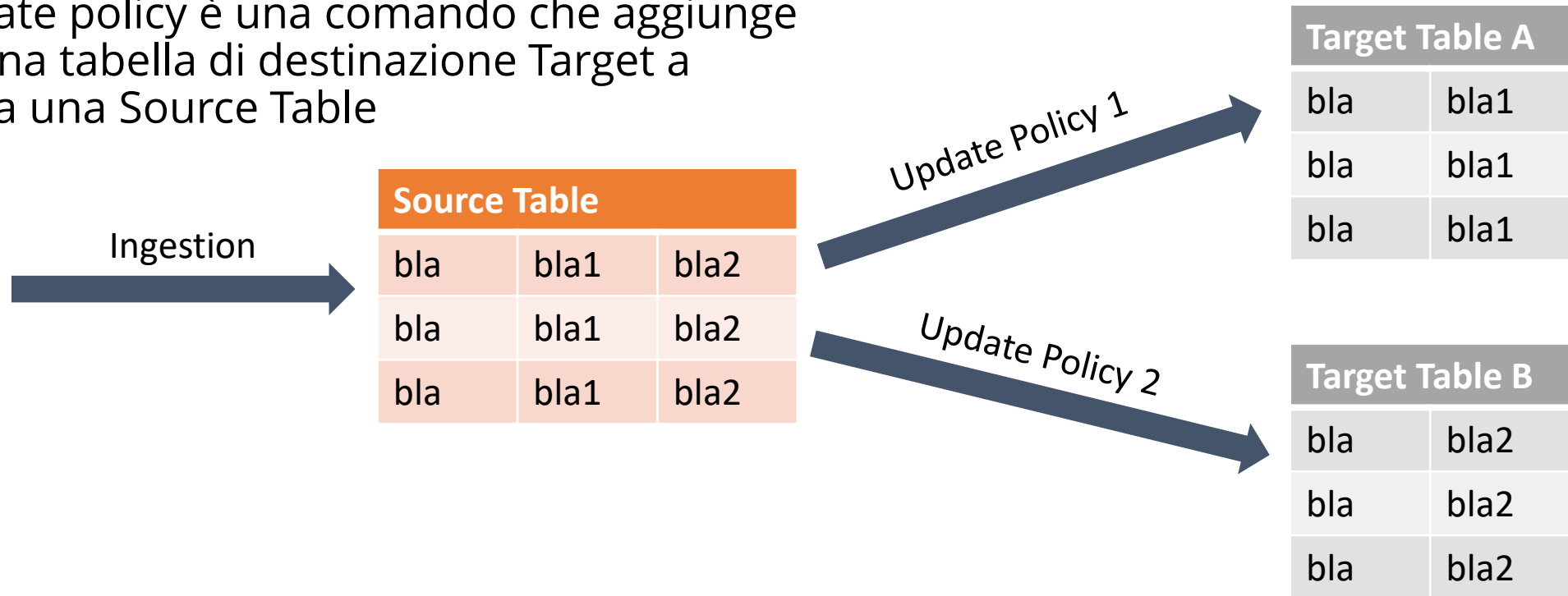
Kusto Query Language

SQL	KQL
SELECT	project , extend, project-away, project-keep ...
WHERE	where , search, ...
JOIN	join kind=inner
UNION	union
GROUP BY	summarize
ORDER BY	sort by , order by, top by
TOP, LIMIT	take

More on <https://learn.microsoft.com/en-us/azure/data-explorer/kusto/query/>

Update Policies

Una update policy è un comando che aggiunge dati ad una tabella di destinazione Target a partire da una Source Table



Materialized View

Una Materialized View è una vista aggregate sui dati di una tabella KQL. I dati sono materializzati anche su disco.

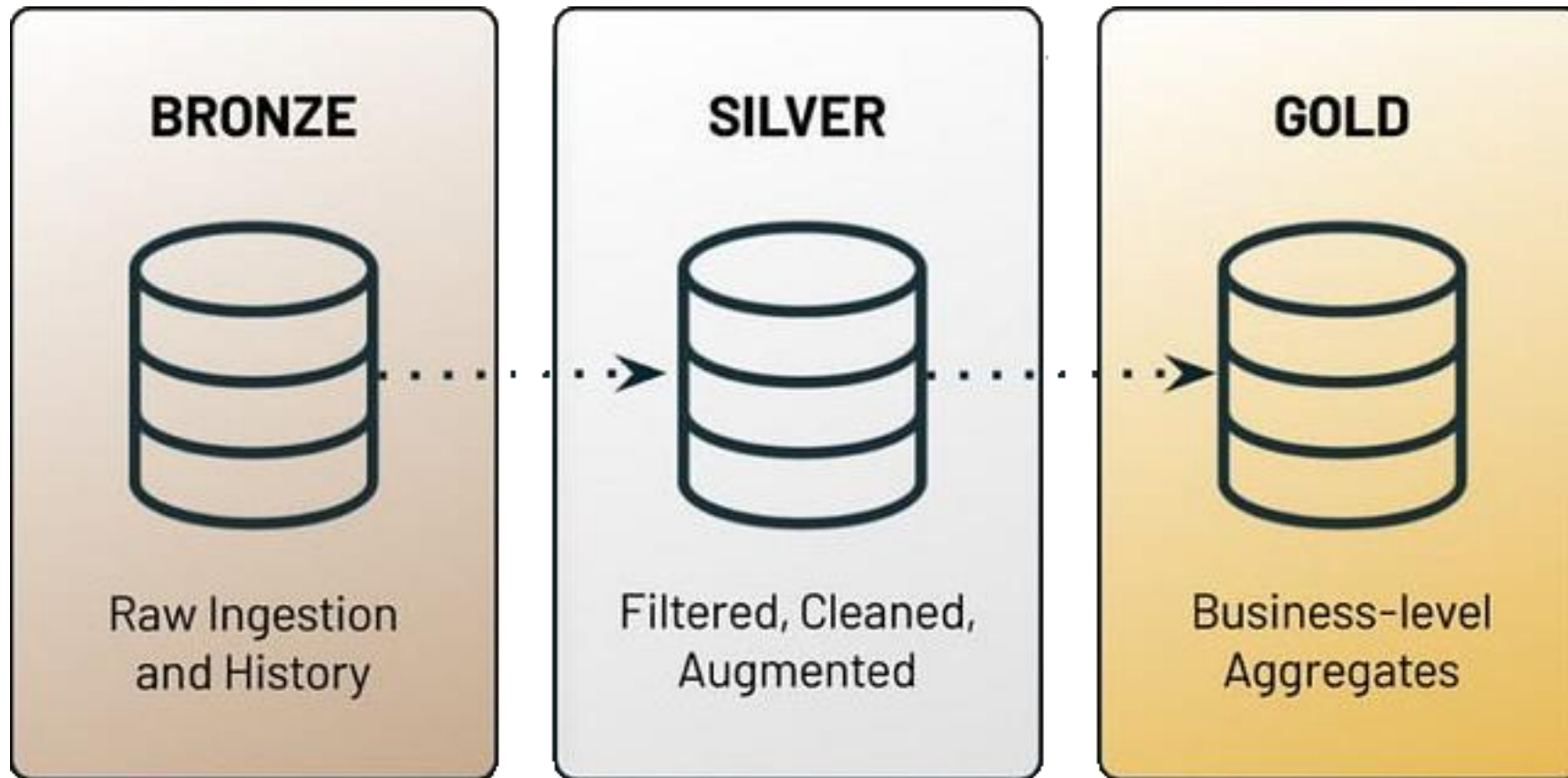
Vantaggi nell'adozione delle Materialized View

Performance
Improvement

Data
Freshness

Cost
Reduction

Medallion Architecture



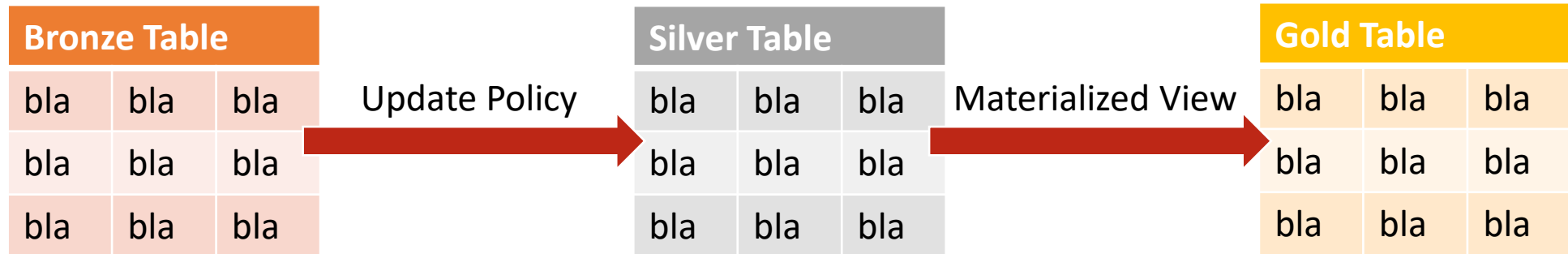
Update Policy o Materialized View?

Update Policy

- Data Transformation
- Data Enrichment

Materialized View

- Data Aggregation



La query di una update policy può invocare stored functions, ma:

- Non può eseguire query incrociate tra diversi eventhouse.
- Non può accedere a dati esterni o tabelle esterne.
- Non può effettuare richieste esterne (utilizzando un plugin).

La query non ha accesso in lettura alle tabelle che hanno abilitata la policy RestrictedViewAccess.

Limitazioni per le Materialized Views

- La tabella sorgente di una **materialized view** deve essere una **tabella in cui i dati vengono inseriti direttamente**
 - Ok update policy o ingest query.
- La tabella sorgente di una materialized view deve avere la **IngestionTime** policy abilitata. (Abilitata di default).
- **NO** operatori come **mv-expand** o il plugin **pivot** che non preservano il valore di **ingestion_time()** e quindi non possono essere utilizzati in una **materialized view** con un **lookback**.
- Una materialized view **non può essere creata sopra un'altra materialized view**, a meno che la prima materialized view non sia del tipo **take_any(*)** aggregation.
- Le materialized views non possono essere definite su **external tables**.

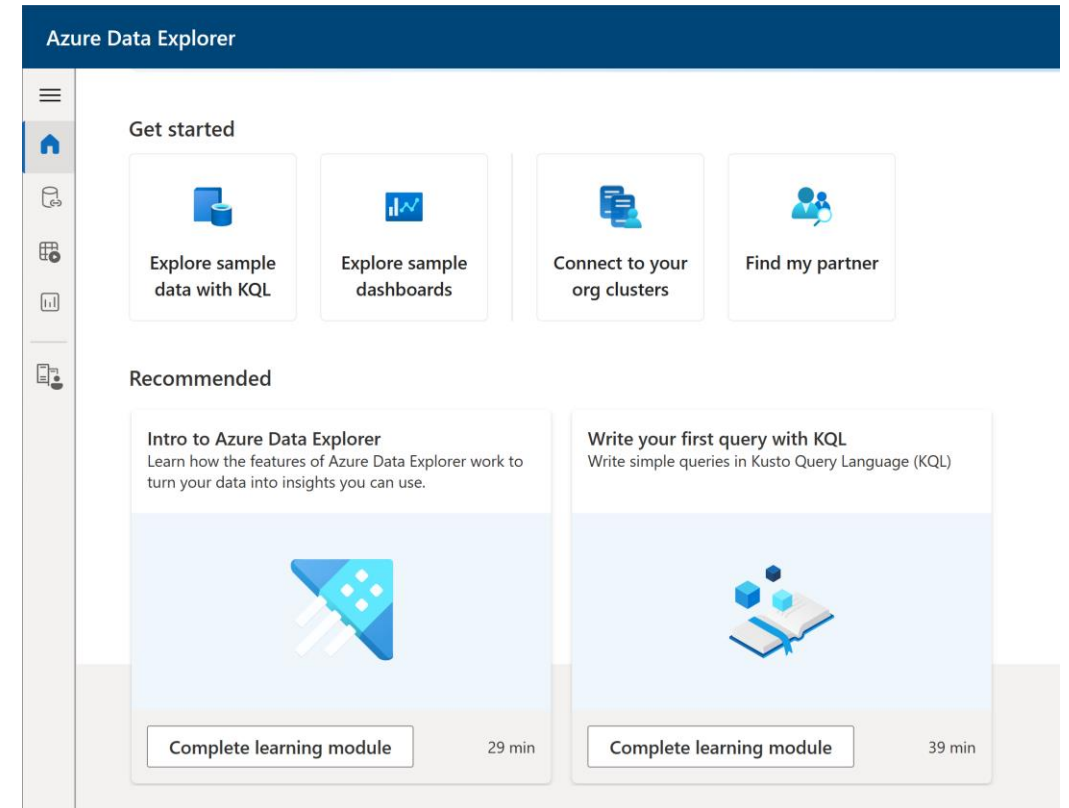


Devo pagare Fabric per far pratica con Kusto?

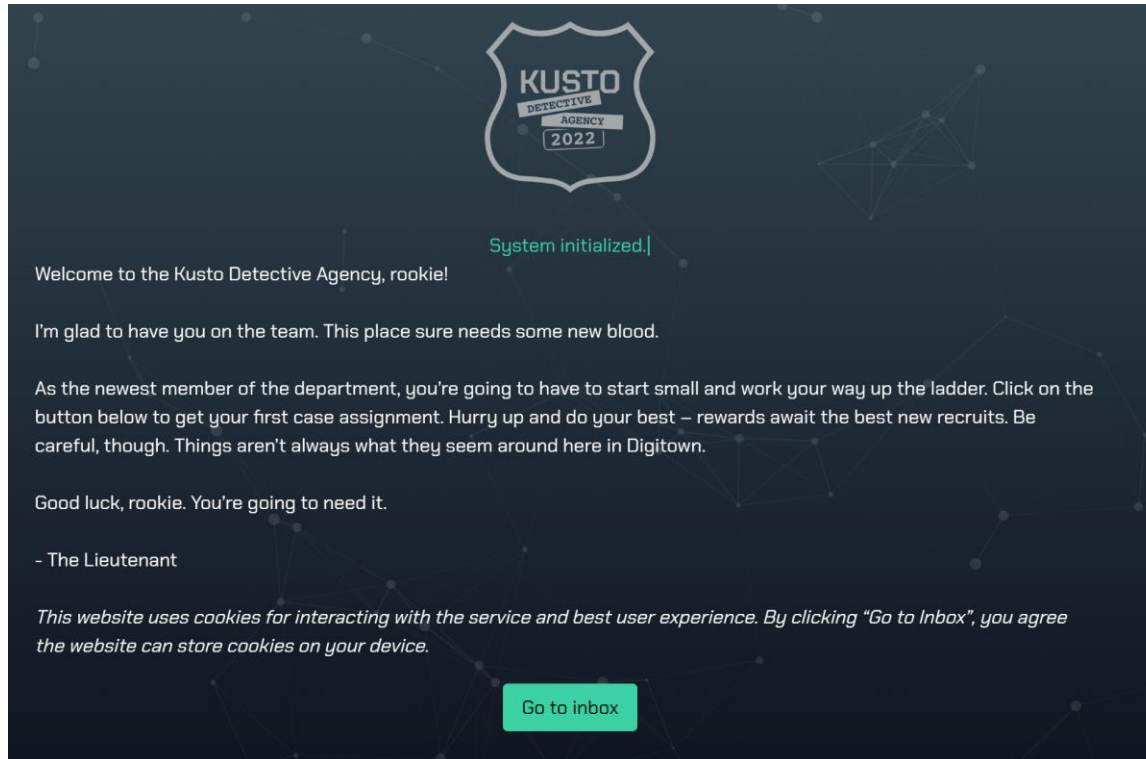
Realtime Analytics is Azure Data Explorer!

Azure Data Explorer supporta dei database “sample” gratuiti su cui è possibile effettuare delle interrogazioni per provare

<https://dataexplorer.azure.com/home>



Devo pagare Fabric per far pratica con Kusto?



Kusto Detective Agency: una gamification per imparare ad usare kusto.

<https://detective.kusto.io/>

About me



Nicola Paro

Solution Architect – beanTech



linktr.ee/nicolaparo

DATA
SATURDAYS

hn0va
community