

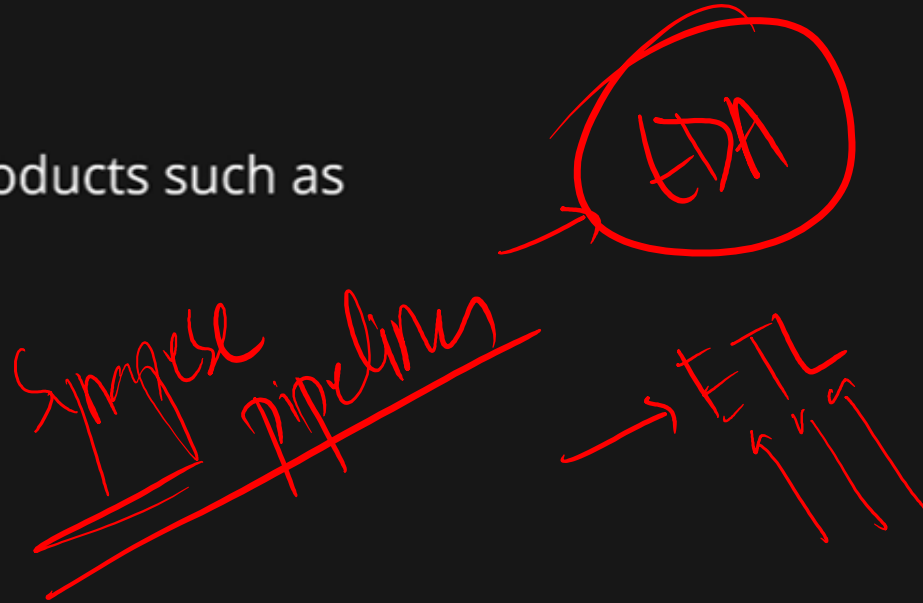
Cloud Computing with Microsoft Azure

www.drmukhan.com

Episode Objective

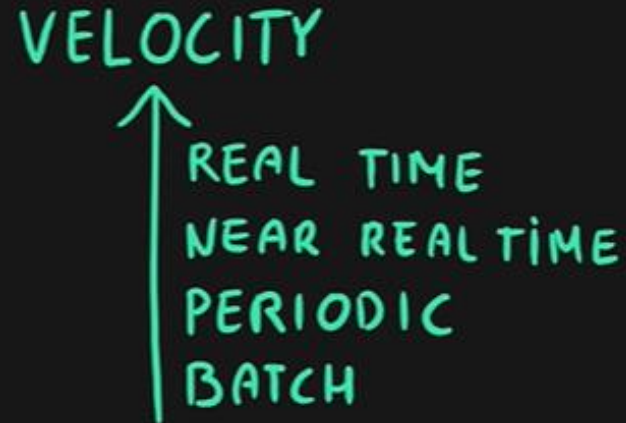
Skills Learned

- Describe Big Data and Analytics products such as
 - Azure Synapse Analytics ✓
 - Azure HDInsight ✓
 - Azure Databricks ✓

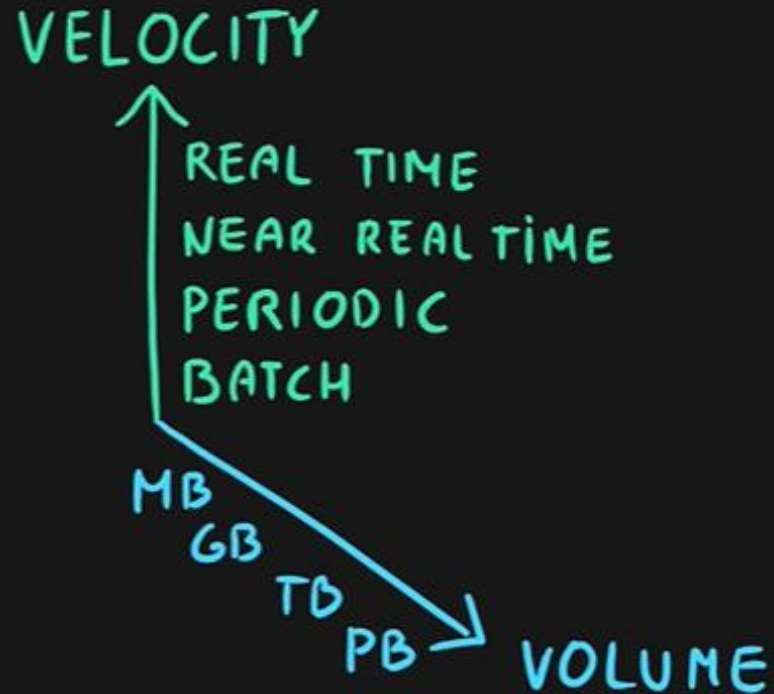


Big Data is a field that of technology that helps with extraction, processing and analysis of information that is too large or complex to deal with traditional software.

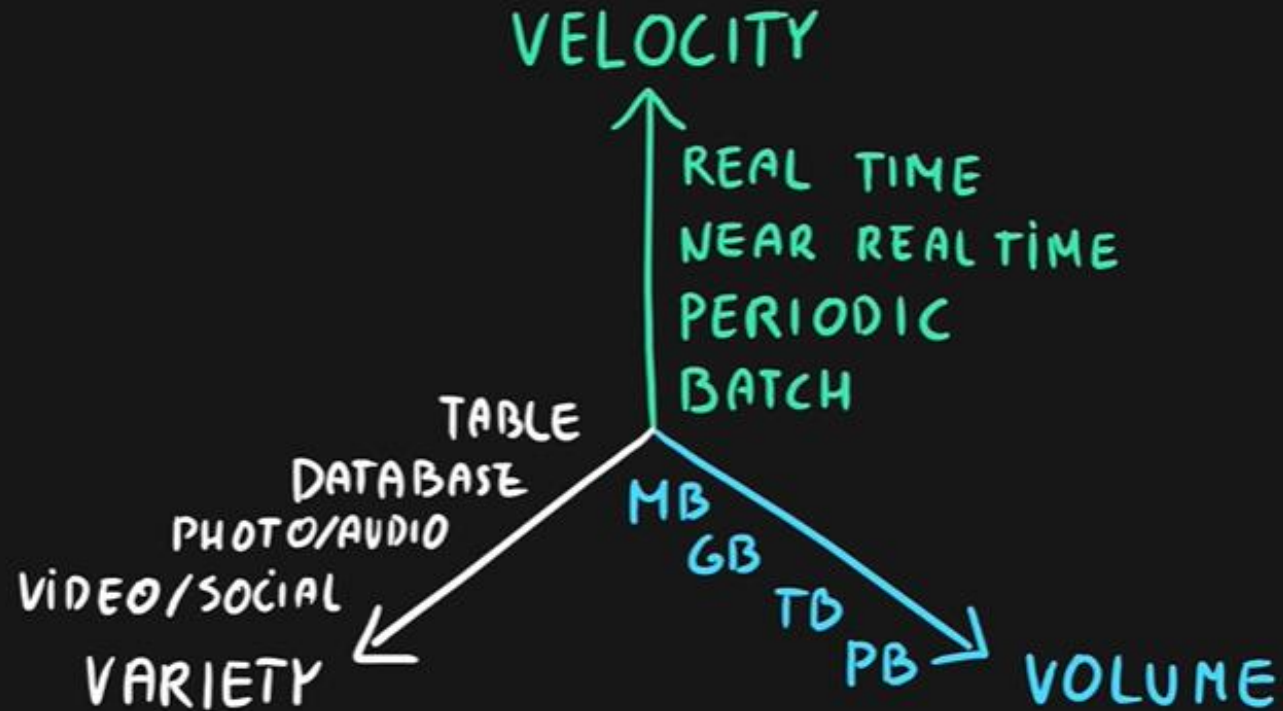
Big Data is a field that of technology that helps with extraction, processing and analysis of information that is too large or complex to deal with traditional software.



Big Data is a field that of technology that helps with extraction, processing and analysis of information that is too large or complex to deal with traditional software.



Big Data is a field that of technology that helps with extraction, processing and analysis of information that is too large or complex to deal with traditional software.





Azure Synapse Analytics

Azure Synapse Analytics



Azure Synapse Analytics



Azure Synapse Analytics



Azure Synapse Analytics

PIPELINES 



Azure Synapse Analytics

PIPELINES 

Spark 



SERVICES



FILES



DATABASES

INGEST



TRANSFORM



STORE



SERVE



DEVELOPMENT

Azure Synapse Analytics

PIPELINES 

Spark 

SYNAPSE SQL 



SERVICES



FILES



DATABASES

INGEST



TRANSFORM



STORE

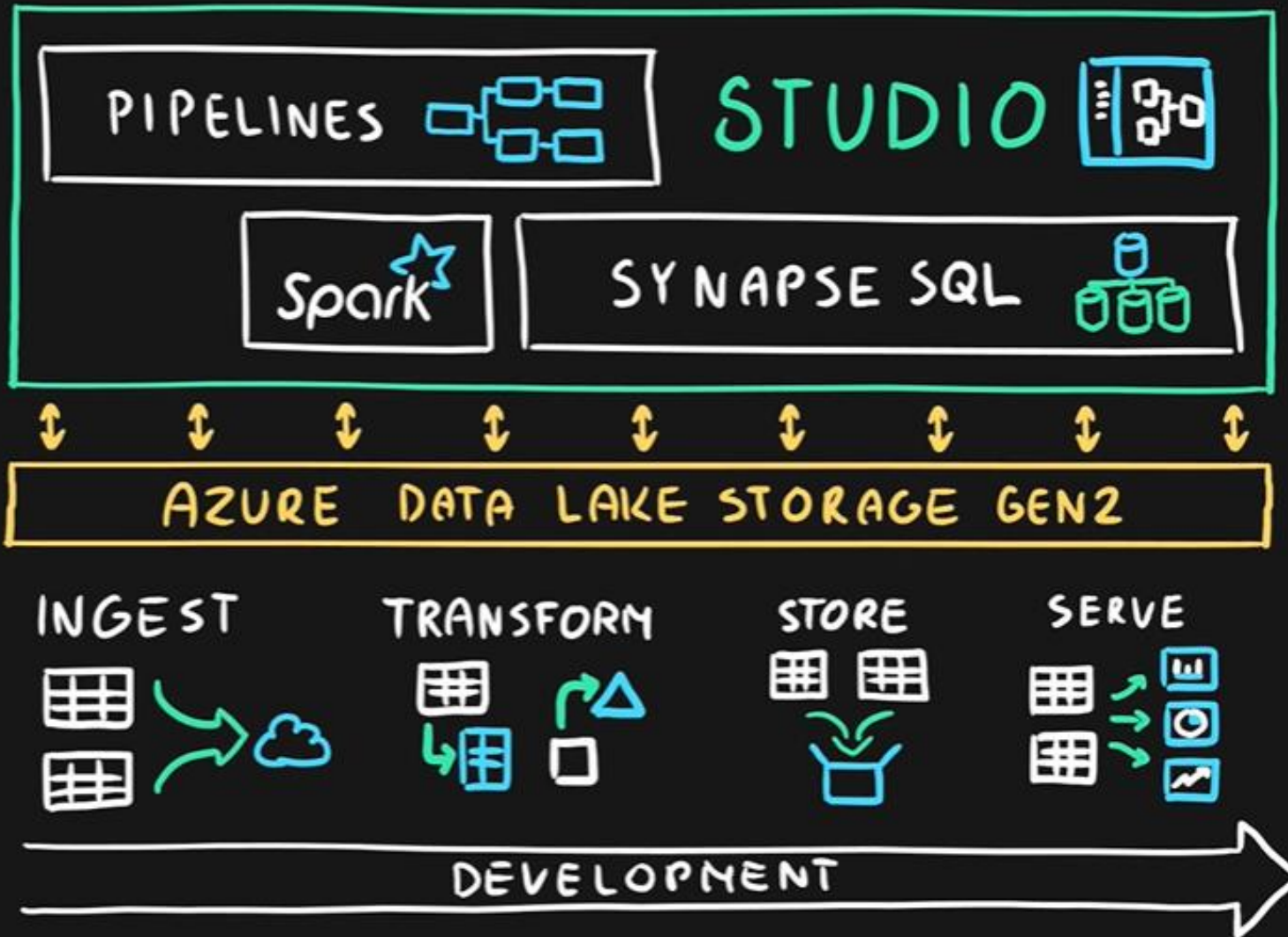


SERVE



DEVELOPMENT

Azure Synapse Analytics



Azure Synapse Analytics

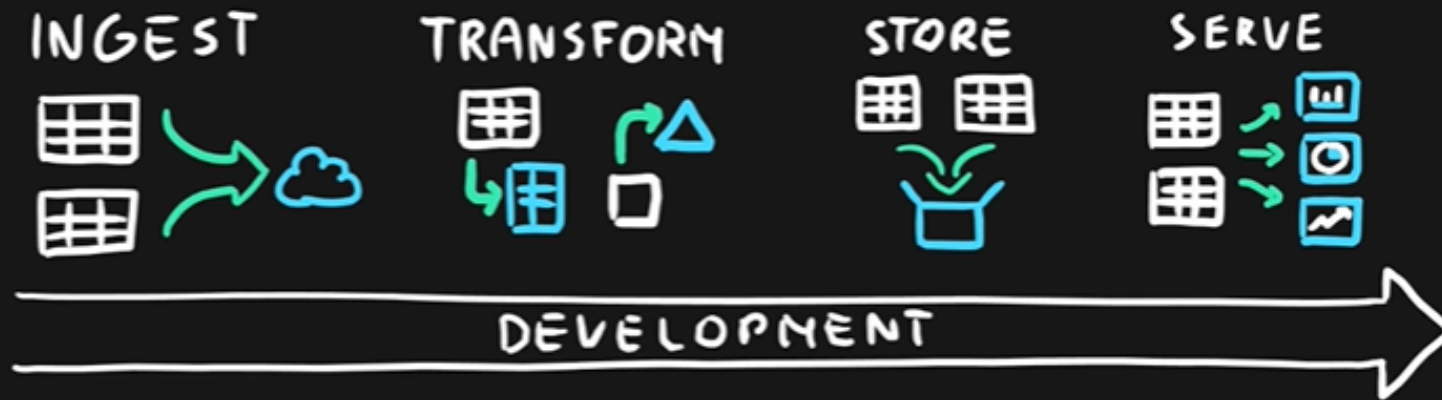
Key Characteristics

- Big data analytics platform (PaaS)
- Multiple components
 - Spark
 - Synapse SQL
 - SQL pools (dedicated – pay for provisioned performance)
 - SQL on-demand (ad-hoc – pay for TB processed)
 - Synapse Pipelines (Data Factory – ETL)
 - Studio (unified experience)



Azure HDInsight

Azure HDInsight



Azure HDInsight



Azure

HDInsight



HBASE



Azure HDInsight

Key Characteristics

- Flexible multi-purpose big data platform (PaaS)
- Multiple technologies supported (Hadoop, Spark, Kafka, HBase, Hive, Storm, Machine Learning)

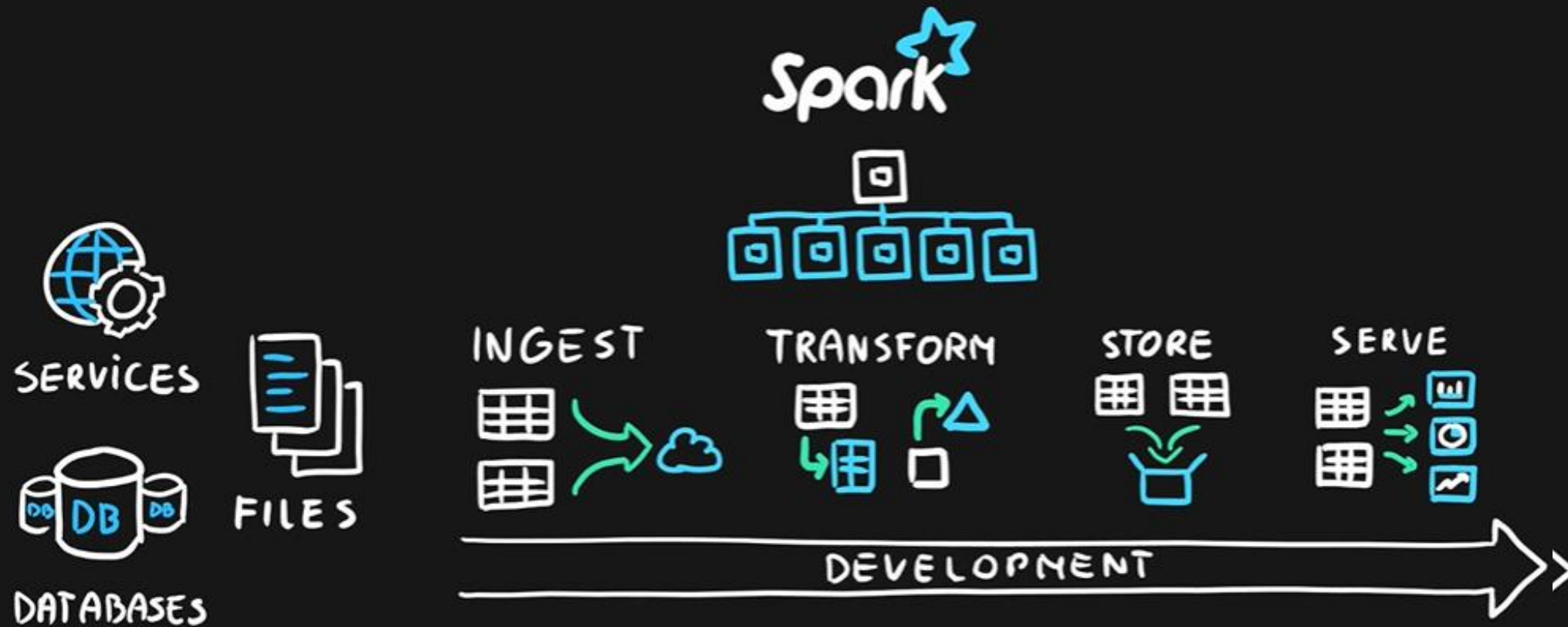


Azure Databricks

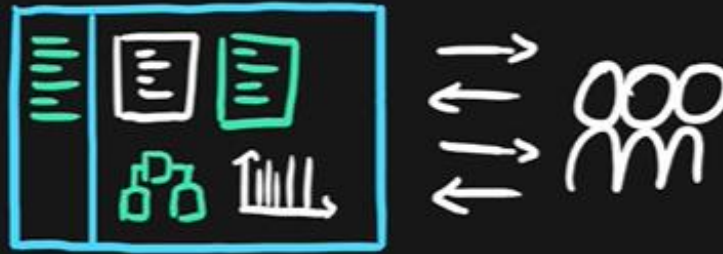
Azure Databricks



Azure Databricks

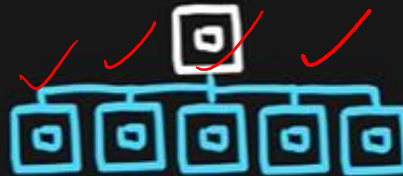


Azure Databricks



— User space
— shared space

Spark



accumulate - pay at the time
your partitions



Azure Databricks

Key Characteristics

- Big data collaboration platform (PaaS)
- Unified workspace for notebook, cluster, data, access management and collaboration
- Based on Spark
- Integrates very well with common Azure data services

Azure Big Data & Analytics Services

Summary

- [Azure Synapse Analytics](#) – Modern workspace for end-2-end enterprise data warehousing analytics with a lot of integrated tools like Data Factory, Spark, SQL, etc.
- [Azure HDInsight](#) – Fully-managed open source analytics service with a lot of supported frameworks & tools like Hadoop, Spark, Kafka, Hive, etc.
- [Azure Databricks](#) – Apache Spark based analytics platform for data transformation and collaborative analytics

THANK YOU!