DP-201
Designing an Azure Data Solution

Exam number: DP-201

Exam title: Designing an Azure Data Solution

Language(s) this exam will be available in: English

Audience (IT professionals, Developers, Information workers, etc.): English

Technology: Microsoft Azure

Exam provider (VUE, Certiport, or both): VUE

Exam Design
Note: This document shows tracked changes that are effective as of June 21, 2019.

Audience Profile

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.

~~Candidates for this exam are~~ Azure data engineers ~~who~~ are responsible for ~~data~~ data-related 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. ~~ingesting, egressing, and transforming data from multiple sources using various services and tools. The Azure data engineer collaborates with business stakeholders to identify and meet data requirements while designing and implementing the management, monitoring, security, and privacy of data using the full stack of Azure services to satisfy business needs.~~

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

Skills measured

Note: This document shows tracked changes that are effective as of June 21, 2019.

Design Azure Data Storage Solutions (~~3040~~-~~3545~~%)

~~Design~~ Recommend an Azure Data solution based on requirements
may include but is not limited to:

- Choose the correct data storage solution to meet the technical and business requirements
- ~~Calculate the storage capacity of data~~
- ~~Recommend tier based on data requirements~~
- ~~Decide~~ Choose the partition distribution type

~~Design Azure Cloud data warehouses~~
~~may include but is not limited to:~~

- ~~Design the data schema~~
- ~~Design for scale~~

Design ~~No-SQL Databases~~non-relational cloud data stores
may include but is not limited to:

- 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 a CosmosDB solution~~
- ~~Select the database platform~~
- ~~Model data storage based on use cases~~
- ~~Select storage types~~

Design ~~Azure SQL Database~~relational cloud data stores
may include but is not limited to:

- Design data distribution and partitions
- Design for scale, including multi-region, latency, and throughput
- Design a solution that uses SQL Database and SQL Data Warehouse
- Design a disaster recovery strategy
- Design for high availability
- ~~Design data load technologies~~
- ~~Design solution architecture~~

~~Design hybrid data scenarios~~
~~may include but is not limited to:~~

- ~~Design hybrid solution~~
- ~~Design data replication and synchronization~~

Design Data Processing Solutions (25-30%)
Design batch processing solutions
may include but is not limited to:
- ~~Design batch processing solutions using Spark~~
- Design batch processing solutions by using 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 ~~big data~~ real-time processing solutions
may include but is not limited to:
- Design for real-time processing by using Stream Analytics and Azure Databricks
- ~~Design for LAMBDA architecture~~
- Design and provision compute resources

~~Design integration solutions~~
~~may include but is not limited to:~~
- ~~Design cloud analytic solutions~~
- ~~Design data processing activities~~
- ~~Design data migration~~

Design for Data Security and Compliance (~~21~~5-~~2030~~%)
Design ~~security for~~ source data access ~~security~~
may include but is not limited to:
- Plan for secure endpoints
- ~~Design network security~~
- Choose the appropriate authentication mechanism , such as access keys, shared access signatures (SAS), and Azure Active Directory (Azure AD)
- ~~Choose the appropriate authorization mechanism~~

Design security for data policies and standards
may include but is not limited to:
- Design for 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

~~Design a data retention policy~~
~~may include but is not limited to:~~
- ~~Design a backup and recovery policy~~
- ~~Evaluate business requirements for data retention~~

~~Design for High Availability and Disaster Recovery (20-25%)~~
~~Design for Optimization~~
~~may include but is not limited to:~~

- Design for optimized HIVE processing
- Design for optimized Data Lake
- Design for optimized SPARK processing
- Design for optimized Azure Stream Analytics
- Design for optimized Data Warehouse
- Design for optimized SQL DB, plan for the data life cycle

Design and implement a disaster recovery strategy
may include but is not limited to:
- Design a backup and recovery policy
- Design an online disaster recovery strategy
- Design for multi-region availability
- Plan for backup and restore

Design for High Availability
may include but is not limited to:
- Design for multi-region availability
- Design for read scale
- Design for multi-master writes