



Analytics - Visualização, Relatórios e Tomada de Decisões com Big Data



# AZURE DATA LAKE e U-SQL

## PARTE 1

**Analytics, Visualização, Relatórios e Tomada de Decisões com Big Data**





# Azure Data Lake e U-SQL – Parte 1

- Azure Data Lake Analytics
- Azure Data Lake Store
- U-SQL
- Jobs do Azure Data Lake Analytics
  - Usando Portal do Azure
  - Usando Visual Studio
  - Usando Azure Data Lake Analytics SDK



# Azure Data Lake Analytics

- Serviço de Análise para Big Data
- Gerenciado, sem infraestrutura
- Dimensionamento instantâneo
- Pago pelo uso
- Plataforma unificada
- Permite análise de dados:
  - Azure Data Lake Store
  - Azure SQL Server
  - Azure SQL Data Warehouse
- Usa U-SQL, R, Python, .NET, etc.



# Azure Data Lake Analytics - Possibilidades

- Preparar grandes volumes de dados para inserção no Data Warehouse
- Processar grandes volumes de dados “obtidos” da Web
- Processamento de dados não-estruturados, como imagens
- Substituir longos processos em lote executados mensalmente por processamento distribuído executado em períodos menores.



# Azure Data Lake Store

- Armazenamento de trilhões de objetos na escala de petabytes
- Dados não estruturados, semiestruturados e estruturados
- Suporta aplicativos que usam HDFS (Hadoop Distributed File System)
- Taxa de transferência escalonável para análises paralelas
- Cobrado por armazenamento e transações (operações de leitura e escrita)
  - <https://azure.microsoft.com/pt-br/pricing/details/data-lake-store>



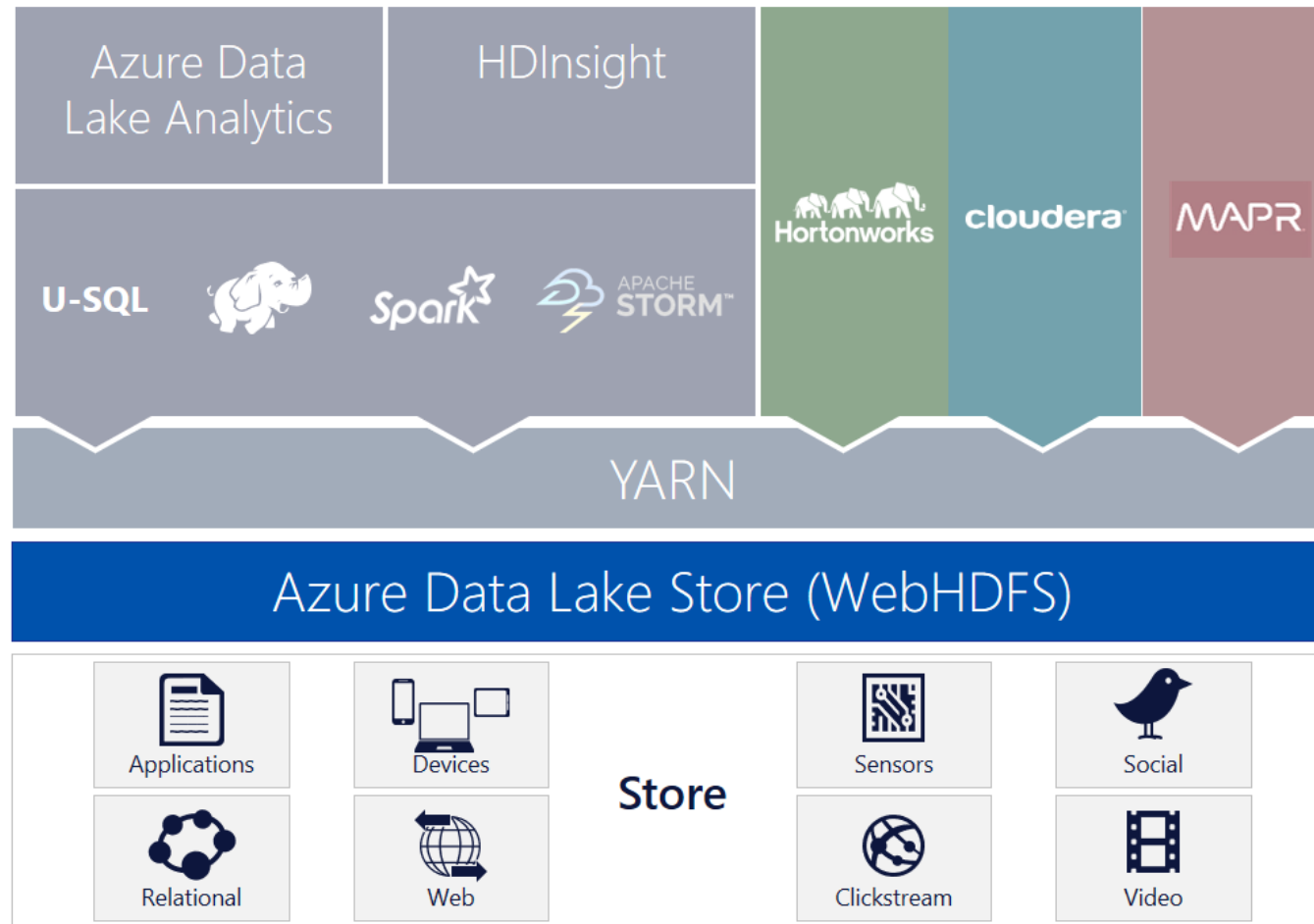
# Azure Blob Storage X Azure Data Lake Store

- Azure Blob Storage:
  - Repositório de dados de uso geral
  - Qualquer tipo de arquivo
- Azure Data Lake Store:
  - Armazenamento otimizado para cargas de trabalho de análise de big data
  - Análise de streaming, dados de Machine Learning, arquivos de log, dados de IOT, grandes conjuntos de dados

<https://docs.microsoft.com/pt-br/azure/data-lake-store/data-lake-store-comparison-with-blob-storage>



# Azure Data Lake Analytics - Arquitetura







# Como serei cobrado?

- Pago sob demanda
- Assinatura “pré-paga” ou “pacotes mensais”
- Cobrado por Horas por Unidade de Análise
- AU significa Azure Data Lake Unit, unidade abstrata de computação envolvendo recursos de CPU e memória.

<https://azure.microsoft.com/pt-br/pricing/details/data-lake-analytics>



# U-SQL

- Combina o poder do SQL com C#
- A execução do código pode ser paralelizada para a escala necessária
- Extensível
- Cargas de trabalho:
  - ETL (Extract, Transform, Load)
  - Análise de Dados
  - Machine Learning



# FIM

**Analytics, Visualização, Relatórios e Tomada de Decisões com Big Data**

