Licensing woes and headaches

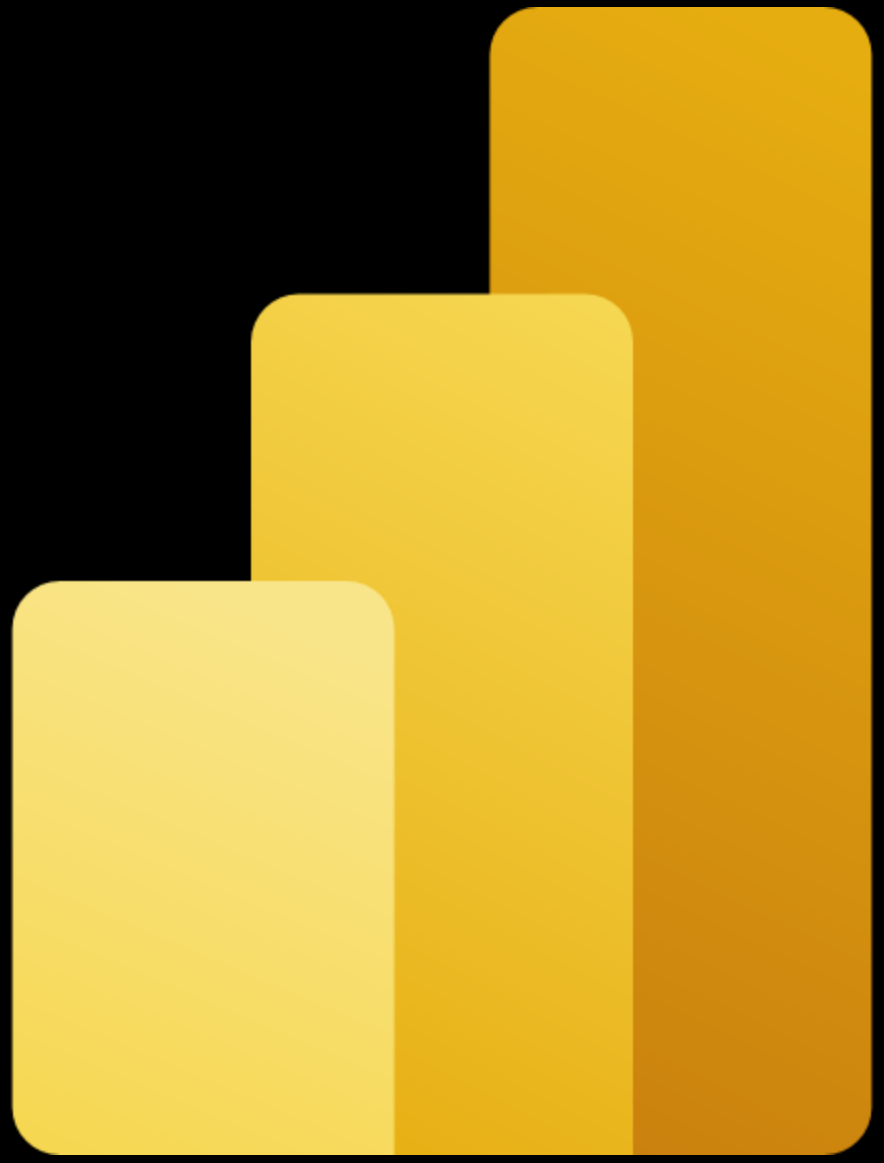
- how to handle Power BI and Fabric

Agenda

- Licensing in Power BI
- The Fabric versions
- What are the capacity limits?
- What to use when?
- Usage patterns

Power BI licensing

- Power BI User based licensing
- Power BI Premium
- Power BI embedded



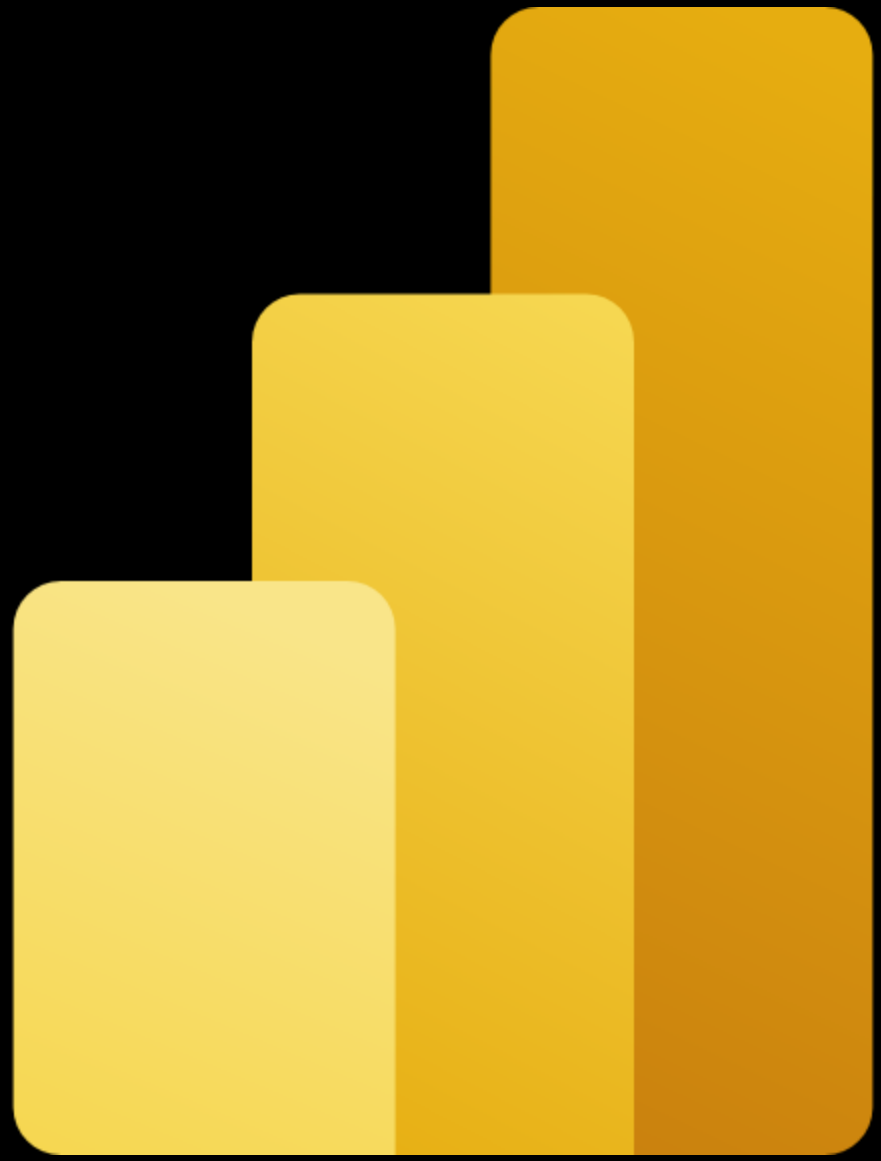
Power BI user-based licensing

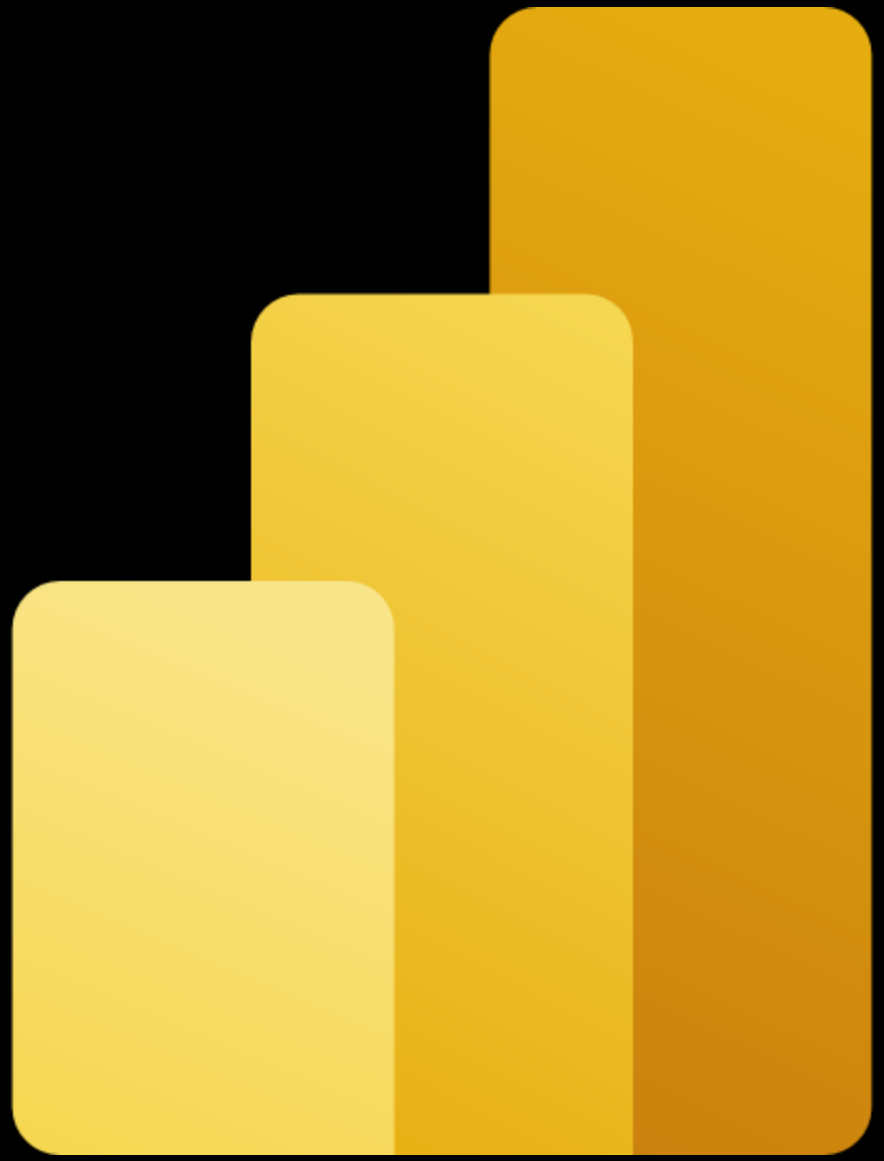
- Power BI Free
- Power BI Pro
- Power BI Premium Per user

- Uses shared capacity

Power BI Premium

- Tenant level subscription with dedicated capacity
- P scale from P1 up to P5
- P1 = 8 vCores, P5 = 128 vCores
- App users don’t need a Pro license



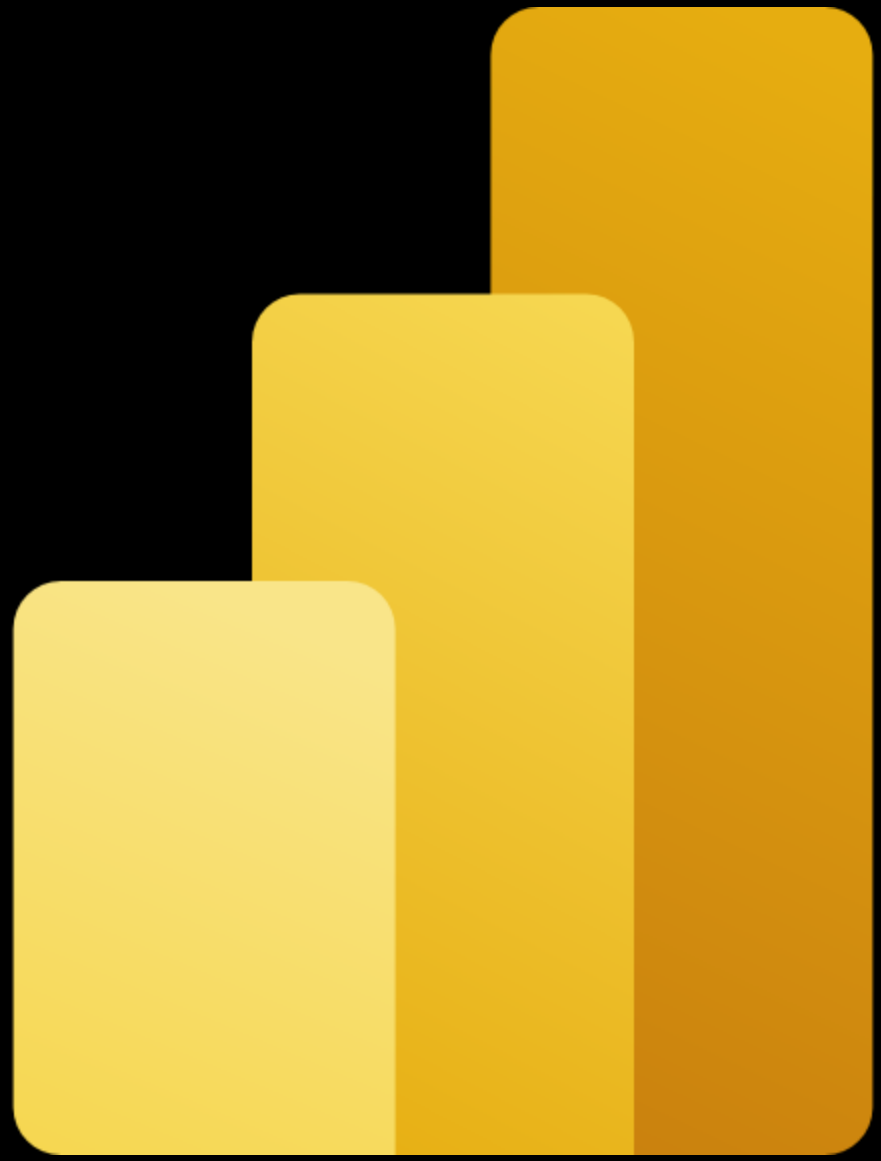


Power BI Premium

- Tenant level subscription with dedicated capacity
- EM is meant for internal embedding
- EM scales from EM1 to EM3
- EM 1 = 1 vCore, EM3 = 4 vCores
- App users need a Pro license
- Users authenticate via Entra ID

Power BI embedded

- Power BI capacity without need for user licenses
- Aimed at external users
- Scaling from A1 up to A8 (P5)
- A1 = 1 vCore, A8 = 128 vCores
- Authenticate with own authentication method
- Non-interactive authentication. Your app uses a *service principal* or a *master user* to authenticate



Fabric licensing

Fabric capacities

- 3 ways to get Fabric capacity:
  - Trial
  - Fabric capacity
  - Power BI Premium capacity

Fabric licensing model is that easy?

Fabric Capacity

- You will see both the term SKU and CU
- SKU = Stock Keeping Units
- CU = Capacity Units

- There are 2 types of SKU
  - Azure SKU
  - Microsoft 365 SKU (discontinued June 2024)

Fabric Capacityic

- Azure SKU
  - Billed per second
  - No commitment
  - Can be paused (storage still being billed)
  - Bought and set up via Azure
- Microsoft 365 SKU
  - Billed monthly or yearly
  - Monthly commitment
  - Actually P SKUs
  - No new customers on P SKU from June 2024

Fabric Capacityic

- SKUs start at F2 and scales to F2048
- F64 = P1 = 8 vCores – or 64 CU
- F1024 = P5 = 128 vCores – or 1024 CU
- F2048 = 256 vCores – or 2048 CU

- F64 and above = no Pro or PPU license for Power BI app users.
- Fabric users still need Pro to develop or access workspaces

So… What separates the SKUs?

- Fabric SKUs have 3 main differentiators
  1. Storage
  2. Compute
  3. Concurrency limits

1. Storage

- Storage is billed seperatly
- Storage does NOT consume CUs
- Transactions however does…
- Higher cost for data outside capacity
- Shortcuts
- Then there’s DirectLake

1. Storage

DirectLake

- Lets Semantic models bypass SQL endpoints
- SKU limitations on several parameters
- Table size
  a) Number of Parquet files
  b) Number of row groups
  c) Rows per table in millions

1. Storage

DirectLake

- Max Model size on disk in GB
- Max Memory in GB
- F2: 1000/1000/300/10/3
- F64: 5000/5000/1500/Unlimited/25
- F1024: 10000/10000/24000/Unlimited/400

2. Compute

Specifically burstable capacity in data warehouse (and SQL endpoints)

- A capacity is a distinct pool of resources
- SKU determines size of said pool
- Baseline and burstable capacity
- Baseline CU = Fx

2. Compute

- Trading in future CU usage
- This is called Smoothing
- CU used / duration / Baseline CU = Scale factor
- Lower capacity has higher Scale factor
- F2 = x32 – F8 and up = x12
- Ingestion is isolated from querying

2. Compute

- Throttling
- Eg. If overuse over 24 hours – throttling starts
- Background rejection normal policy
  - All new jobs rejected
- Exception for warehouse : data modeling
- Other throttle policies:
  - Overage protection
  - Interactive delay
  - Interactive rejection

2. Compute

- How to handle throttling:
  - Upgrade SKU
  - Find and fix overuse source(s)
  - Wait until throttling stopped
  - Pay your way out… (pause and restart)
  - Create alerts for overusage to avoid throttling

- Coming soon: Surge protection

3. Concurrency limits

Specific for Spark

- 1 CU = 2 Spark vCore
- FIFO queue
- Burstable CU for 3x baseline CU
- Throttle = rejection

- Currently – default spark pool medium size…

So what does this all mean?

- How can we know what things *actually* costs?
- Let us look at some samples…

Eventstream Capacity

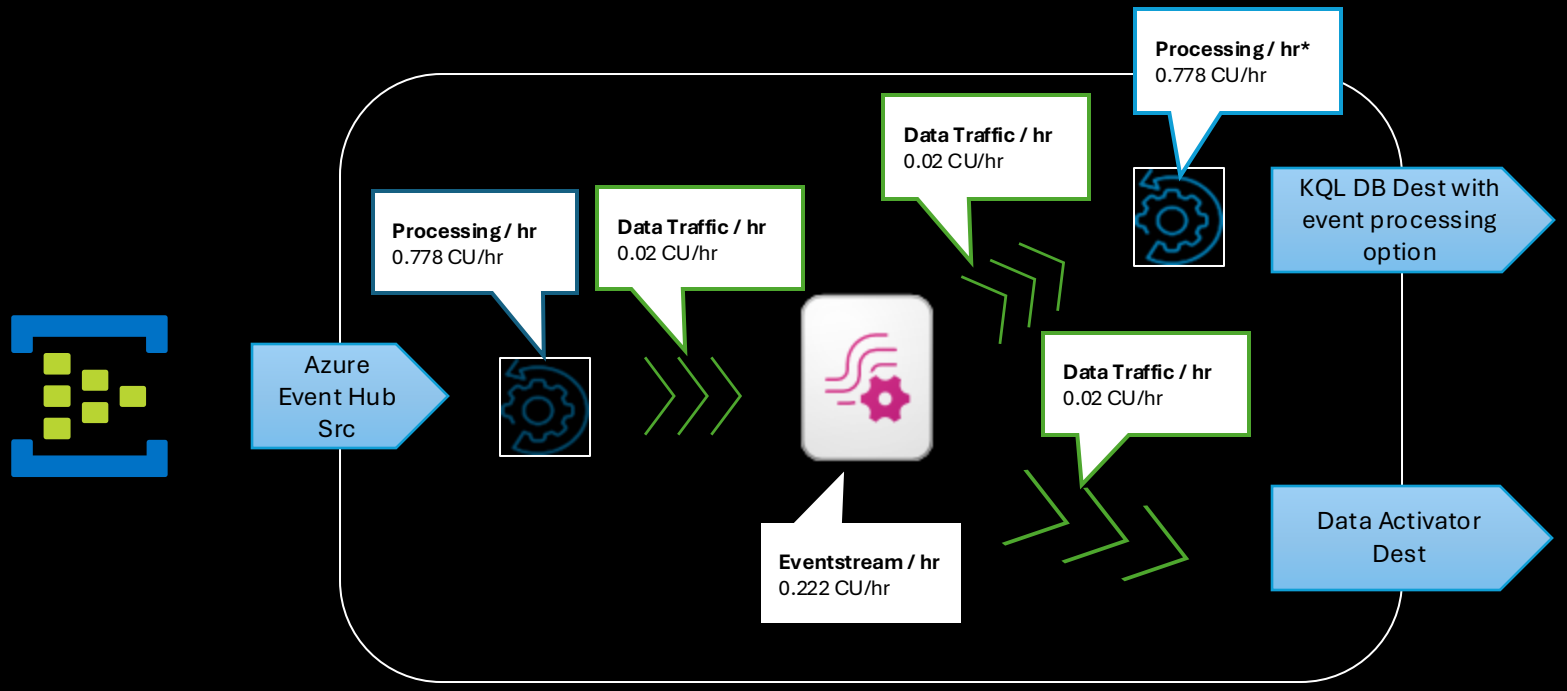
OPERATION IN CAPACITY METRICS APP	DESCRIPTION	OPERATION UNIT OF MEASURE	FABRIC CONSUMPTION RATE
Eventstream Per Hour	Flat charge* (per eventstream)	Per hour	0.222 CU per hour
Eventstream Data Traffic per GB	Data ingress & egress volume	Per GB	0.342 CU per hour per GB
Eventstream Processor Per Hour	Computing resources consumed by the processor	Per hour	0.778 CU per CPU hour**
OneLake Standard Storage	Used for extended retention (Includes 24-hour retention)	Per GB	$0.023 per GB***
Eventstream Connectors Per vCore hour	Computing resources consumed by the connector	Per hour	0.611 CU per vCore hour

*Eventstream is charged if Eventstream is not idle (i.e. data is not ingested or egressed for at least 2 hours)
**Processor base rate 2.333 CU per hour, starting at 1/3 CPU hour (0.778)
***Refer to OneLake Storage pricing for more detail - Microsoft Fabric - Pricing | Microsoft Azure

Monitor Microsoft Fabric event streams capacity consumption - Microsoft Fabric | Microsoft Learn

Eventstream usage scenario

An eventstream is ingesting from an Azure Event Hub source that is coming in at **1 MB/minute (0.059GB/hr)**. The data is transformed and filtered before sending to a KQL DB destination and also routed to a Data Activator destination.



Total CU/hr: **1.838** CU

*Processing can cost more depending on transformation complexity (e.g. aggregation)

**Does not include downstream costs for KQL DB and Data Activator

Eventhouse Usage Scenerio

Eventhouse
4.25 Cores
100 GB in Cache
500 GB Total Storage

Capacity via Activity
Queries or Command or Ingestion
80% active would need 3.4 CUs and require min 4 CUs

Storage

100 GB at Premium Storage (~$0.20/GB)

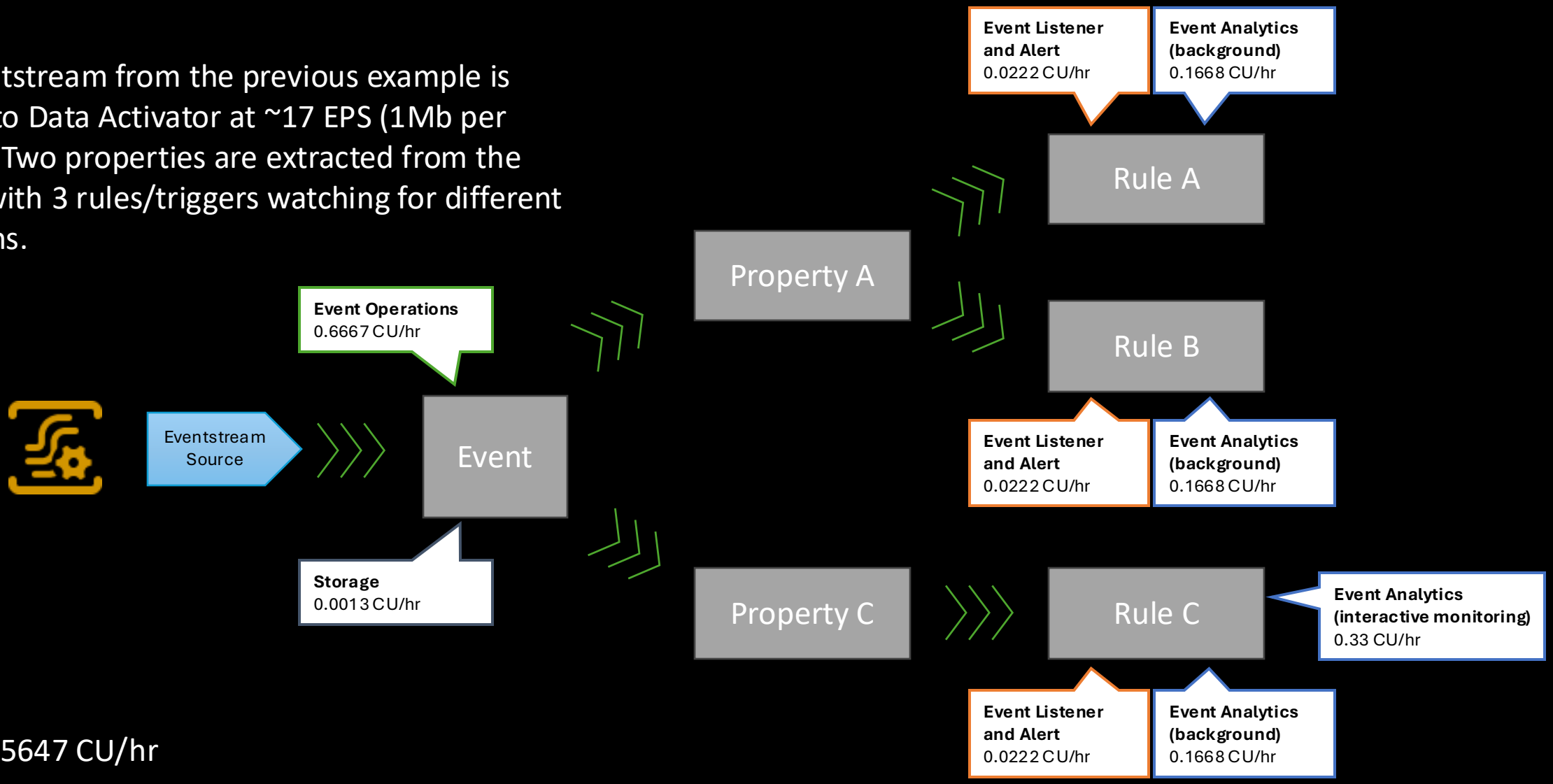
500 GB at Standard Storage (~$0.02/GB)

Data Activator – Consumption Meters

BILLING METER	DESCRIPTION	OPERATION UNIT OF MEASURE	FABRIC CONSUMPTION RATE
Real-Time Intelligence – Event Listener & Alert	Flat charge per rule (trigger)	Per hour	0.0222 CU
Real-Time Intelligence – Event Operations	Data ingress	Per M events*	11.111 CU
Data Activator – Event Analytics	Events processed to evaluate rules, run queries etc.	Per M events	2.78 CU
Storage		Per GB/ month	1.333 CU

***Event size 0-1Kb = 1 event, 1-2Kb = 2 events, 2-3Kb = 3 events etc.**

The eventstream from the previous example is sending to Data Activator at ~17 EPS (1Mb per minute). Two properties are extracted from the events, with 3 rules/triggers watching for different conditions.



Total = 1.5647 CU/hr

What to use when

- If predictability important -> Pre-purchase
- If complicated cost splitting -> many small F
- If need for more power -> F+
- How to calculate SKU needs?
- Coming soon: Fabric SKU calculator

Usage patterns

- No Premium capacity yet?
- Existing Premium customers – do all your reports need to be on Premium?
- One big shared capacity or several smaller?
- Fabric as front of a larger ecosystem?

Feedback

Johan Ludvig Brattås

Director, Deloitte

/johanludvig

@intoleranse

jbrattas@deloitte.com

GitHub

Chronic volunteer

Co-organizer – DataSaturday Oslo
President – MDPUG Oslo
Frequent voulenteer in general

When not geeking out over new tech

Teaching coeliacs how to bake gluten free
Baking
Hiking
Gardening