

Designing and managing a cost-effective data platform in Azure Synapse Analytics



Mathias Halkjær

Principal Architect
- Data & Analytics
Fellowmind



@HalkjaerM



[linkedin.com/in/mhalkjaer](https://www.linkedin.com/in/mhalkjaer)



www.fluxbi.com



Erwin de Kreuk

Principal Consultant - Data & AI
InSpark



@ErwinDeKreuk



[linkedin.com/in/ErwinDeKreuk](https://www.linkedin.com/in/ErwinDeKreuk)



[ErwinDeKreuk.com](https://www.ErwinDeKreuk.com)

github.com/edkreuk



Erwin de Kreuk

Principal Consultant - Lead Data & AI - InSpark



www.linkedin.com/in/erwindekrek



@erwindekrek



linkedin.com/in/erwindekrek



erwindekrek.com



github.com/edkrek

Erwin de Kreuk
Principal Consultant



Microsoft®
Most Valuable
Professional



MATHIAS HALKJÆR

Principal architect, Data & Analytics, **Fellowmind**

Data enthusiast with a great passion for the synergy between data analytics and cloud data platforms. Principal Architect for Fellowmind working with building up our clients' data capabilities, implementing data platform technologies and all-in-all maximize the value they get out of their data.

- **Microsoft MVP**, Data Platform
- **MSc in Product Development and Innovation**
- **Certified Associate in Project Management (CAPM)®**
- **Sergeant**, Royal Danish Air Force
- Parent, Foodie, Casual Gamer, Tech-aficionado

FAVORITE TOOLS:



www.linkedin.com/in/mhalkjaer

Mathias Halkjær
Principal Architect



Microsoft®
Most Valuable
Professional





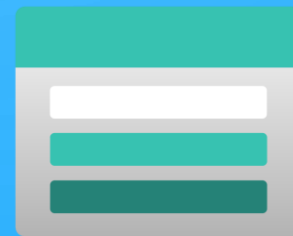
WHY ?



Today's menu:

Agenda

1. Storage
2. Compute
3. Pipeline Execution
4. Reservation
5. Microsoft Fabric
6. Architecture
7. Managing costs





01

STORAGE



An abstract graphic on the left side of the slide, consisting of a dark grey background with a network of thin, light blue lines. These lines form a complex, branching pattern that resembles a circuit board or a neural network, with several small circles at the end of the lines.

01

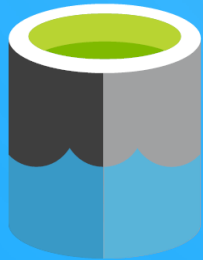
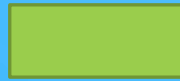
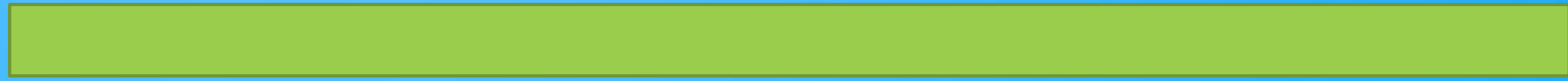
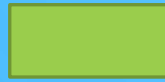
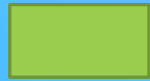
17 MAY 2024

STORAGE



Storage

- Data Lake
- Dedicated SQL Pools
- Azure SQL
- Fabric Storage



Data Lake

Price:
€1.6 - €18.4



Dedicated SQL Pool

Price:
€21.54



Azure SQL

Price:
€210.0



Storage accounts (Data Lakes)

Architecture (E.g., Medallion Architecture)

Tiers

- Premium
- Hot
- Cool
- Cold (preview)
- Archive

Live Cycle Management

- Storage/Read/Write
- Data Lake Gen 2 versus blob storage
- Environments (dev/test/pre-prod/prod)

Redundancy

- LRS
- ZRS
- GRS

Premium	Hot	Cool	Cold	Archive
Unlimited - \$0.18 per GB	First 51,200 GB / month - \$0.019 per GB Next 460,800 GB / month - \$0.018 per GB Over 512,000 GB / month - \$0.017 per GB	Unlimited - \$0.01 per GB	Unlimited - \$0.00443 per GB	Unlimited - \$0.002 per GB



02

17 MAY 2024

COMPUTE

CHOOSING COMPUTE



Serverless

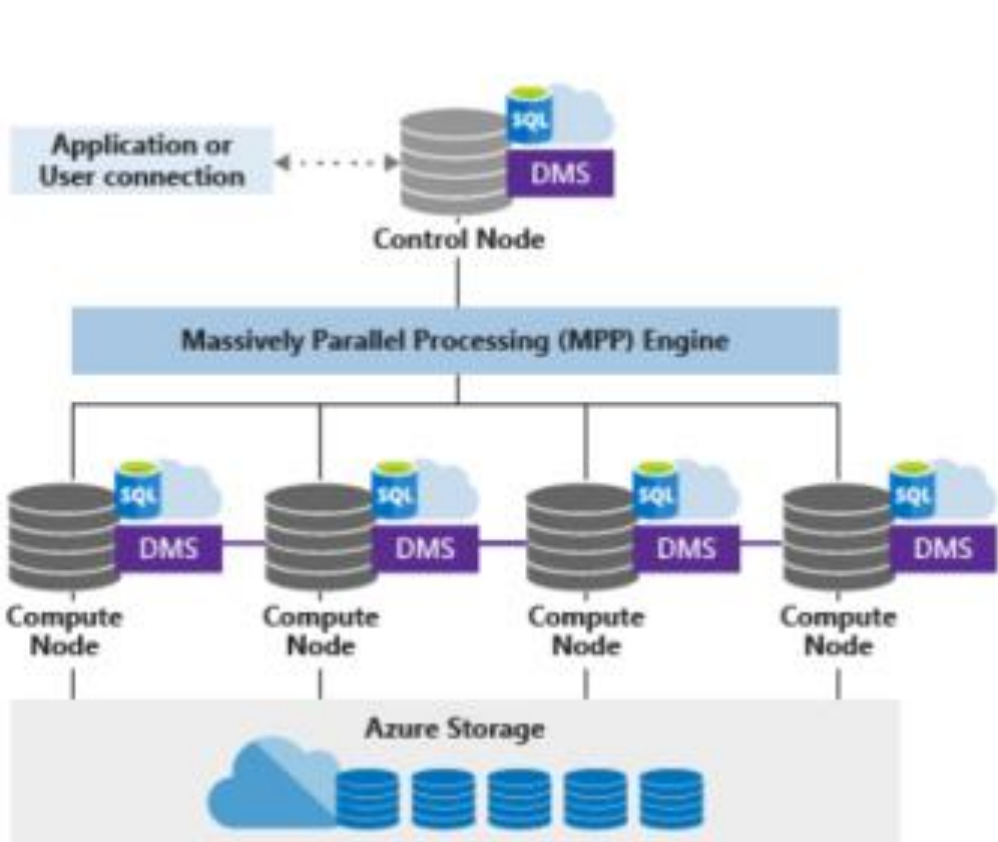


Dedicated

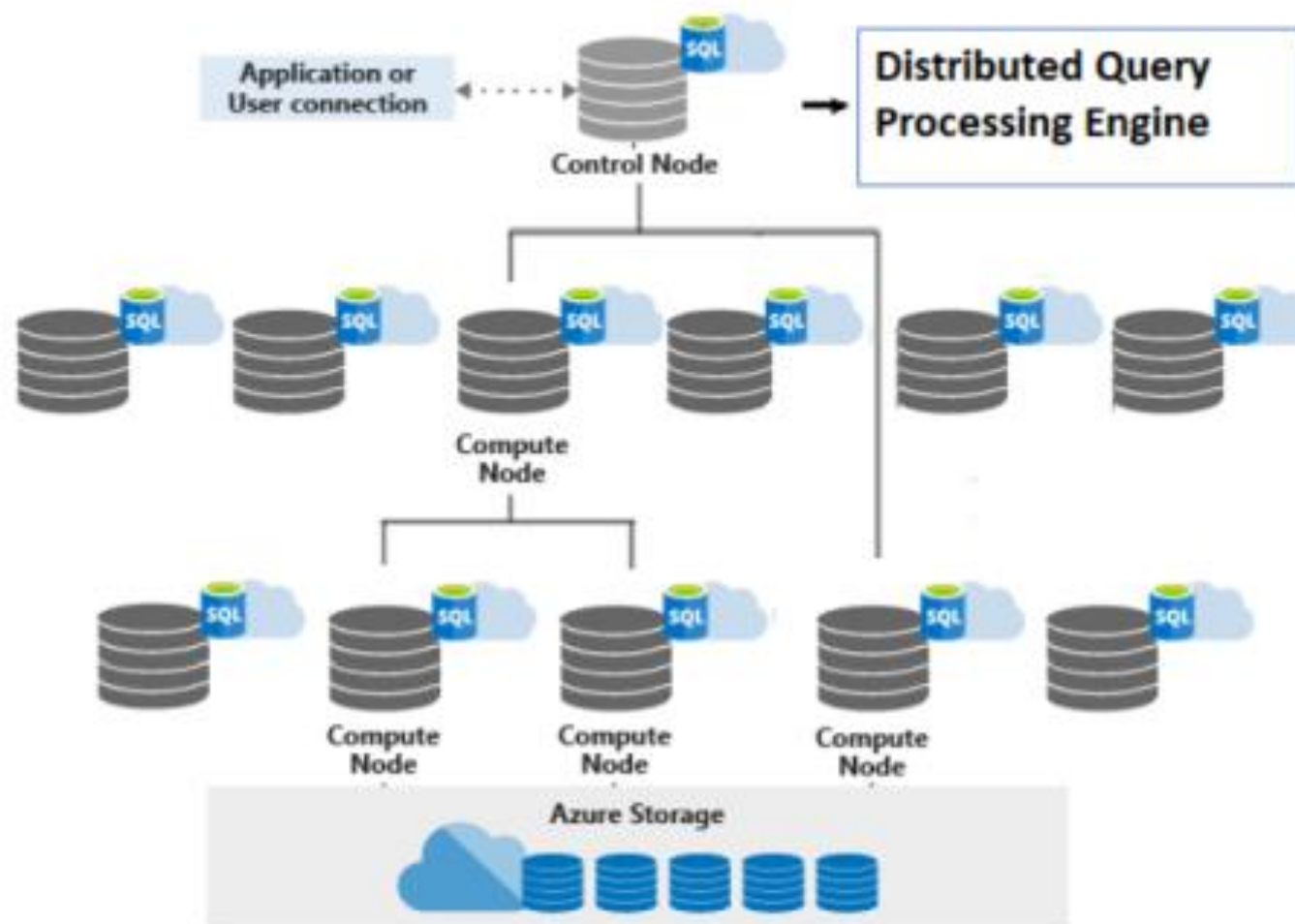


Spark

Dedicated SQL pool



Serverless SQL pool



SERVERLESS SQL POOL

- Partition (Pruning)
- Optimize queries
- Vacuum/optimize
- Consuming Serverless from other applications (e.g. Power BI)
- Watch out for DirectQuery or similar live connections



Serverless

DEDICATED SQL POOL

- Minimum 1 TB, 60 million rows
- Consider serverless
- No small updates

- Data storage is charged at the rate of €20.827 per TB per month of data stored (€0.03/1 TB/ includes the size of your data warehouse and 7-days of incremental snapshot storage.

- [Pricing - Azure Synapse Analytics | Microsoft Azure](#)

Data storage and snapshots

Data storage is charged at the rate of €20.827 per TB per month of data stored (€0.03/1 TB/hour). Data storage includes the size of your data warehouse and 7-days of incremental snapshot storage.



Dedicated

Service Level	DWU	Pay as you go
DW100c	100	€998,144/month
DW200c	200	€1,996,288/month
DW300c	300	€2,994,432/month
DW400c	400	€3,992,575/month
DW500c	500	€4,990,719/month
DW1000c	1000	€9,981,438/month
DW1500c	1500	€14,972,156/month

SPARK

- Run notebooks (or mapping data flows) in a managed synapse spark environment

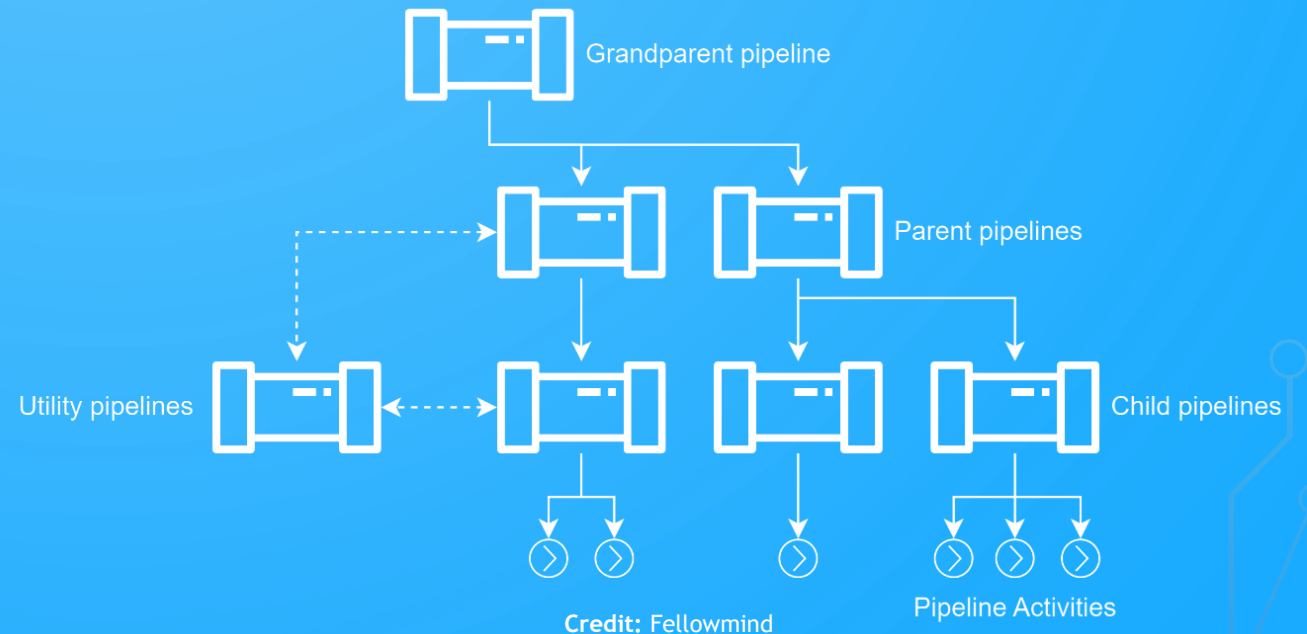


Spark

Type	Price	Free Quantity
Memory Optimized	€0.145 per vCore-hour	120 free vCore-hours per month until June 30, 2023*
GPU accelerated (public preview)	€0.158 per vCore-hour	N/A

RESOURCE SCALING

- What is resource scaling?
- Different methods
- Overall orchestration



An abstract graphic on the left side of the slide, consisting of a dark grey vertical band. Within this band, there is a complex network of thin, light blue lines that resemble a circuit board or a neural network. These lines are interconnected with small, light blue circles of varying sizes, some of which are positioned at the ends of the lines, creating a sense of connectivity and flow.

03

17 MAY 2024

EXECUTION



Pricing overview

Type	Azure Hosted Managed VNET Price	Azure Hosted Price	Self Hosted Price
Orchestration Activity Run	€0.937 per 1,000 runs	€0.937 per 1,000 runs	€1.405 per 1,000 runs
Data Movement	€0.235/DIU-hour	€0.235/DIU-hour	€0.094/hour
Pipeline Activity Integration Runtime	€0.937/hour (Up to 50 concurrent pipeline activities)	€0.005/hour	€0.002/hour
Pipeline Activity External Integration Runtime	€0.93633/hour (Up to 800 concurrent external pipeline activities)	€0.00024/hour	€0.00010/hour

ANY QUESTIONS?



An abstract graphic on the left side of the page, consisting of a dark grey vertical band. Within this band, there is a complex network of thin, light blue lines that resemble a circuit board or a neural network. These lines are interconnected with small, light blue circles of varying sizes, some of which are positioned at the ends of the lines, creating a sense of connectivity and flow.

04

17 MAY 2024

RESERVATIONS

RESERVATIONS

- Reserved Capacity (SQL POOL)
- Choose the right size
- Purchase are made per 100 cDWU
- Can be used on multiple pools
- 1 year 37% savings / 3 years 65% saving
- Use more=> pay as you go pricing
- Running 24/7 or smart scaling in DTAP
- You can refund reservations, but scope cannot exceed \$50,0000

	DW1000C	DW2000C
Pay as you go	€13.674/hour	€27.347/hour
1 year reserved	€8.6138/hour	€17.2276/hour
3 year reserved	€4.7857/hour	€9.5713/hour

RESERVATIONS

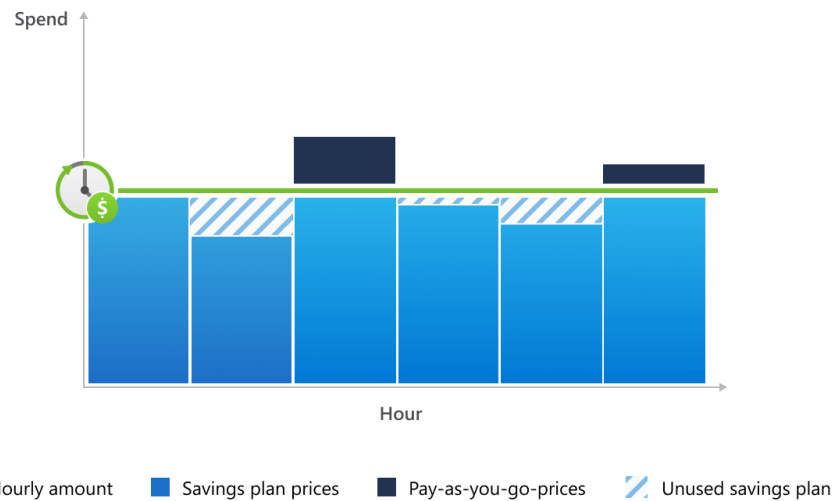
- Reserved Capacity (SQL POOL)
- Pre-Purchase Plans (SYNAPSE)
- Synapse Commit Units (SCUs)
- Upfront payment, 12 month period
- Up to 28% savings
- No refund

Tier	SCU's	Discount %	Price
1	5.000	6%	€4,255.898
3	24.000	11%	€19,341.695
6	360.000	28%	€234,708.200

RESERVATIONS

- Reserved Capacity (SQL POOL)
- Pre-Purchase Plans (SYNAPSE)
- Reserved Capacity (SHIR)

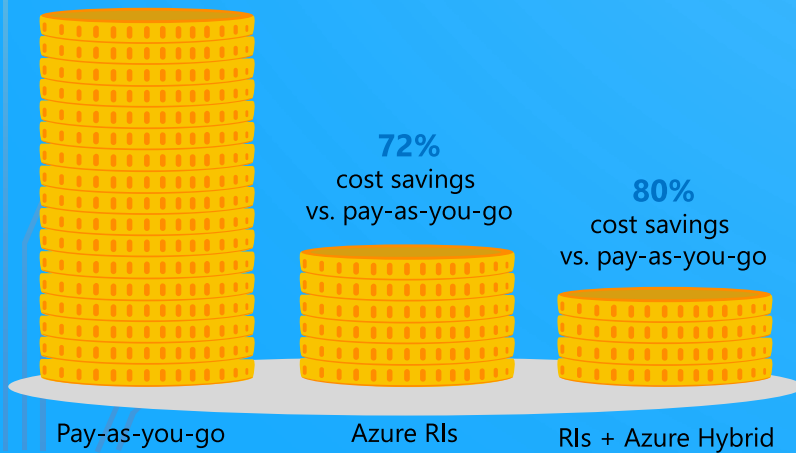
- Saving plans 1 or 3 years
- If your usage is \leq to your hourly amount:
 - Get lower, savings plan prices.
 - Pay the full hourly amount, even if usage is less.
 - Incur no additional costs—usage is covered by the plan.
- If your usage $>$ your hourly amount:
 - Usage up to your hourly amount is billed at lower prices and included in the cost of your plan.
 - Additional usage is billed at pay-as-you-go prices and invoiced separately.



EXISTING LICENSE

- Reserved Capacity (SQL POOL)
- Pre-Purchase Plans (SYNAPSE)
- Reserved Capacity (SHIR)

Save up to **80%** with RIs and Azure Hybrid Benefit



- Reserved Instance 1 or 3 years
 - Based on Cores
- Azure Hybrid (use your own license)
- You can also refund a reservation, up to \$50,000 USD in a 12 month rolling window

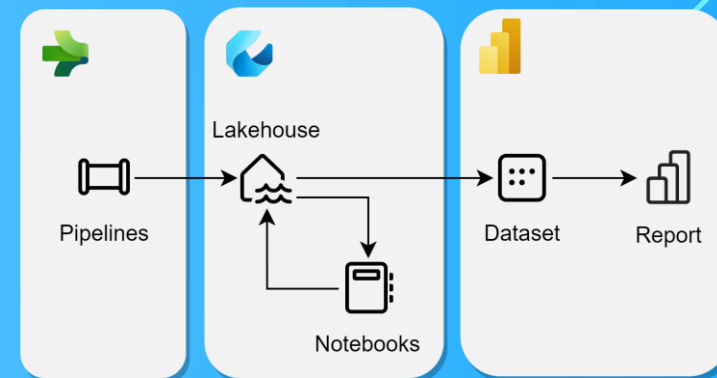


05

17 MAY 2024

MICROSOFT FABRIC

Capacity & pricing



Microsoft Fabric offers a variety of purchasable capabilities divided into SKUs, each providing unique computing power quantified by Capacity Units (CU).

Fabric features two SKU types:

- **Azure** – Billed per second with no commitment.
- **Microsoft 365** – Billed monthly or yearly, with a monthly commitment

SKU	Capacity unit (CU)	Pay-as-you-go (hourly)	Pay-as-you-go (monthly)
F 2	2	\$0.36	\$262.80
F 4	4	\$0.72	\$525.60
F 8	8	\$1.44	\$1,051.20
F 16	16	\$2.88	\$2,102.40
F 32	32	\$5.76	\$4,204.80
F 64	64	\$11.52	\$8,409.60
F 128	128	\$23.04	\$16,819.20
F 256	256	\$46.08	\$33,638.40
F 512	512	\$92.16	\$67,276.80
F 1024	1024	\$184.32	\$134,553.60
F 2048	2048	\$368.64	\$269,107.20

Free report viewers

An abstract graphic on the left side of the page, consisting of a dark grey vertical band. Within this band, there is a complex pattern of thin, light blue lines that resemble a circuit board or a network diagram. These lines are interconnected with small, light blue circles of varying sizes, some of which are positioned at the ends of the lines, creating a sense of connectivity and flow.

06

17 MAY 2024

ARCHITECTURE

ARCHITECTURE

Azure Service Management

- Landing zones
- Policies & Platform management
- Etc.

Solution

- Storage
- Ingestion
- Transform
- Master data management
- Data Quality
- Serving
- Orchestration
- Archive & History

Core Azure Tenant

- Tenant administration
- Identity & Access (Azure AD) Etc.

Network

- VPN
- Network
- Private links
- Managed VPN
- Firewall
- Security, Encryption in transit

Reliability

- DevOps
- Back-up
- Governance
- SOP's

Illustration credit: **Fellowwind**



Adding more networking or security

Managed VNets

- Needs be configured during setup

Execution

Azure Managed Vnet

€ 280

Managed virtual network

Choose whether to set up a dedicated Azure Synapse-managed virtual network for your workspace. [Learn more](#)

Managed virtual network ⓘ

☐ Enable ☒ Disable

i To control public network access to your Synapse workspace, you must enable managed virtual network.



Adding more networking or security

Self-hosted Integration Runtime

- Needed to connect to on-prem data sources
- Required when Data Exfiltration protection is Enabled

= €490.96

Average per month
(€0.00 charged upfront)

Allow outbound data traffic only to approved targets ⓘ

☒ Yes ☐ No

ⓘ Private endpoints will be allowed to target resources in approved Azure AD tenants only. The Azure AD tenant of the current user will be included by default and is not listed below.



Adding more networking or security

Deny public Endpoints

- No Power BI connection => Power BI Data Gateway

= €490.96
Average per month
(€0.00 charged upfront)



public access allowed/**Denied**

public access allowed/**Denied**

Public network access

Choose whether to permit public network access to your workspace. You can modify the firewall rules after you enable this setting. [Learn more](#)

Public network access to workspace endpoints

☐ Enabled ☒ Disabled

i Disabling this setting will require you to use private endpoints to connect to your workspace. [Create a private endpoint](#)

▷ Azure Private Link
Enabled for the entire organization

▷ Block Public Internet Access
Enabled for the entire organization



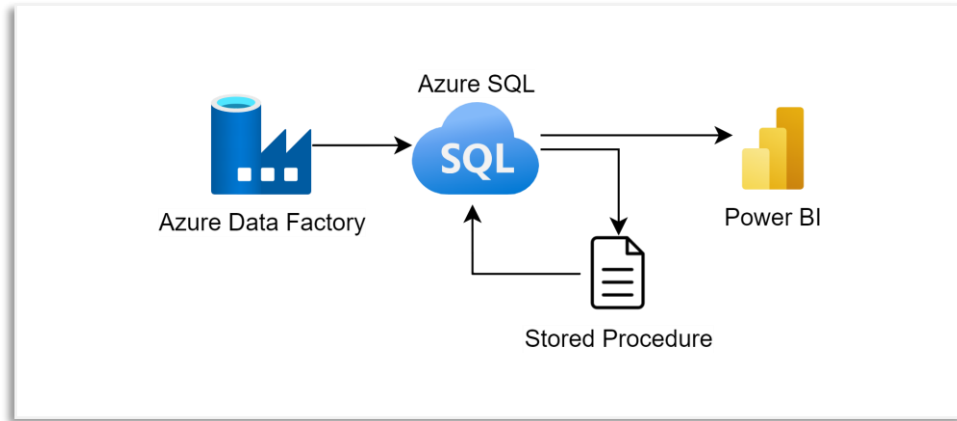
Storage/Compute

- Explain storage and compute decoupling

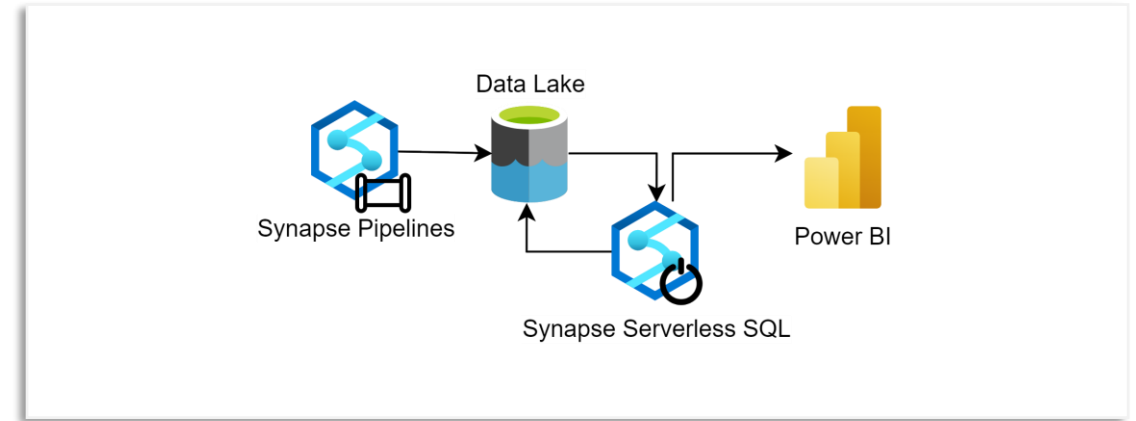


Architectural blueprint(s)

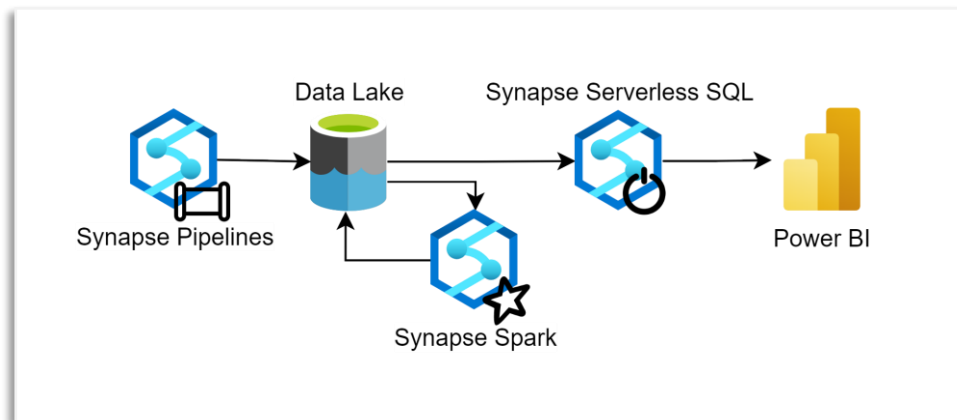
CLOUD DATA WAREHOUSE



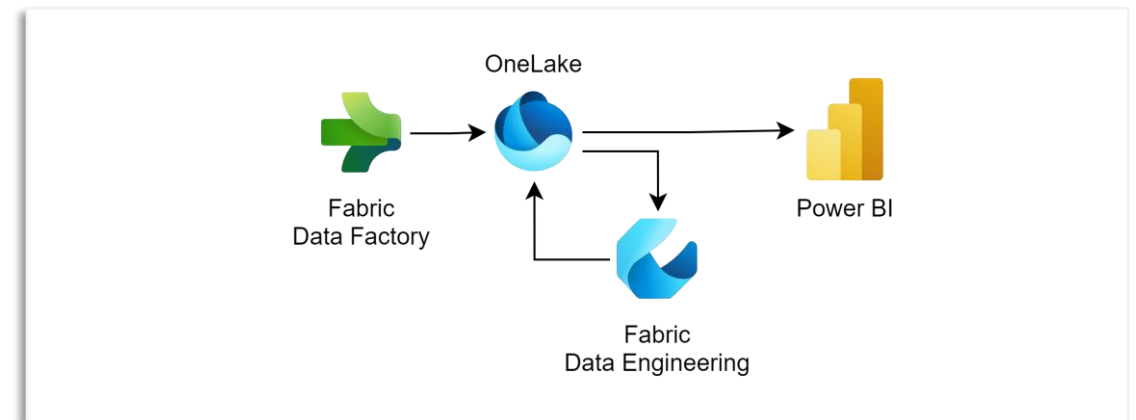
LOGICAL DATA WAREHOUSE



CLOUD LAKEHOUSE



FABRIC LAKEHOUSE



An abstract graphic on the left side of the slide, consisting of a dark grey vertical band. Within this band, there is a complex network of thin, light blue lines that resemble a circuit board or a neural network. These lines are interconnected with small, light blue circles of varying sizes, some of which are positioned at the ends of the lines, creating a sense of connectivity and flow.

07

17 MAY 2024

MANAGING COSTS



Managing

- Budgets
- Monitoring

ANY
QUESTIONS
LEFT?



The image features a solid blue background. In the four corners, there are decorative, light blue line art elements that resemble circuit traces or stylized neural network connections. These elements consist of thin lines that branch out and terminate in small circles. The top-left and bottom-left corners have more complex, multi-branching patterns, while the top-right and bottom-right corners have simpler, more linear patterns.

FEEDBACK FEEDBACK FEEDBACK

Thank you for attending!



Mathias Halkjær

Principal Architect - Data & Analytics
Fellowmind



@HalkjaerM



[linkedin.com/in/mhalkjaer](https://www.linkedin.com/in/mhalkjaer)



www.fluxbi.com



Erwin de Kreuk

Principal Consultant - Lead Data & AI
InSpark



@erwindekrek



[linkedin.com/in/erwindekrek](https://www.linkedin.com/in/erwindekrek)



erwindekrek.com



github.com/edkrek

