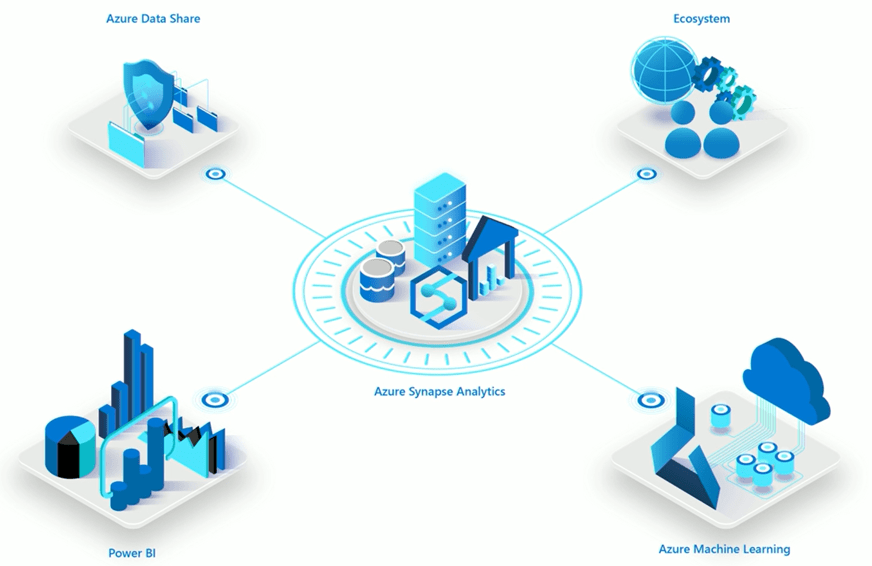
**Azure Assignment**

**Azure Synaps**

****

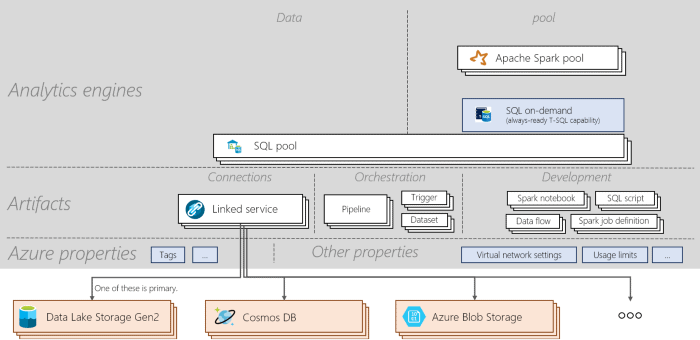
Azure Synapse Analytics is a cloud-based analytics service provided by Microsoft Azure. It enables the integration of big data and data warehousing technologies in a single, unified platform. Azure Synapse allows users to ingest, prepare, manage, and serve data for business intelligence and machine learning purposes.

Key features of Azure Synapse include:

1. **Integration**: It integrates various data processing capabilities such as data ingestion, data preparation, data warehousing, big data analytics, and data serving.
2. **Unified Workspace**: Users can work with both structured and unstructured data in a single workspace, eliminating the need to switch between multiple tools or services.
3. **SQL Pool**: Azure Synapse provides a SQL-based analytics service for data warehousing, formerly known as Azure SQL Data Warehouse. It allows users to run complex queries on large datasets.
4. **Spark Pool**: Users can run Apache Spark jobs to analyze and process big data in real-time or batch mode. This feature provides scalability and flexibility for handling large-scale data analytics workloads.
5. **Data Integration**: Azure Synapse supports data integration from various sources, including Azure Blob Storage, Azure Data Lake Storage, Azure SQL Database, and more.
6. **Security and Compliance**: It offers robust security features to protect data at rest and in transit, including role-based access control (RBAC), encryption, and compliance certifications such as GDPR, HIPAA, and SOC.
7. **Serverless SQL Pool**: This feature allows users to run on-demand queries without the need to provision or manage resources explicitly, enabling cost-effective ad-hoc analysis.
8. **Machine Learning Integration**: Azure Synapse integrates with Azure Machine Learning, enabling data scientists to build, train, and deploy machine learning models using the data stored in Synapse.

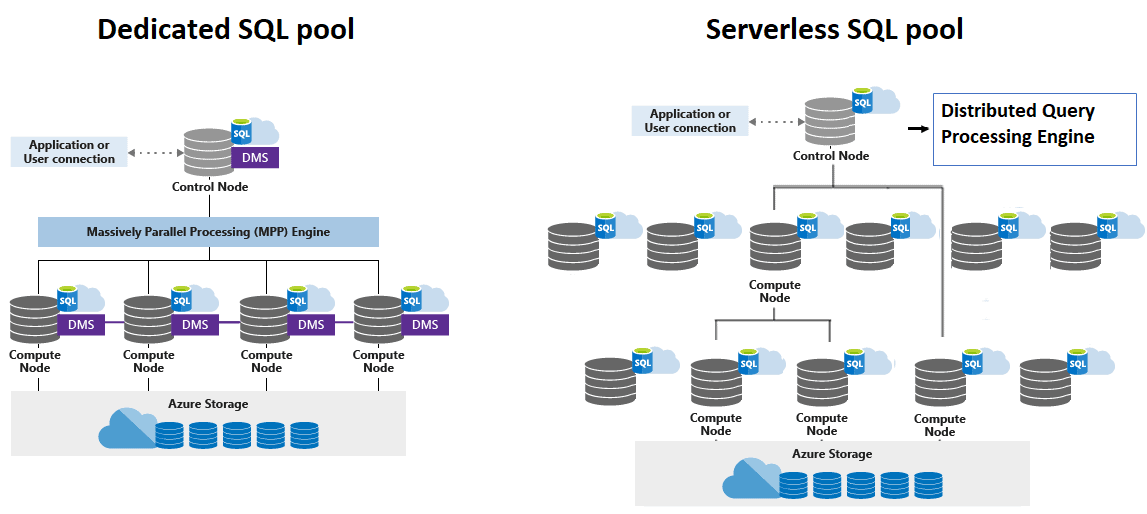
Overall, Azure Synapse Analytics provides a comprehensive platform for organizations to perform advanced analytics and derive valuable insights from their data assets.

**Architecture**



**Synapse SQL:** It is the ability to do T-SQL based analytics in Synapse workspace. It consists of two consumption models: dedicated and serverless.

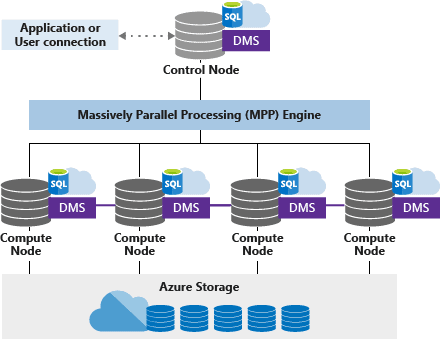
* **Dedicated SQL pools** are used for dedicated models and a workspace can have any number of these pools.
* **Serverless SQL pools** are used for serverless models and every workspace has one of these pools.



**Apache Spark for Synapse:** Serverless Apache Spark pools are created and used in the Synapse workspace to use Spark analytics. It consists of the following components:

* Apache Spark for Synapse
* Apache Spark pool
* Spark application
* Spark session
* Notebook
* Spark job definition

## **Features of Azure Synapse Analytics**

* Azure Synapse offers cloud data warehousing, dashboarding, and machine learning analytics in a single workspace.
* It ingests all types of data, including relational and non-relational data, and it lets you explore this data with SQL.
* Azure Synapse uses **massively parallel processing** or MPP database technology, which allows it to manage analytical workloads and also aggregate and process large volumes of data in an efficient manner.
* It gives you the ability to query massive data stores using either an on-demand serverless deployment (which scales automatically as needed to handle any processing or load) or provisioned resources.
* It is compatible with a wide range of scripting languages like Scala, Python, .Net, Java, R, SQL, T-SQL, and Spark SQL.
* It facilitates easy integration with Microsoft and Azure solutions like Azure Data Lake, Azure Blob Storage, and more.
* It includes the latest security and privacy technologies such as real-time data masking, dynamic data masking, always-on encryption, Azure Active Directory authentication, and more.