Exam DP-201:
Designing an Azure Data Solution

Introduction

Designing an Azure Data Solution (DP-201) certification course is ideal for professionals who strive to upscale their skills in developing Azure data solutions. Moreover, this certification allows the candidate to validate their skill set in order to design marvelous data infrastructures, keeping in mind the cost-effectiveness as well as data security and compliance. The core aspects of Databricks integrated with Azure are also taught to the candidate during the training.

Course Outline

MODULE 1: Design Azure data storage solutions (40-45%)

- Recommend an Azure data storage solution based on requirements
  - choose the correct data storage solution to meet the technical and business requirements
  - choose the partition distribution type

- Design non-relational cloud data stores
  - design data distribution and partitions
  - design for scale (including multi-region, latency, and throughput)
  - design a solution that uses Cosmos DB, Data Lake Storage Gen2, or Blob storage
  - select the appropriate Cosmos DB API
  - design a disaster recovery strategy
  - design for high availability

- Design relational cloud data stores
  - design data distribution and partitions
  - design for scale (including latency, and throughput)
  - design a solution that uses Azure Synapse Analytics
  - design a disaster recovery strategy
  - design for high availability

MODULE 2: Design data processing solutions (25-30%)

- Design batch processing solutions
  - design batch processing solutions that use Data Factory and Azure Databricks
  - identify the optimal data ingestion method for a batch processing solution
  - identify where processing should take place, such as at the source, at the destination, or in transit

- Design real-time processing solutions
  - design for real-time processing by using Stream Analytics and Azure Databricks
  - design and provision compute resources

MODULE 3: Design for data security and compliance (25-30%)

- Design security for source data access
  - plan for secure endpoints (private/public)
  - choose the appropriate authentication mechanism, such as access keys, shared access signatures (SAS), and Azure Active Directory (Azure AD)

- Design security for data policies and standards
  - design data encryption for data at rest and in transit
  - design for data auditing and data masking
  - design for data privacy and data classification
  - design a data retention policy
  - plan an archiving strategy

- plan to purge data based on business requirements

Prerequisites

- Basic Computer Knowledge.
- Design an Azure Data solution
- Design Azure Cloud data warehouses
- Design No-SQL Databases
- Design Azure SQL Database
- Design hybrid data scenarios
- Design batch processing solutions
- Design big data real-time processing solutions
- Design integration solutions
- Design source data access security
- Design security for data policies and standards
- Design a data retention policy
- Design for Optimization
- Design and implement a disaster recovery strategy
- Design for High Availability

Target Audience

Candidates for this exam are Microsoft Azure data engineers who collaborate with business stakeholders to identify and meet the data requirements to design data solutions that use Azure data services.

Azure data engineers are responsible for data-related design tasks that include designing Azure data storage solutions that use relational and non-relational data stores, batch and real-time data processing solutions, and data security and compliance solutions.

Candidates for this exam must design data solutions that use the following Azure services: Azure Cosmos DB, Azure Synapse Analytics, Azure Data Lake Storage, Azure Data Factory, Azure Stream Analytics, Azure Databricks, and Azure Blob storage.

Duration

16 Hours