Exam DP-200: Implementing an Azure Data Solution –
Skills Measured

The content of this exam will be updated on March 31, 2020. Please continue scrolling to the red line section below to view the changes.

Audience Profile

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 SQL Database, Azure Synapse Analytics (formerly Azure SQL DW), Azure Data Lake Storage, Azure Data Factory, Azure Stream Analytics, Azure Databricks, and Azure Blob storage.

Skills Measured

NOTE: The bullets that appear below each of the skills measured are intended to illustrate how we are assessing that skill. This list is not definitive or exhaustive.

NOTE: In most cases, exams do NOT cover preview features, and some features will only be added to an exam when they are GA (General Availability).

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 data store
- provide access to data to meet security requirements
- implement for high availability, disaster recovery, and global distribution

Implement relational data stores
- configure elastic pools
- configure geo-replication
- provide access to data to meet security requirements
- implement for high availability, disaster recovery, and global distribution
- implement data distribution and partitions for Azure Synapse Analytics
- implement PolyBase

**Manage data security**

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

**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 windowing functions
- implement event processing by using Stream Analytics
- ingest and query streaming data with Azure Data Explorer

**Monitor and Optimize Data Solutions (30-35%)**

**Monitor data storage**

- monitor relational and non-relational data sources
- implement Blob storage monitoring
- implement Data Lake Storage monitoring
- implement SQL Database monitoring
- implement Azure Synapse Analytics monitoring
- implement Cosmos DB monitoring
- implement Azure Data Explorer 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
• optimize Stream Analytics
• optimize Azure Synapse Analytics
• optimize SQL Database
• manage the data lifecycle

See below changes as of March 31, 2020

Audience Profile

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 SQL Database, Azure Synapse Analytics (formerly Azure SQL DW), Azure Data Lake Storage, Azure Data Factory, Azure Stream Analytics, Azure Databricks, and Azure Blob storage.

Skills Measured

NOTE: The bullets that appear below each of the skills measured are intended to illustrate how we are assessing that skill. This list is not definitive or exhaustive.

NOTE: In most cases, exams do NOT cover preview features, and some features will only be added to an exam when they are GA (General Availability).
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 data store
- provide access to data to meet security requirements
- implement for high availability, disaster recovery, and global distribution

Implement relational data stores

- configure elastic pools
- configure geo-replication
- provide access to data to meet security requirements
- implement for high availability, disaster recovery, and global distribution
- implement data distribution and partitions for Azure Synapse Analytics
- implement PolyBase

Manage data security

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

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
- implement Copy Activity within Azure Data Factory
- create linked services and datasets
- create pipelines and activities
- implement Mapping Data Flows in Azure Data Factory
- 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 windowing functions
Monitor and Optimize Data Solutions (30-35%)

Monitor data storage

- monitor relational and non-relational data sources
- implement Blob storage monitoring
- implement Data Lake Storage monitoring
- implement SQL Database monitoring
- implement Azure Synapse Analytics monitoring
- implement Cosmos DB monitoring
- implement Azure Data Explorer 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
- optimize Stream Analytics
- optimize Azure Synapse Analytics
- optimize SQL Database
- manage the data lifecycle