Microsoft Fabric: The Data and AI Catalyst

Accelerate your evolution

In the dynamic landscape of modern business, data represents a valuable asset waiting to be harnessed. There are many tools and services available from the Data and AI market players, but with its General Availability (GA) launch on 15 Nov 2023, Microsoft has changed the playing field by providing Microsoft Fabric as a SaaS solution to support organisations in extracting value from their data.

So, what is Microsoft Fabric?

At its core, Microsoft Fabric is a SaaS offering, seamlessly integrating individual PaaS tools and services under one cohesive umbrella. It is a holistic end-to-end platform for data and AI solutions, covering all facets of the data lifecycle, from ingestion, processing and storage to analysis and visualisation. It comes with enterprise-grade data security, governance and compliance features ensuring that organisations can safeguard sensitive information, maintain regulatory compliance and confidently manage their data assets. These features include data encryption at rest and in transit, Role-Based Access Control (RBAC), auditing and logging and Data Loss Prevention (DLP).

Key components of Microsoft Fabric

Fabric brings together components from the realm of Azure Data Lake Storage, Power BI, Azure Synapse Analytics, Azure Data Factory and Azure Data Explorer and integrates them into a cohesive solution ensuring they work together seamlessly. These components are notably but not exhaustively:

OneLake

Central to Fabric's capabilities is the robust and secure data storage infrastructure that underpins the platform. It is built on top of Azure Data Lake Storage Gen 2 and provides a single logical data lake for the entire organisation. Only one such lake is created for each Fabric tenant. It supports any type of data, structured, semi-structured and unstructured. All Fabric data items are automatically stored in open delta parquet format, allowing all the Fabric services and compute engines to use the same data across the toolsets. Since all Fabric components can access and interact with the data, it reduces or even removes the need to replicate/duplicate data.

Data Factory

It provides the data integration portion of the platform for the ingestion, preparation, transformation and movement of data across the platform.

There are two primary features under the Data Factory umbrella: Dataflows (built on the Power Query experience) and Data Pipelines (akin to the Azure Data Factory experience). Fabric Data Factory seamlessly integrates these features providing advanced workflow and orchestration capabilities.

Synapse Data Engineering

Data engineering allows for the ingestion, preparation, storage, analysis and interaction with large volumes of data of all kinds. It is underpinned by Spark and the Lakehouse architecture. It uses developer-friendly notebooks (using languages like python, Scala, R or SparkSQL), Spark Jobs and data pipelines for data exploration, transformation and analysis. Lakehouses are provisioned with a read-only SQL endpoint that can be queried using T-SQL.

Synapse Data Warehouse

This aligns to the more 'traditional' data warehouse everyone is familiar with, although this one is a fully-managed SaaS solution. It natively supports the delta parquet format used by OneLake and allows for data interaction using a SQL based interface. It also provides cross-query capabilities across the data in OneLake without the need for data duplication.

Power BI

The well-known and very widely used business analytics service that gives users the ability to create and share interactive reports and dashboards to create actionable insights from the data. With Power BI in Fabric, it is part of the integrated platform and includes capabilities like DirectLake mode where Power BI can directly load parquet-formatted files from the data lake without the need for Lakehouse or Warehouse intermediaries.

Synapse Data Science

Here data scientists can leverage the same secured data in OneLake to create Machine Learning (ML) models and Experiments. Using the Lakehouse item and notebooks with data integration pipelines, data can be explored, experiments and models built and trained, and results used to enrich the data for use in visualisation tools like Power BI.

Synapse Real-Time Analytics

It is a fully managed data analytics platform that has been optimised for streaming and time-series based data. It utilises a KQL database and Kusto query language that is search performant across structured, semi-structured and unstructured data. And the bonus is that it is fully integrated into Fabric, providing a solution for high velocity low latency data volumes that can be accessed and utilised by other Fabric services.

Eventstream also forms part of the real-time analytics. This provides for the ingestion of event-based data from sources such as Azure Event Hubs and Azure IoT Hubs and the transformation and routing of that data to various destinations like a Lakehouse or KQL database.

So why use Microsoft Fabric?

The transition from individual PaaS tool offerings to the combined SaaS offering of Fabric presents an evolution of the data and AI landscape. Disparate PaaS offerings, each addressing specific needs, while providing best-of-breed capabilities, often resulted in fragmented workflows, interoperability challenges and increased complexity. Moving to a SaaS offering such as Fabric, presents an opportunity to benefit from the combined tools and services under one umbrella in terms of:

Simplicity and Integration

It simplifies the procurement, deployment and management of data and AI services. Instead of dealing with multiple platforms and the complexity that comes with it, the Fabric platform provides a single unified experience, accelerating time-to-value and focusing resources on business outcomes.

It allows for data integration across domains, clouds and accounts without physical data movement through the use of shortcuts, virtual links to other data storage locations that allow for data interaction as if the data was physically in Fabric, reducing the latency normally associated with data ingestion.

Scalability and Performance

As Fabric is a SaaS offering, it leverages the inherent scalability and performance of the Azure cloud, enabling responses to evolving demands. There are currently 11 Fabric capacity options available, allowing for scalability as compute needs grow. Coupled with the automatic Bursting (temporarily using more compute than purchased) and Smoothing (spreading the evaluation of over-usage compute) it provides an easier management experience.

Consistent User Experience

Fabric provides a consistent user experience across all the tools and services provided. Transitioning between personas and related tools and services is seamless and provides an interface familiarity that teams can leverage to become more efficient.

Operational Efficiency

Instead of maintaining and updating multiple disparate platforms, Fabric provides centralised management, automated provisioning and monitoring through its integrated toolsets and APIs, allowing administrators to automate and reduce the complexity and time required for manual operations.

Conclusion

In the ever-evolving world of data and AI platforms, Microsoft Fabric presents a compelling opportunity to set out on the journey of extracting value and reaping the benefits from your invaluable data assets.
How can Fujitsu help you to get started on the journey?

For organisations that are just starting on their data and AI journey, and for those that may already be on it but are looking for help with the next steps, Fujitsu Data and AI have the expertise to assist.

We have created the Fujitsu Fabric Accelerator, a set of best practices, templates, notebooks and 'in the trenches' know-how that are brought to the client engagement to assist with rapidly standing up the foundational capability to deliver value-add data analytics and reporting.

Once established, Fujitsu can further help you build out the platform to ensure you get the most out of your investment.

Email us today or call 03 9924 3000 to discuss how one of our Fabric specialists can help your business to reap the benefits and extract maximum value from your data.

Contact
Fujitsu Data & AI
+61 3 9924 3000