ASCENT

A GUIDE TO GREAT.

Whitepaper.

Composable IT Architectures for Data Management and Analytics using Microsoft Fabric.

Mohammed Brückner

Contents.

About the author.

linkedin.com/in/mbrueckner
mohammed.brueckner@ascent.io

Mohammed Brückner

Mohammed is an experienced and passionate technologist with a proven track record of solving complex business challenges across a range of industries for more than 17 years. His specialist subject is the evolving world of Platform Economies - the business models and ecosystems that are built around digital platforms. He is also the author of *IT's Not Magic, It's Architecture and The DALL-E Cookbook For Great AI Art: For Artists. For Enthusiasts*.

Introduction: Embracing Agility in Fabric.

Composable architectures have made remarkable progress in achieving agility and scalability. These architectures are modular and flexible, allowing organizations to quickly adjust to changing business needs and scale their data management and analytics capacities with ease. They are sometimes also called 'composable application architectures'.

There are many advantages of using a composable approach. Organizations can have a unified data landscape that merges different data sources into a consistent and synergetic whole. This merger makes end-to-end data analytics more efficient and faster to roll out. Microsoft Fabric's strong security and governance features ensure that data is safe and meets regulatory requirements, even as it moves through the different parts of the architecture.

However, moving to a fully composable IT infrastructure is not easy. A clear transition plan is required to handle the difficulties of integrating new technologies with existing systems. The plan must focus on realistic use cases that show the real benefits of composable architectures, and incorporate best practices learnt from successful implementations, providing a guide for others to emulate.

Also important is the development of a culture of innovation within the organization that promotes constant learning and exploration, creating an environment where new ideas push the enterprise towards higher levels of agility and performance.

Danish insurer, AP Pension*, transformed its data management with Microsoft Fabric and Azure Databricks. This integration streamlined data processes, improved data quality and enhanced analytics, enabling better decision-making and customer service.

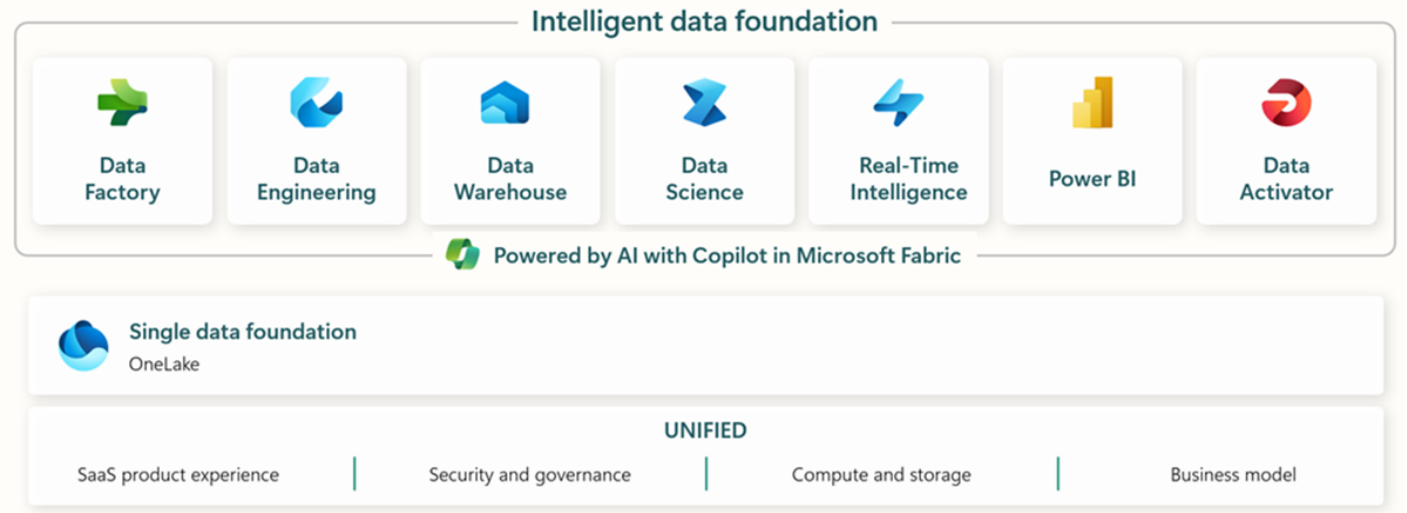
* For further details on the AP Pension case study, please refer to the Microsoft customer story titled "AP Pension revolutionizes its data governance and analytics with Microsoft Fabric" (https://customers.microsoft.com/en-us/story/1771984633872247465-ap-pension-azure-databricks-insurance-en-denmark)

Fabric's Composable Approach to Data Management.

Microsoft Fabric is an opinionated end-to-end analytics solution that transcends the complexities and potential limitations of a custom-crafted solution based on individual components. It deeply integrates the core building blocks of Azure, makes connectivity a core concern and adheres to loose coupling in its design, as you would find as part of any composable application. Fabric is the realization of Microsoft's longstanding Intelligent Data Platform vision, providing a 'single pane of glass' and tight interoperability across transactional and analytic workloads.

"The architectural elegance of Microsoft Fabric is in the innovation that has enabled the platform to reduce complexity while increasing capability."

– **Murray Foxcroft**, Ascent CTO.



Ultimately, Fabric aspires to be the ultimate one-stop shop for all data needs, seamlessly handling data integration, analysis, and storage for a holistic and efficient data management experience.

The lifeblood of any data platform is its ability to move and manipulate data securely, reliably and with agility. Fabric excels in this arena, providing a diverse palette of tools, including:

Data Workflows: Leveraging the power and familiarity of Apache Airflow, Fabric enables the creation and management of Python-based data pipelines. These pipelines, structured as Directed Acyclic Graphs (DAGs), provide a clear and intuitive way to define complex data workflows.

Data Factory: Fabric bridges the gap between the familiar realm of Microsoft's products and services, external SaaS (like Salesforce, ServiceNow, SAP, etc), on-premise implementations and technical capabilities (databases, storage buckets, event hubs, etc).

With an impressive repertoire of over 170 connectors, Fabric's Data Factory acts as a universal translator, seamlessly moving data between diverse environments. Whether data resides in multi-cloud ecosystems, on-premise systems, or within virtual networks, Data Factory ensures smooth and reliable data transit. Data Factory pipelines are the workhorses that automate data movement and manipulation. They can call APIs of microservices, run in containerized environments for flexibility, reach into corporate networks and leverage Azure Service Bus queues or topics for asynchronous processing and loosely coupled data flows between stages. This orchestration power allows you to build powerful pipelines, for example retrieving sales data from Dynamics 365 and combining it with customer data from SharePoint to create a holistic view for data-driven decisions.

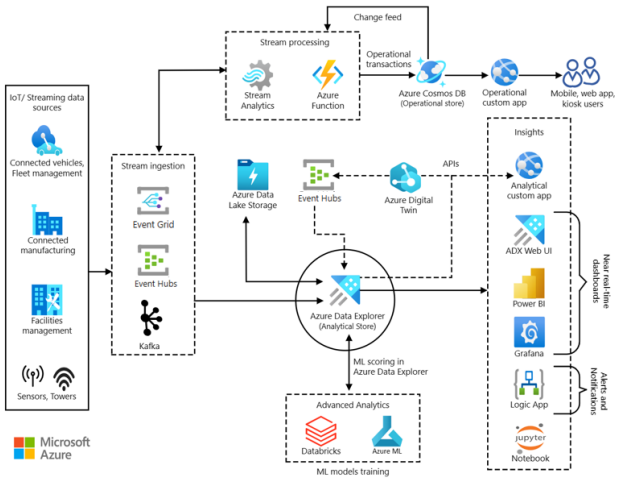
Lakehouse Management: Fabric empowers users to directly manipulate data within their Lakehouse. Pipelines can be crafted to orchestrate Spark jobs to enable batch processing, and notebooks provide a flexible environment for data preparation, and transformation.

Underpinning all data exchange mechanisms are Fabric's REST APIs. These APIs provide a robust foundation for automation and embedding Fabric's capabilities into broader workflows and applications. Fabric therefore itself may become the eponymous fabric of a modular and re-usable application/system architecture.

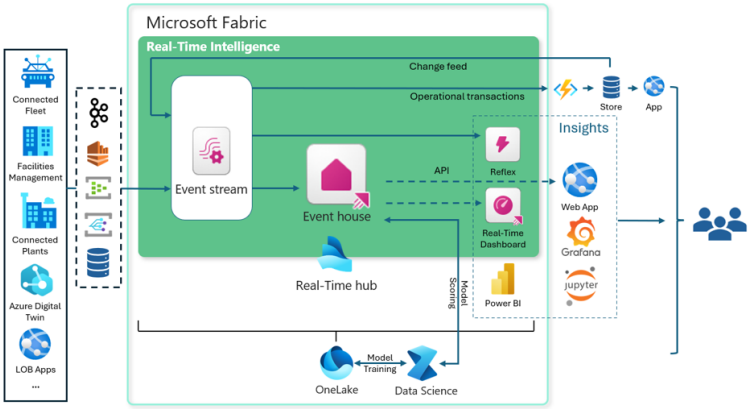
The cornerstone of Fabric data orchestration is the already mentioned Data Factory, meticulously conducting the movement and transformation of data throughout its lifecycle. From the moment data enters the system to the moment it reaches its destination, ready for analysis, Data Factory ensures each step is executed with precision.

Real-Time Intelligence: A New Layer of Sophistication.

Azure PaaS-based solution architecture



Fabric Real-time Intelligence solution architecture



Microsoft Fabric's eventhouse is a crucial component of Microsoft Fabric's Real-Time Analytics suite, an advanced database workspace designed to handle and analyze large volumes of event-based data in real-time. It enables efficient management of streaming data for various applications, including telemetry, log data, IoT data, and financial records, providing another layer of sophistication in Microsoft Fabric. This new workload, announced at Build 2024, empowers enterprises to harness the power of high-volume, real-time data flowing into the platform via Event Hubs and Streams, with minimal latency. It also eliminates the need for complex data pipelines in certain scenarios.

Real-Time Hub acts as a central nervous system, ingesting, processing, and routing events from various data sources across the organization to any data store within Fabric. Real-time dashboards empower multiple users with immediate insights, while Microsoft Copilot integration within Fabric can further optimise the process by generating queries and deriving insights from the streaming data.

Further Enrichment: Azure Databricks.

Further enriching its data processing capabilities, the Microsoft Build 2024 conference revealed that Fabric will integrate with Azure Databricks better than ever - Azure Databricks Unity Catalog tables will be directly accessible within the Fabric portal, allowing users to configure and manage them directly in OneLake. This integration allows Fabric users to tap into the power of lakehouse data connected to Azure Databricks, providing streamlined cross-platform data access and analysis.

Fabric will also support industry standards like Apache Iceberg, allowing for better integration with Snowflake data storage. This makes Snowflake just as compatible as Azure Databricks. Free storage mirroring is the 'icing' on the cake here, making it cheaper to include external data lakes that use industry standard like Iceberg and Delta.

Central to Fabric's data management philosophy is OneLake – a unified data lake that serves as a central repository for all data assets. This centralized approach eliminates data silos and simplifies data management by providing a single source of truth.

Built-in metadata services further enhance OneLake's capabilities. These services provide critical contextual insights and schema intelligence, enabling sophisticated analytics and robust governance models.

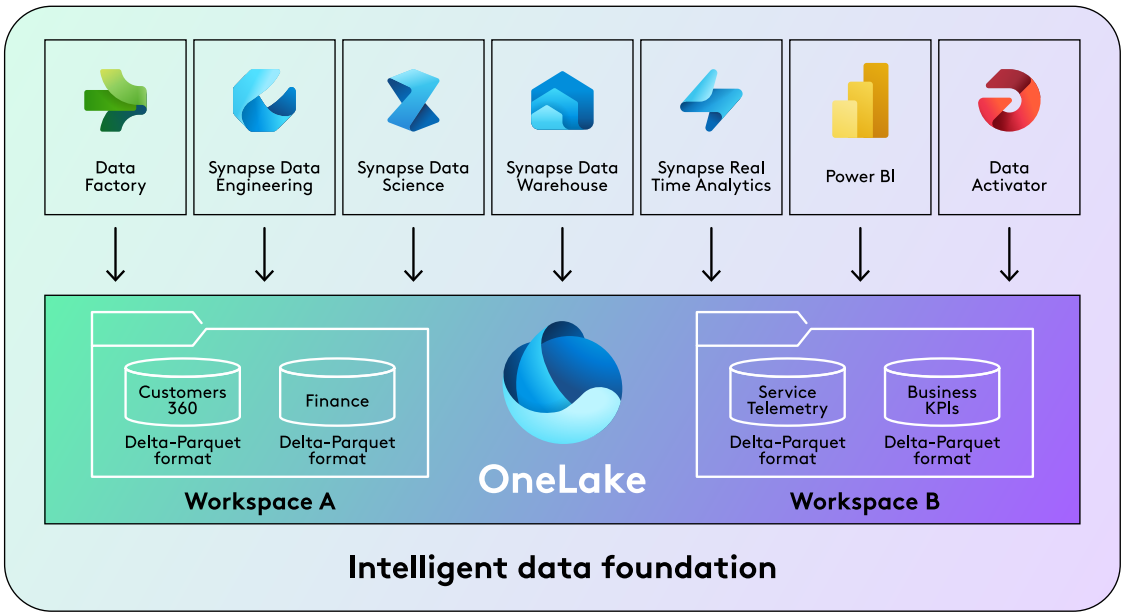
OneLake: The Source of Data Insights in Fabric.

OneLake is an intelligent data storage ecosystem that is part of Microsoft Fabric's composable data management framework. It goes beyond the boundaries of traditional isolated systems by providing a unified and open lake architecture.

Grounded in the Delta Lake format for effective analytical workload processing across different systems, OneLake improves compatibility and reduces complexity through standardized storage practice and flexible connectivity.

OneLake also has built-in metadata services that provide contextual insights and schema intelligence that are vital for advanced analytics and governance models.

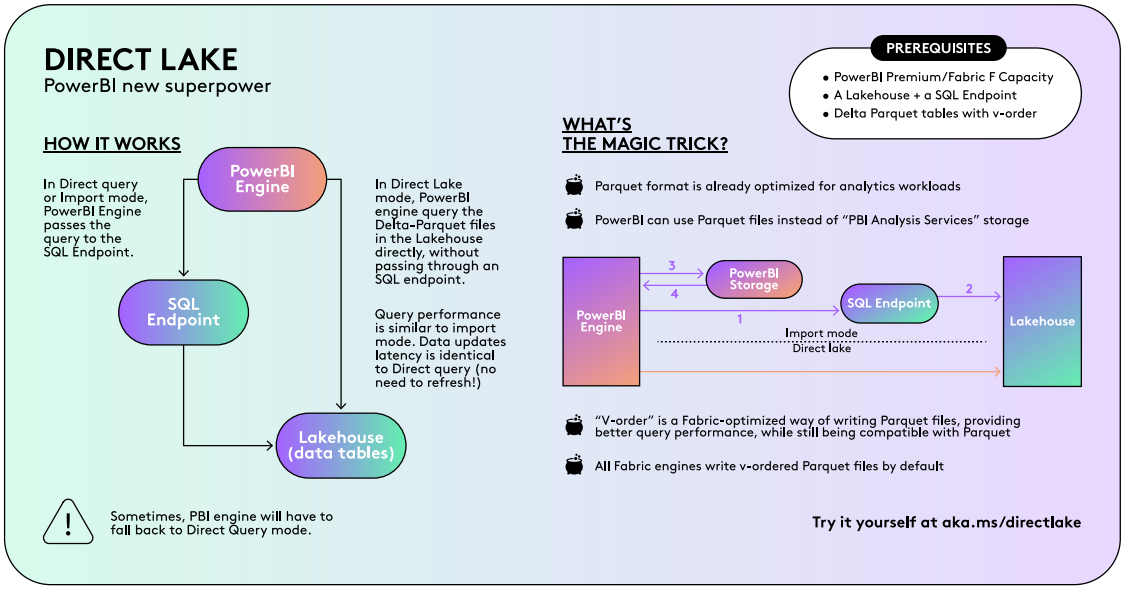
This platform is key in creating powerful pipelines that use deep insights from data assets to improve machine learning workflows and ensure strict compliance with regulatory frameworks.



To maximize OneLake's potential, certain best practices are recommended for the entire architecture dealing with this central go-to storage:

- **Product-Led Approach:** Embrace a product mindset to build and manage data applications in OneLake, focusing on incremental delivery aligned with agile methodologies.
- **Domain Modeling:** To organize your data aligned to your business.
- **Lifecycle Management:** Implement source control and CI/CD practices for quality management, coauthoring, versioning and well managed incremental feature releases and enhanced data application quality.

One of the most interesting capabilities that OneLake empowers is **Direct Lake**, as illustrated below:



This dramatically simplifies the traditional complexities of data architecture by enabling real-time direct access and manipulation of data stored in its Lakehouse structure.

By connecting Microsoft Fabric's OneLake and Power BI, Direct Lake offers these advantages:

- Unlike import mode in Power BI that requires a data refresh, Direct Lake provides near real-time access to the data in OneLake
- Although resembling DirectQuery, Direct Lake does not penalise massive datasets thanks to the use of optimized Parquet files specifically designed for efficient querying by Power BI's Vertipaq engine
- Power BI struggles with importing/ manipulating extremely large datasets. Direct Lake changes that, and the same goes for the need of copying and storing data within the Power BI model itself, reducing storage requirements and streamlining analysis.

All of this optimizes analytic workflows with low-latency interactions while maintaining strict governance and security through Azure's advanced permission controls, striking a balance between accessibility and compliance for enterprises handling large datasets.

Microsoft Fabric Governance and Security.

Use Fabric and supporting services to:

- Establish a robust governance framework
- Integrate security practices
- Connect Databricks Unity Catalog or Microsoft Purview for data management, discovery and classification
- Monitoring and observability are essential for high availability and proactive issue resolution, necessitating comprehensive monitoring and logging.

Collaboration and convergence.

Fabric is the platform that unites people, technology and data in a central managed environment. It handles different types of data, such as operational, analytical, time series and real time in one place, allowing you to manage your entire data estate comprehensively. Moreover, Fabric's Domains feature is essential for modern data architecture, offering a structured, secure, and scalable way to handle and analyze data and align it with your business architecture. Using these domains, organizations can improve their data management abilities, ensure compliance, and generate business value through advanced analytics and insights across the business.

Conclusion: A Brighter Future with Composable Fabric.

Composable architecture is a cultural and technological shift. It requires commitment to cross-functional cooperation — a culture of experimentation — and clear practices for governance and security protocols.

Fabric offers ways to streamline and simplify your data landscape and it is developing quickly. You don't need to discard your existing data platform investment and replace it with Fabric, but for organizations with mature data platforms, Fabric can provide a single interface for your organization's data consumers, enhancing traditional self-service beyond 'build your own reports'.

Fabric is technically highly impressive, and arguably destined to be the most comprehensive, capable data platform available to businesses. At the time of writing, it is still a new product, but its roadmap and evolution will be rich and rapid.

Organizations may be hesitant about Fabric now, but if they keep an open mind, follow its rising star and revisit subject matter expert opinion every few months, they'll transform the business value their data assets offer.

https://learn.microsoft.com/en-gb/fabric/release-plan/

ASCENT

Why choose Ascent?

As Microsoft Azure Data & AI and Cloud experts with advanced specializations and engineering centered around well-architected principles, we are well placed to help you adopt modern data management and analytic techniqes.

We can help with strategy, governance and readiness, additionally providing guidance on applicable Microsoft funding programs that help you to reduce the cost of your data journey.

Together we will:

- Speed up implementation times using highly skilled engineers and efficient deployment models
- Reduce unnecessary overhead so you can focus on what makes your business unique
- Take full advantage of the modular nature of Microsoft Fabric's composable architecture
- Build governance, security, and reliability into your new platform.

Let Ascent's expertise and experience help you:

- Transform from traditional infrastructures to more dynamic ecosystems that meet current digital demands
- Strategize how best to extract the most from your data, from both business and technical perspectives
- Drive tangible results faster thanks to our expertise that covers key areas such as governance mechanisms and strategies for holistic integration between different components of the modern enterprise landscape.