



envision
VIRTUAL SUMMIT SERIES
DATA & ANALYTICS

Microsoft Fabric

The Magic of a Unified Analytic Reporting Platform

Tuesday, September 12, 2023



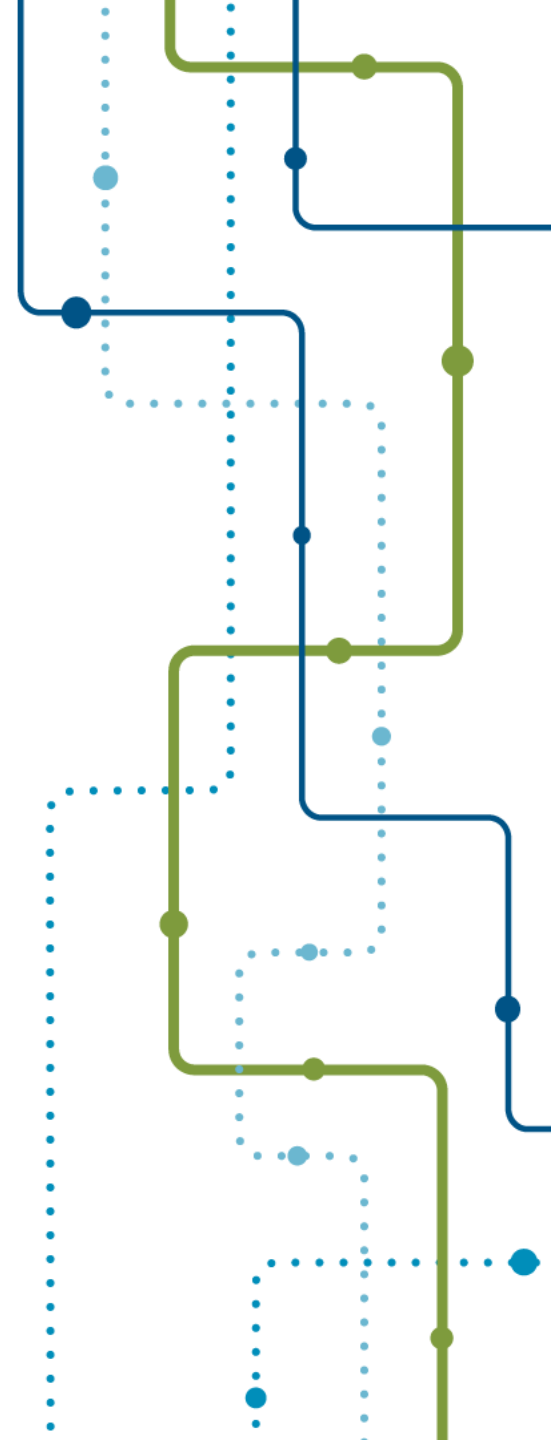
Paul Turley
Director & BI&A Competency Lead



PRESENTED BY 

Session abstract

Get ready to be blown away by Microsoft's sensational lake house architecture, powered up with the incredible capabilities of Power BI! This mind-blowing innovation is set to revolutionize the entire industry. Imagine a single, unified platform that lets you create a comprehensive end-to-end data warehouse and analytics solution, right there in Power BI itself! Brace yourself as BI & DW developers unlock the full potential of cutting-edge technologies like SPARK, delta lake, parquet, Python notebooks, pipelines, Kusto, SQL, and so much more, all seamlessly integrated within the Power BI platform. It's a dazzling symphony of possibilities, where you can orchestrate mind-boggling enterprise solutions using just one extraordinary product. Don't miss out on witnessing the magic firsthand and discover the limitless wonders of this groundbreaking new platform. Get ready to have your expectations shattered and your imagination set on fire!



Look, up in the cloud:

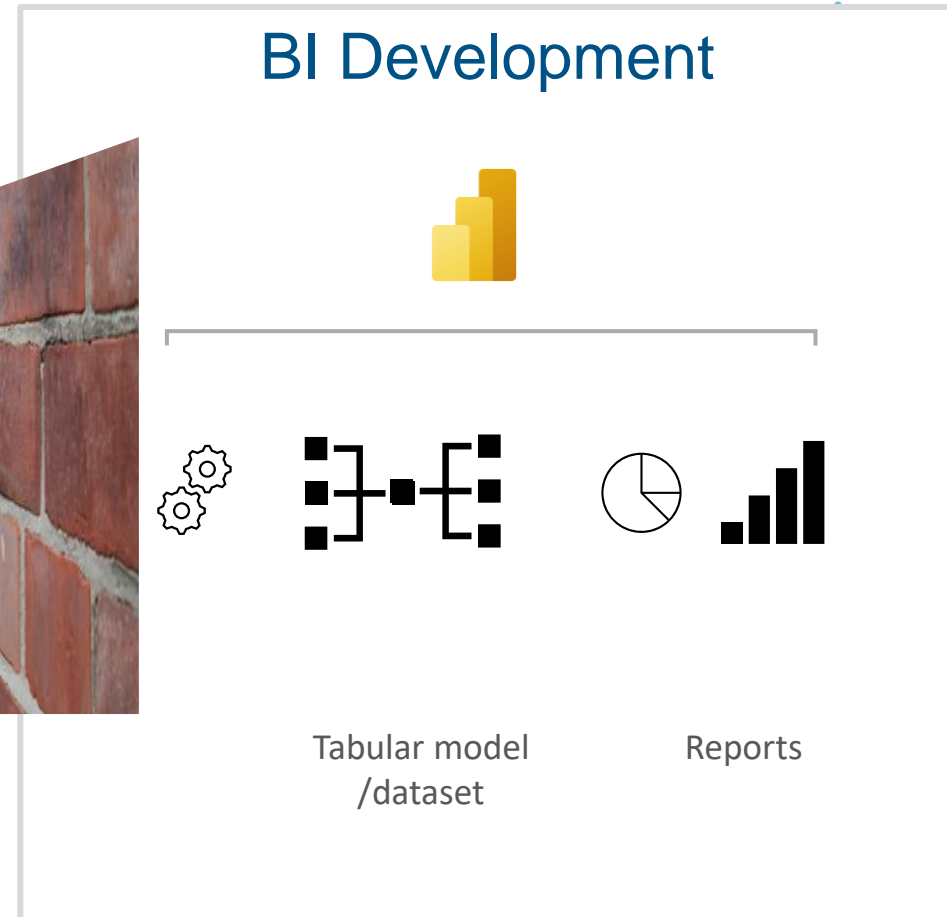
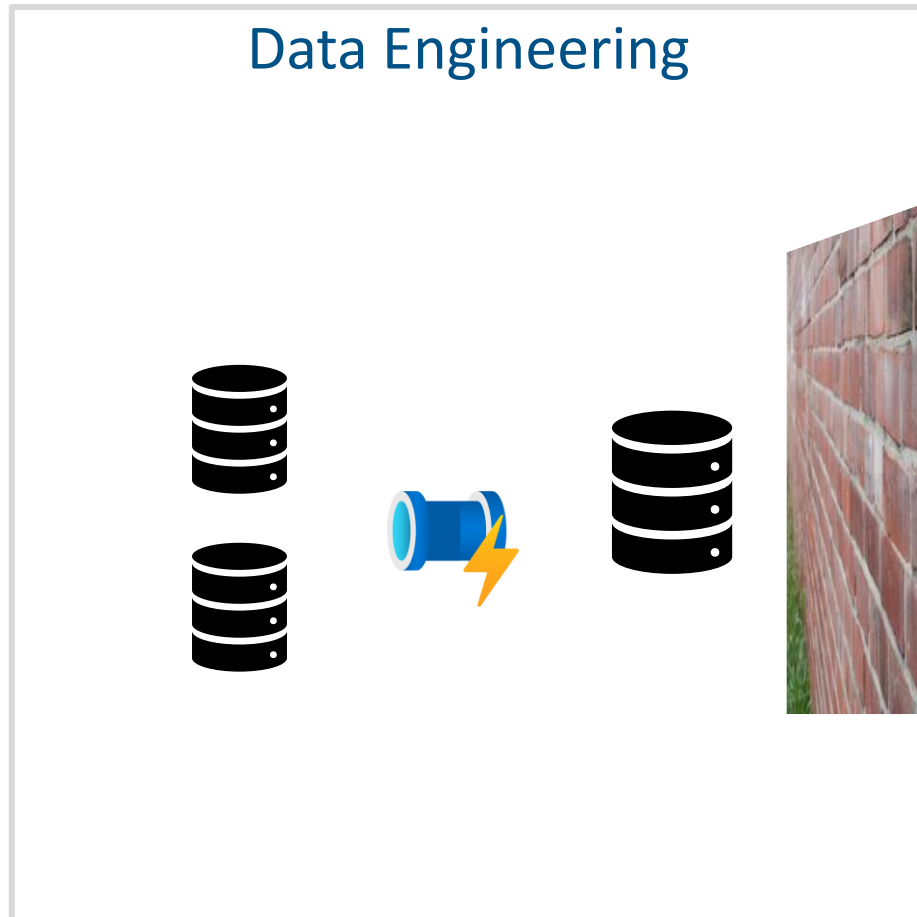
It's Azure... It's Power BI...

...no,

It's Fabric!



Conventional Power BI Solutions

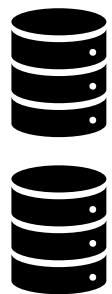


Microsoft Fabric



Data Engineering

BI Development



Source systems
& APIs



Pipelines



Lakehouse



Python
notebooks



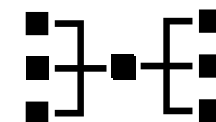
Data science



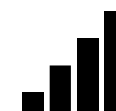
Clusters & compute



Delta Lake



Tabular model
/dataset



Reports



Synapse Data Warehousing



Synapse Data Engineering



Data Factory



Synapse Data Science



Synapse Real Time Analytics



Power BI



T-SQL



Spark

Serverless compute



KQL



Analysis Services

OneSecurity



Warehouse

Delta – Parquet Format



Lakehouse

Delta – Parquet Format



OneLake



Kusto DB

Delta – Parquet Format



Dataset

Delta – Parquet Format

Fabric Concepts

- Workspaces
- Lakehouse
- OneLake
- Files (Source & Parquet)
- Delta tables
- Shortcuts
- Pipelines
- Notebooks
- Direct Lake





Microsoft Fabric

Data analytics for the era of AI

Lakehouse | OneLake | Direct Lake

Complete Analytics Platform

Everything, unified

SaaS-ified

Secured and governed

Lake Centric and Open

OneLake

One copy

Open at every tier

Empower Every Business User

Familiar and intuitive

Built into Microsoft 365

Insight to action

AI Powered

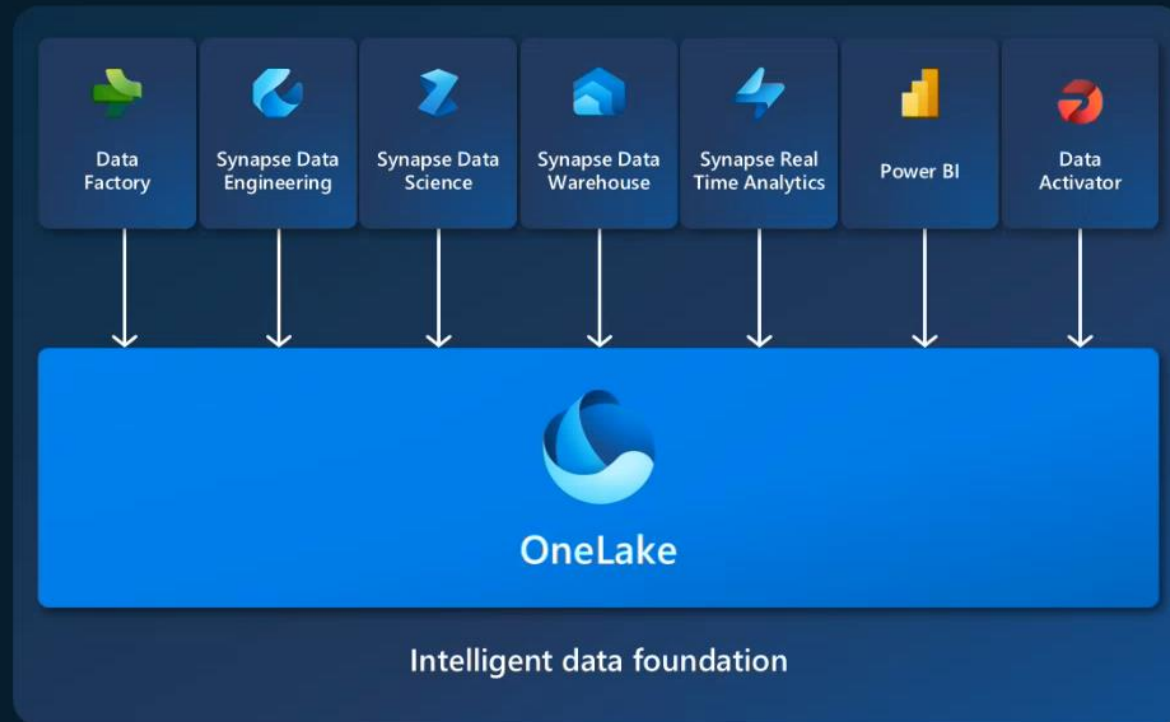
Copilot accelerated

GPT on your data

AI-driven insights

OneLake for all Data

“The OneDrive for Data”



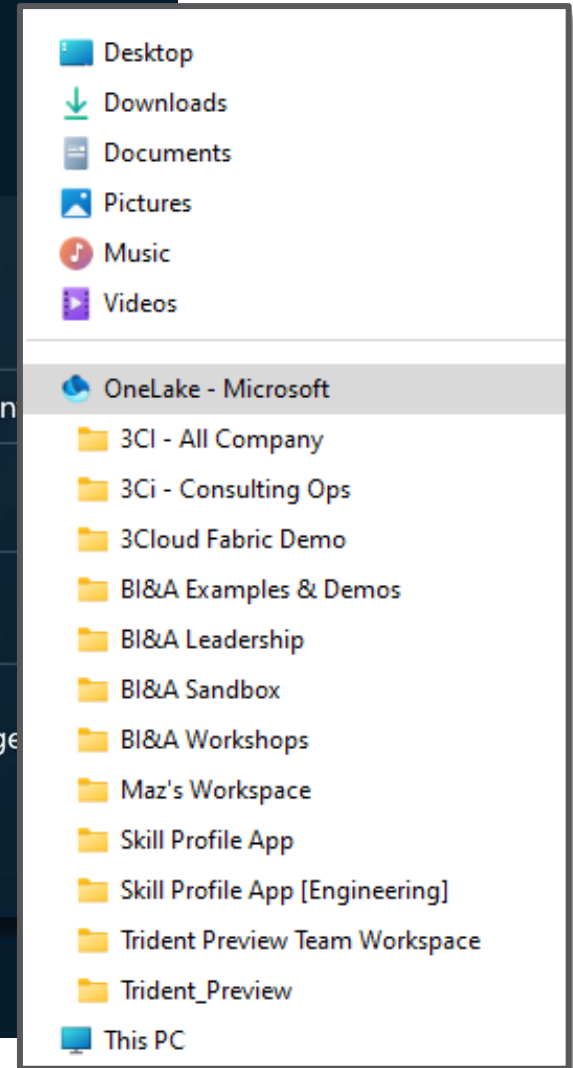
A single SaaS lake for the whole organization

Provisioned automatically with the tenant

All workloads automatically store their data in the OneLake workspace folders

All the data is organized in an intuitive hierarchical namespace

The data in OneLake is automatically indexed for discovery, MIP labels, lineage, PII scans, sharing, governance and compliance

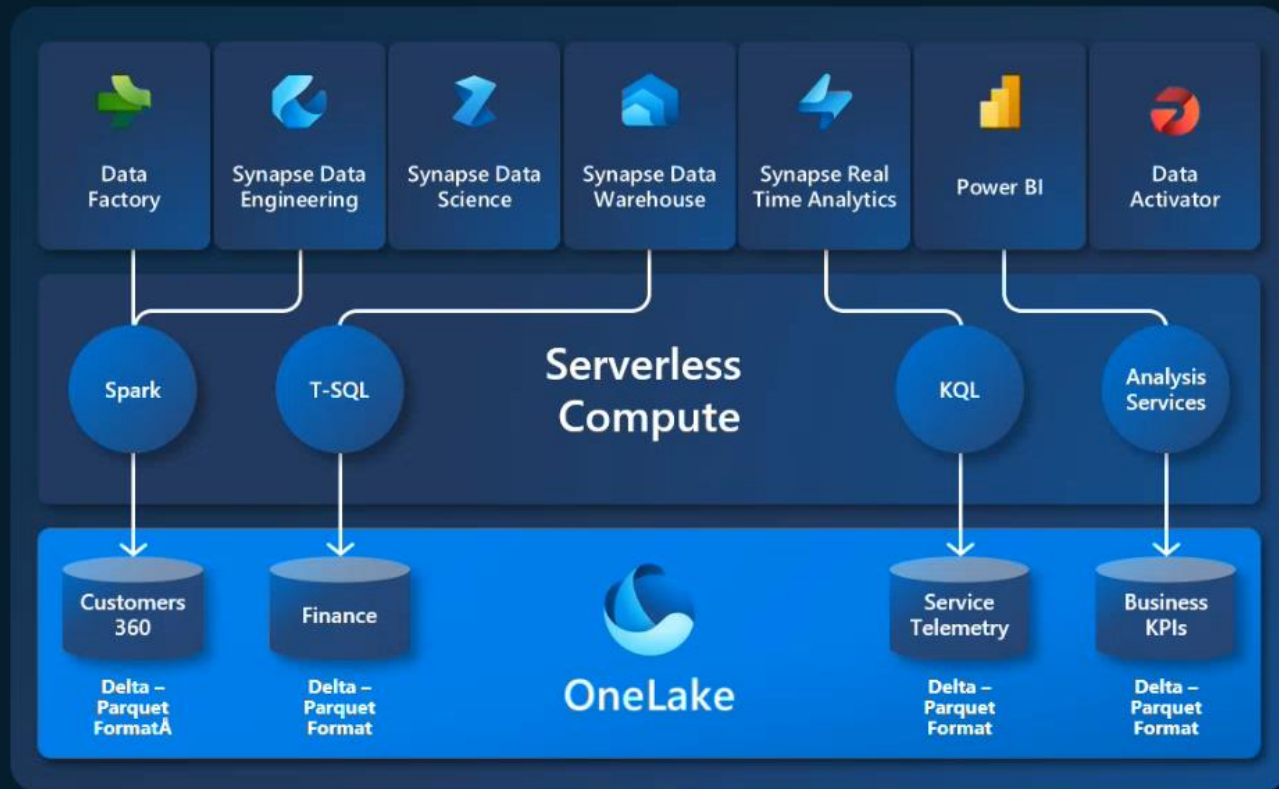


Accessing a Fabric workspace:

- Power BI portal <https://app.powerbi.com/groups/6f4ea7cf-61a2-4efa-b021-62735e3882bc/list?experience=power-bi>
- OneLake File Explorer <C:\Users\PaulTurley\OneLake - Microsoft\BI&A Workshops>

One Copy for all computes

Real separation of compute and storage



All the compute engines store their data automatically in OneLake

The data is stored in a single common format

Delta – Parquet, an open standards format, is the storage format for all tabular data in Analytics vNext

Once data is stored in the lake, it is directly accessible by all the engines without needing any import/export

All the compute engines have been fully optimized to work with Delta Parquet as their native format

Shared universal security model is enforced across all the engines

One Copy for all computes

Universal security makes it real



All the compute engines store their data automatically in OneLake

The data is stored in a single common format

Delta – Parquet, an open standards format, is the storage format for all tabular data in Analytics vNext

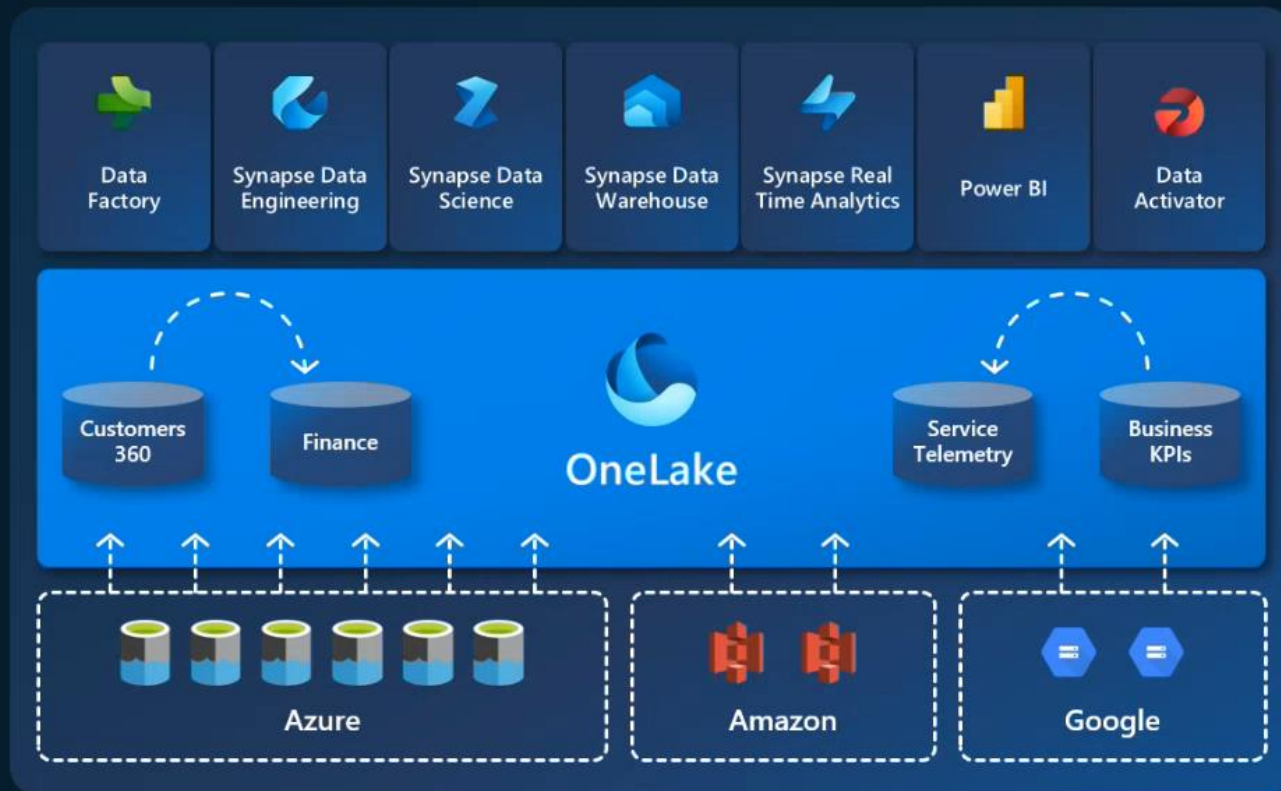
Once data is stored in the lake, it is directly accessible by all the engines without needing any import/export

All the compute engines have been fully optimized to work with Delta Parquet as their native format

Shared universal security model is enforced across all the engines

Taking One Copy to the Next Level

Shortcuts



Sharing data in OneLake is as easy as sharing files in OneDrive, removing the needs for data duplication

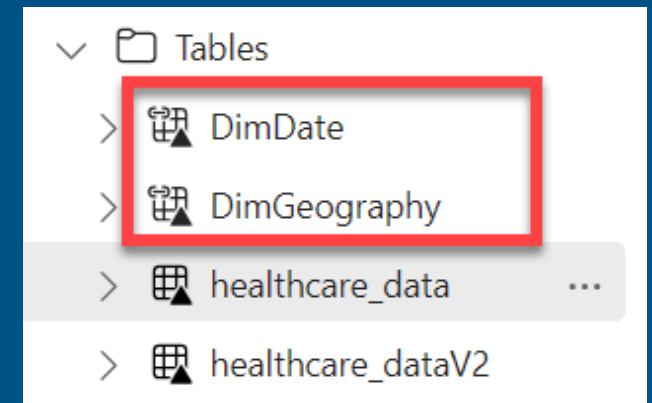
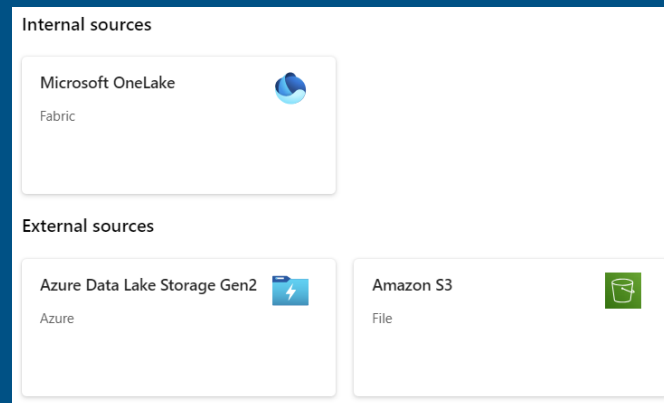
With [shortcuts](#), data throughout OneLake can be composed together without any data movement

Shortcuts also allow instant linking of data already existing in Azure and in other clouds, without any data duplication and movement, making [OneLake a multi-cloud data lake](#)

With support for industry standard APIs, OneLake data can be directly accessed by any application or service

Shortcuts

- Access remote sources as if stored in the Lakehouse
- No need to copy or stage data
- Processed in the target workspace region
- Additional sources are planned



Notebooks

Can be authored using:

- PySpark Python
- Spark Scala
- Spark SQL
- Spark R



Direct Lake

- Native Power BI dataset storage moves from SSAS file storage to Parquet delta files
- Columnar compression will apply “VertiPaq” in-memory scans
- Only one copy of data in delta “source” tables rather than cached copy using Import mode
- Performance is on-par with Import storage mode

Upgrade to Microsoft Fabric at your own pace



Continue building on Synapse Gen2, Azure Data Factory, Azure Data Explorer, Azure Databricks.

1



Mount existing Synapse Gen2, Azure Data Factory, Azure Data Explorer to Microsoft Fabric, at zero cost/risk.

2



Upgrade to full Microsoft Fabric experience with tooling and support from Microsoft.

3

What Does Fabric Look Like?

The screenshot shows the Microsoft Fabric Power BI interface. The top navigation bar includes 'Work', 'Power BI', and a search bar. The main content area is titled 'BI&A Workshops' and displays a 'New' section with various data assets categorized into Data Activator, Data Engineering, Data Factory, Data Science, and Data Warehouse. A blue callout box is overlaid on the interface, containing the text 'Add assets to a workspace ...simple!'.

BI&A Workshops Create for lunch n learn workshops and internal training.

+ New Upload Create app Manage access ...





Name	Type	Pa
ContosoDW_Import_01	Data pipeline	Pa
ContosoDW_Lakehouse	Dataset (default)	BI
	SQL endpoint	BI
	Lakehouse	Pa
Healthcare_Samples	Dataset (default)	BI
Healthcare_Samples	SQL endpoint	BI
Healthcare_Samples	Lakehouse	Pa
Import Store and Geography	Data pipeline	Pa
Pauls_Healthcare_Sample_Generator	Notebook	Pa
Sample Healthcare Data v2	Notebook	Pa

Demo

Problem/Solution

Healthcare testing data

- Lakehouse is easier to access than local files or traditional database
- Python notebook
- Python code generated by ChatGTP
- Generated table includes 10,566 realistic sample rows


	Healthcare_Samples	Dataset (default)
	Healthcare_Samples	SQL endpoint
	Healthcare_Samples	Lakehouse
	Sample Healthcare Data v2	Notebook

Demo

Problem/Solution

Low-cost, assessable training data

- Lakehouse is easier to access than local files or traditional database
- SQL Server database tables transformed to Lakehouse using pipelines
- SQL Endpoint represents data as if it were SQL Server
- Lakehouse views abstract queries
- “It just works!”

	ContosoDW_Import_01	Data pipeline
	ContosoDW_Lakehouse	Dataset (default)
	ContosoDW_Lakehouse	SQL endpoint
	ContosoDW_Lakehouse	Lakehouse
	Import Store and Geography	Data pipeline

Demos

1. Migrating SQL Server tables to Fabric delta tables
 - Connect using SSMS
 - Create views
2. Query & create objects using SQL endpoint in SSMS
3. Import delimited files collection
 - Resulting Power BI dataset implements DirectLake storage
 - Use OneLake Explorer to add and browse files
4. Generate & transform data frame using Python notebook

Demo links:

- [Workspaces & new objects, personas](#)
- [Import tables using pipelines](#)
- [Airline transformations notebook](#)
- [Healthcare sample generator notebook](#)

Resources

- My blog: SqlServerBi.blog
- Guy In A Cube: GuyInACube.com
- Fabric portal: Fabric.Microsoft.com

- **Paul Turley**
paul@intelligentbiz.net
@paul_turley
linkedin.com/in/pturley

