Exam DP-200: Implementing an Azure Data Solution

Introduction

Data Engineering on Microsoft Azure (Exam DP-203) certification course is ideal for professionals who aspire to become Azure Data Engineers. You will successfully acquire the knowledge and skills to become capable of integrating, transforming, and consolidating the data from various structured and unstructured data systems into structures that are suitable for building analytics solutions. In addition to that, you will understand the data through exploration, and build, maintain, secure, and work with data processing pipelines by using different tools and techniques.

** The Exam DP-200 has been replaced by the Exam DP-203 (beta).

** Further please note that exam DP-200 will retire on June 30, 2021, at 11:59 PM CST (Central Standard Time), after which you wouldn't be able to take this exam.

Course Outline

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

- Implement non-relational data stores
    - implement a solution that uses Cosmos DB, Data Lake Storage Gen2, or Blob storage
    - implement data distribution and partitions
    - implement a consistency model in Cosmos DB
    - provision a non-relational datastore
    - provide access to data to meet security requirements
    - implement for high availability, disaster recovery, and global distribution

- Implement relational data stores
    - provide access to data to meet security requirements
    - implement for high availability and disaster recovery
    - implement data distribution and partitions for Azure Synapse Analytics
    - implement PolyBase

- Manage data security
    - implement data masking
    - encrypt data at rest and in motion

MODULE 2: Manage and develop data processing (25-30%)

- Develop batch processing solutions
    - develop batch processing solutions by using Data Factory and Azure Databricks
    - ingest data by using PolyBase
    - implement the integration runtime for Data Factory
    - create linked services and datasets
    - create pipelines and activities
    - create and schedule triggers
    - implement Azure Databricks clusters, notebooks, jobs, and autoscaling
    - ingest data into Azure Databricks

- Develop streaming solutions
    - configure input and output
    - select the appropriate built-in functions
    - implement event processing by using Stream Analytics

MODULE 3: Monitor and optimize data solutions (30-35%)

- Monitor data storage
    - monitor relational and non-relational data stores
    - implement Blob storage monitoring
    - implement Data Lake Storage Gen2 monitoring
    - implement Azure Synapse Analytics monitoring
    - implement Cosmos DB monitoring
    - configure Azure Monitor alerts
    - implement auditing by using Azure Log Analytics

- Monitor data processing
    - monitor Data Factory pipelines
    - monitor Azure Databricks
    - monitor Stream Analytics
    - configure Azure Monitor alerts
    - implement auditing by using Azure Log Analytics

- Optimize Azure data solutions
    - troubleshoot data partitioning bottlenecks
    - optimize Data Lake Storage Gen2
    - optimize Stream Analytics
    - optimize Azure Synapse Analytics
    - manage the data lifecycle

Prerequisites

- Candidates who wish to take up the Implementing an Azure Data Solution DP-200 certification exam should have a minimum working experience with Microsoft Azure. It would be great if the candidate has already cleared the Microsoft Azure Fundamentals AZ-900 examination.

- A participant should also be able to implement data solutions across Azure services that include Azure Cosmos DB, Azure Synapse Analytics, Azure Data Factory, Azure Stream Analytics, Azure SQL Database, and more.

Target Audience

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

Azure data engineers are responsible for data-related implementation tasks that include provisioning data storage services, ingesting streaming and batch data, transforming data, implementing security requirements, implementing data retention policies, identifying performance bottlenecks, and accessing external data sources.

Candidates for this exam must be able to implement data solutions that use the following Azure services: Azure Cosmos DB, Azure Synapse Analytics (formerly Azure SQL DW), Azure Data Factory, Azure Stream Analytics, Azure Databricks, and Azure Blob storage.

Duration

24 Hours