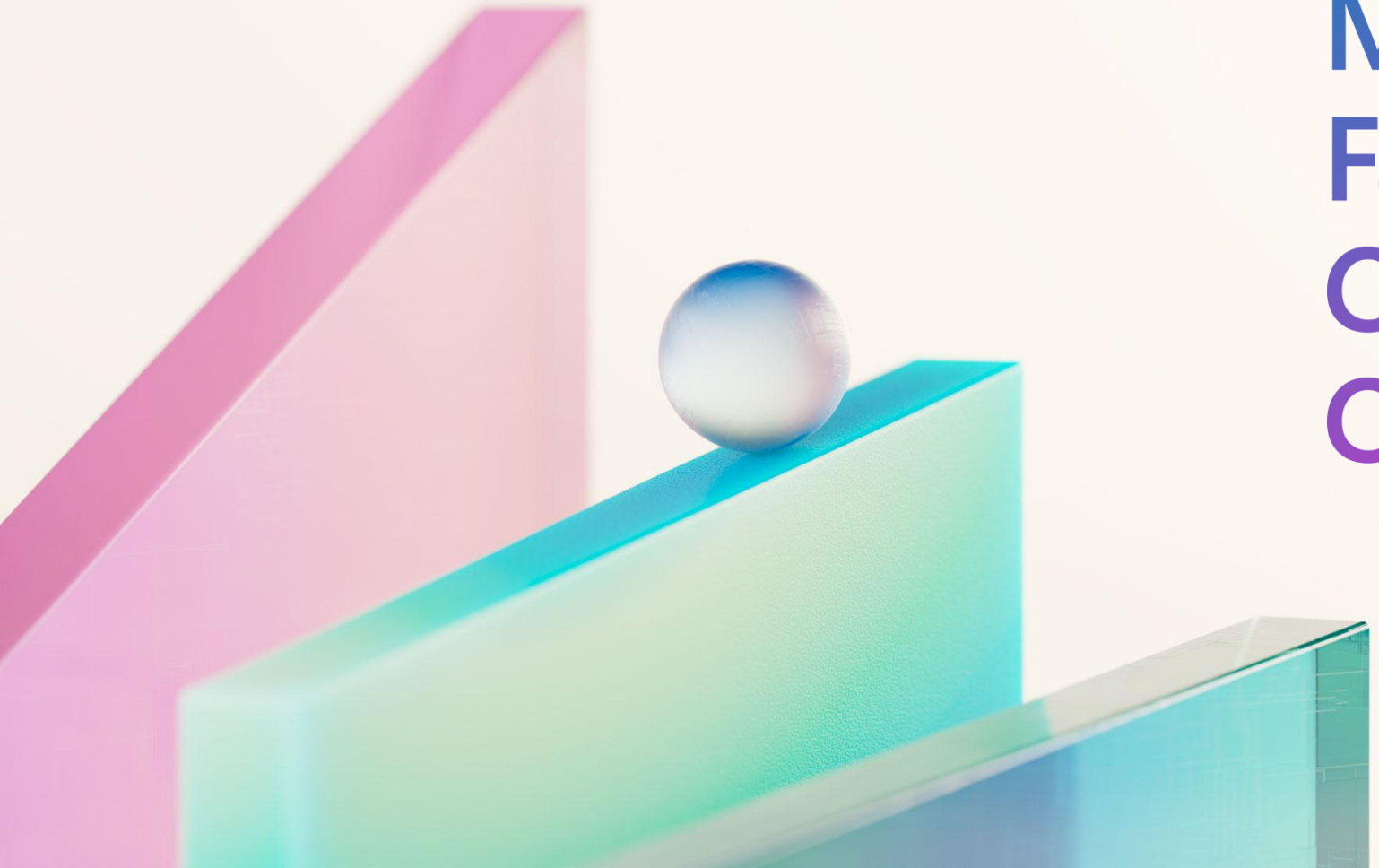
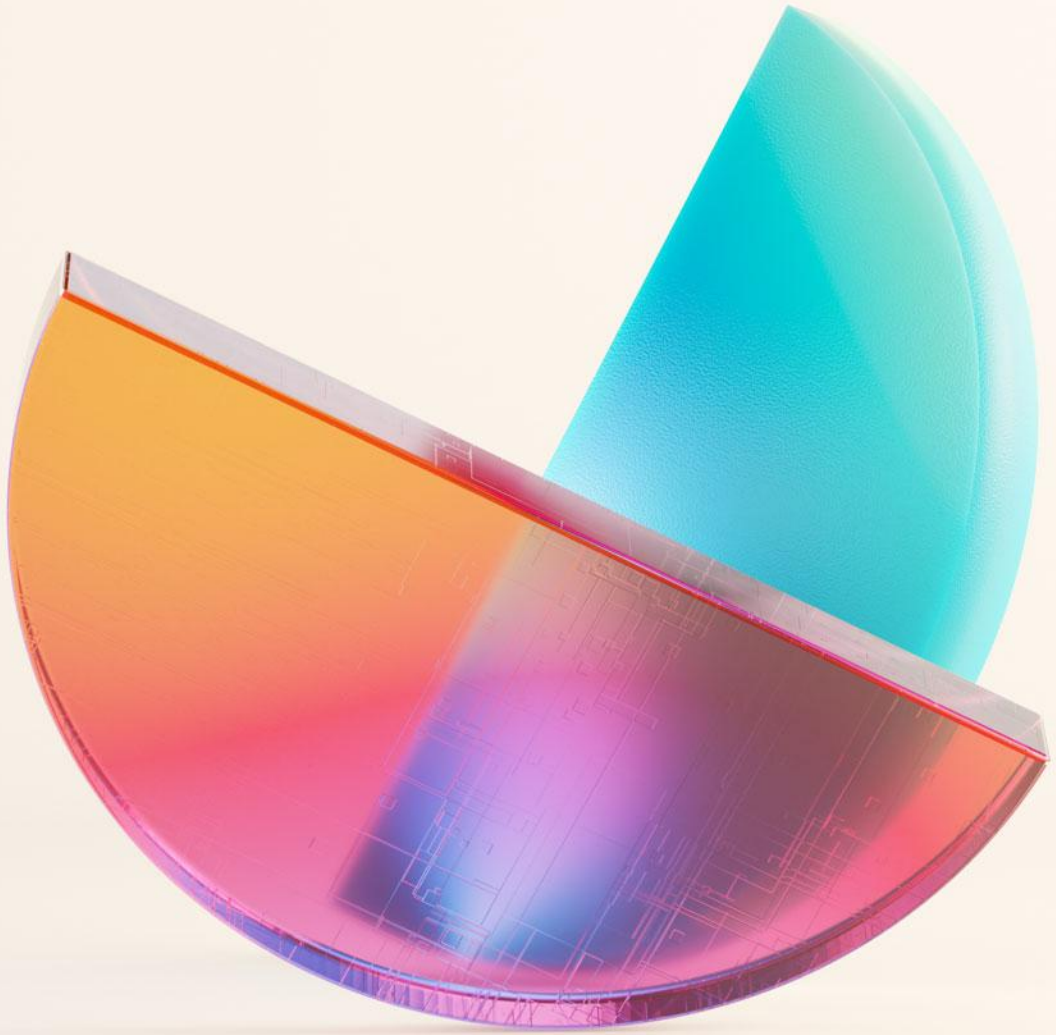
Microsoft



Microsoft Fabric Community Conference

Microsoft

Microsoft Fabric
Community Conference



Unlocking Synergy: Exploring Snowflake and Fabric Integration with Iceberg Tables & Mirroring

Matthias Nohl | b.telligent

Tim Spannagel | b.telligent

Your Speakers

Matthias Nohl

Strategic Partner Manager &
Management Consultant @b.telligent

linkedin.com/in/matthiasnohl

@mnohlimits

Tim Spannagel

Team Lead &
Principal Consultant @b.telligent

linkedin.com/in/tim-spannagel

@TSpannagel

Agenda

01
Introduction

02
Integration with Snowflake Mirroring

03
Snowflake Mirroring Demo

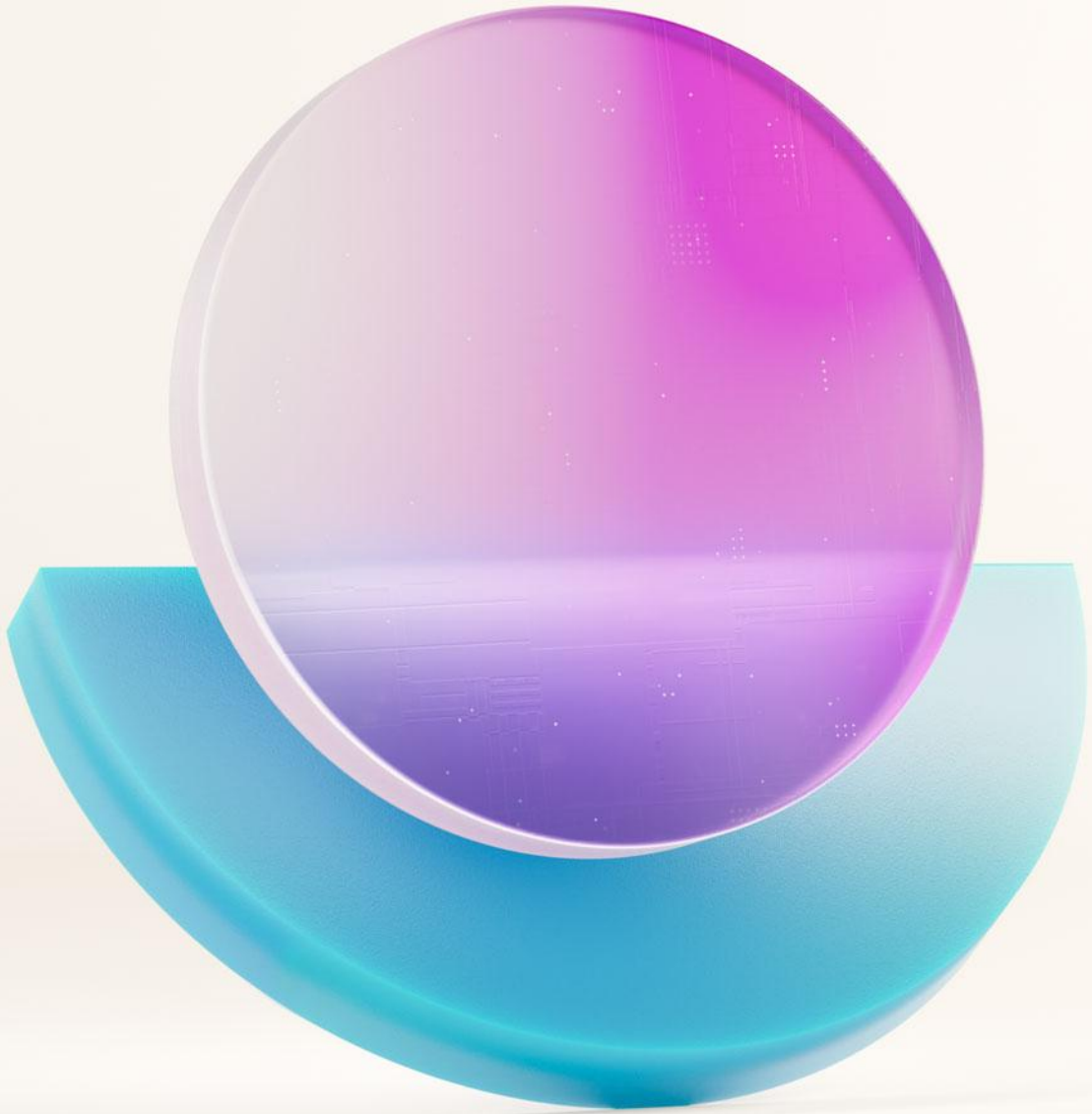
04
Integration with Snowflake Iceberg Tables

05
Snowflake Iceberg Tables Demo

06
What should be used and how?

07
Q&A

Microsoft Fabric
Community Conference



Introduction

Microsoft Fabric

Data Factory

Data Engineering

Data Warehouse

Data Science

Real-Time Intelligence

Power BI

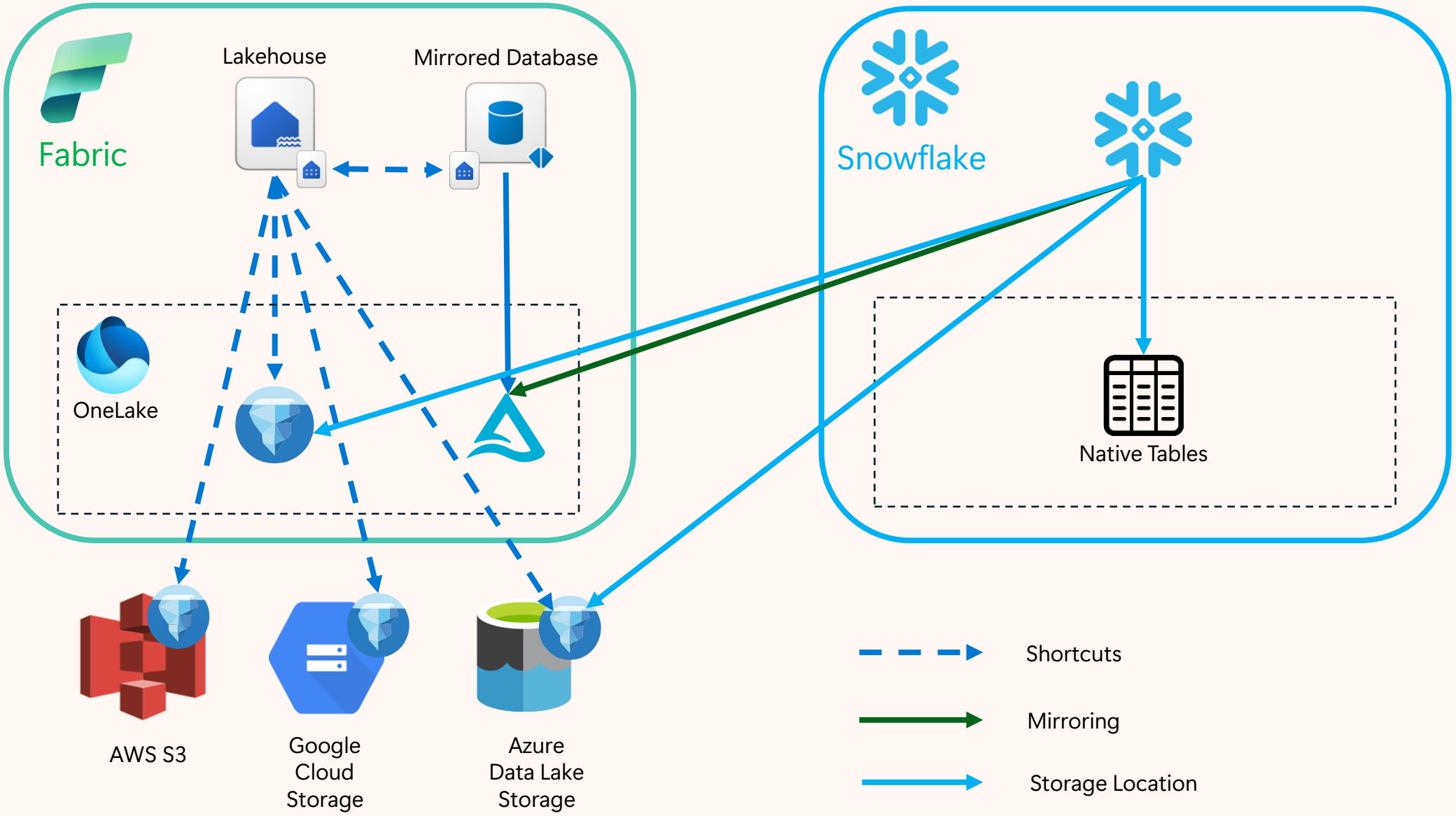
Partner & Industry workloads

Copilot in Fabric

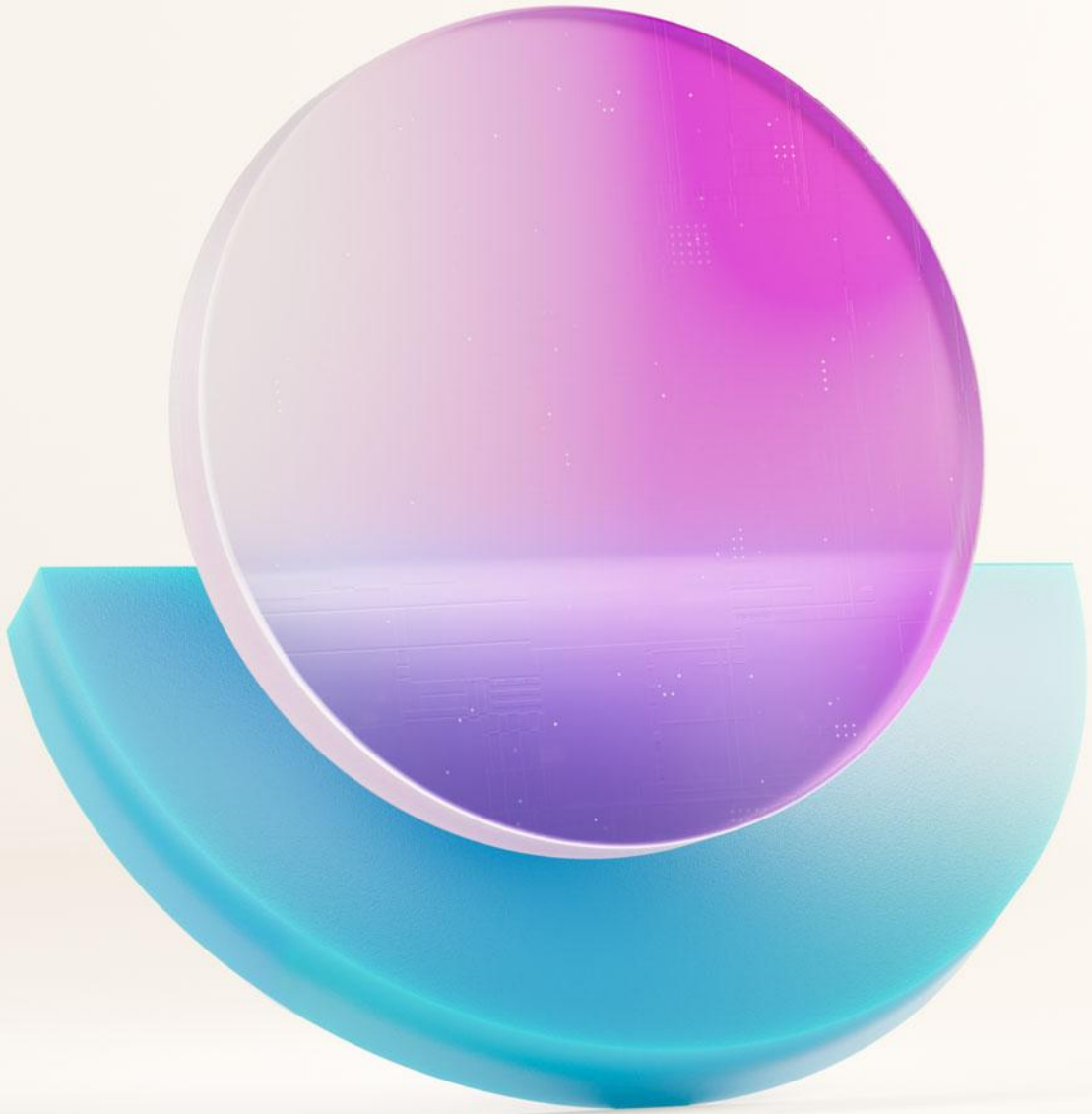
OneLake

Microsoft Purview

Integration Snowflake & Microsoft Fabric



Microsoft Fabric
Community Conference



Integration with Snowflake Mirroring

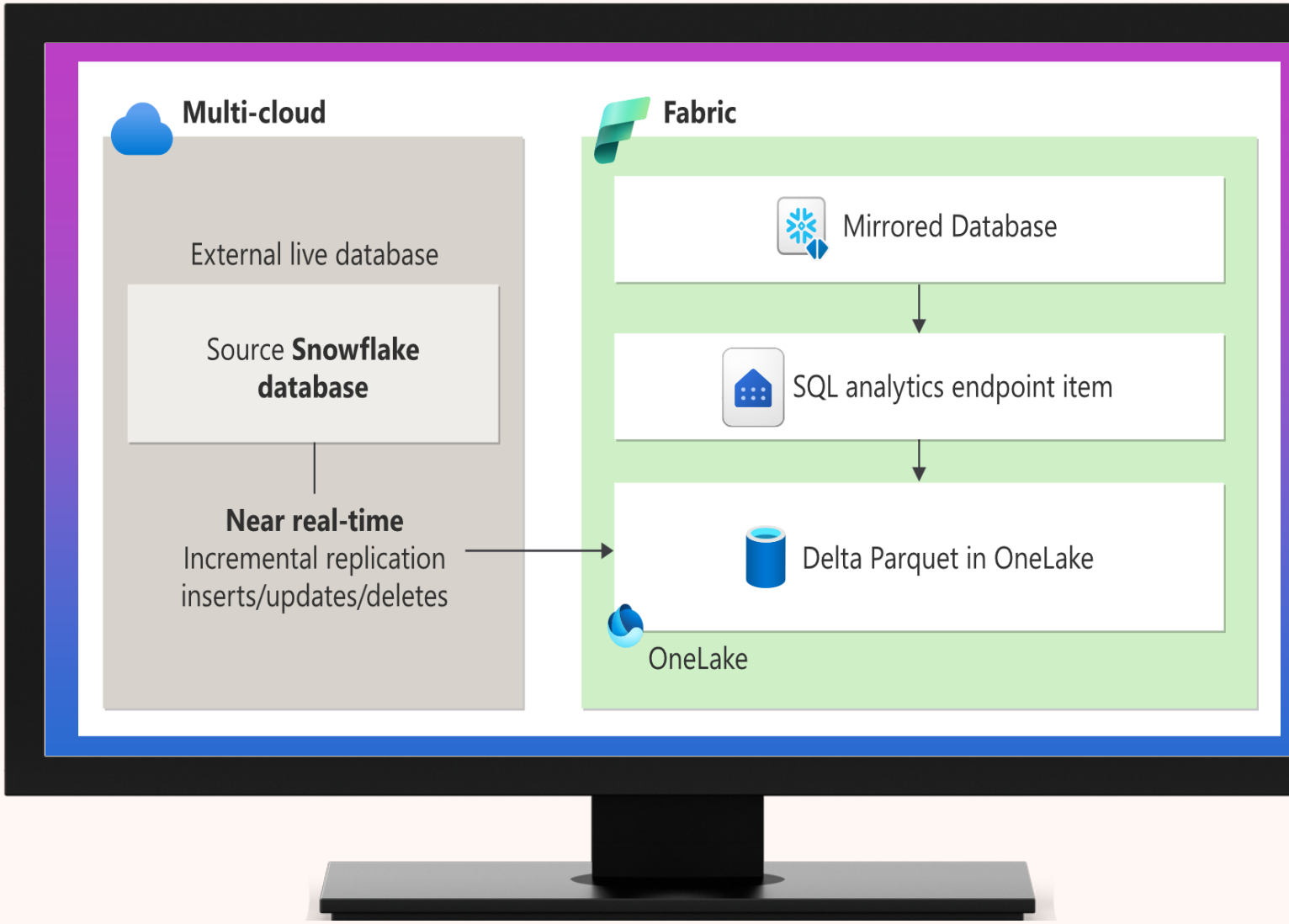
Mirroring Snowflake in Microsoft Fabric

Fabric items:

- Mirrored database
- A SQL analytics endpoint
- A default semantic model

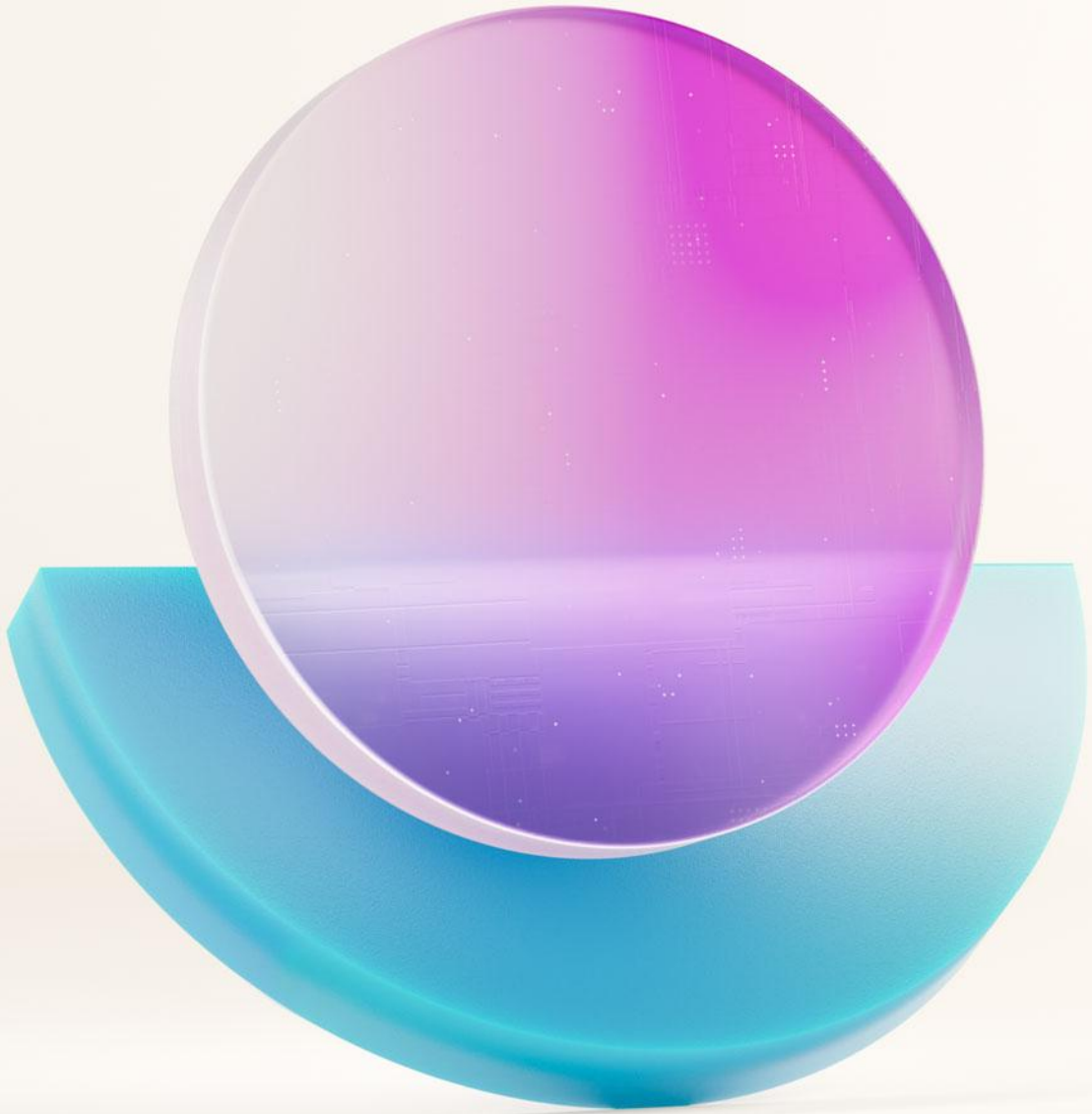
Cost considerations:

- Fabric compute used to replicate your data into Fabric OneLake is free. The Mirroring storage cost is free up to a limit based on capacity.
- Snowflake virtual warehouse compute charges
- Snowflake services compute charges



Source: Microsoft Fabric Mirrored Databases From Snowflake - Microsoft Fabric | Microsoft Learn

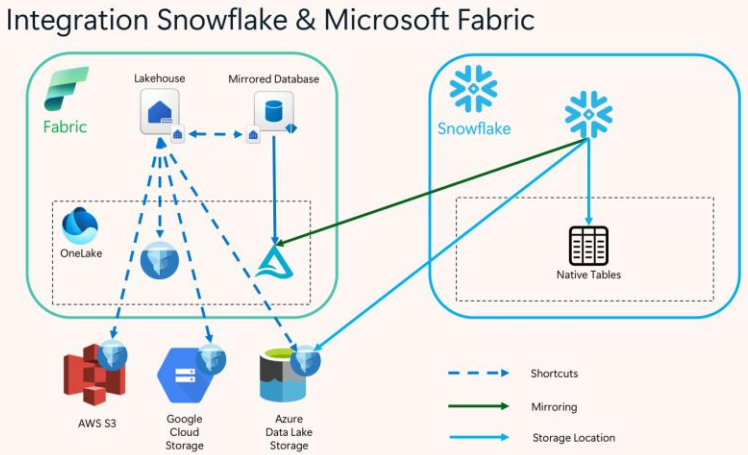
Microsoft Fabric
Community Conference



Snowflake Mirroring Demo

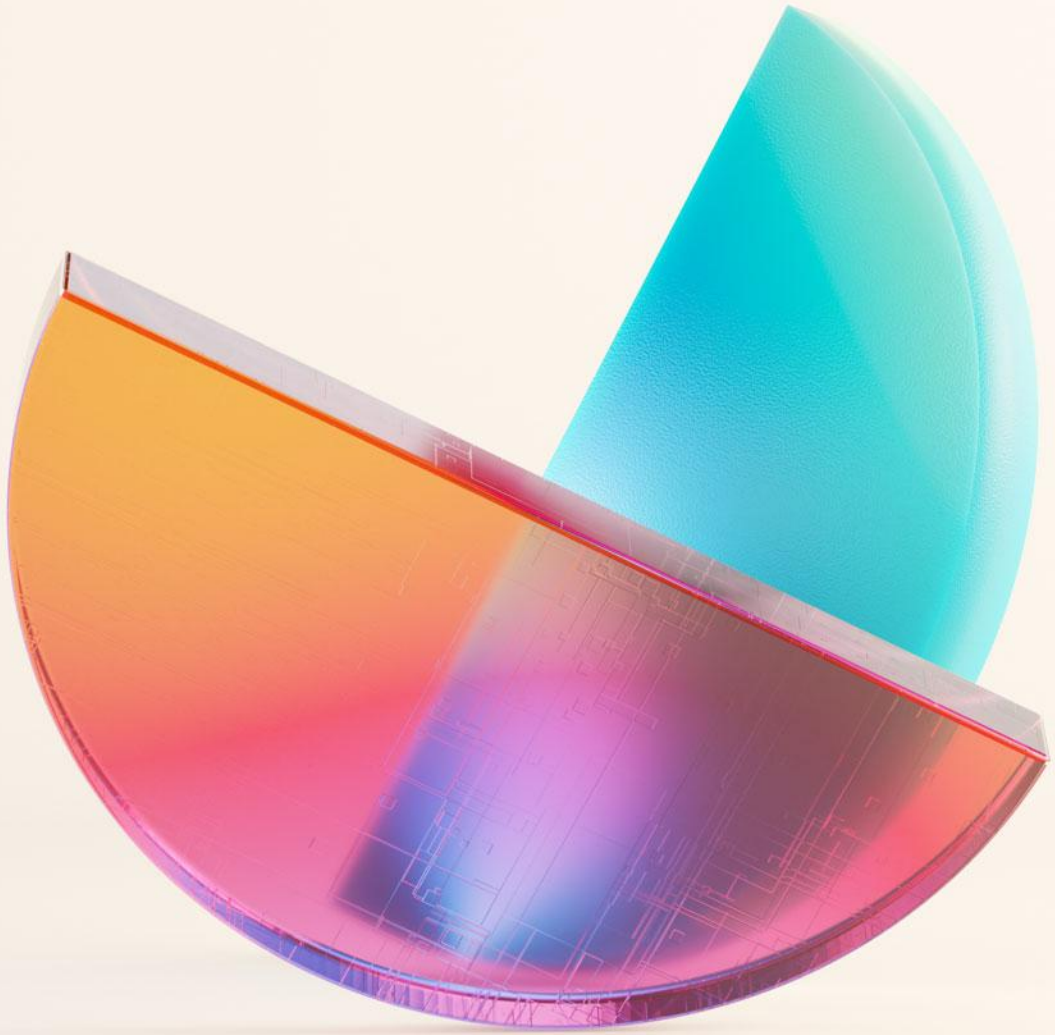
Step-by-Step

1) Create or use existing Warehouse for Compute in Snowflake
2) Create or use existing Database
3) Create Database Role in Snowflake / Grant Permissions
4) Create Functional Roles in Snowflake / Grant Permissions
5) Create Technical User in Snowflake / Grant Permissions
6) Create Tables
7) Enable Change Tracking on Tables
8) Create Fabric „Mirrored Snowflake“ item
9) Start Sync in Fabric
10) Build Semantic Model
11) Build Report



Microsoft

Microsoft Fabric
Community Conference



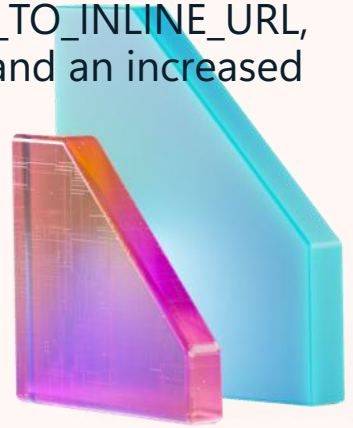
Limitations (1/2)

- **Database level limitations**
  - If there are no updates in a source table, the replicator engine starts to back off with an exponentially increasing duration for that table, up to an hour. The same can occur if there's a transient error, preventing data refresh. The replicator engine will automatically resume regular polling after updated data is detected.
  - Only replicating native tables are supported. Currently, External, Transient, Temporary, Dynamic tables aren't supported.
  - Source schema hierarchy is replicated to the mirrored database. For mirrored databases created before this feature enabled, the source schema is flattened, and schema name is encoded into the table name. If you want to reorganize tables with schemas, recreate your mirrored database.
  - Mirroring supports replicating columns containing spaces or special characters in names (such as , ; { } ( ) \n \t =). For tables under replication before this feature enabled, you need to update the mirrored database settings or restart mirroring to include those columns.
  - The maximum number of tables that can be mirrored into Fabric is 500 tables. Any tables above the 500 limit currently can't be replicated.
  - If you select Mirror all data when configuring Mirroring, the tables to be mirrored over will be determined by taking the first 500 tables when all tables are sorted alphabetically based on the schema name and then the table name. The remaining set of tables at the bottom of the alphabetical list will not be mirrored over.
  - If you unselect Mirror all data and select individual tables, you're prevented from selecting more than 500 tables.

Source: Limitations for Microsoft Fabric Mirrored Databases From Snowflake - Microsoft Fabric | Microsoft Learn

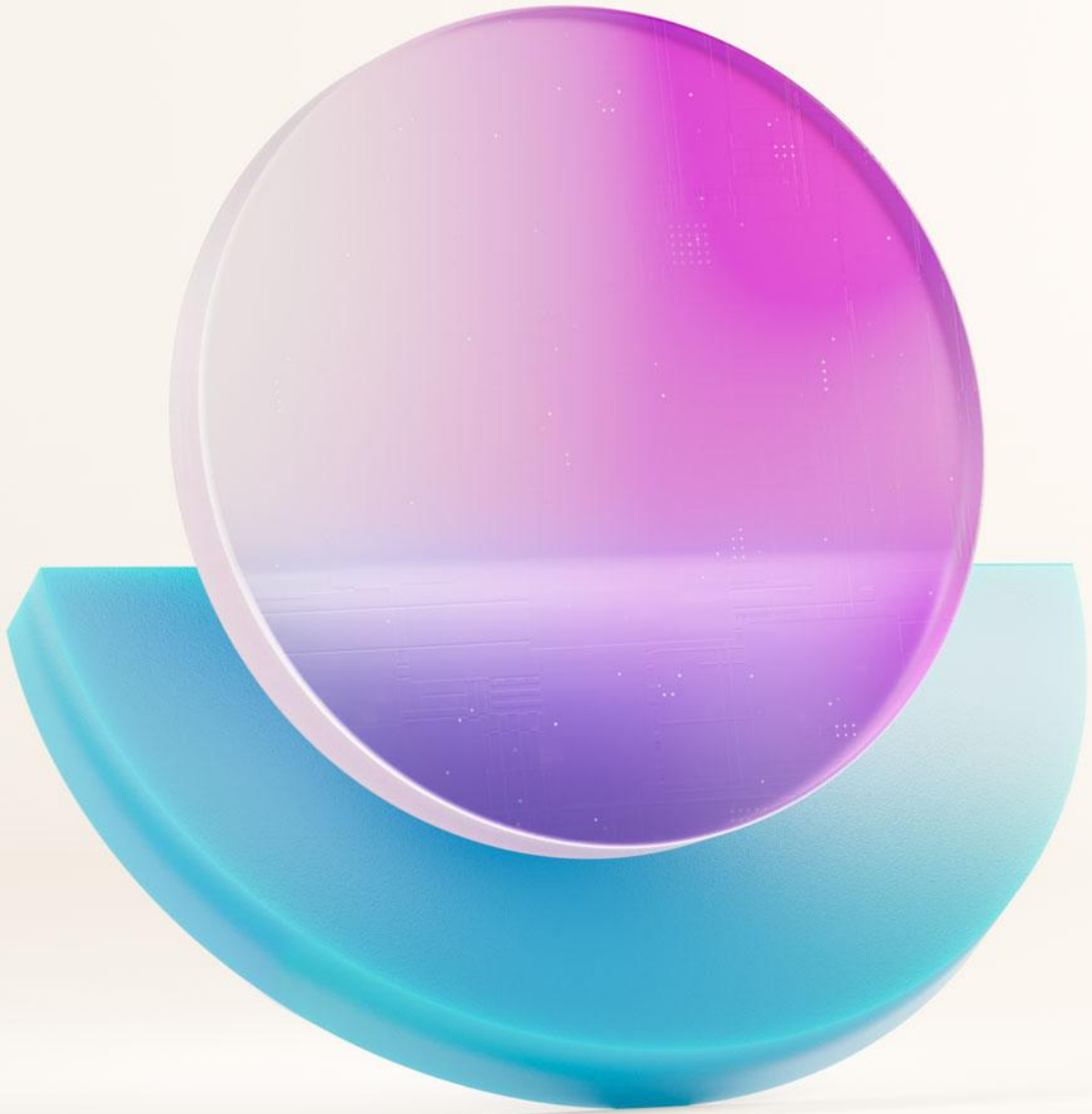
Limitations (2/2)

- **Network and firewall**
  - Mirroring does not support Snowflake instances behind a virtual network or private networking
- **Security**
  - Authentication via username/password and Entra
  - Sharing recipients must be added to the workspace. To share a dataset or report, first add access to the workspace with a role of admin, member, reader, or contributor.
- **Performance**
  - If you're changing most the data in a large table, it's more efficient to stop and restart Mirroring. Inserting or updating billions of records can take a long time.
  - Some schema changes aren't reflected immediately. Some schema changes need a data change (insert/update/delete) before schema changes are replicated to Fabric.
  - When mirroring data from Snowflake to a customer's OneLake, we usually stage the data to improve performance. However, if data exfiltration from Snowflake through inline URL is disabled via PREVENT_UNLOAD_TO_INLINE_URL, direct read from Snowflake could be required. This approach can lead to slower replication times and an increased risk of connection timeouts, particularly for large datasets.
- **Region availability limitation**



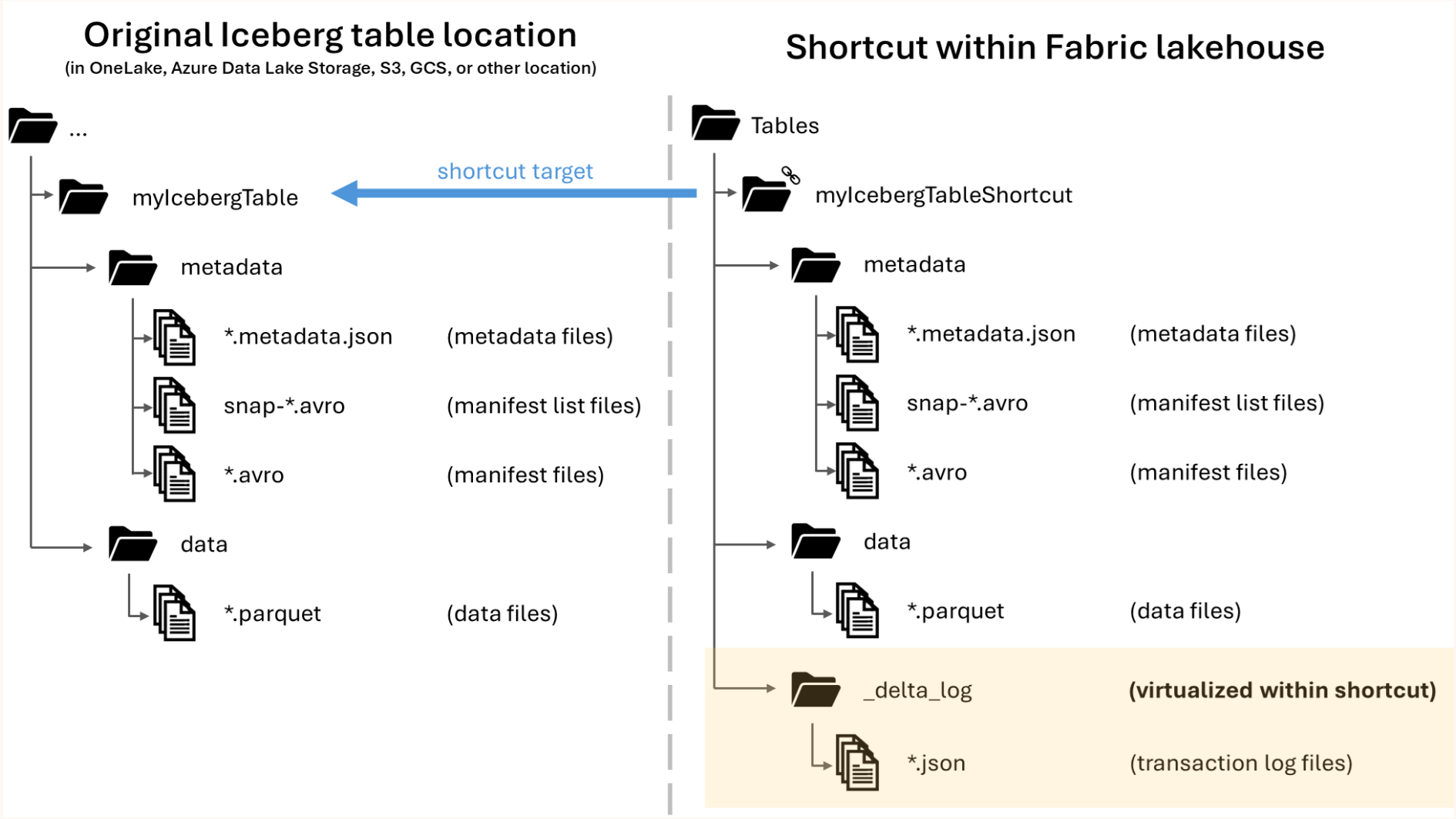
Source: Limitations for Microsoft Fabric Mirrored Databases From Snowflake - Microsoft Fabric | Microsoft Learn

Microsoft Fabric
Community Conference



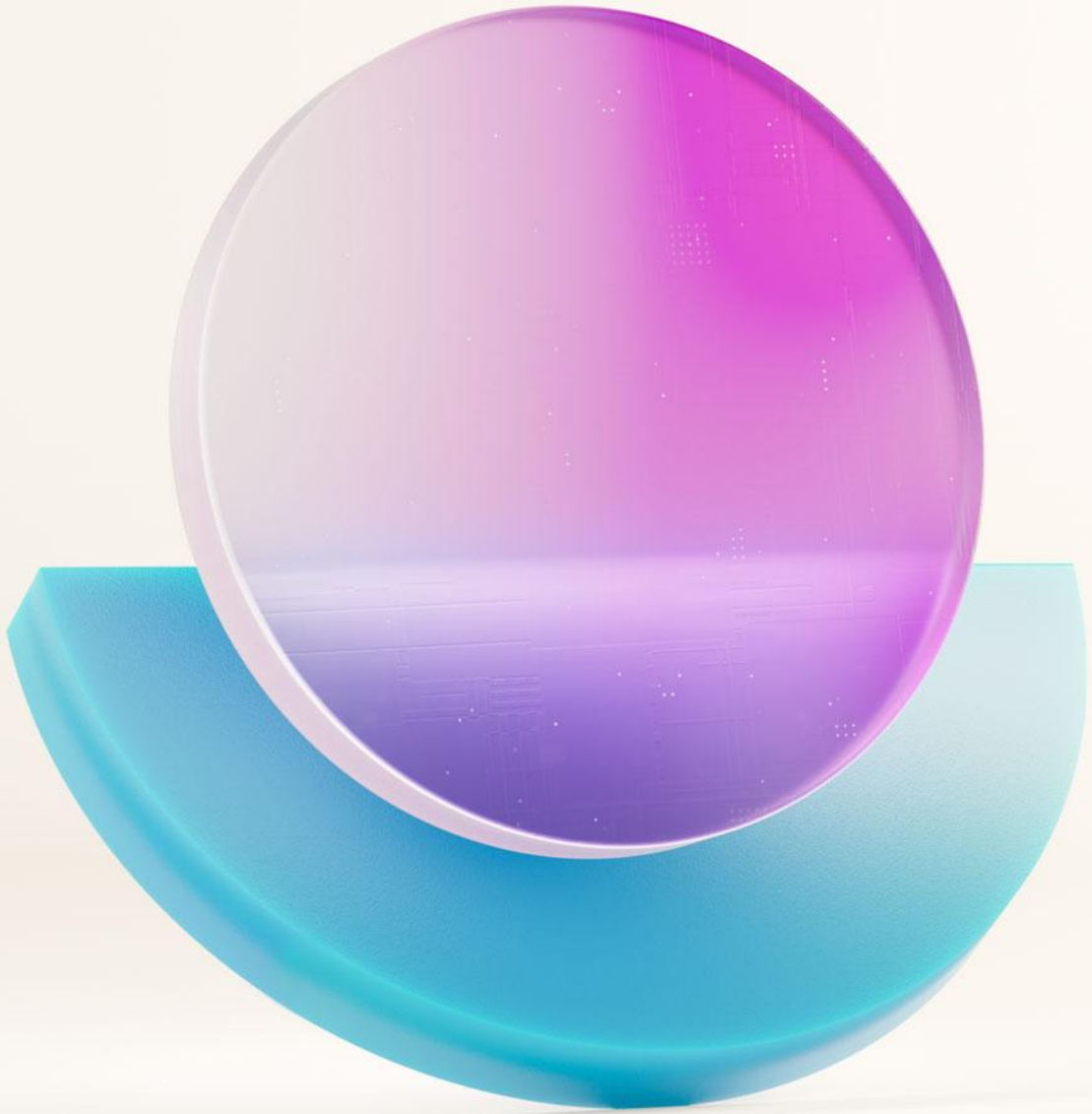
Integration with Snowflake Iceberg Tables

Iceberg Table Integration via Fabric Shortcuts in a Lakehouse



Source: Create shortcuts to Iceberg tables - Microsoft Fabric | Microsoft Learn

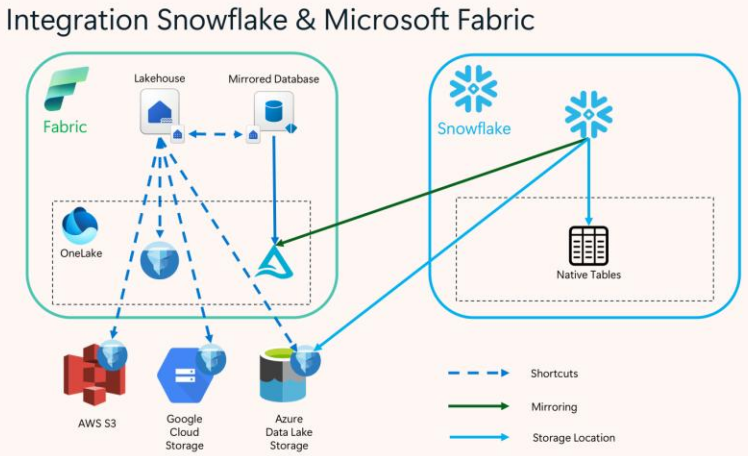
Microsoft Fabric
Community Conference



Snowflake Iceberg Tables Demo

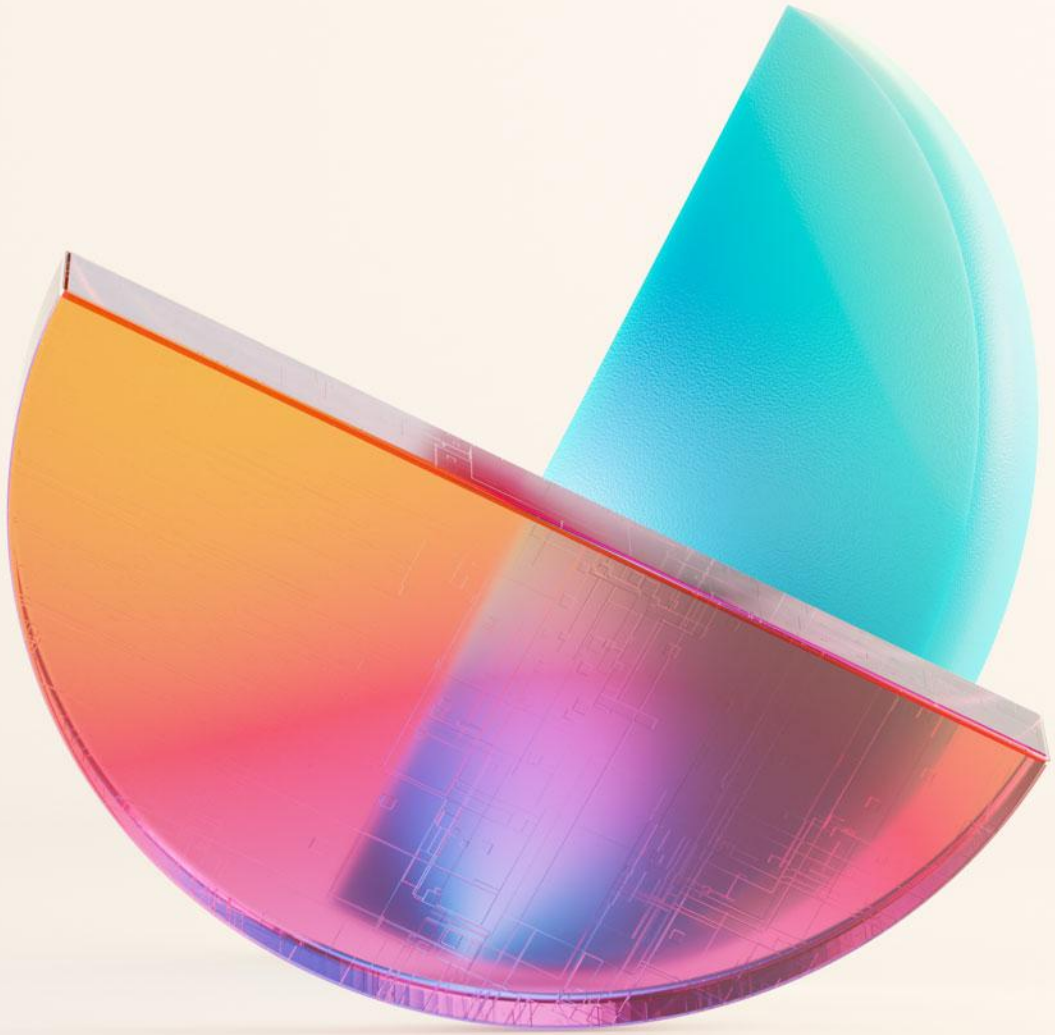
Step-by-Step

1) Create External Volume (ADLS Gen2 / OneLake) in Snowflake
2) Create Service Principal in Azure via Consent URL
3) Add Service Principal as „Storage Blob Data Contributor“ / Fabric Workspace „Contributor“
   OneLake: Allow API access for this SP
4) Write Iceberg Tables within Snowflake
5) Create Shortcut to ADLS Gen2 / OneLake
6) Build Semantic Model
7) Build Report



Microsoft

Microsoft Fabric
Community Conference

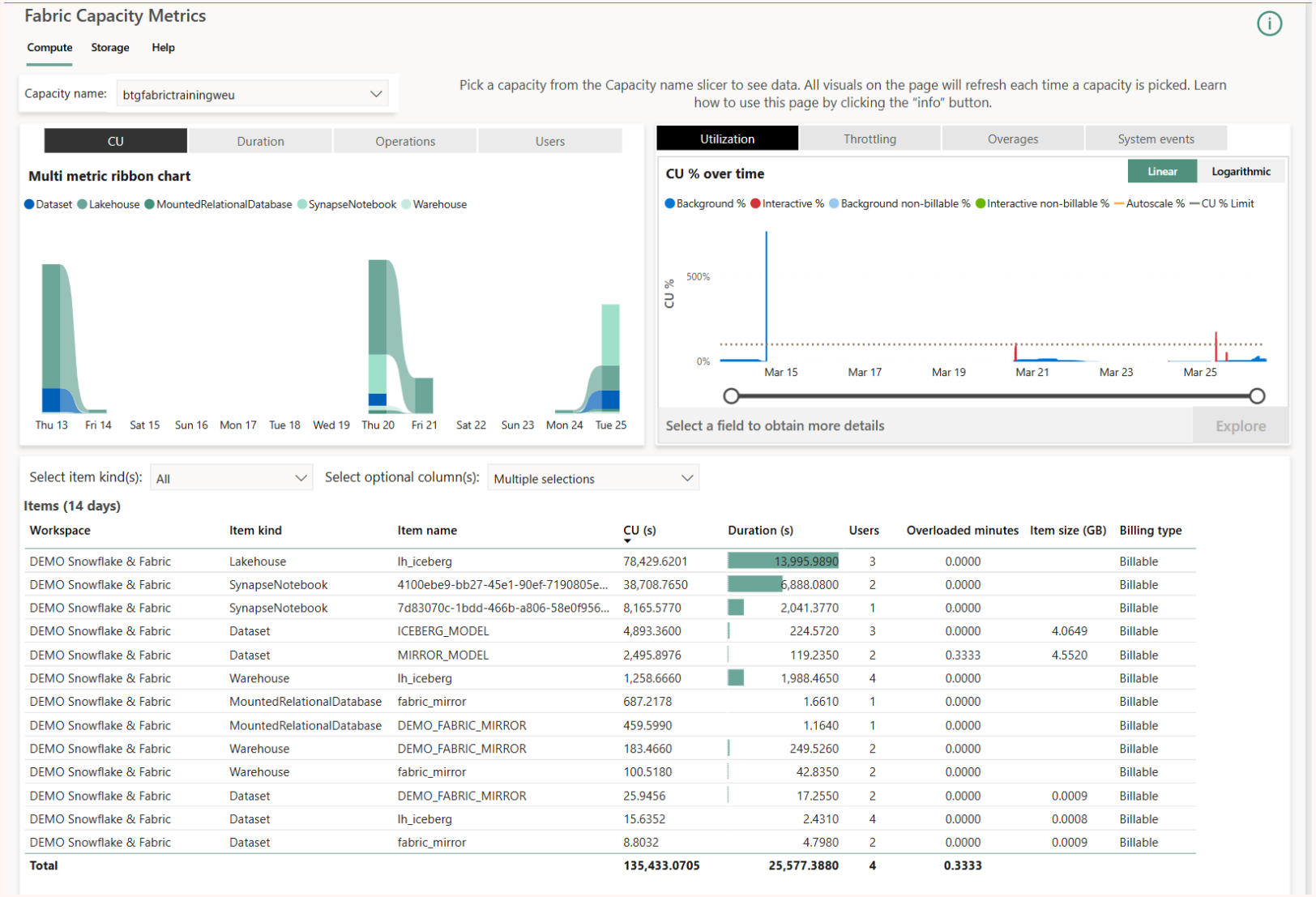


Limitations

- Supported data types
- Type width issue
- Iceberg table metadata storage isn't portable
- Iceberg table folders must contain only one set of metadata files
- Metadata changes not immediately reflected
- Region availability limitation
- Private links not supported
- OneLake shortcuts must be same-region
- Iceberg tables must be copy-on-write (not merge-on-read)
- Certain Iceberg partition transform types are not supported

Source: https://learn.microsoft.com/en-us/fabric/onelake/onelake-iceberg-tables#limitations-and-considerations

Microsoft Fabric Capacity Metrics



Workspace	Item kind	Item name	CU (s)	Duration (s)	Users	Overloaded minutes	Item size (GB)	Billing type
DEMO Snowflake & Fabric	Lakehouse	lh_iceberg	78,429.6201	13,995.9890	3	0.0000		Billable
DEMO Snowflake & Fabric	SynapseNotebook	4100ebe9-bb27-45e1-90ef-7190805e...	38,708.7650	6,888.0800	2	0.0000		Billable
DEMO Snowflake & Fabric	SynapseNotebook	7d83070c-1bdd-466b-a806-58e0f956...	8,165.5770	2,041.3770	1	0.0000		Billable
DEMO Snowflake & Fabric	Dataset	ICEBERG_MODEL	4,893.3600	224.5720	3	0.0000	4.0649	Billable
DEMO Snowflake & Fabric	Dataset	MIRROR_MODEL	2,495.8976	119.2350	2	0.3333	4.5520	Billable
DEMO Snowflake & Fabric	Warehouse	lh_iceberg	1,258.6660	1,988.4650	4	0.0000		Billable
DEMO Snowflake & Fabric	MountedRelationalDatabase	fabric_mirror	687.2178	1.6610	1	0.0000		Billable
DEMO Snowflake & Fabric	MountedRelationalDatabase	DEMO_FABRIC_MIRROR	459.5990	1.1640	1	0.0000		Billable
DEMO Snowflake & Fabric	Warehouse	DEMO_FABRIC_MIRROR	183.4660	249.5260	2	0.0000		Billable
DEMO Snowflake & Fabric	Warehouse	fabric_mirror	100.5180	42.8350	2	0.0000		Billable
DEMO Snowflake & Fabric	Dataset	DEMO_FABRIC_MIRROR	25.9456	17.2550	2	0.0000	0.0009	Billable
DEMO Snowflake & Fabric	Dataset	lh_iceberg	15.6352	2.4310	4	0.0000	0.0008	Billable
DEMO Snowflake & Fabric	Dataset	fabric_mirror	8.8032	4.7980	2	0.0000	0.0009	Billable
Total			**135,433.0705**	**25,577.3880**	**4**	**0.3333**		

Source: Fabric operations - Microsoft Fabric | Microsoft Learn
OneLake consumption - Microsoft Fabric | Microsoft Learn
Map each REST operation to a price - Azure Blob Storage | Microsoft Learn

Microsoft Fabric Capacity Metrics

Fabric Capacity Metrics

Compute Storage Help

Capacity name: btgfabrictrainingweu

Pick a capacity from the Capacity name slicer to see data. All visuals on the page will refresh each time a capacity is picked. Learn how to use this page by clicking the "info" button.

CU Duration Operations Users | Utilization Throttling Overages System events

Operation name	CU (s)	Duration (s)	Users	Billing type
Lakehouse Operations	54,592.3370	13,647.9840	2	Billable
OneLake Iterative Read via Proxy	13,453.8492	26.8730	1	Billable
OneLake Other Operations	3,852.2034	24.7440	1	Billable
OneLake Read via Proxy	2,695.5846	29.4330	1	Billable
OneLake Write via Proxy	1,860.8300	0.6840	1	Billable
Lakehouse Table Maintenance	1,060.9865	265.2410	1	Billable
OneLake Write via Redirect	563.7342	0.2370	1	Billable
OneLake Read via Redirect	234.1872	0.2260	1	Billable
OneLake Other Operations Via Redirect	115.9080	0.5670	1	Billable
Total	**78,429.6201**	**13,995.9890**	**3**	**Billable**

Lakehouse

Operation name	CU (s)	Duration (s)	Users	Billing type
OneLake Iterative Read via Proxy	174.3072	0.5860	1	Billable
OneLake Other Operations Via Redirect	96.5016	0.3830	1	Billable
OneLake Write via Proxy	83.4750	0.0940	1	Billable
OneLake Read via Redirect	63.6896	0.0600	1	Billable
OneLake Write via Redirect	41.6256	0.0410	1	Billable
Total	**459.5990**	**1.1640**	**1**	**Billable**

MountedRelationalDatabase

DEMO Snowflake & Fabric	Dataset	DEMO_FABRIC_MIRROR	25.9456	17.2550	2	0.0000	0.0009	Billable
DEMO Snowflake & Fabric	Dataset	lh_iceberg	15.6352	2.4310	4	0.0000	0.0008	Billable
DEMO Snowflake & Fabric	Dataset	fabric_mirror	8.8032	4.7980	2	0.0000	0.0009	Billable
Total			**135,433.0705**	**25,577.3880**	**4**	**0.3333**		

Source: Fabric operations - Microsoft Fabric | Microsoft Learn
OneLake consumption - Microsoft Fabric | Microsoft Learn
Map each REST operation to a price - Azure Blob Storage | Microsoft Learn

Microsoft Fabric Capacity Metrics

Fabric Capacity Metrics

Export data

Compute Storage Help

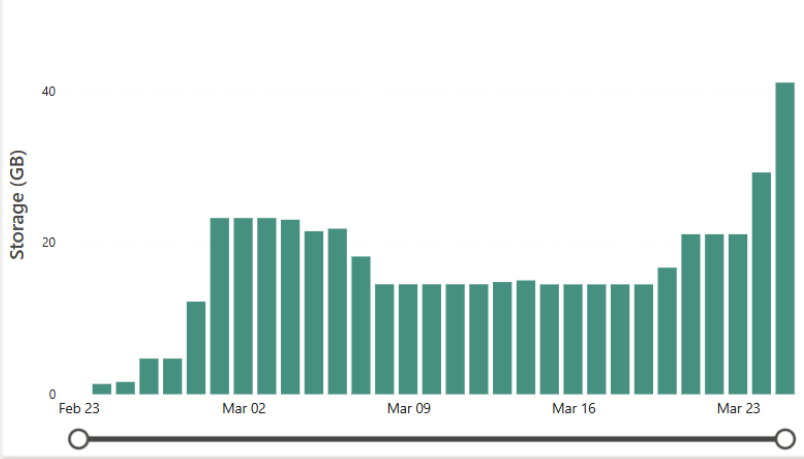
Capacity name: btgfabrictrainingweu | Date range: 2/23/2025 3/25/2025 | Experience: All | Storage type: All

1	49.85	17.17
Workspaces	Current storage (GB)	Billable storage (GB)

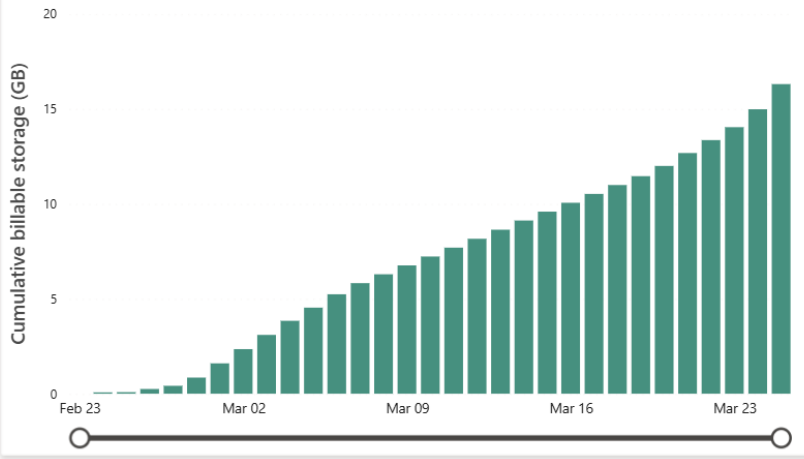
Top 10 workspaces by billable storage (GB) %

Workspace name	Workspace Id	Operation name	Deletion status	Billing type	Current storage (GB)	Billable Storage (GB)	Billable storage %
DEMO Snowflake & Fabric	[redacted]	OneLake Mirroring Storage	Active	Non-billable	39.55	14.17	82.53%
DEMO Snowflake & Fabric	[redacted]	OneLake Storage	Active	Billable	10.30	3.00	17.47%

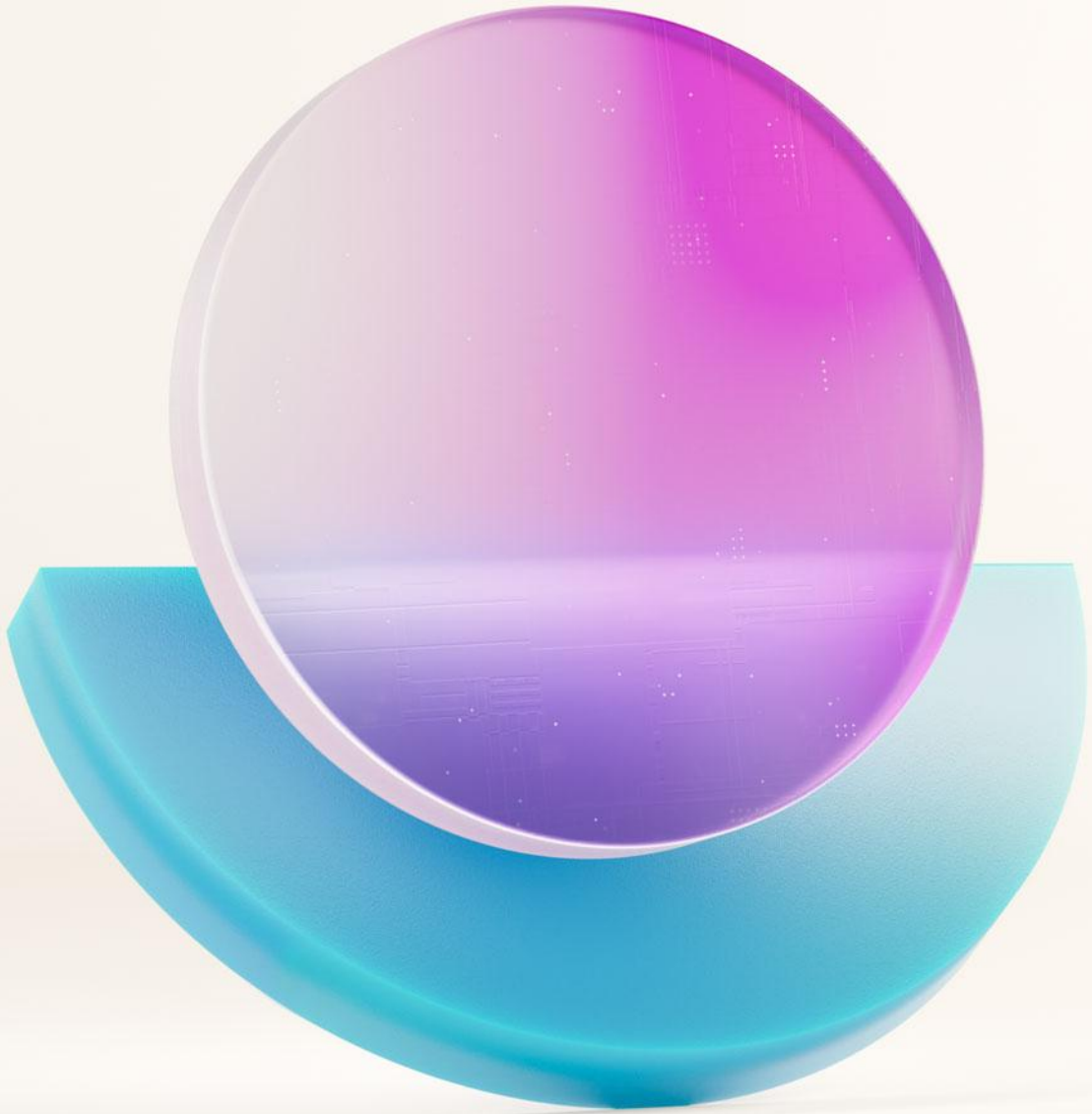
Storage (GB) by date



Cumulative billable storage (GB) by date



Microsoft Fabric
Community Conference



What should be used and how?

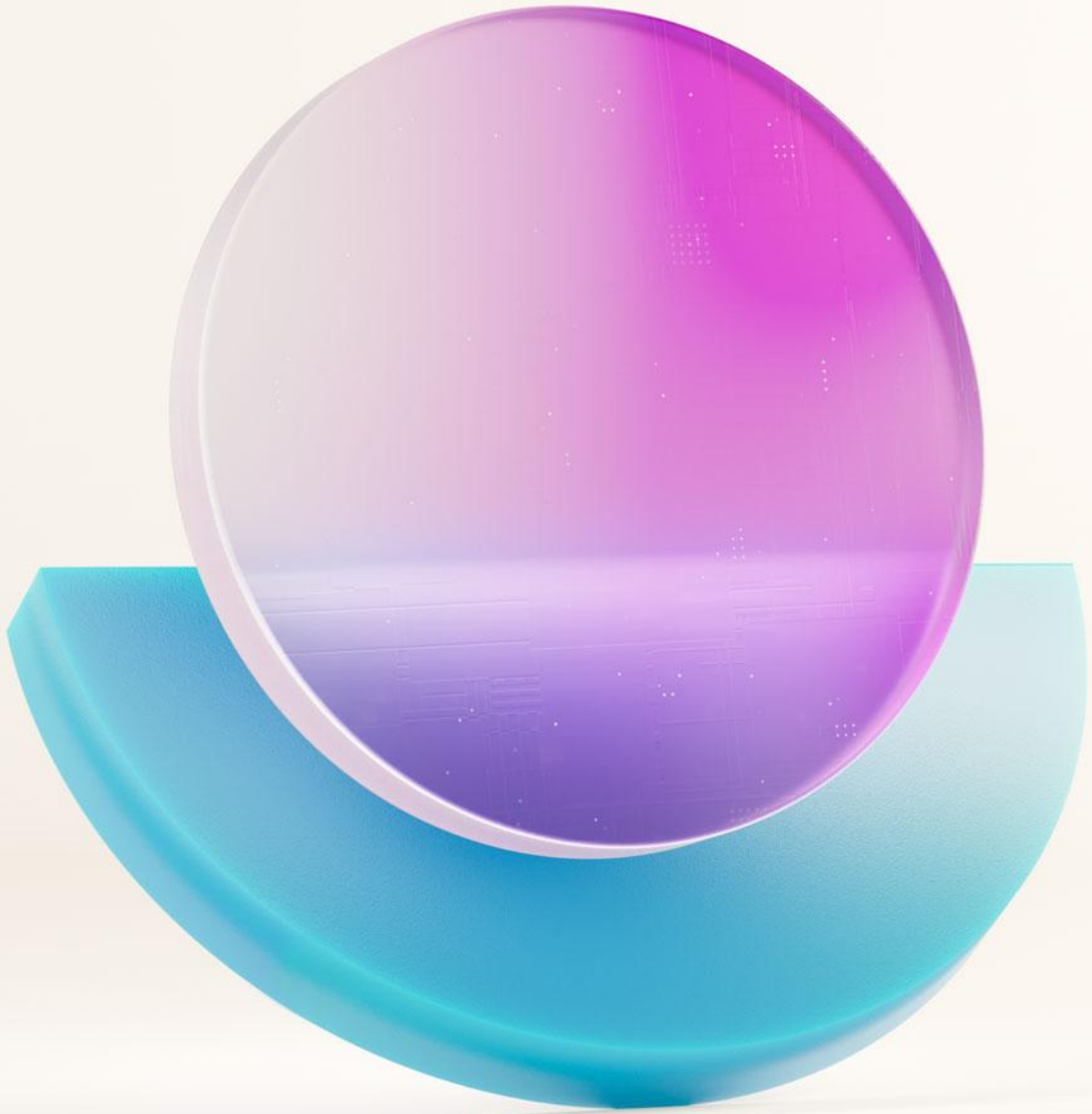
What should be used and how?

- Power BI Reporting on Top of Snowflake
  - Quick Win: Mirroring
  - Why? Direct Lake / V-Order Tables

- Data Mesh Concepts
  - DataMart / Gold Layer in Snowflake and expand analytics via Iceberg Tables to Fabric
  - Why? Enablement of Domain Driven Architectures

- Gen AI
  - Using REST Call from a Fabric Notebook to call Open AI API endpoints

Microsoft Fabric
Community Conference



Q&A

Q&A

1st question…What's next?

- Automatic conversion of Delta Lake formatted tables to Iceberg.
- Converting tables that are directly written to the Tables area of Fabric lakehouses – no shortcut needed.
- Schema-level shortcuts – one shortcut, multiple Iceberg tables.
- Deeper integration with Snowflake including a dedicated Snowflake data item in Fabric to automatically sync Iceberg and Delta tables.

Our wish: Bi-directional Access!

Source: Unlock the power of your Iceberg data in OneLake | Microsoft Fabric Blog | Microsoft Fabric

Microsoft

Microsoft Fabric
Community Conference

Thank You
DANKE SCHÖN ❤️

Feedback is appreciated in Whova!