



Microsoft Fabric

Alisa Benešová
B0m33bdt MS Fabric

7.11 2024

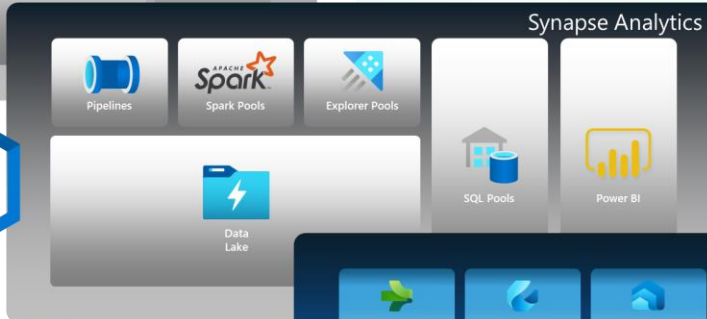
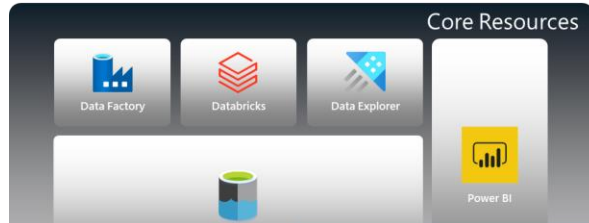
- 
1. What is MS Fabric?
 1. Timeline
 2. Azure Stack before & After
 3. Basic architecture
 2. OneLake
 3. MS Purview
 4. Fabric Experiences
 5. Fabric Engines
 6. Fabric GUI
 7. DEMO

What is MS Fabric

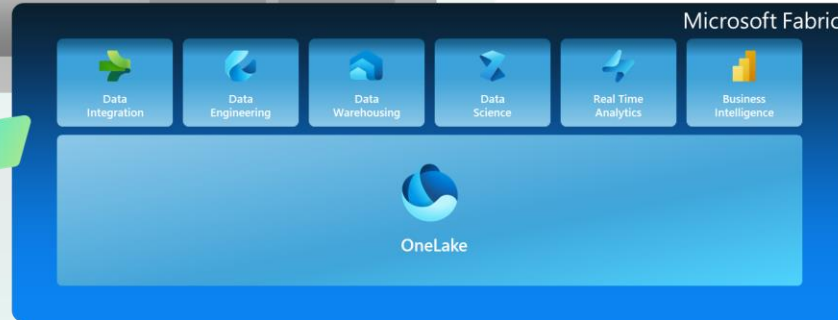
- Public preview 2023, development since 2019.
- Azure cloud based.
- All-in-one analytics solution that integrates data engineering, data warehousing, real-time analytics, and more under one platform.
- Promises to simplify data workflows and unite collaboration across analytics, data science, and business intelligence.
- Key components
 - OneLake, Microsoft Purview, Data Factory, Synapse, Power BI, Azure OpenAI...

2017 Azure Databricks

Dbx founded 2013, Snowflake founded 2012



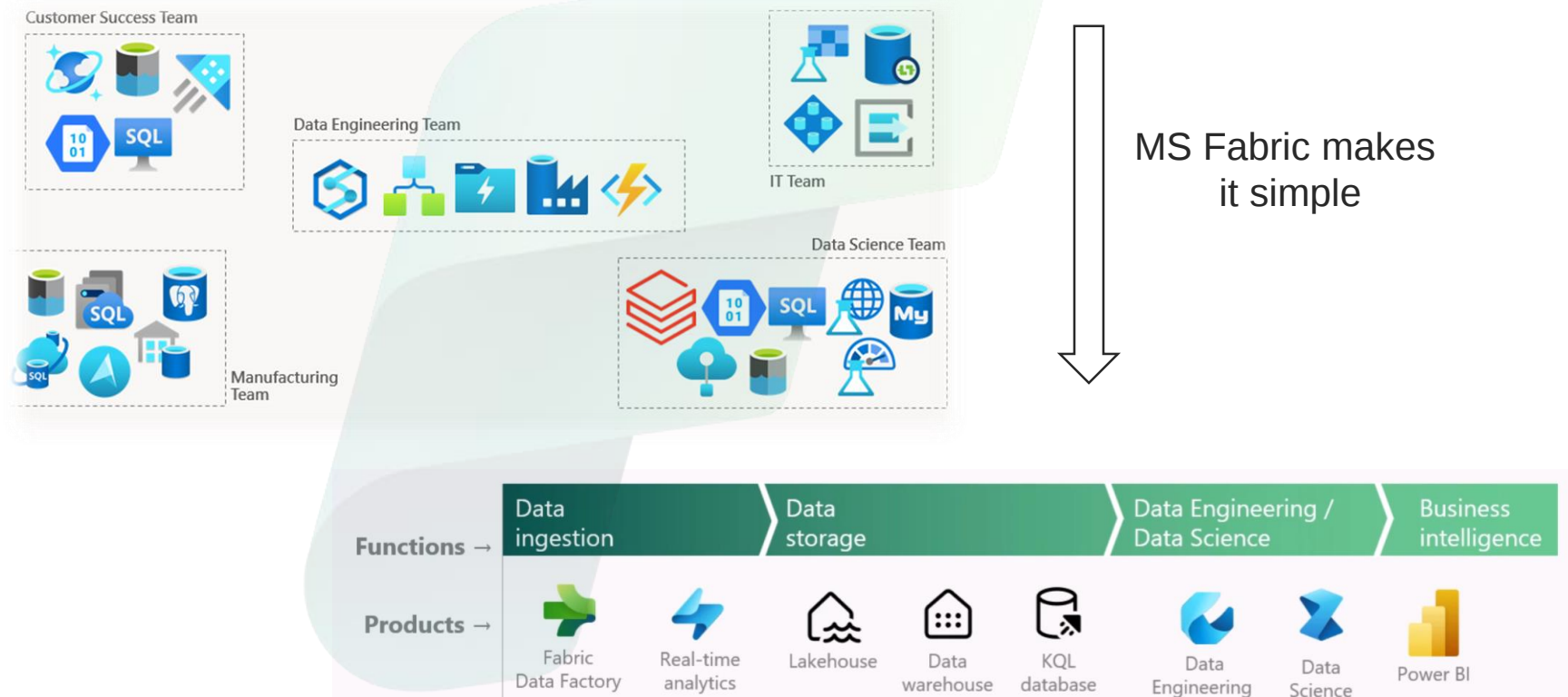
2019 Azure Synapse Analytics



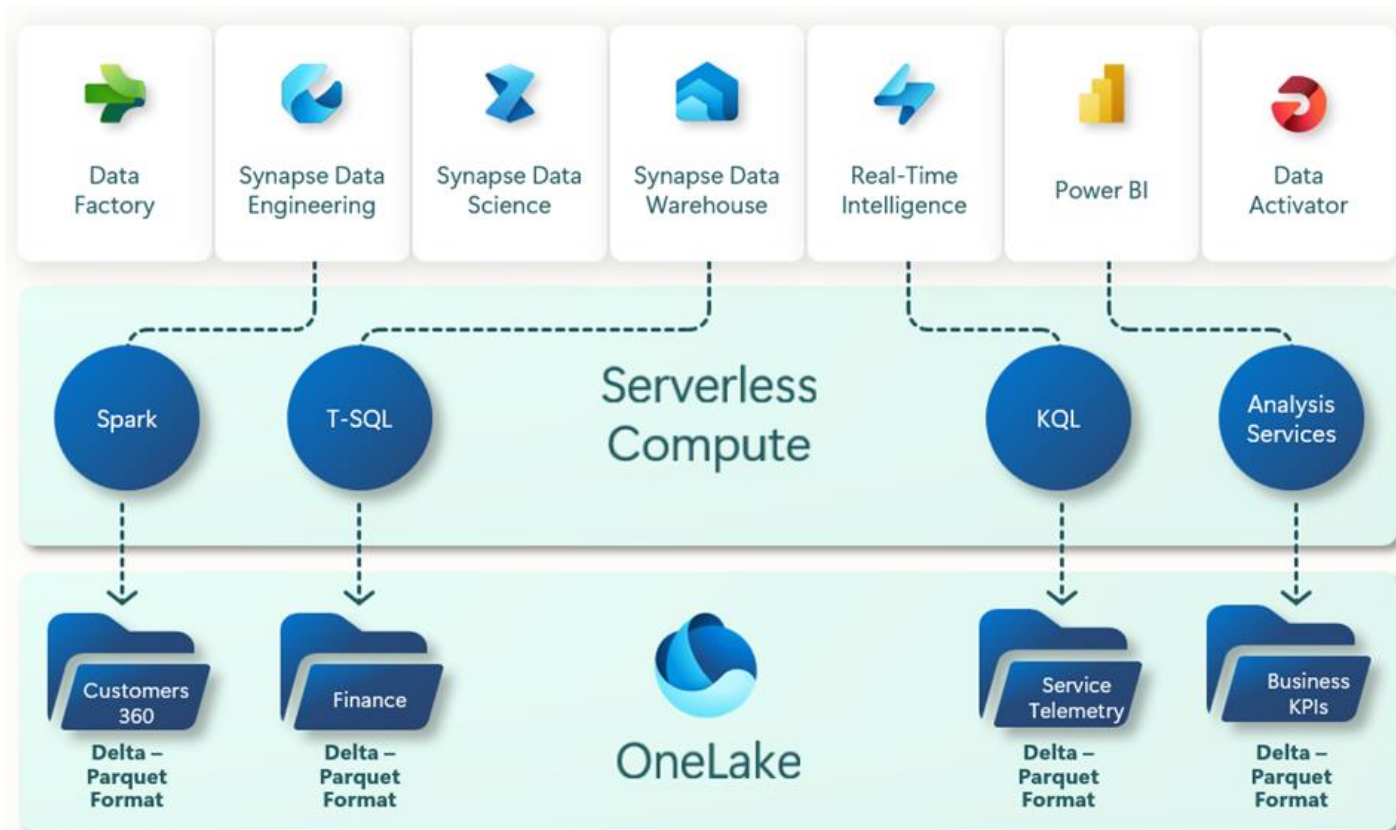
2023 MS Fabric

Azure Stack before and now

< PROFINIT >



Basic architecture



One Lake

< PROFINIT >

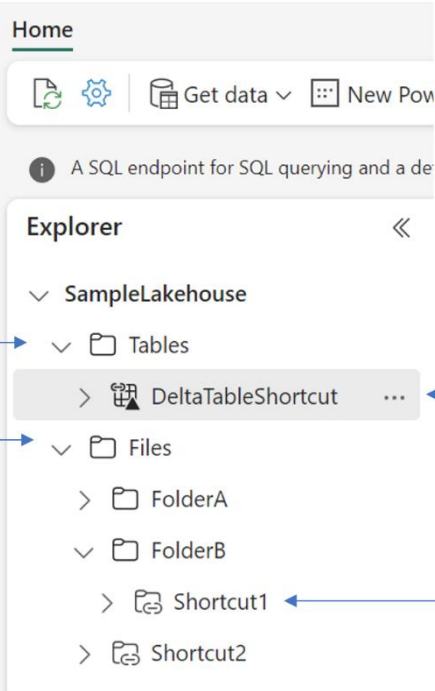
> Data Lake

- Foundation layer for all data, one data, no silos
- Unified governance, security, discoverability
- Supports Delta, Iceberg
- Prquet files, Versioning, schema enforcement, ACID, time travel

> Shortcuts

- Virtual copies to sources on other clouds, on premise, other services
- KQL





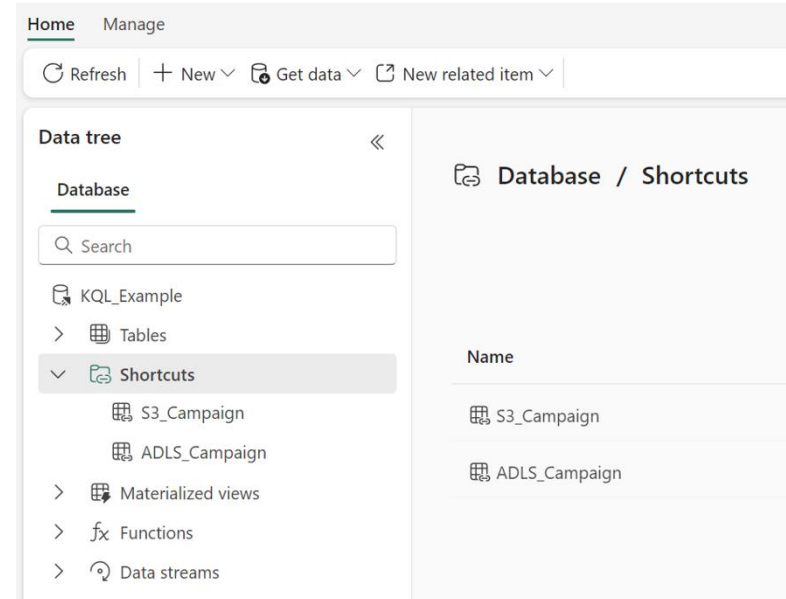
Pyspark>>

```
df = spark.read.format("delta").load("Tables/DeltaTableShortcut")
```

SQL>>

```
SELECT TOP (100) * FROM [SampleLakeh].[dbo].[MyShortcut]
```

KQL>> external_table('MyShortcut') | take 100



Microsoft Purview

> Purview

- Based on Apache Atlas, but not fully opensource, free version has some [limitations](#)
- Fully managed, unified platform for data governance, compliance, security
- Across multicloud, on-premise, SaaS

> PurviewData Map

- Register and scan your data for sensitive data

> Purview Catalog

- Manage data integrity, access policies, data quality,...
- (similar to UnityCatalog)



Fabric 7 experiences

> Data Factory

- ETL, primarily focus on moving and transforming your data (into Fabric)
- Alt: Azure Data Factory, Synapse Pipeline, PowerBI Flow

> Synapse Data Engineering

- Collect, store, process, and analyze large volumes of data
- Notebooks (VScode), Lakehouse, DataPipelines, Spark jobs
- Alt: Azure Data Lake Services (ADLS Gen2), Databricks, Snowflake

> Synapse Data Science

- Supports the entire data science workflow, from data exploration, preparation and cleansing to experimentation, modeling, model scoring and serving of predictive insights to BI reports.
- Notebooks, Experiments using MLFlow, MLModels
- Alt: Azure Machine Learning, Synapse notebooks, Databricks notebooks

Fabric 7 experiences, pt2

> Synapse Data Warehouse

- Provides DWH like transactional tables, schemas, views, stored procedures... using TSQL
 - built on top of [Polaris engine](#) (paper, Distributed SQL engine)
- Alt: Synapse SQL Serverless/Dedicated or Snowflake or MSSQL etc.

> Synapse Real Time Analytics

- Set of tools to ingest, manage and analyze real-time event data
- Eventstream (no code tool), KQL Databse (built on KQL engine), KQL Queryset
- Alt: Azure Data Explorer, Azure Event Hubs, Azure Stream Analytics

Fabric 7 experiences, pt3

> Power BI

- Reports, Semantic Models (known as Dataset)
- Alt: Tableau, Looker, Toughspot

> Data Activator

- no-code experience or automatically taking actions when **patterns** or **conditions** are detected in changing data
- Reflex item (**objects**, its trackable **property**, **trigger** over property and **action** when trigger is triggered)

Fabric 4 engines

> The SQL engine

- Used by: Data Warehouse experience
- Languages: T-SQL

> The Spark engine

- Used by: Data Engineering experience, Data Science experience
- Languages: Python (PySpark), R (SparkR), Scala & SQL (SparkSQL).

> The KQL engine

- Used by: Real-time Analytics experience
- Languages : KQL

> The Analysis Services engine

- Used by: Power BI experience
- Languages: DAX, Power Query

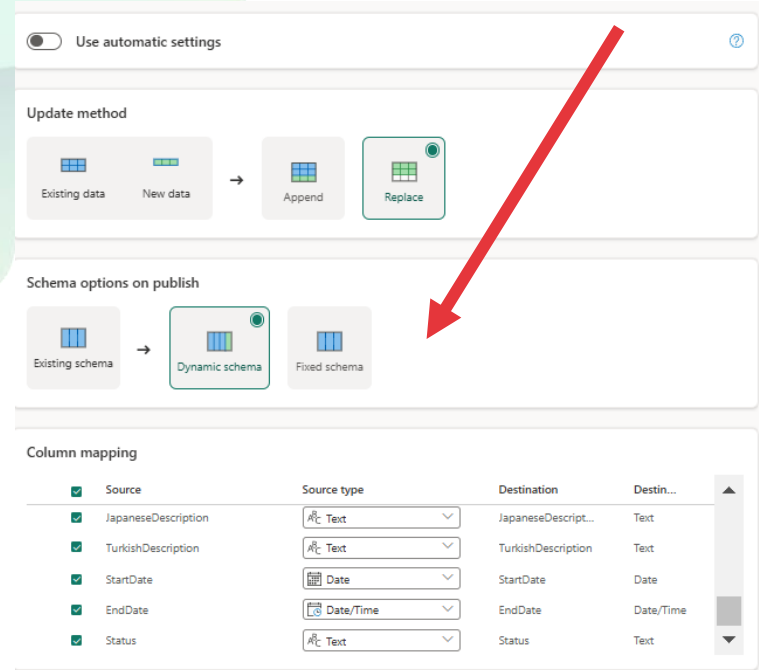
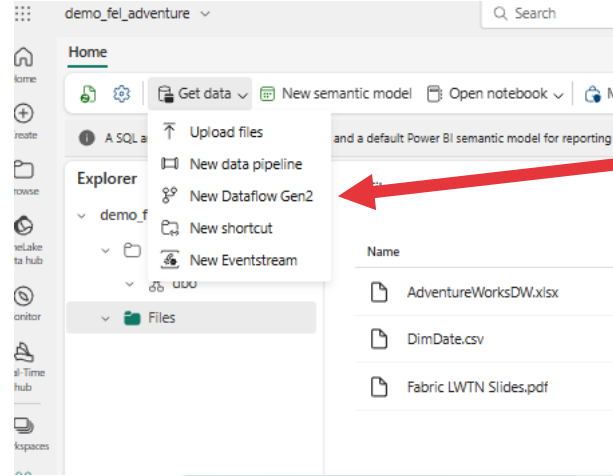
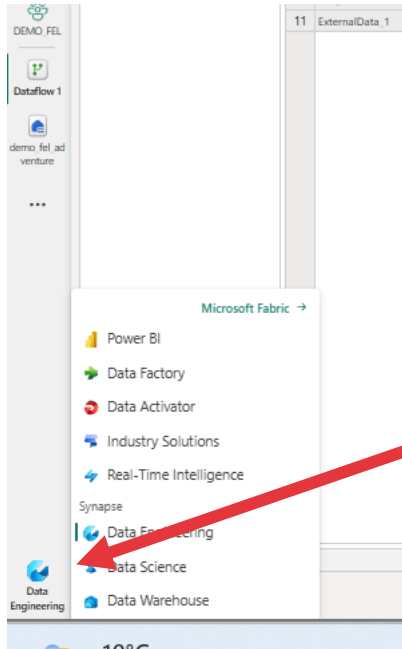
Fabric GUI... Complicated...

< PROFINIT >

How to link data

Customize datatypes

Choose your experience



DEMO

< PROFINIT >



MS Fabric? Databricks? Snowflake?

	MS fabric	Databricks	Snowflake
Primary focus	Unified analytics for Microsoft-Azure	Data engineering, data science, and ML	Cloud-based data warehousing
Best for	Microsoft-centric organizations	Machine learning, large-scale analytics	Enterprise segment, SQL-based analytics, data warehousing

	MS Fabric SaaS	Databricks PaaS	Snowflake SaaS
PROS	MS/Azure unified	Multicloud, optimized for AI/ML, OpenSource focus	DWH focused, strong security&governance, multicloud
CONS	Vendor lockin, Cost complexity, still new -> immature	Less focused on BI, data governance maturity	Less ML/AI, vendor lockin

Děkujeme za pozornost



Profinit EU, s.r.o.
Tychonova 2, 160 00 Praha 6

Tel.: + 420 224 316 016, web: www.profinit.eu

 LINKEDIN
linkedin.com/company/profinit

 TWITTER
[@profinit_EU](https://twitter.com/profinit_EU)

 FACEBOOK
facebook.com/Profinit.EU

 YOUTUBE
Profinit EU, s.r.o.