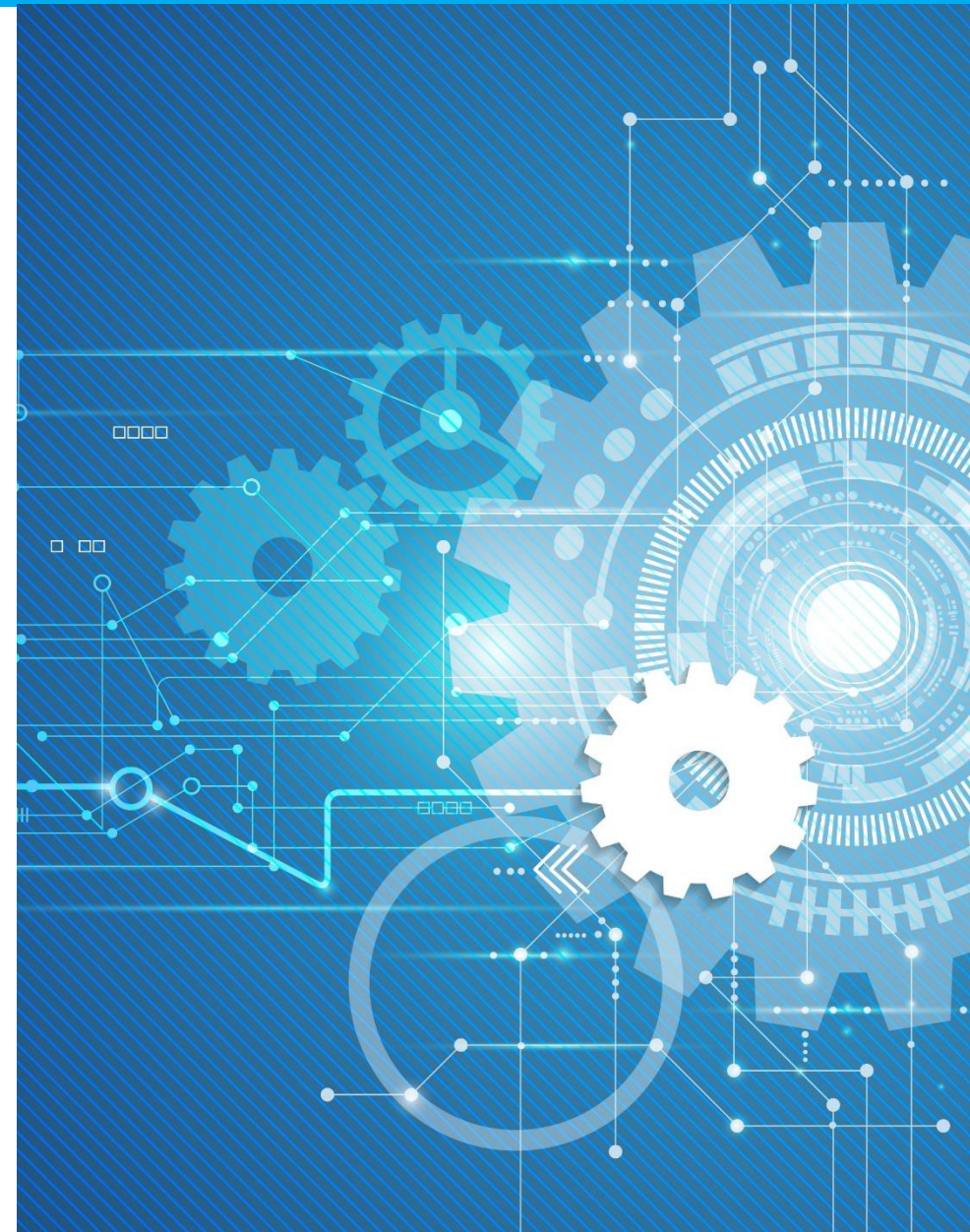


Microsoft Fabric 1:1 Session

Agenda

- Introductions
- The Driving Factors
- The History
- Fabric Overview
- Overview Demo
- Working with Fabric
- Demo
- Fabric Readiness
- Thoughts on Adoption
- Q&A



Spyglass MTG Overview

Expertise, Experience & Excellence

Spyglass is a Specialized Expert Consulting firm focused on helping clients select, architect, implement, migrate to and manage their Microsoft Technology.

- **A Women Owned/Women Led Company**
- **Headquartered in Lincoln, RI**
- **25+ years providing Microsoft Solutions to US clients**
- **25+ years average experience on our team**
- **18+ years as a Gold Partner**
- **5 Solutions Badges**
- **5 Advanced Specializations**
- **FastTrack Ready Partner**
- **90% of our business is repeat business**
- **80% of our new business comes from Microsoft**



Competencies and Specializations



Microsoft
Advanced Specializations

- Teamwork Deployment
- Adoption & Change Management
- Analytics on Azure
- Azure Data Warehouse Migration
- Infra & Database Migration to Microsoft Azure



Microsoft
PowerBI Partner



Microsoft
Solutions Partner

- Infrastructure Azure
- Data & AI Azure
- Digital & App Innovation Azure
- Modern Work
- Security



Microsoft
FAST TRACK
READY PARTNER

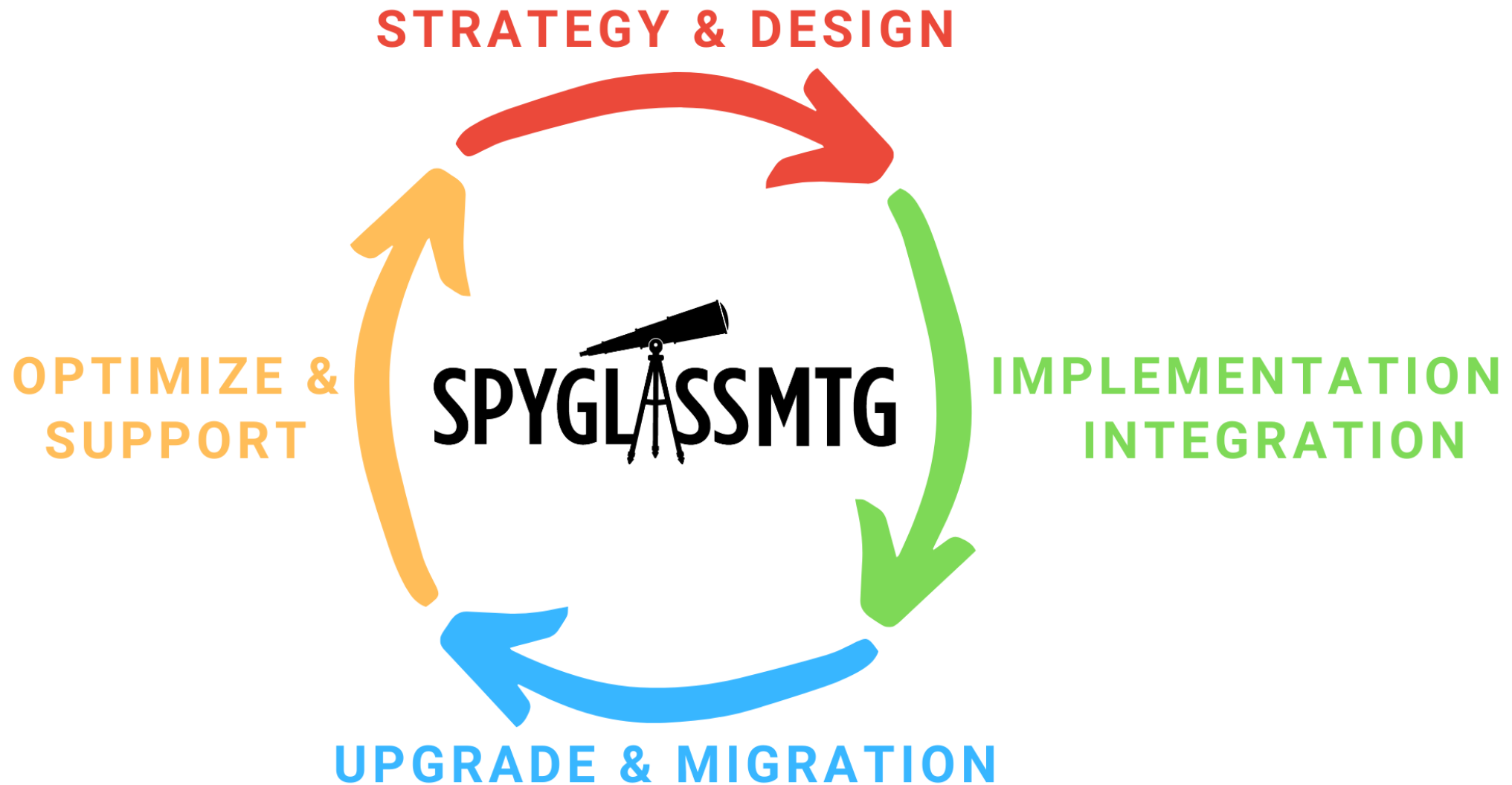


PARTNER SUCCESS SERIES
OEA Advanced
Partner



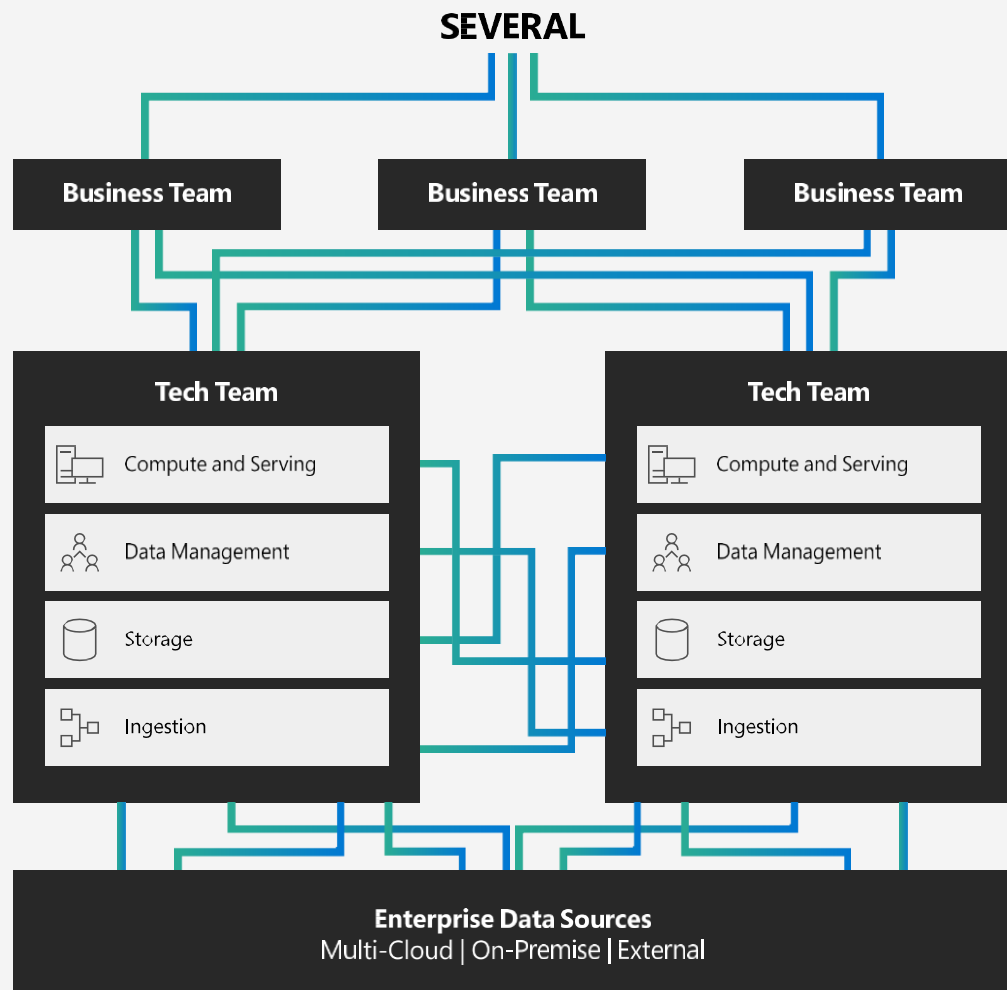
Microsoft
Cloud Solution Provider

Full Lifecycle Microsoft Consulting





The History



An organically evolved data estate?

The most common challenge in enterprise data estates is that data has organically evolved. Companies have data estates with a ton of data and infrastructure redundancy. There are often multiple teams operating silos of data that are not truly connected. The platforms required to transform data into actionable insights are inherently very technical and fragmented. Usually, teams of engineers and developers are needed to deploy and manage these platforms. Organizations have many siloed cloud or

on-premises data sources from different vendors in different formats that hold critical information. It is very difficult to find deep and accurate insights without a single source of truth. Stitching together the unique analytics tools organizations need is complicated. Costs associated with procuring and managing these capabilities can be exorbitant. And there is a significant risk associated with the lack of governance.

- **Every analytics project has many subsystems**
- **Every subsystem needs a different class of product**
- **Products often come from multiple vendors**
- **Integration across products is complex, fragile and expensive**

Data Mesh, Data Fabric, Data Hub

There are three data estate architectures and concepts that organizations are applying to the modernization of their data estate. The data mesh, data fabric, and data hub.

Data Mesh

A series of domains assigned to individual LOBs that enables access to the data they need with maximum autonomy by upholding the four principles of a data mesh.

Data Fabric

A system for automating data management tasks, such as unifying and cleaning disparate sources as well as authorizing data access, that helps a business make the most of its existing data sources without needing to migrate them.

Data Hub

An open and governed lakehouse platform that provides the storage foundation for serving data to multiple domains efficiently, so domain users can easily consume it.

Data Landing Zones

Data landing zone architecture illustrates the layers, their resource groups, and services each resource group contains. The architecture also provides an overview of all groups and roles associated with your data landing zone, plus the extent of their access to your control and data planes.

Data Management

The data management landing zone is a management function and is central to cloud-scale analytics. It's responsible for the governance of your analytics platform.

Data Integration

Each data landing zone has several layers that enable agility for the service data integrations and data products it contains. You can deploy a new data landing zone with a standard set of services that let the data landing zone begin ingesting and analyzing data.

Data Products

Data products are data served as product and computed, saved, and served by polyglot persistence services, which can be required by certain use cases. The process of creating and serving a data product can require services and technologies that aren't included in the data landing zone core services.

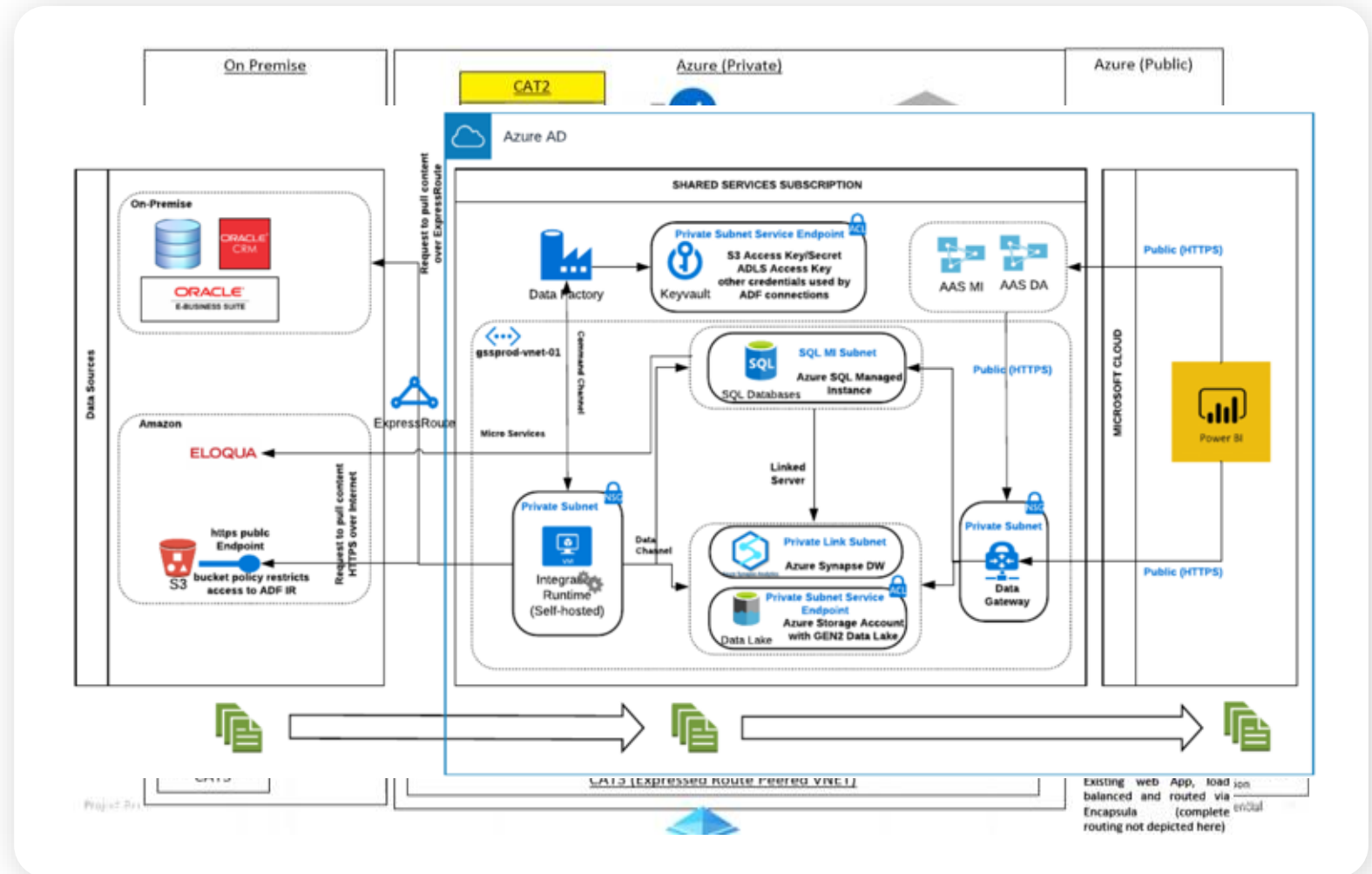
Analytics is complex and fragmented

Every project has many subsystems

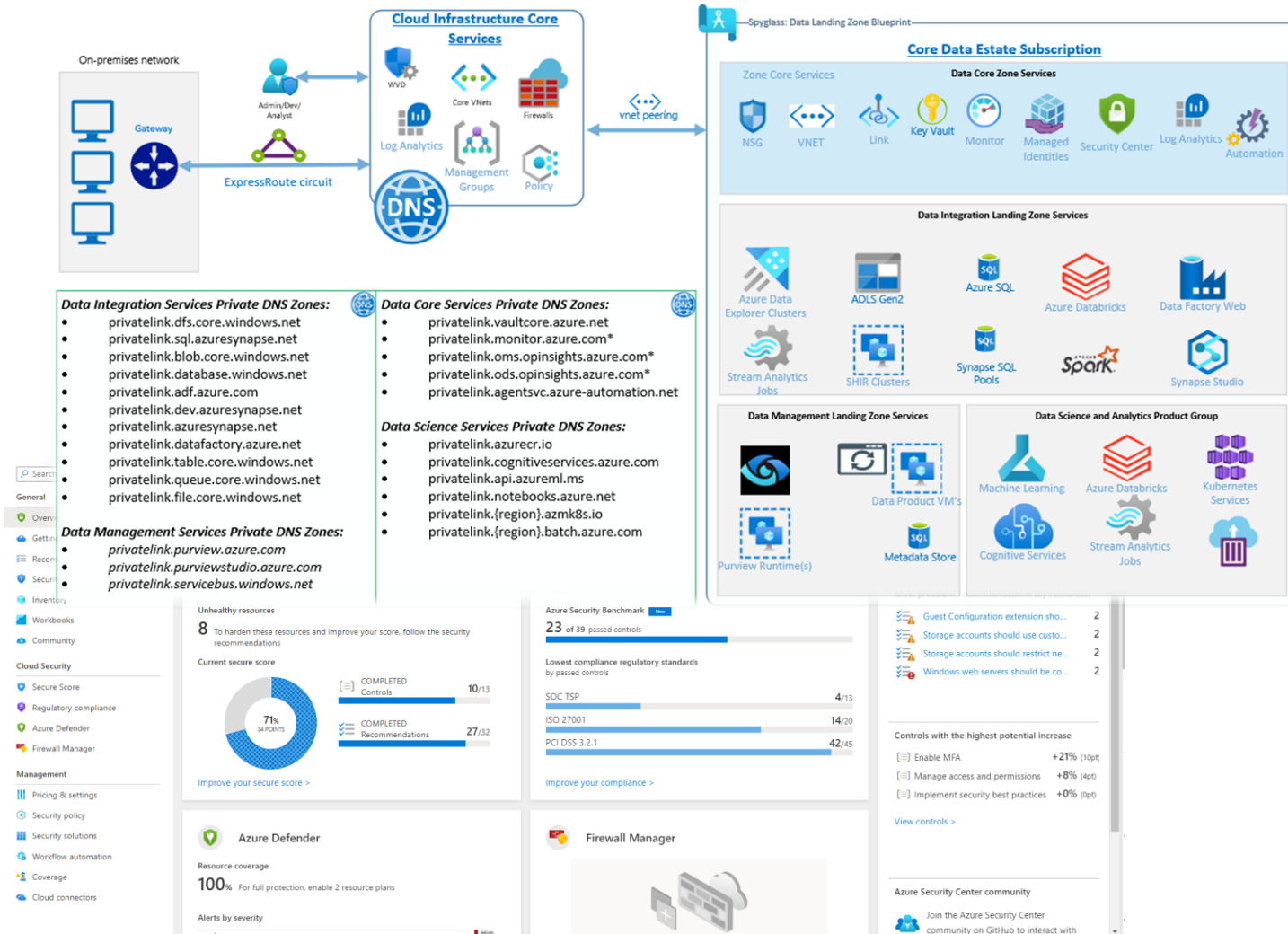
Every subsystem need a different class of product

Products often comes from multiple vendors

Integration is complex, fragile and expensive



Data Landing Zone Example



Layer

Core services

Resource Examples

- Network
- Monitoring
- Shared integration runtimes
- CI/CD Agents
- Security Agents

Core Data services

- Metastore for Data Services
- Data Lake Services
- Data Storage
- Shared integration runtimes
- Apache Spark
- Azure Synapse Analytics

Data Applications

- Azure Machine Learning
- Open AI
- Cognitive Services
- Data Share
- Stream Analytics

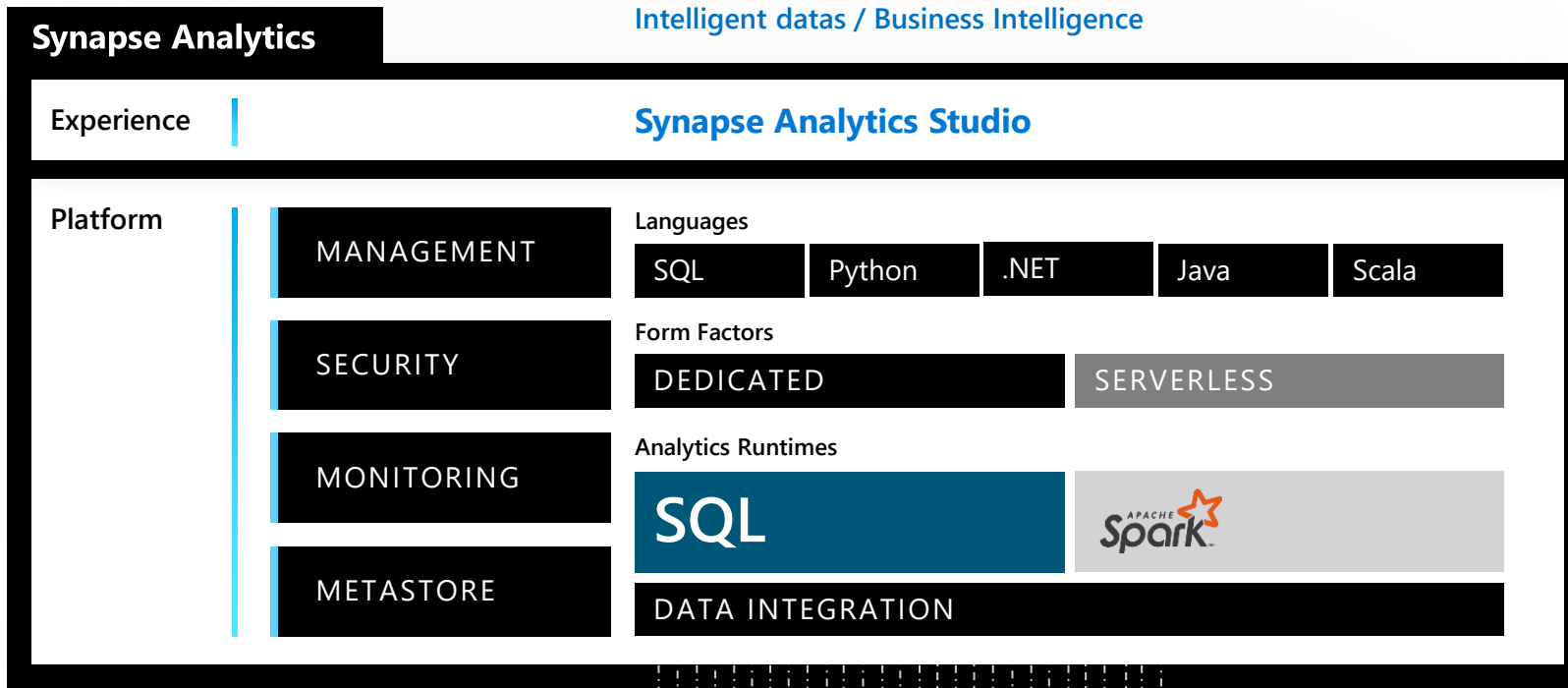
Visualization

- Power BI
- Data Flows

Synapse Analytics: The first unified, cloud native platform for analytics

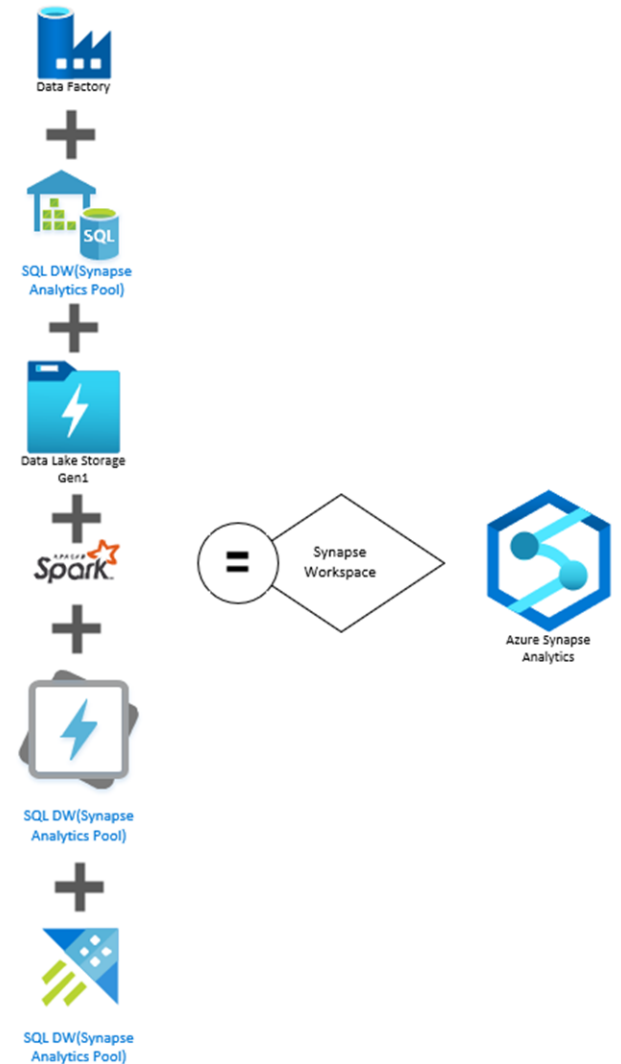
Synapse Workspaces are the integration of a number data service within a legacy Azure Data Landing Zone. What was once a reference architecture of many components has been consolidated into a single data solution.

Artificial Intelligence / Machine Learning / Internet of Things
Intelligent datas / Business Intelligence



Azure
Data Lake Storage

Common Data Model
Enterprise Security
Optimized for Analytics

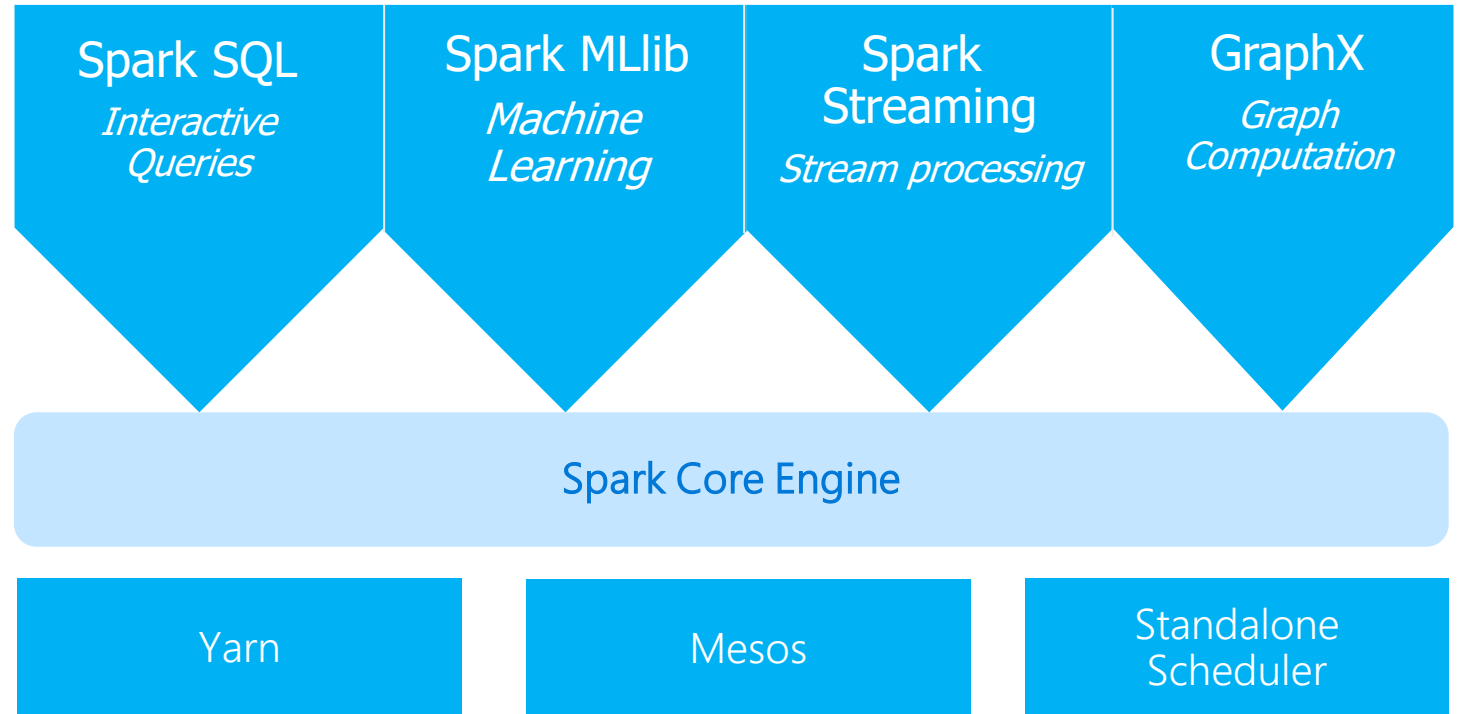


Synapse Provides Apache Spark

A unified, open source, parallel, data processing framework for Big Data Analytics

Spark Unified:

- Batch Processing
- Interactive SQL
- Real-time processing
- Machine Learning
- Deep Learning
- Graph Processing

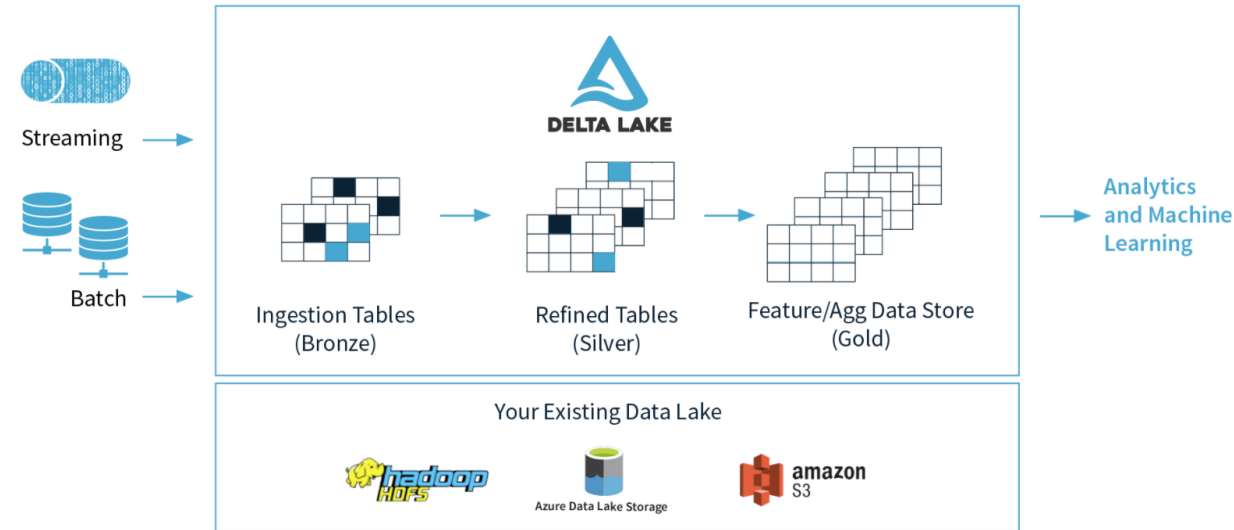


<http://spark.apache.org>

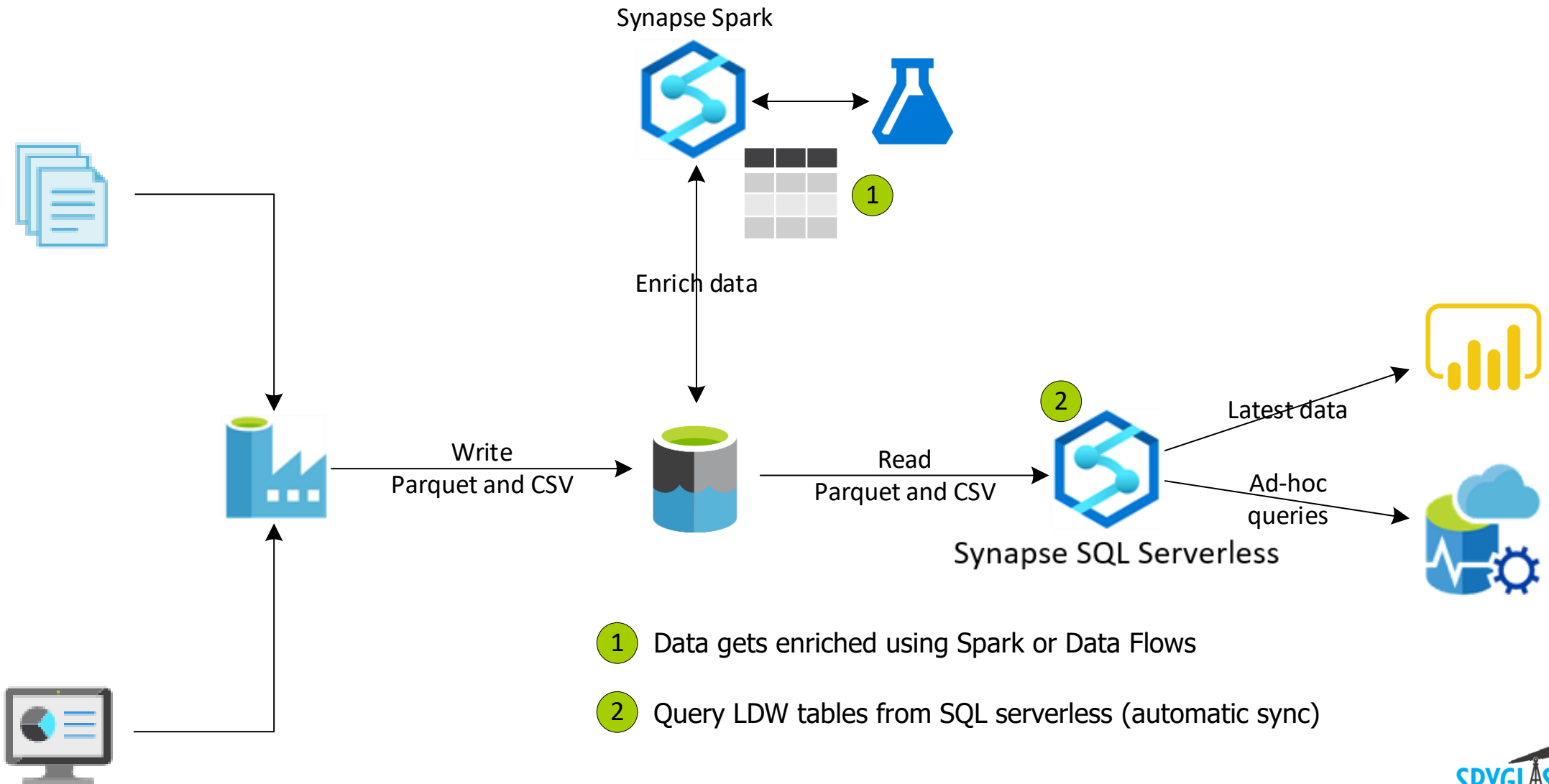
Synapse Adopts Delta Lake

An open-source storage layer that brings ACID transactions to Apache Spark and big data workloads.

- ACID Transactions
- Scalable Metadata Handling
- Time Travel (data versioning)
- Open Format
- Unified Batch and Streaming Source and Sink
- Schema Enforcement
- Schema Evolution
- Audit History
- Updates and Deletes



Usage pattern evolved – Enrich in Spark, serve with SQL



Power BI: the bridge between data and decisions

Unified self-service & enterprise BI platform



Announcing



Microsoft Fabric

The data platform for the era of AI

Next Gen: Microsoft Fabric

A unified analytics solution for the era of AI

Enabling the art of the possible, exceeding customer expectations, and delivering transformative business value just got easier with the introduction of Microsoft Fabric.

Microsoft Fabric brings together **the best parts of data mesh, data fabric, and data hub** to provide **a one-stop-shop for data integration, data engineering, real-time analytics, data science, and business intelligence** needs without compromising the privacy and security of your data. By joining top companies using Microsoft Fabric, you will enable teams to experience an end-to-end, highly integrated, single solution that is easy to understand, onboard, create, and operate.

With this new solution, you will establish a **unified source of truth** by bringing together all analytics workloads in a lake-first foundation. Teams will also be able to reduce the time and effort it takes to uncover impactful insights through **democratized access** to data. And this can be done confidently through a secure, **governed solution**.



Data Integration

This solution offers comprehensive data migration and integration to enable a lake-first pattern. Azure Data Factory connectors enable data integration, while Azure Synapse Link connectors enable “no code” and “always synchronized” data integration for operational databases.



OneLake

All data is ingested into a data lake on Azure Data Storage Gen—**a cost- and performance-optimized data lake storage service**—for the most demanding business intelligence, machine learning, and artificial intelligence workloads.



Analytics

Data scientists can bring their preferred compute frameworks, languages, runtimes, and tools into the data lakehouse and further enhance the data through feature engineering and statistical techniques.



Business Intelligence

Best-in-class integrated solutions to responsibly democratize business intelligence with self-serve tools and experiences for data analysts and data citizens.



Governance

Microsoft Purview then provides a single pane governance solution to help effectively scan and manage your data estate—even as it grows and scales.

A unified SaaS-based solution that stores all organizational data where analytics workloads operate

Microsoft Fabric combines Data Factory, Synapse Analytics, Data Explorer, and Power BI into a single, unified experience, on the cloud. The open and governed data lakehouse foundation is a cost-effective and performance-optimized fabric for business intelligence, machine learning, and AI workloads at any scale.

It is the foundation for migrating and modernizing existing analytics solutions, whether this be data dataliances or traditional data warehouses.

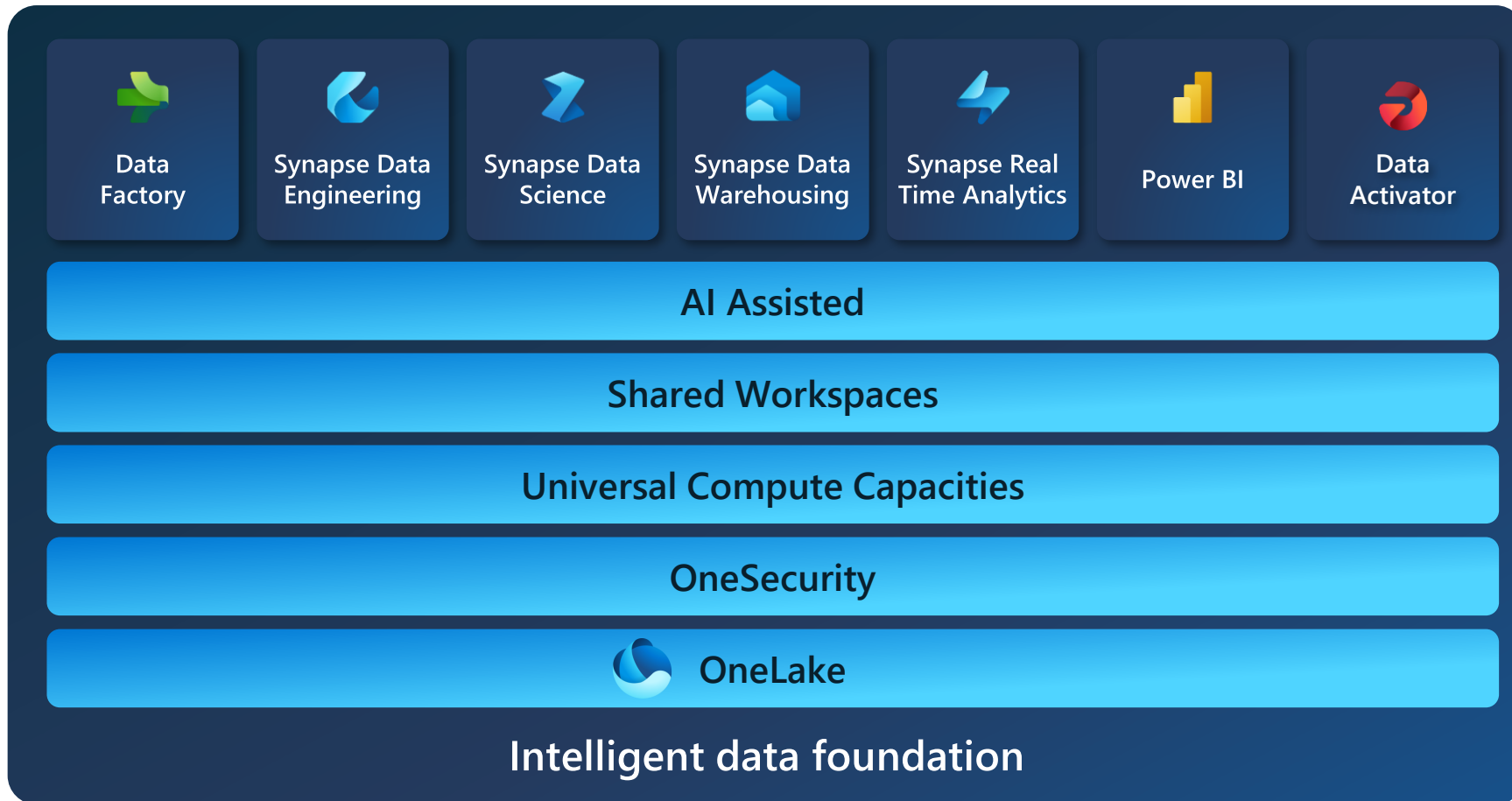
By establishing connectivity and integration, organizations can transform their unstructured and siloed data into a valuable strategic asset through:

- Data modernization backed by the Microsoft Azure Cloud
- Cloud native applications at any scale
- Responsible, powerful AI to make more informed decision-making
- Analytics and insights at a faster rate
- Responsible machine learning and artificial intelligence
- Governance backed by Microsoft Purview





Microsoft Fabric

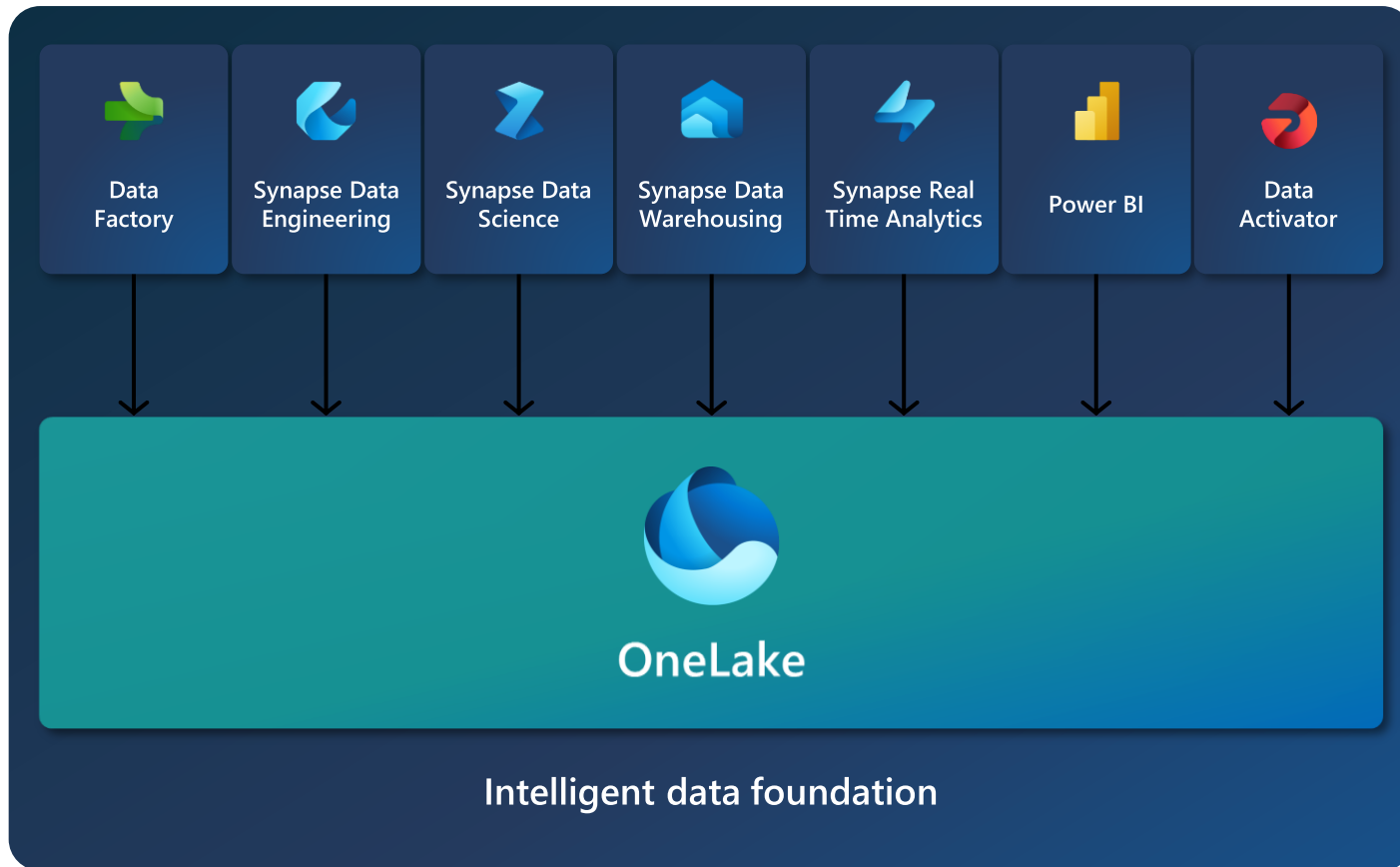


Single...

- Onboarding and trials
- Sign-on
- Navigation model
- UX model
- Workspace organization
- Collaboration experience
- Data Lake
- Storage format
- Data copy for all engines
- Security model
- CI/CD
- Monitoring hub
- Data Hub
- Governance & compliance

OneLake for all Data

“The OneDrive for Data”



A single SaaS lake for the whole organization

Provisioned automatically with the tenant

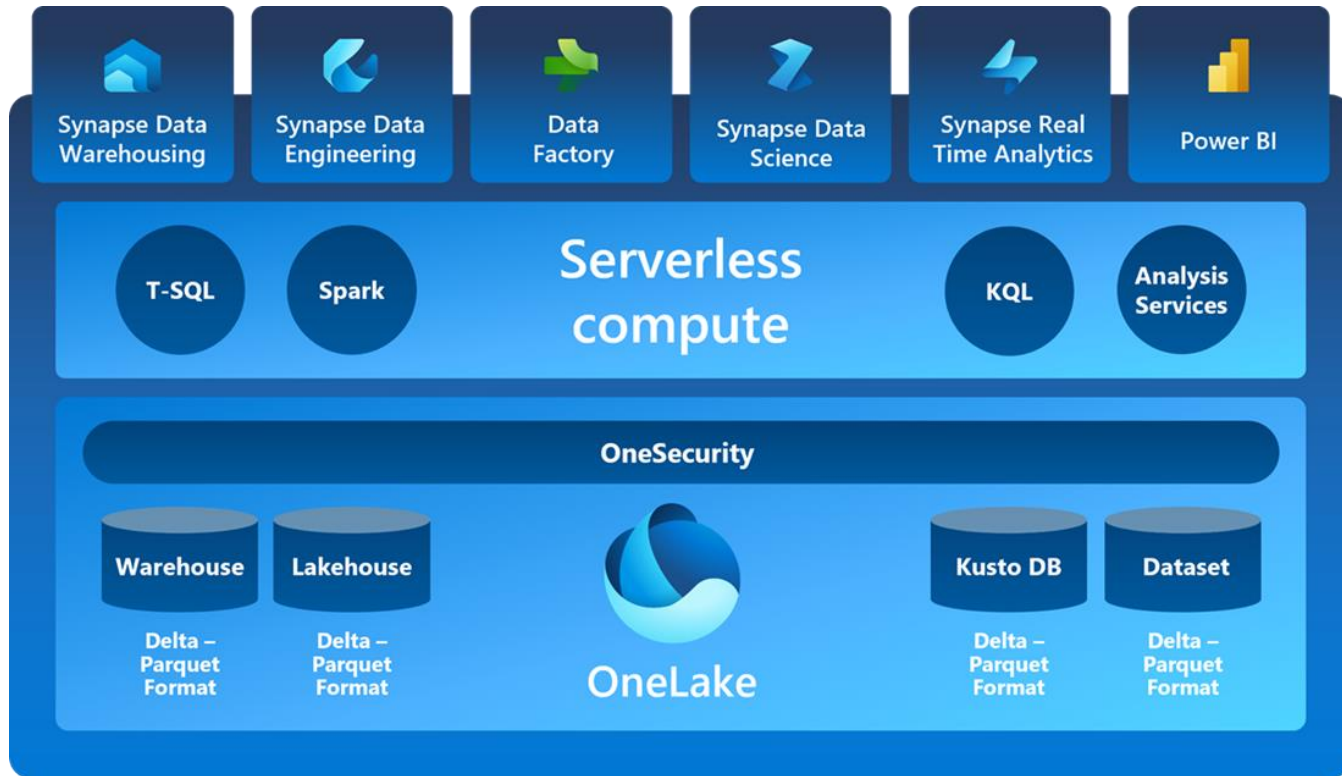
All workloads automatically store their data in the OneLake workspace folders

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

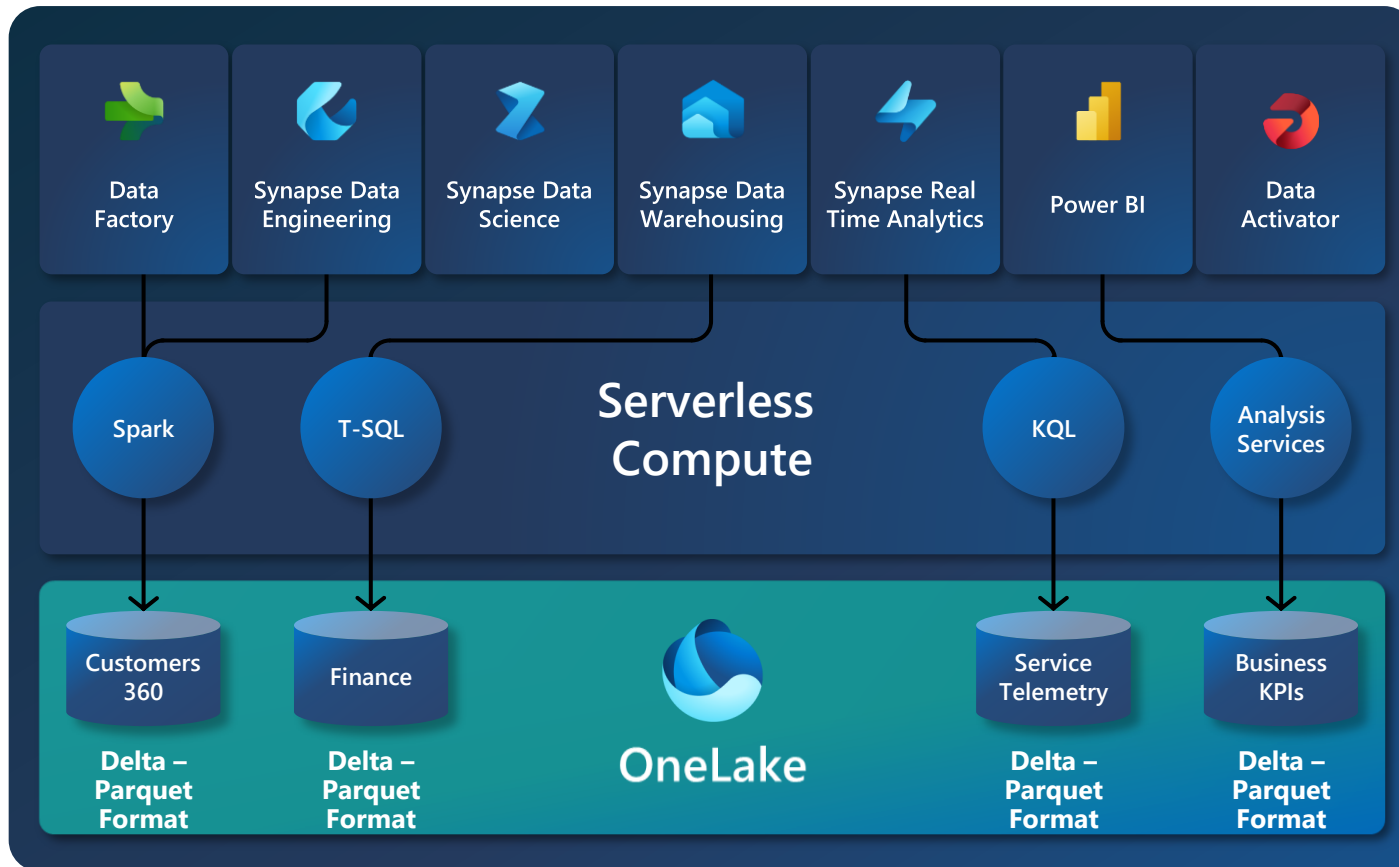
Next Gen: Microsoft Fabric

A unified analytics solution for the era of AI



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

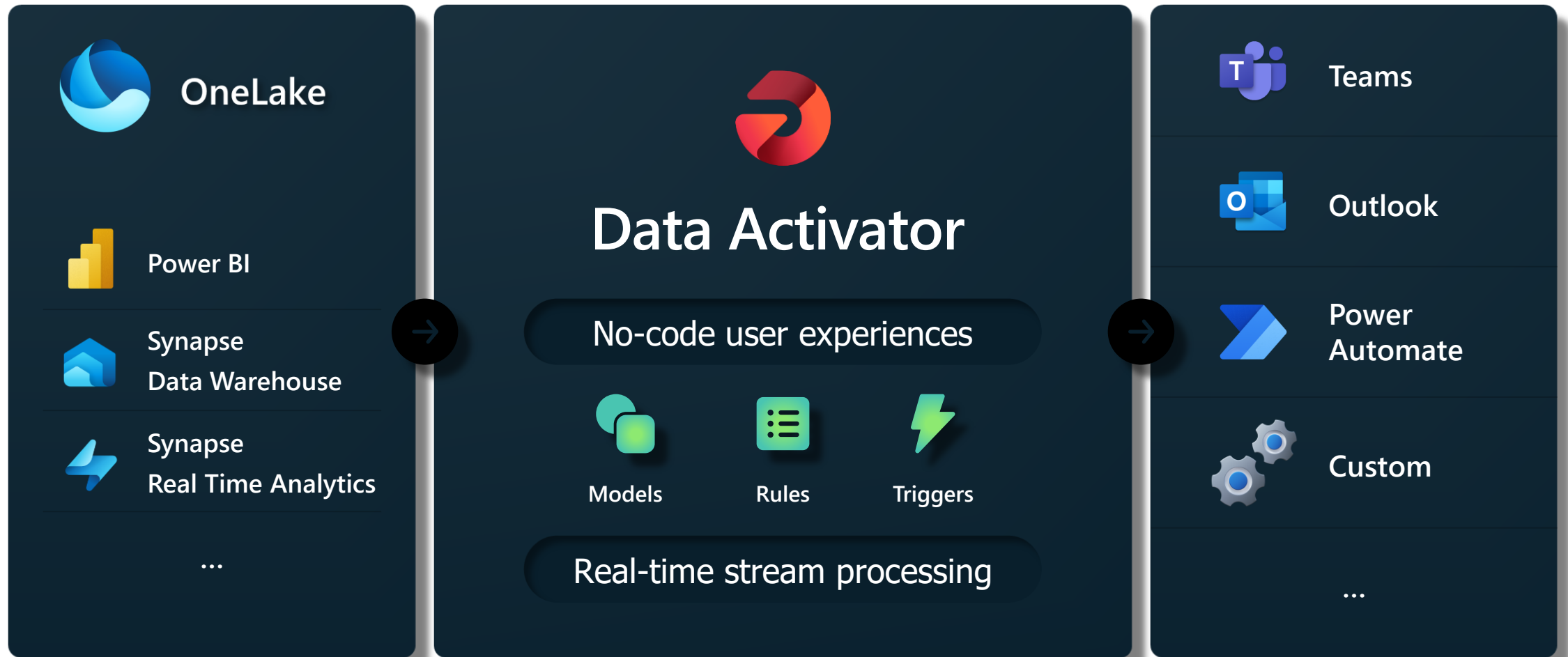
Delta – Parquet, an open standards format, is the storage format for all tabular data in Analytics vNext

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

Introducing Data Activator



SaaS

"It just works"

5x5

Frictionless onboarding

Instant Provisioning

Quick results w/ Intuitive UX

Success by default

Minimal knobs

Auto optimized

Auto Integrated

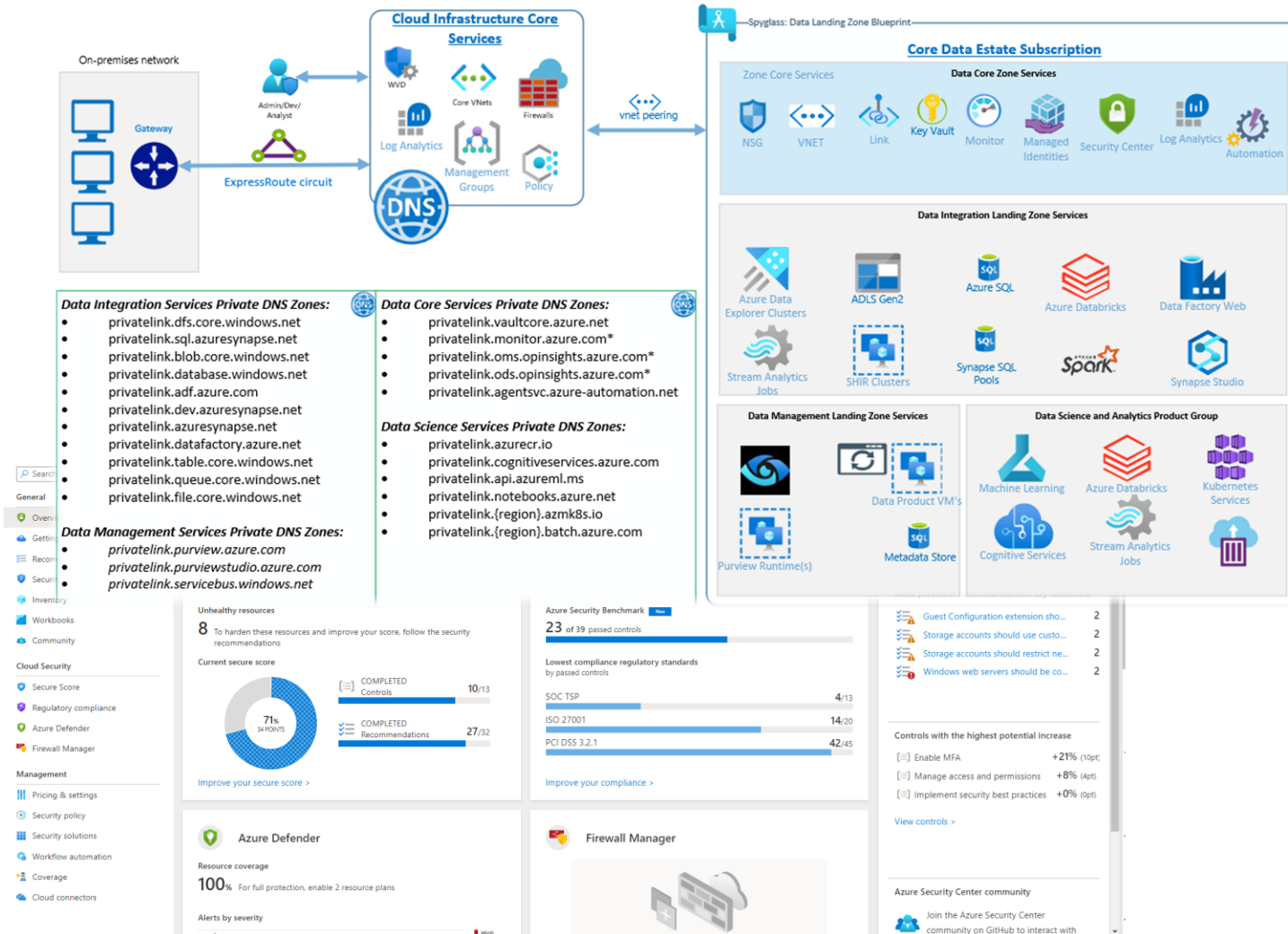
Centralized administration

Tenant-wide governance

Centralized
security management

Compliance built-in

Fabric is the SaaS Data Landing Zone



Layer

Core services

SaaS Fabric Capacity

- Network
- Monitoring
- Shared integration runtimes
- CI/CD Agents
- Security Agents

Core Data services

- Metastore for Data Services
- Data Lake Services
- Data Storage
- Integration runtimes
- Apache Spark
- Synapse Analytics

Data Applications

- Machine Learning
- Open AI
- Cognitive Services
- Data Share
- Stream Analytics

Visualization

- Power BI
- Data Flows

Demo: Deploy a Landing Zone in Fabric



Working With Fabric

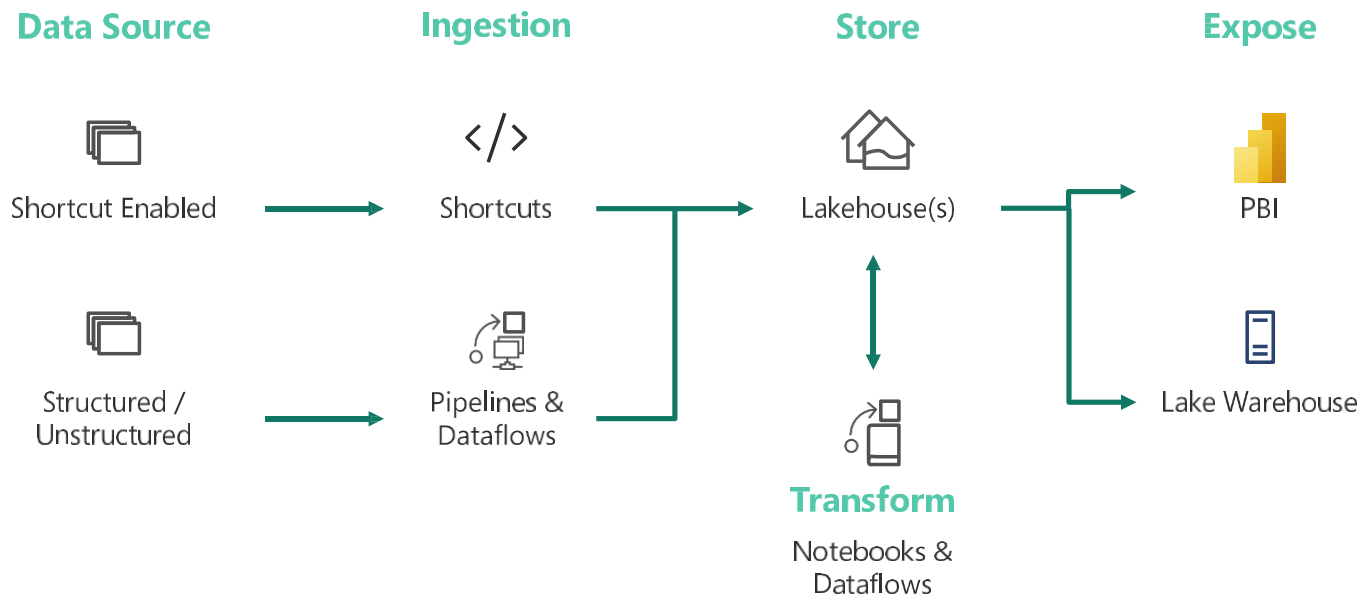
Lakehouse

The Microsoft Fabric Lakehouse analytics scenario makes it so that data can be ingested into OneLake with shortcuts to other clouds repositories, pipelines, and dataflows in order to allow end-users to leverage other data.

Once that data has been pulled into Microsoft Fabric, users can leverage notebooks to transform that data in OneLake and then store them in lakehouses with medallion structure.

From there, users can begin to analyze and visualize that data with Power BI using the see-through mode or SQL endpoints.

The Data Lakehouse scenario



Build and implement an end-to-end lakehouse for **your organization:**

1. Create a Microsoft Fabric workspace
2. Quickly create a lakehouse – an optional module to implement medallion architecture (Bronze, Silver, and Gold)
3. Ingest, transform and load data into the lakehouse – bronze, silver and gold zones as delta lake tables for medallion architecture
4. Explore OneLake, OneCopy of your data across lake mode and warehouse mode
5. Connect to your lakehouse using TDS/SQL endpoint
6. Create Power BI reports using DirectLake – to analyze sales data across different dimensions
7. Orchestrate and schedule data ingestion and transformation flow with Pipeline
8. Cleanup resources by deleting the workspace and other items



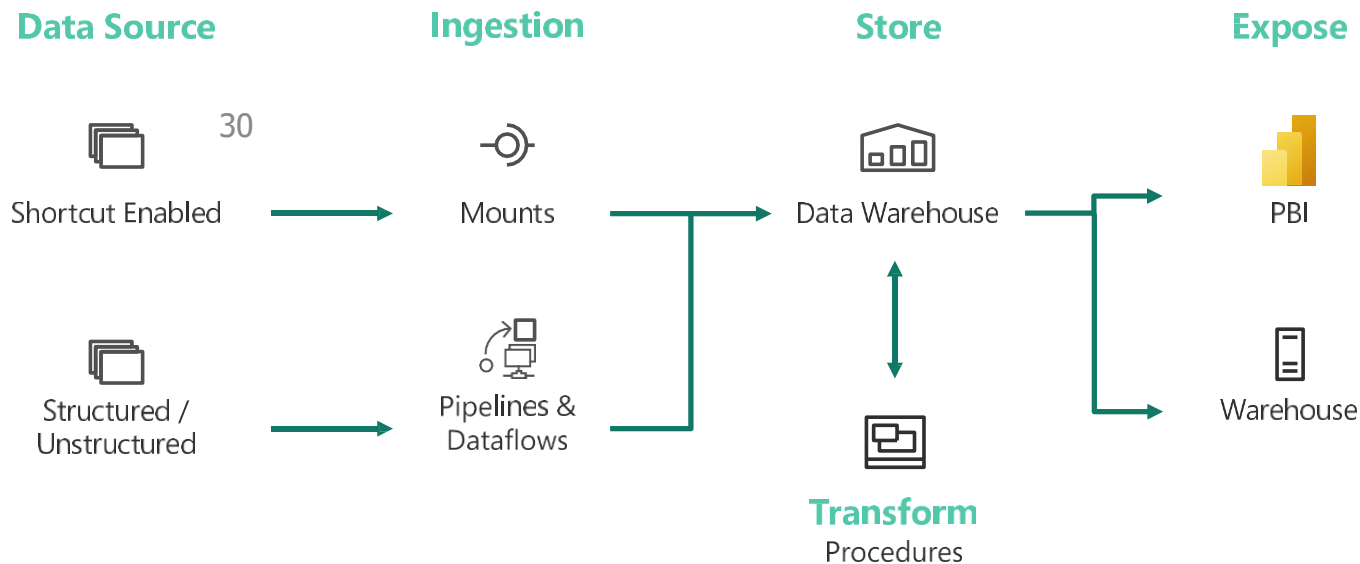
Data Warehouse

The Data Warehouse analytics scenario takes existing sources that are mounted, while pipelines and dataflows can bring in all other data that is needed.

IT teams can then define and store procedures to transform the data, which is stored as Parquet/Delta Lake files in OneLake.

From there, business users can analyze and visualize data with Power BI, again using the see-through mode or SQL endpoints.

The Data Warehouse scenario



Build and implement an end-to-end data **warehouse** for **your organization**:

1. Enable Microsoft Fabric in your tenant
2. Create a Fabric workspace
3. Quickly create a data warehouse
4. Ingest data from source to the data warehouse dimensional model
5. Transform the data to create aggregated datasets using T-SQL
6. Perform orchestration, data ingestion, and data transformation with pipelines
7. Query the data warehouse using T-SQL and a visual query editor
8. Create Power BI report using DirectLake mode to analyze the data in place
9. Cleanup resources by deleting the workspace and other items



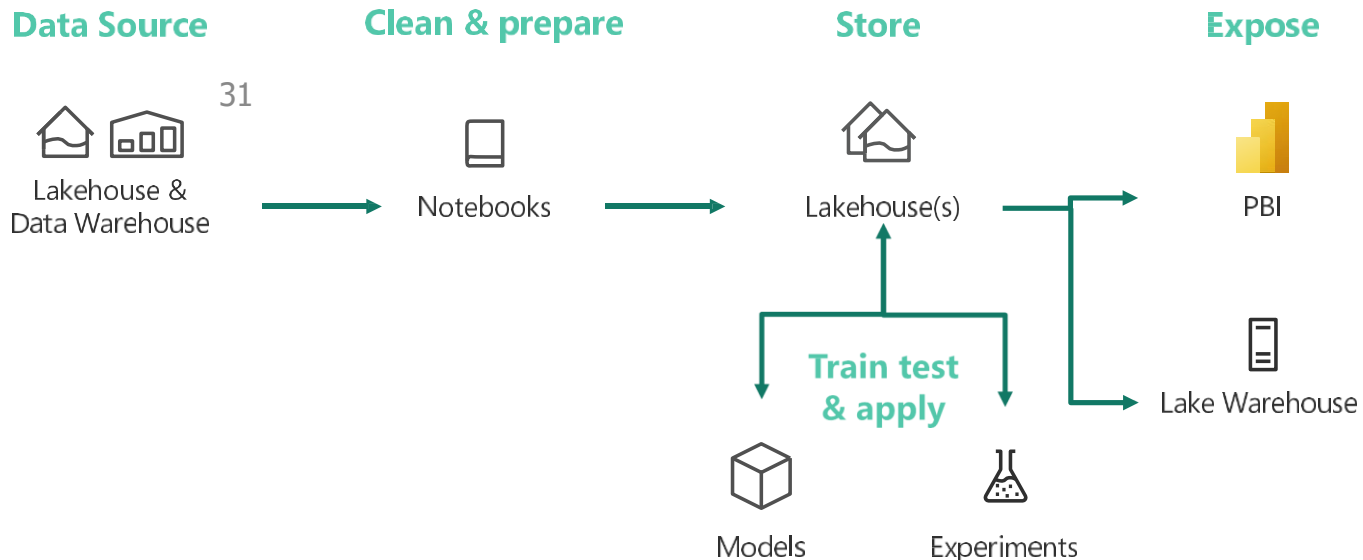
Data Science

The Data Science analytics scenario can be ingested similarly to the Lakehouse and Data Warehouse paths.

Once the data is ingested, it is cleaned and prepared using notebooks and then stored in the lakehouse with medallion structure.

After the data is cleaned and stored, machine learning models can be trained and tested directly on the lakehouse.

The Data Science scenario



In this scenario, consider performing the following activities:

1. Use the Microsoft Fabric notebooks for data science scenarios
2. Ingest data into Microsoft Fabric lakehouse using Apache Spark
3. Load existing data from the lakehouse delta tables
4. Clean and transform data using Apache Spark
5. Create experiments and runs to train a machine learning model
6. Register and track trained models using MLflow and the Microsoft Fabric UI
7. Run scoring at scale and save predictions and inference results to the lakehouse
8. Visualize predictions in PowerBI using DirectLake

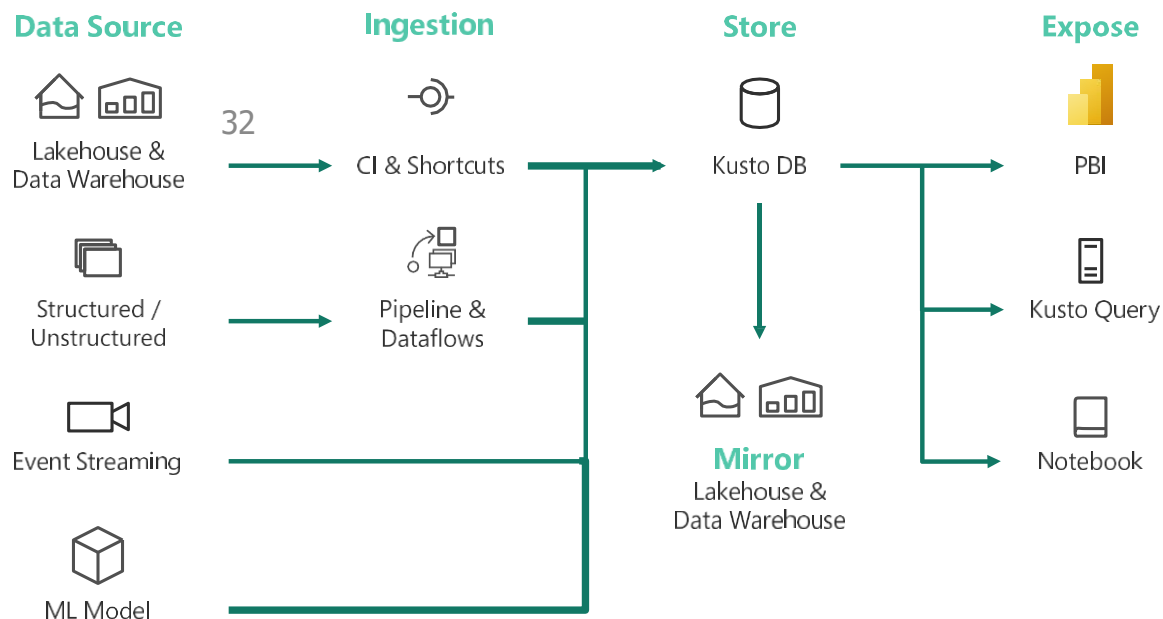


Real Time Analytics

Unlike the Data Science, Lakehouse, and Data Warehouse analytics scenarios, streaming data can be ingested into the Microsoft Fabric in several ways to achieve real time analytics.

Users can leverage Event Hub, IoT Hub, pipelines, dataflows, notebooks, or open-source products like Kafka, Logstash, and more.

The end-to-end Real-Time Analytics scenario



Once ingested into Microsoft Fabric, streaming data can be stored in Kusto DB and mirrored into the lakehouse. After the data has been stored, machine learning models can be trained and tested directly on the lakehouse with experiments.

Like the other scenarios, business users can analyze and visualize the data with Power BI using the see-through mode or SQL endpoints. Data can also be exposed through KQL or notebooks using Spark.

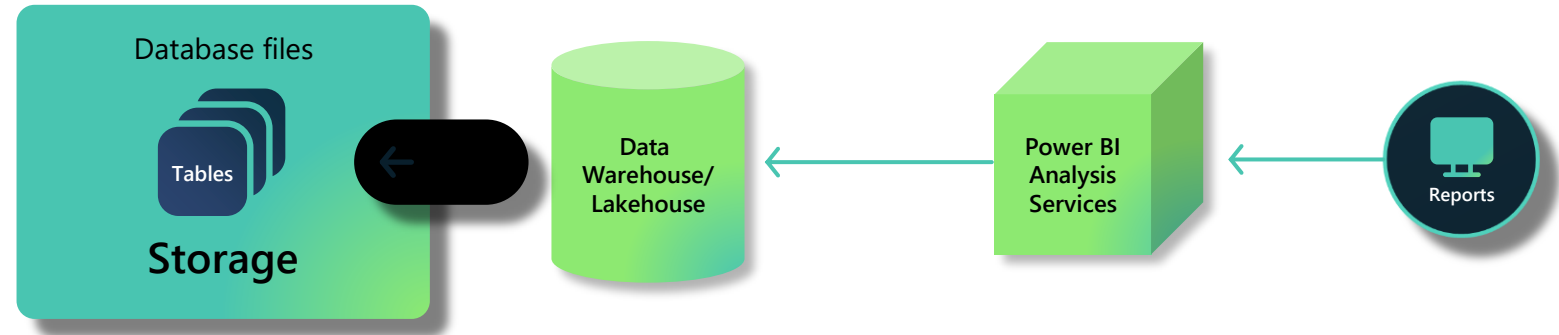
In this scenario, consider performing the following activities:

1. Create a KQL Database
2. Create Eventstream
3. Stream data from Eventstream to KQL Database
4. Check your data with sample queries
5. Save queries as a KQL Queryset
6. Create a Power BI report
7. Create a OneLake shortcut



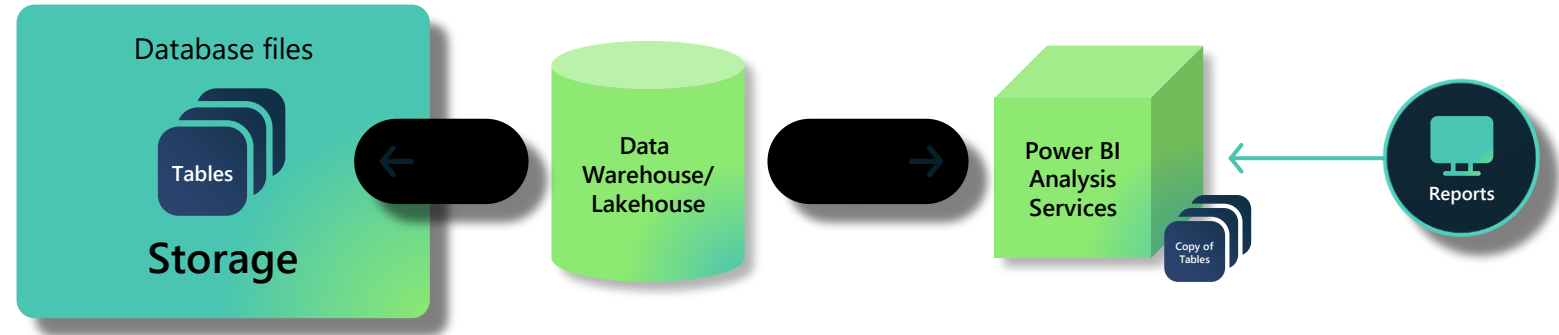
"Direct Query Mode"

Slow, but real time



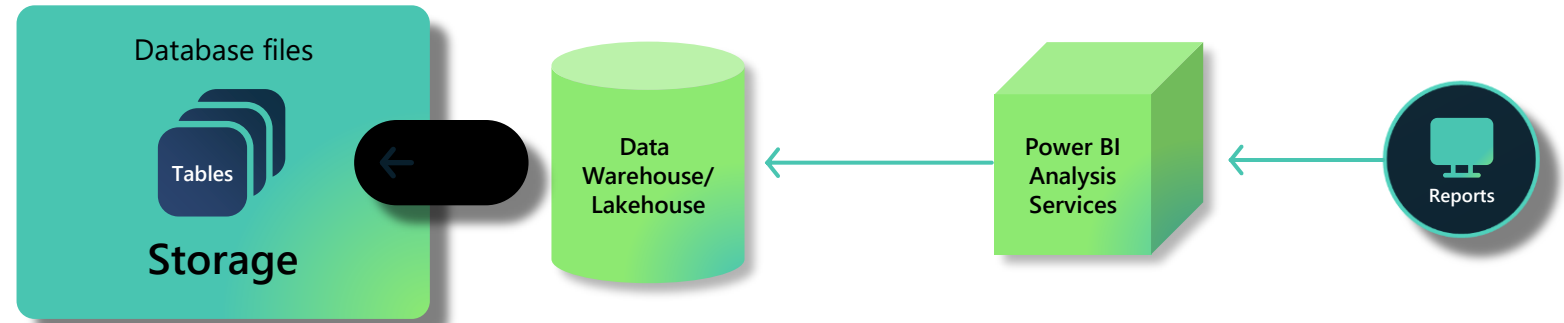
"Import Mode"

Latent & duplicative but fast



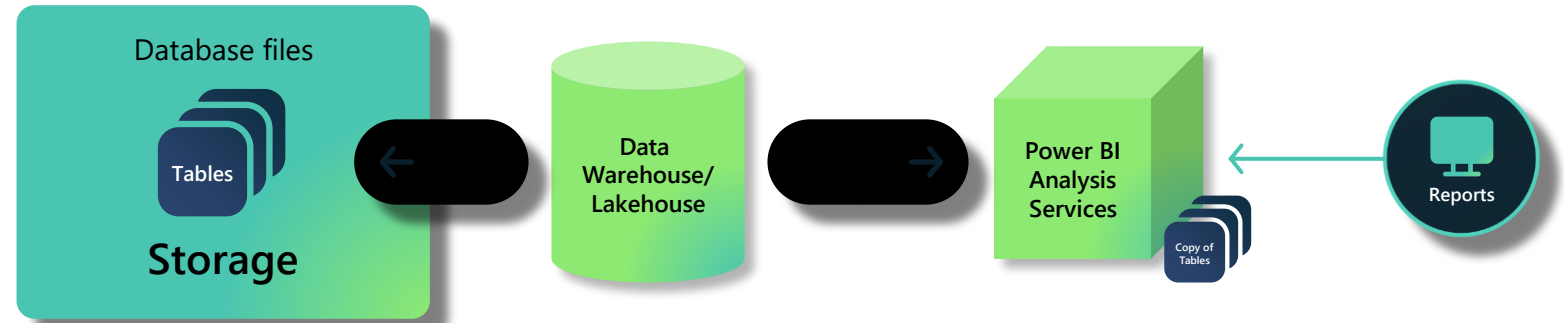
"Direct Query Mode"

Slow, but real time



"Import Mode"

Latent & duplicative but fast

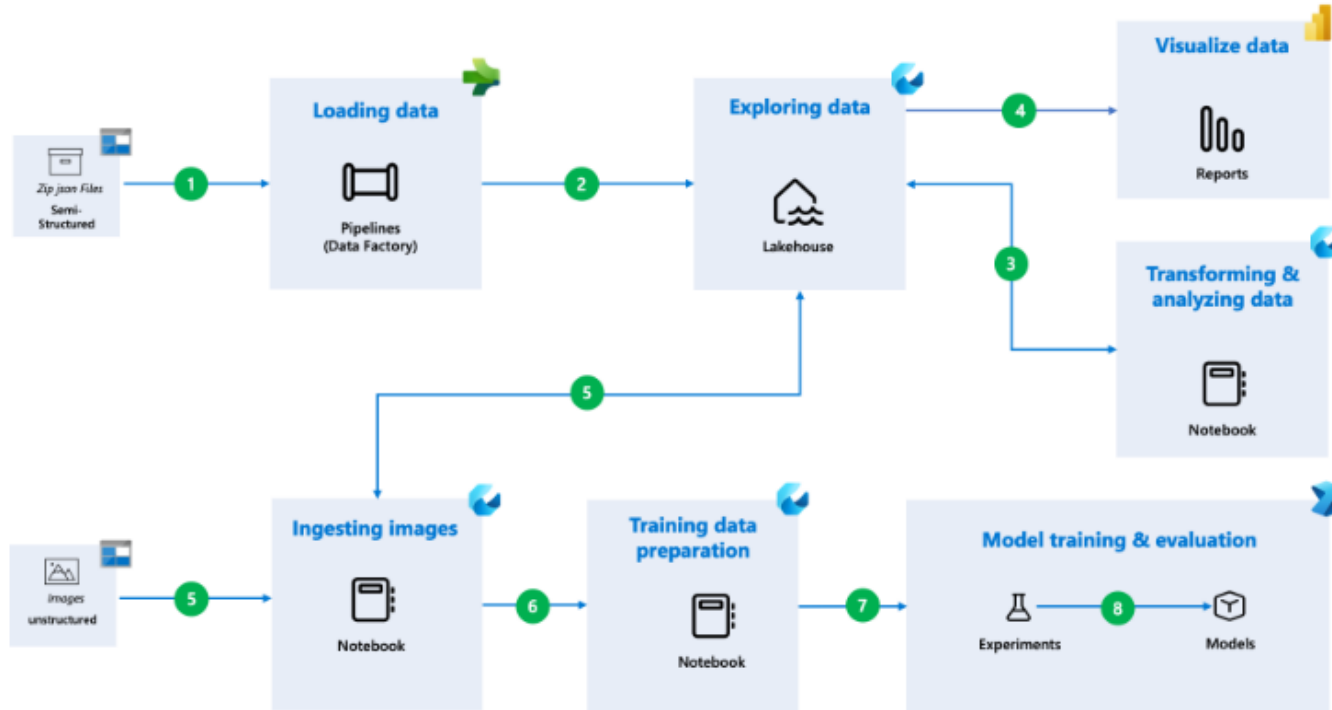


"Direct Lake Mode"

Perfect!



Demo: Working with Fabric



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

Get Fabric Ready

- Data Strategy
- Modernize Existing Workloads
 - Delta Lake
 - Spark Workloads
 - Medallion (Bronze, Silver, Gold)
- Governance & Adoption Plan
 - Envision
 - Experiment
 - Scale
- Develop, plan and migration strategy to Microsoft Fabric
 - Assess & Evaluate
 - Plan & design
 - Migrate
 - Monitor & Govern
 - Optimize & Modernize

Key Capabilities are the Bridge

- Capacities
 - A Microsoft Fabric subscription consists of **tenants, capacities, and workspaces** and can be organized in various ways according to your organizational needs.
- One Lake
 - **OneLake is a single, unified, logical data lake for the whole organization.** Like OneDrive, OneLake comes automatically with every Microsoft Fabric tenant and is designed to be the single place for all your analytics data.
- Delta Lake
 - In order to achieve seamless data access across all compute engines in Microsoft Fabric, **Delta Lake is chosen as the unified table format**
 - **HDInsight**
- Shortcuts
 - Allow users to reference data without copying it. It unifies data from different lakehouses, workspaces, or **external storage, such as ADLS Gen2 or AWS S3**
- Service Tags & Gateways
 - **Azure service tags** with Microsoft Fabric to enable an **Azure SQL Managed Instance (MI) to allow incoming connections from the Microsoft Fabric.**
- Mirroring
 - **Snowflake**
 - **SQL**
- OneLake Integrates with Azure Data Services
 - **Synapse**
 - **Databricks**
 - **HDInsight**

Fabric Adoption

Envision

- Get started
- Project team
- Executive sponsorship
- Adoption strategy and vision
- Success criteria
- Change management plan
- Readiness
- Training strategy

Experiment

- Champions
- Community
- Rewards and recognition
- Communication plan
- Awareness
- Ongoing support

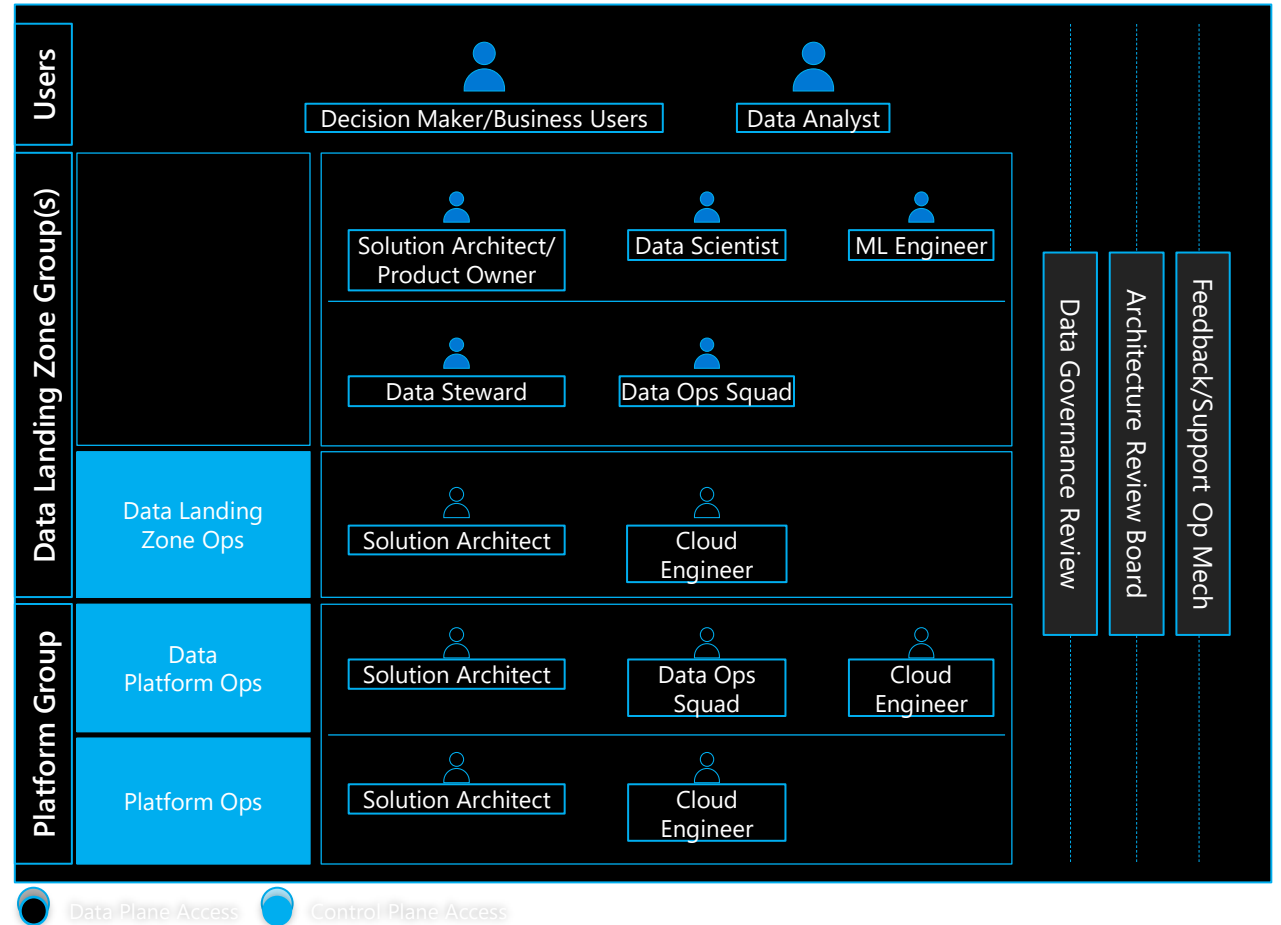
Scale

- Center of Excellence (CoE)
- Adoption maturity
- Operational plan
- Success tracking
- Change management
- Activity calendar

Modernize your teams and operations

Prescriptive guidance to create agile, vertical, cross-domain teams.

- Define personas, roles, and responsibilities.
- Develop a skilling plan to build capacity to drive your strategy.
- Enable communication between leadership, business, IT, and delivery teams.
- Engage end users continually.



Assemble your team

Any successful project begins by bringing the right team together. Your project stakeholders should include representatives across your organization who have accountability for project success.

Role	Responsibilities	Department
Executive Sponsor	Communicate high-level vision and values of the Microsoft Fabric to the company	Executive Leadership
Success Owner	Ensure the business goals are realized from the adoption process	Any department
Champions	Help evangelize Fabric. Create a circle of influence and can feed back to the Adoption team what works/what doesn't work.	Multiple departments
Training Lead	Manage and communicate training content about Fabric – can be internal or external vendor	IT or other
Department Leads (Stakeholders)	Identify how specific departments will use the Fabric and encourage engagement	Any department (management)
Communication Lead	Oversee company-wide communications about the Fabric	Corp Communications, IT or other
Fabric Admin Team	Responsible for establishing an environment strategy, setting up <u>data loss prevention (DLP) policies</u> , and managing users, <u>capacity, and licensing</u> . They also make data available to makers through connectors, integration, or migration.	IT
Fabric Nurture Team	Organizes data-in-a-day events and hackathons, provides mentorship to makers, ensures new makers get off to a good start, and generally evangelizes the platform. Provide business change management.	Multiple departments

Note: Though we recommend having each of these roles fulfilled, in smaller organizations some of these roles may be carried out by the same person.

Roles and Responsibilities



Data Strategy

Align Fabric strategy with Digital Transformation strategy

Drive innovation

Key Decision makers

Business Value

Adoption and Change management

Empower new ways of working



Fabric Admin team

Environment Strategy

DLP Policies

Data integration / migration guidelines

Guardrails

License management



Fabric Nurture team

Data-in-a-day / Hackathons

Solution envisioning

Community of Practice

Evangelize Platform



Fabric Re-Usable Assets & Governance

Automation / Azure DevOps

Data Architecture reviews

Library of common utilities, patterns and templates

Data Catalog



User support team

Help Desk guidance / FAQ

Ticketing process

Support Process & Resources

Common Questions & Answers

- Does the provisioning of the lake database also require provisioning of One Lake location?
 - No. One lake is not an option. It's the foundational storage mechanism for Fabric. Like Power BI, it's a 'managed' deployment of Azure Data Lake Gen 2.
- What do we do with our existing data lake/ pipelines/ notebooks / data flows? Does it need additional migration? If so, what's the effort/complexity?
 - Pipelines and notebooks will have a migration path to Fabric. The tech is fundamentally the same from a code base perspective. Likewise, bridges such as shortcuts allow your existing data lake to be integrated into Fabric.
- Azure ML with fabric
 - Machine Learning in Fabric is not Azure ML, it's Synapse ML but improved. The integration of Azure ML is much the same as it is with Synapse, it's still a separate service that can be leveraged and it better for data scientists.
- When to use between Synapse Data engineering or Data warehousing
 - Data engineering and data warehousing are just two experience pages within Fabric. They both provide similar options for working with data. I'd guide you to think about, when to use data warehouse vs. lakehouse instead.

Common Questions & Answers

- From the data engineering perspective, what new features are available in Fabric but not in Synapse?
 - Fabric is built on Power BI's framework, which means Dataflows Gen 1 and Gen 2 (the power query version) are available for use within Fabric as a data engineer.
 - Likewise, Spark Dataflows are not in Fabric.
 - Fabric also provides relational model development, in which you can establish relationships in the warehouse for which datasets have a native understanding of.
 - CoPilot will also change how engineers develop.
 - Performance differences in using TSQL stored procs vs PySpark
 - Spark and TSQL are different compute architectures. Much like today, it's a choice of building a lake house or a dedicated warehouse. TBH, I'm not aware of any benchmarking between the lakehouse and warehouse.
 - The initial GA of Fabric is highly focused on features and capabilities, performance is slated for the next release, however, so far performance has been very good overall.
 - How does monitoring work in Fabric?
 - The Admin monitoring workspace is designed to provide admins with monitoring capabilities for their organization.
 - User activities are logged in the Power BI activity log and in the unified audit log.
- For Fabric items: Currently, only create, read, update, and delete operations.

Common Questions & Answers

- What is the pricing model for Fabric/Fabric Azure SKUs and the purpose?
 - [Announcing Microsoft Fabric capacities are available for purchase | Microsoft Fabric Blog | Microsoft Fabric](#)
 - [Microsoft Fabric licenses - Microsoft Fabric | Microsoft Learn](#)
- Auto scale of SKUs like Power BI?
 - [Out of box autoscaling is not available yet but is being worked on in the following months.](#)
 - [However, you can install the metric app and monitor usage to help make an informed decision on proper sku selection.](#)

- Scale of Direct Query on Delta Lake? Does this transfer the resource usage from memory of Power BI to Fabric and eliminate the row limit or memory limit issues?
 - [Instead of leveraging Direct Query/Import mode it would be better to leverage Direct Lake which doesn't have the same limitations.](#)

[Learn about Direct Lake in Power BI and Microsoft Fabric - Power BI | Microsoft Learn](#)

- Is the long-term vision to migrate all data estates to Fabric?
 - [To be determined and migration communication will hopefully be announced soon.](#)

Common Questions & Answers

- What's the difference between an Azure Synapse Dedicated Pool and a Fabric Data Warehouse?
- In Azure Synapse we have dedicated and serverless options that come with their own advantages and disadvantages. In Fabric we re-wrote the DW engine as one product that includes the advantages of both:
 - Elasticity of serverless
 - Indexing, caching, and performance from Dedicated Pool.
 - Separation of compute and storage
 - The new engine is optimized for Delta/Parquet
 - no need to create heap tables because it's fast
- The new architecture allows for countless new features that were not available in Dedicated Pool
- What scenarios might not a fit for Fabric? i.e. Operational data stores?
 - In general, Fabric's sweet spot is in the Lakehouse and OLAP areas. This won't replace OLTP scenarios as of today.
- Can you configure specific workloads and their resource consumption? Such as limit data flow and pipelines, provide more for data warehouses?
 - Given the nature of SaaS there isn't a lot of options to get granular with workload management at this time.
 - You can create multiple Fabric Capacities at different SKUs and designate them to desired workspaces to help control costs.

Common Questions & Answers

- How does monitoring work in Fabric?
 - The Admin monitoring workspace is designed to provide admins with monitoring capabilities for their organization.
 - User activities are logged in the Power BI activity log and in the unified audit log.
 - For Fabric items: Currently, only create, read, update, and delete operations.
- How does Fabric and Purview work together?
 - Microsoft Purview works with Microsoft Fabric so users can discover and manage Microsoft Fabric items in Microsoft Purview applications. The integration currently allows you to take advantage of these applications:
 - Microsoft Purview Information Protection
 - Microsoft Purview Data Loss Prevention (DLP)
 - Microsoft Purview Audit
 - Note: Soon you'll be able to use more applications, like the [Microsoft Purview Data Catalog](#)

Closing & Questions