



When Two Become One

# Fabric SQL Data Warehouse

Future Data Driven  
Summit 2023

Andy Cutler  
September 27th



09/27/2023

**Future Data Driven  
Summit 2023**

Free, online,  
Community event!

**5 Tracks! 10 hours!  
Over 50 speakers!**





# Andy Cutler



Independent BI/DW Consultant

Azure Data Platform & Power BI

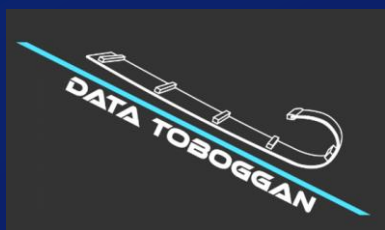
[datahai.co.uk](http://datahai.co.uk)

[serverlesssql.com](http://serverlesssql.com)

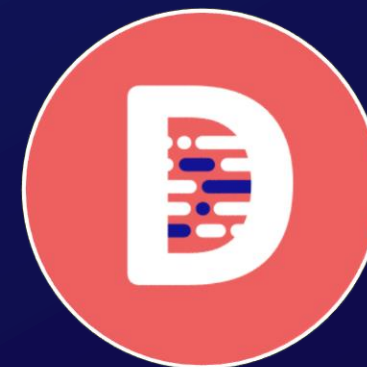
[twitter.com/MrAndyCutler](https://twitter.com/MrAndyCutler)

[linkedin.com/in/andycutler/](https://linkedin.com/in/andycutler/)

[datatoboggan.co.uk](http://datatoboggan.co.uk)



SCAN ME



# Thank you for organising!



# Agenda



**Exploring the Timeline**



**Warehouses and Lakehouses in Fabric**



**Demo**



# Timeline

...a 15 year journey

# History

## 2008-2010

All started with DATAlegro (bought in 2008)



Installed on commodity appliances (Dell hardware)

Evolved into SQL Server 2008 Parallel Data Warehouse

PDW was created to compete with other large-scale data warehouse products like Oracle, Teradata.

DATAlegro were visionaries!

### MAGIC QUADRANT

Figure 1. Magic Quadrant for Data Warehouse Database Management Systems, 2006



# History

2015

Azure SQL Data Warehouse

Azure hosted version of  
Parallel Data Warehouse

Gen1 and Gen2 versions

Easier to purchase and use!



Platform-as-a-Service (PaaS)

# History

## 2019

Azure Synapse Analytics

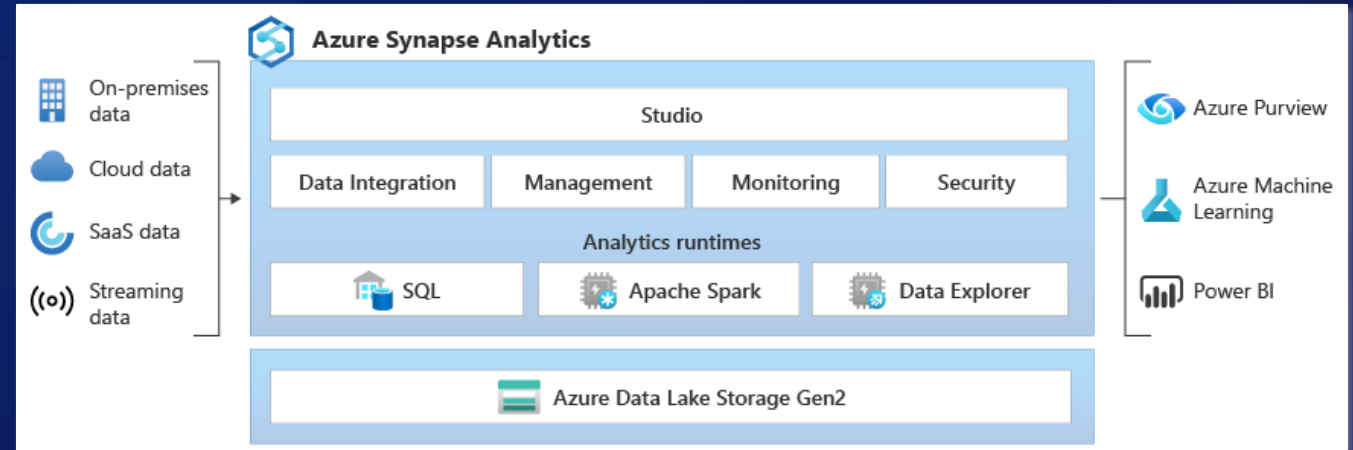
Renamed Azure SQL Data Warehouse to Dedicated SQL Pools

Introduced a new SQL engine called Serverless SQL Pools

Added Data Factory features as Pipelines

Added Spark and Data Explorer

Added Power BI integration...sort of



Platform-as-a-Service (PaaS)





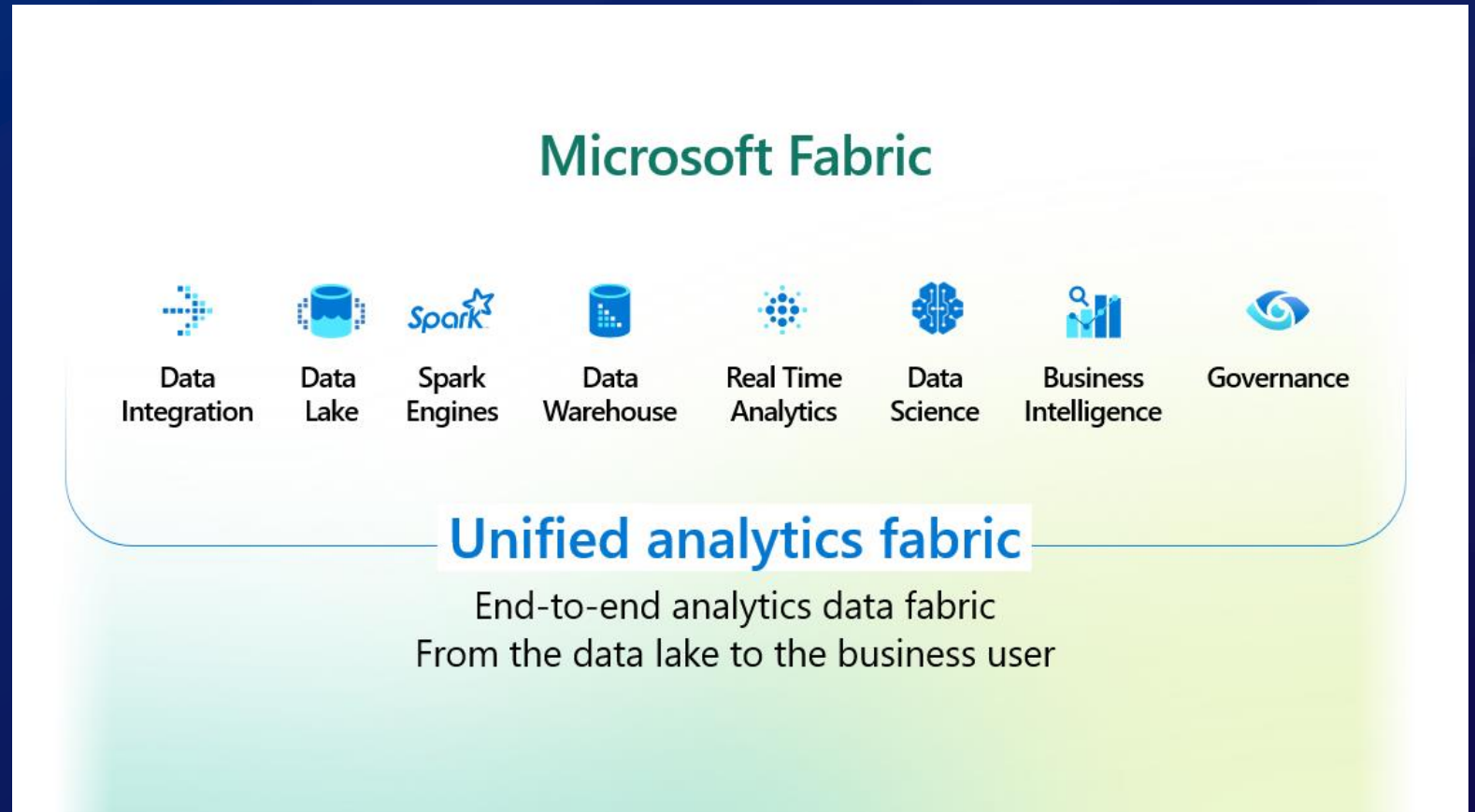
# History

2023

Microsoft Fabric

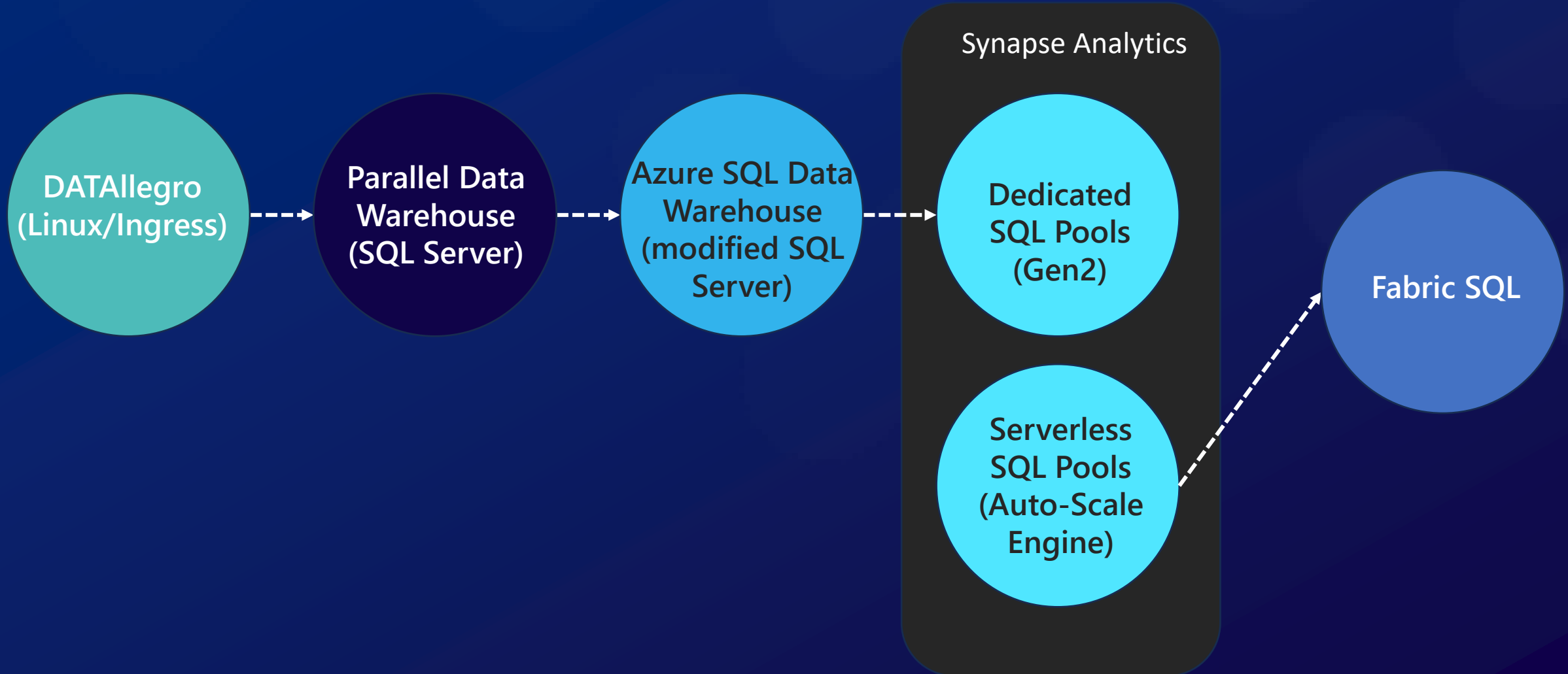
“Another” integration of several services...?

Unified SQL Engine



Software-as-a-Service (SaaS)

# SQL Engine Timeline





# Fabric

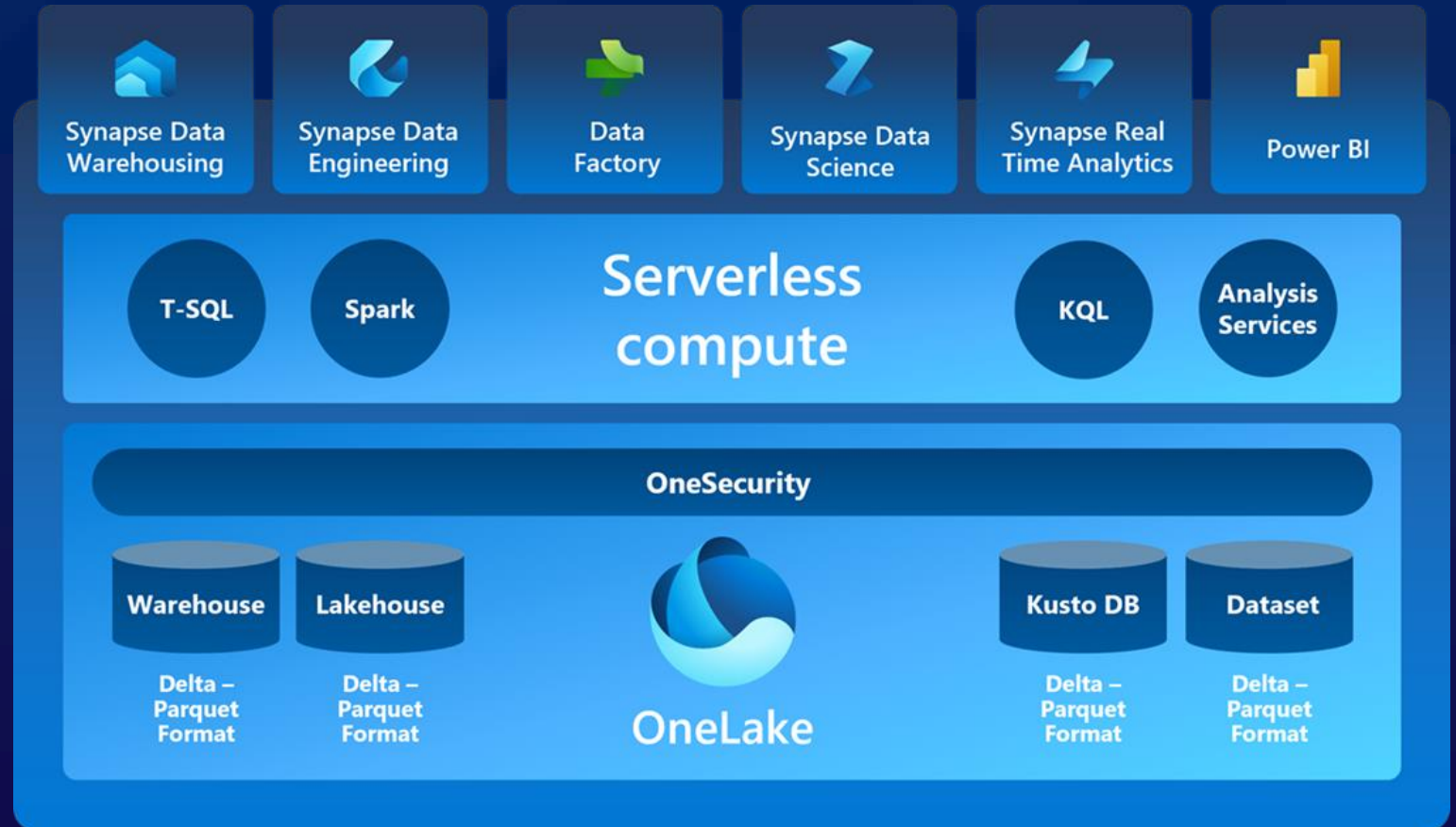
# What is Fabric?

Adds "significant" functionality to a Power BI tenant

Build out your Lakehouse or Warehouse architecture

Supports Real-time data streaming & Machine Learning

Storage underpinned by open-source Delta Lake format



<https://github.com/microsoft/Fabric-Readiness>

# Lakehouse



Persona



Workspace

Strunctured, Unstructured, Semi-Structured data

Data Engineer

Uses Spark (Python, Scala, R)

Folders and files (and databases too...)

Bronze, Silver, and Gold Layer

New DirectLake connection from Power BI Datasets

# Warehouse



Persona



Workspace

Structured data

SQL Data Warehouse Developer

T-SQL for reading and writing

Tables, Schemas, and Databases

Can be used in the Gold layer

DirectQuery and Import modes only (currently)

# What Powers the Warehouse?

An auto-scale out, fault tolerant, distributed compute engine

Enhanced version of Serverless SQL Pools engine

Native storage uses Delta Lake

Transaction Log

Metadata



Name	Status	Date modified	Type
📁 _delta_log	☁️	07/09/2023 13:52	File folder
📁 6FF257A9-AC3F-476F-BA07-C0A1D1AE5F29	☁️	07/09/2023 13:52	File folder
📁 68248786-5E1E-4F6B-9EFB-1676652BA62E	☁️	18/09/2023 09:32	File folder

Parquet file format

Compressed "columnar"



Name	Status	Date modified	Type	Size
🔹 0B4F17F5-B996-4682-A506-C308AA33A1B7.parquet	☁️	18/09/2023 09:35	Apache Parquet Vi...	9 KB
🔹 0B81A77B-A3CB-4EE5-B317-E3E8238FA63C.parquet	☁️	18/09/2023 09:35	Apache Parquet Vi...	3 KB
🔹 0B624A7C-BD19-4EA8-980A-B80BEC5E5182.parquet	☁️	18/09/2023 09:33	Apache Parquet Vi...	5 KB
🔹 0B6244C8-99C4-414E-9F42-3F133911C3C4.parquet	☁️	18/09/2023 09:35	Apache Parquet Vi...	9 KB

# Loading Options

## Code

SQL COPY INTO...

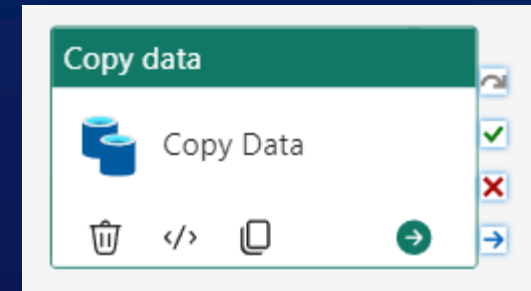
SQL CREATE TABLE AS...SELECT

SQL INSERT INTO...VALUES / SELECT

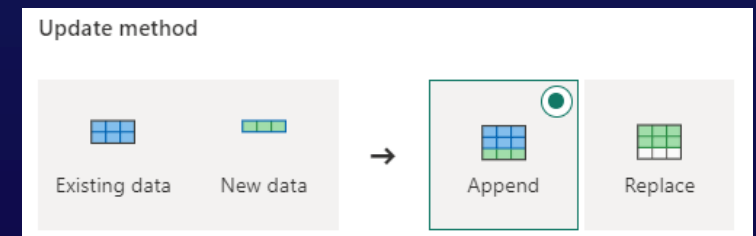
Pyodbc (from Notebooks)

## Low/No-Code

Pipelines



Dataflows Gen2  
(Powerquery FTW!)



# Demo



**Thank You**

# References

- [Azure SQL Data Warehouse is now Azure Synapse Analytics | Azure Blog | Microsoft Azure](#)
- [Parallel Data Warehouse \(PDW\) benefits made simple | James Serra's Blog](#)
- [Parallel Data Warehousing \(PDW\) Explained | James Serra's Blog](#)
- [Microsoft Closes Acquisition of DATAlegro – Stories](#)
- [DATAlegro Placed in Visionaries Quadrant for Data \(globenewswire.com\)](#)
- [Data warehouse appliance – Wikipedia](#)
- [Microsoft Closes Acquisition of DATAlegro – Stories](#)
- [Ted Kummert: Microsoft Business Intelligence Conference – Stories](#)
- [Microsoft Business Intelligence Conference 2008 – Stories](#)
- [Fabric decision guide - choose a data store - Microsoft Fabric | Microsoft Learn](#)
- [Data Warehousing in Microsoft Fabric - Microsoft Fabric | Microsoft Learn](#)

# Demo

- Open Fabric
- Enable Fabric with specific user group
- Show Azure capacities
- Create Workspace
- Create Warehouse and show what's inside (go grab the URL)
- Open SSMS to show objects
- Open ADS project and deploy project
- Create constraints and foreign keys (copy from synapse workbook)
  - Show relationships in power bi
  - Show connection method and talk about import (workspaces that are not fabric)
- Populate with data in various ways (perf between each method?)
  - COPY INTO – DimDate
  - INSERT INTO...VALUES – DimEvent
  - PIPELINE DimProducts
  - Copy from Lakehouse FactWebTelemetry
    - Create lakehouse and create shortcut to lakehouse
- Clone table (while fact table loading)
- Create report in Power BI
- Show transactions
- Show Capacity Usage Metrics