Research and write about - Big Data service in Azure

Microsoft Azure focuses on the AI and analytics on it cloud service. Leveraging the cloud computation, different analytical and storage services are built which can be used in big data solutions. Azure is one of the most popular cloud service providers offering various range of services is cloud. Some of the services with which Big data solution can be achieved are

Azure Synapse Analytics

It is SQL Data warehouse and provides the service of loading data from disparate data sources. The data collected can be both relational and non-relational either on-premises or in cloud. With addition in collecting data, it helps to analyze data using SQL. It provides workspace and user-friendly dashboards with appealing and insightful visualization.

Azure Databricks

Apache Spark is reliable and efficient tool for processing huge amount of structured and unstructured datasets. Relying on the advantages of the Apache Spark, Databrick offers its services. It supports languages like Python, Java, Scala, R and other AI/ML libraries such as TensorFlow and Pytorch. It eliminates the trouble of managing the spark in local data center and supports auto-scaling and auto termination features.

Azure HDInsight

Although the use of the Apache Hadoop has declined in this decade, Hadoop ecosystem is still considered to be powerful and allows to process and store the distributed data sets among the clusters. HDInsight allows user to create and manage the cluster quickly and scale them up as per the requirement. It can be integrated with the services like Data Factory and Data Lake to perform analytics on the data already lying there. Apache Spark, Hive, Storm, HBase, Apache Kafka are some of the tools that are bundled in HDInsight.

Azure Data Factory

Azure Data Factory is the ETL service. ETL stands for Extraction, Transformation and Load. An ETL process takes the large amount of datasets and convert it into the format that is suitable for the analysis. Data Factory helps to build ETL strategy and perform ETL operation without any code or configuration. It does so by using visual editor. It provides connectors to almost 90 data sources which includes Amazon S3 and Google Big Query as well.

Azure Machine Learning

It provides the large bundle of the Pre trained and configured machine learning library and model. These algorithms can be used to work on the real-world data sets. It speeds up the model creation and build pipelines. Beside that it provides comparison using different algorithm so that user can choose which algorithm to use for his/her use case.

Some more Big data services by azure are

- Azure Stream Analytics, Data lake Analytics, Azure Analysis services