Microsoft Fabric
Unveiling Symphony of Connectivity

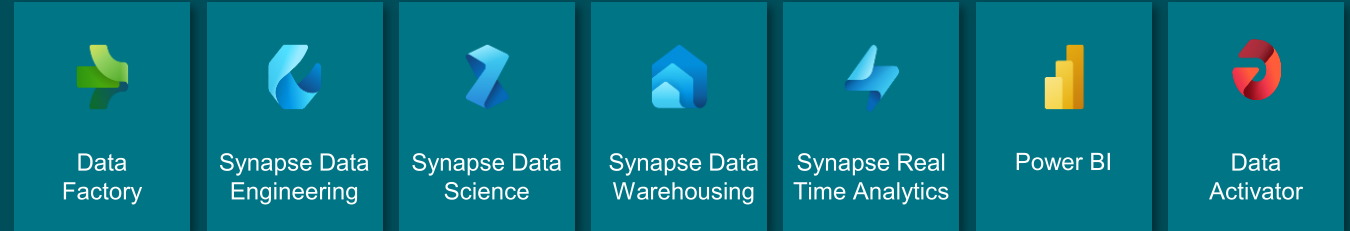# What is Microsoft Fabric?

Value Proposition

Microsoft Fabric is an all-in-one analytics solution for enterprises

Single integrated environment for data engineering, analytics, and data science

A unified data lake that allows you to retain the data

Centralized administration and governance across all experiences

Data Factory	Synapse Data Engineering	Synapse Data Science	Synapse Data Warehousing	Synapse Real Time Analytics	Power BI	Data Activator

AI Assisted

Shared Workspaces

Universal Compute Capacities

OneSecurity

OneLake

Intelligent data foundation

Persistent

Important Drivers for Microsoft Fabric Adoption and Key Use Cases

Business Drivers

\ Need of a 360-degree view of business entity to create impactful applications

\ Quicker time to insights to improve operational efficiency / enable better customer outcomes

Organizational Drivers

\ Self-service data access capabilities increasing business agility and speed

\ Unified data format shared between data engineers and data consumers improves collaboration fostering better alignment

Data Management Drivers

\ Offers diverse data transformation/ delivery methods: ETL, data virtualization, data streaming, change data capture, and APIs for enterprise data access.

\ Unified governance layer

Microsoft Fabric use cases

Enterprise AI implementation

\ Enterprise AI demands a solid foundation of trusted data.

\ Microsoft Fabric ensures seamless integration and governance, enabling AI builders to effortlessly acquire, prepare, and organize data for swift access

Data product strategy

\ Data products requires a balance between streamlined data management and flexible scalability.

\ Leveraging Microsoft Fabric automates the creation and maintenance tasks associated with these products.

Mergers & acquisitions

\ A Microsoft Fabric foundation allows seamless addition of new data sources, enhancement of data catalogs, and removal a subset of data without changing the architecture (carve outs / spell offs / spin-offs).

Unified data governance

\ Microsoft Fabric provides a centralized place for managing data security and privacy.

\ It can help apply consistent data masking, encryption, and anonymization policies across all data sources.

Persistent

Our Microsoft Fabric Offerings in Brief

Unified Data Storage with OneLake

\ OneLake as a single data foundation across all Microsoft Fabric workloads.

\ Shortcuts for virtualized access to external data sources without duplication.

\ Delta Lake format ensures open and optimized data management.

Lakehouse Architecture with Microsoft Fabric Lakehouse

\ Integrated SQL & Spark engines for structured and unstructured data.

\ Direct Lake mode for Power BI eliminates refresh delays.

\ Supports T-SQL & Spark for diverse analytics workloads.

Real-time Data Processing with Microsoft Fabric

\ Low-latency event ingestion from multiple sources like IoT, Kafka, and Event Hubs.

\ Direct integration with KQL databases for instant insights.

\ Real-time transformation and enrichment using built-in capabilities.

AI & ML-Powered Analytics with Microsoft Fabric

\ Integrated notebooks (Spark & Python) for advanced AI/ML workflows.

\ MLflow for model tracking, training, and deployment within Microsoft Fabric.

\ Prebuilt AI models & AutoML enable faster insights generation.

Business Intelligence at Scale with Microsoft Fabric Power BI

\ Direct Lake mode for high-speed analytics without data duplication.

\ Copilot-powered insights enable AI-assisted report generation.

\ Self-service BI capabilities empower business users with governed access.

Comprehensive Data Governance with Microsoft Fabric & Purview

\ End-to-end data lineage tracking across Microsoft Fabric workloads.

\ Centralized policy enforcement for data security & compliance.

\ Role-based access control (RBAC) & fine-grained permissions.

Cost-Optimized Data Management with Microsoft Fabric

\ Serverless, pay-as-you-go model minimizes operational overhead.

\ Unified billing across Azure & Microsoft Fabric workloads for better cost control.

\ Data deduplication & compression improve storage efficiency.

Cross-Cloud & Hybrid Data Connectivity with Microsoft Fabric

\ Federated queries enable real-time access to external data sources.

\ Supports hybrid cloud scenarios for enterprises with multi-cloud strategies.

\ Interoperability with Databricks & Snowflake for cross-platform analytics.

Persistent

When to Consider Microsoft Fabric and When Not to

👍

Highly interrelated data but it is not unified
Challenges in data discovery/data sharing due to data silos between different teams
Multiple data assets/ data lakes/ data platforms/ data warehouses in the organization (Complex governance)
Data is stored across multi cloud environments/ hybrid environments

👎

Not enough data available in terms of the variety and velocity
Current architecture and data landscape is not very complicated
Limited data integration needs
Have working centralized Data Platform aggregating all data across the enterprise

Persistent

Microsoft Fabric Vs Traditional Data Stack Design

	Microsoft Fabric	Traditional DIY stack
Integration & Ingestion	• Microsoft Fabric built in capabilities : Data Pipelines, Data Flows, and Copy commands, cross-warehouse ingestion • Ease of integration with variety of source system	• Tangled stack of tools, adhoc pipelines, integration code
Data Transformation	• Multiple compute options like Spark, Synapse SQL, Databricks. • Universal Compute - a single pool of computing power that can be used for multiple different analytics tasks	• Flexibility of choice of transformation tools like Informatica but may involve additional investment around licenses and skills
Storage	• Microsoft Fabric recommends OneLake as a single, unified, logical data lake that is built on ADLS Gen2 leveraging open-source Parquet and Delta formats	• Traditional Datalake can support multiple formats to store data but effectively storing and managing data could be a challenge
Reporting	• Enterprise-grade reporting features supplied using Power BI	• Flexibility with choice of reporting tool
Data Virtualization	• Microsoft Fabric shortcuts allows users to reference data across clouds and OneLake without copying it. This feature acts as a virtualization layer to existing file data	• May need to invest in external data virtualization tools like Denodo, Actifio, Dremio, Atscale
Data Governance	• Microsoft Purview hub is a centralized place which helps Fabric administrators and users manage and govern their Fabric data estate	• Lacks unified data governance layer due to breadth of tools & technologies in overall architecture
Cost	• 'Pay as you Go' based on CUs • Eliminates the complexities and costs associated with separate billing for individual services	• Separate costing for individual tools & services can lead to challenge in cost apportionment/optimization

Persistent

Microsoft Fabric & Databricks

	Microsoft Fabric + Databricks (Unity Catalog + ETL)
Integration & Ingestion	• Raw data ingestion in Fabric OneLake through Fabric's data factory • Databricks ingestion can happen through ADF in Data Lakehouse raw layer
Data Transformation	• Databricks is more proven and matured tool for data transformation for both Data engineering and AI/ML related use cases. • Databricks also provides capabilities such as Delta Live tables etc. which are typically meant for incremental CDC processing tailor made for medallion architecture • Compute heavy transformation jobs can continue to run on Databricks • Phased approach of developing and migrating specific use cases on Fabric based on cost benefit analysis
Storage	• Fabric has 'OneLake' built on Azure Data Lake Storage (ADLS) Gen2 as a primary storage layer on a global workspace • Databricks can have separate ADLS Gen2 and can push/pull data into OneLake
Reporting	• Fabric has tight Integration with Power BI via OneLake • For Databricks use cases, transformed data can be accessed in Fabric OneLake through shortcut
Data Virtualization	• Fabric shortcuts allows users to reference data across clouds and OneLake without copying it • Lakehouse Federation from Databricks (In Public Preview)
Data Governance	• Microsoft Purview hub is a centralized place which helps Fabric administrators and users manage and govern their Fabric data estate • Unity catalog can be used to create metastores with Data governance rules for Databricks workloads

Persistent

Microsoft Fabric Vs Snowflake Vs Databricks

	Microsoft Fabric	Snowflake	Databricks
Summary	All-in-one analytics platform providing a unified environment for data engineering, data science, machine learning, and business intelligence	A cloud-based platform designed to modernize data warehousing catering to caters to various data applications, including data warehousing, data lakes, data engineering, data science	Unified analytics platform, built on top of Apache Spark providing variety of features for data processing, data warehousing, and machine learning
Data Ingestion	Data Pipelines, Data Flows, and Copy commands, cross-warehouse ingestion	Snowpipes, Streaming, COPY command and third-party tools like Mitillion, Fivetran	Auto Loader, Copy Into and network partners like Azure Data Factory, Fivetran, Qlik, Infoworks, StreamSets and Syncsort
Storage	OneLake as a single, unified, logical data lake that is built on ADLS Gen2 leveraging open-source Parquet and Delta formats	Scalable Cloud blob storage, hosts data, tables, and query results, is designed to scale independently from compute resource	Allow data to be held in various formats on cloud object storage (S3, ADLS, or Google Cloud storage) with a separate processing layer
Data Processing	Multiple compute options like Spark, Synapse SQL, Kusto DB. Universal Compute - a single pool of computing power that can be used for multiple different analytics tasks	Known for high-performance SQL-based data warehousing solution, with support for complex queries, transactional processing	Variety of data processing capabilities beyond SQL, including real time stream processing, machine learning using the power of Apache Spark
Reporting	Enterprise-grade reporting features supplied using Power BI	Basic support for building dashboards and charts	Basic support for building dashboards and charts
Data Sharing	Not a default option with Microsoft Fabric. Need to use Azure Data Share service for data sharing	Provides data sharing with other Snowflake accounts using secure data sharing. Data monetization capabilities using the marketplace	Delta Sharing for sharing datasets across organizations, without the need for the other organization to have Databricks
AI/ML Support	Microsoft Fabric is empowered by leading-edge AI capabilities, including Co-Pilot, Chat GPT, and AI-driven insights	Snowflake Cortex (native ML environment)	Built-in support for popular ML frameworks, distributed ML libraries, and collaborative UI
Verdict	**If you're looking for an all-encompassing analytics ecosystem that integrates seamlessly with Azure services, Microsoft Fabric is your go-to platform** **It is good choice for enterprises that require a unified environment for data engineering, machine learning, and business intelligence**	**Snowflake is best suited for SQL-like business intelligence applications and provides better performance.** **It is designed for scalable and efficient handling of large volumes of data primarily focused on data warehousing and analytics for structured data using ETLs**	**If focus is on big data processing and machine learning with optimized Apache Spark performance, Databricks is the platform to go for.** **It offers a cloud-agnostic approach, available on AWS, Azure, and Google Cloud, and provides specialized features like Delta Lake for data reliability and MIflow**

Persistent

Build enterprise data platform (OneLake) using Microsoft Fabric & AI

Customer is a global leader in above-ground storage tank solutions provider

Challenges

\ Siloed Data - Operational data scattered across different systems (Salesforce, ServiceNow, ADP etc.)

\ No seamless flow of data due to lack of integration

\ Reactive approach due to lack of access to real-time unified data – Loss of Revenue

\ Limited scalability due to on-prem legacy system

\ Absence of built-in AI use cases hindering customers ability to take informed decisions

## Solution

\ Built a modern cloud data platform to enable seamless data integration, processing and analysis to drive operational excellence and innovation

\ Leveraged Microsoft Fabric for scalability, reliability and cost-effectiveness

\ Medallion architecture for data refinement and enrichment

\ Establish a semantic data layer to provide a unified view of data for downstream applications

\ Unified Data Governance – Data Security, Data Lineage, Data Quality etc.

Unified Data Platform
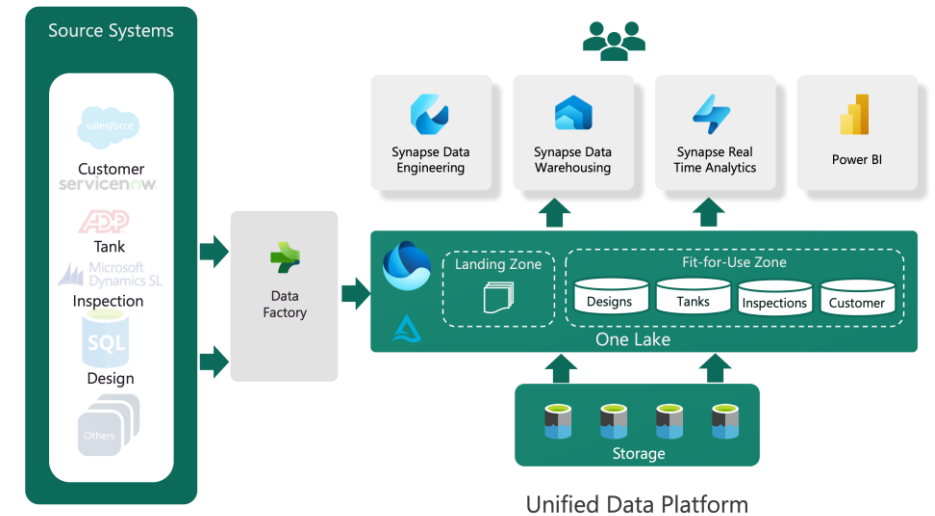
Benefits
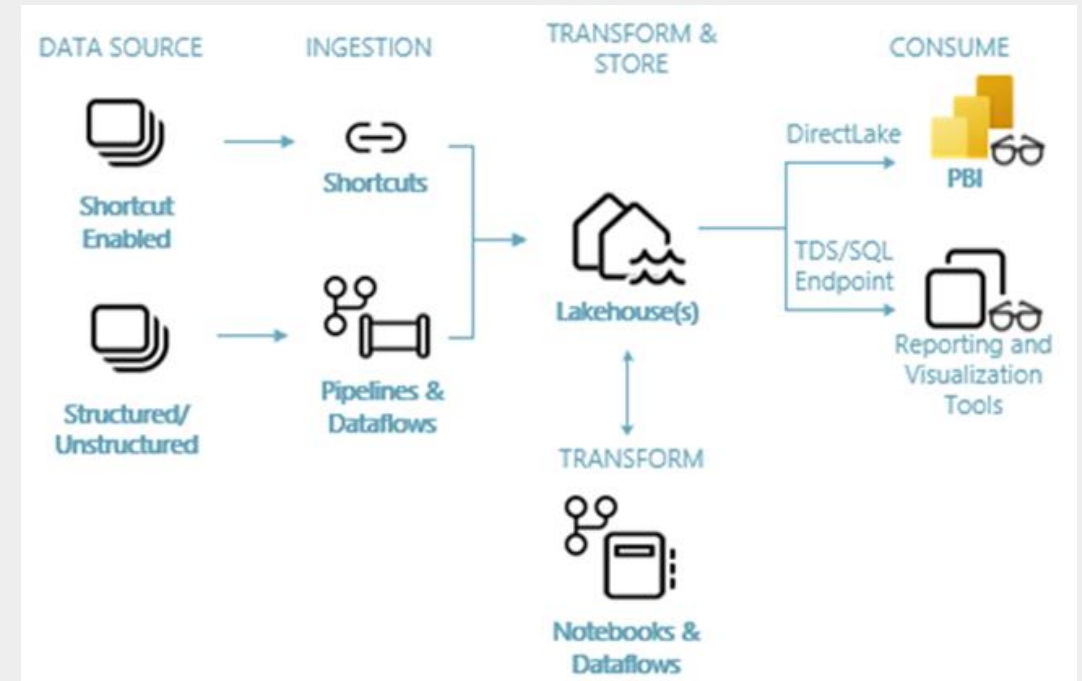
\ Centralized Datawarehouse for analytical workloads

\ Serves both data professionals & citizen developers

\ Operational efficiency and Cost optimization

Persistent
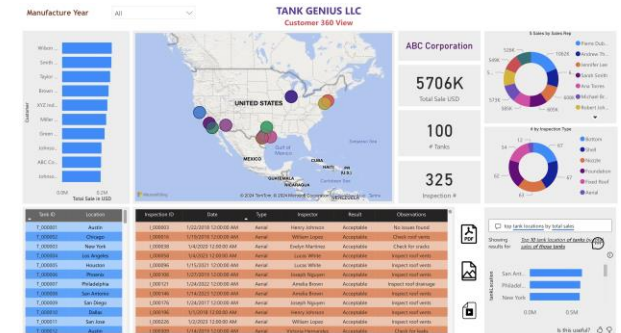
Microsoft Fabric – Next Steps

Next Steps –Demonstrate Proof-of-Value through use case implementation

\ Identify business use case for pilot implementation

\ Provision Microsoft Fabric Resources

\ Test Data (schema/tables/files etc.) preparation in ADLS Gen2

\ Build cloud data platform
  - o Data Ingestion
  - o Data transformation - Medallion Architecture – Delta table creation – Bronze, Silver, Gold layer
  - o Build sample dashboard & reports
  - o Access management & security
  - o Data governance

\ Evaluate overall
  - o Ease of implementation (OneLake shortcuts, Dataflow, Copilot)
  - o Performance analysis (Dynamic scaling - Serverless Pool)
  - o Governance and cost etc.

\ Prepare short-term and long-term roadmap based on the assessment result (cost and performance benefits)

Persistent

THANK YOU

Persistent