



**Microsoft**

**Exam Questions DP-201**

Designing an Azure Data Solution

## NEW QUESTION 1

- (Exam Topic 1)

HOTSPOT

You need to ensure that security policies for the unauthorized detection system are met. What should you recommend? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

| Setting               | Value   |
|-----------------------|---|
| Audit log destination | <div> <div>▼</div> <div> Storage queue<br/> Event Hub<br/> Event Grid<br/> Blob storage </div> </div> |
| Detection app service | <div> <div>▼</div> <div> Function App<br/> Web App<br/> API App </div> </div>                         |

- A. Mastered
- B. Not Mastered

**Answer:** A

### Explanation:

Box 1: Blob storage

Configure blob storage for audit logs.

Scenario: Unauthorized usage of the Planning Assistance data must be detected as quickly as possible. Unauthorized usage is determined by looking for an unusual pattern of usage.

Data used for Planning Assistance must be stored in a sharded Azure SQL Database. Box 2: Web Apps

SQL Advanced Threat Protection (ATP) is to be used.

One of Azure's most popular service is App Service which enables customers to build and host web applications in the programming language of their choice without managing infrastructure. App Service offers auto-scaling and high availability, supports both Windows and Linux. It also supports automated deployments from GitHub, Visual Studio Team Services or any Git repository. At RSA, we announced that Azure Security Center leverages the scale of the cloud to identify attacks targeting App Service applications.

References:

<https://azure.microsoft.com/sv-se/blog/azure-security-center-can-identify-attacks-targeting-azure-app-service-ap>

## NEW QUESTION 2

- (Exam Topic 1)

STION NO: 5 HOTSPOT

You need to design the authentication and authorization methods for sensors.

What should you recommend? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

| Requirement    | Method  |
|----------------|---|
| Authentication | <div> <div>▼</div> <div> HMAC header<br/> Resource Token<br/> Azure Managed Identity<br/> Storage account connection string </div> </div> |
| Authorization  | <div> <div>▼</div> <div> Custom RBAC role<br/> Cosmos DB user<br/> Azure Active Directory user<br/> IoT device identity </div> </div>     |

- A. Mastered
- B. Not Mastered

**Answer:** A

### Explanation:

Sensor data must be stored in a Cosmos DB named treydata in a collection named SensorData Sensors must have permission only to add items to the SensorData collection

Box 1: Resource Token

Resource tokens provide access to the application resources within a Cosmos DB database.

Enable clients to read, write, and delete resources in the Cosmos DB account according to the permissions they've been granted.

Box 2: Cosmos DB user

You can use a resource token (by creating Cosmos DB users and permissions) when you want to provide access to resources in your Cosmos DB account to a client that cannot be trusted with the master key.

#### References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data>

### NEW QUESTION 3

- (Exam Topic 2)

You need to design the encryption strategy for the tagging data and customer data.

What should you recommend? To answer, drag the appropriate setting to the correct drop targets. Each source may be used once, more than once, or not at all.

You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

**Answer Area**

| Encryption methods          | Solution component      | Encryption method |
|-----------------------------|-------------------------|-------------------|
| Encryption at rest          | Tagging data            |                   |
| Transparent data encryption | Processed customer data |                   |
| Azure Key Vault             |                         |                   |

- A. Mastered
- B. Not Mastered

**Answer:** A

#### Explanation:

All cloud data must be encrypted at rest and in transit. Box 1: Transparent data encryption

Encryption of the database file is performed at the page level. The pages in an encrypted database are encrypted before they are written to disk and decrypted when read into memory.

Box 2: Encryption at rest

Encryption at Rest is the encoding (encryption) of data when it is persisted. References:

<https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/transparent-data-encryption?view=> <https://docs.microsoft.com/en-us/azure/security/azure-security-encryption-atrest>

### NEW QUESTION 4

- (Exam Topic 3)

You plan to use Azure SQL Database to support a line of business app.

You need to identify sensitive data that is stored in the database and monitor access to the data. Which three actions should you recommend? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Enable Data Discovery and Classification.
- B. Implement Transparent Data Encryption (TDE).
- C. Enable Auditing.
- D. Run Vulnerability Assessment.
- E. Use Advanced Threat Protection.

**Answer:** CDE

### NEW QUESTION 5

- (Exam Topic 4)

A company stores data in multiple types of cloud-based databases.

You need to design a solution to consolidate data into a single relational database. Ingestion of data will occur at set times each day.

What should you recommend?

- A. SQL Server Migration Assistant
- B. SQL Data Sync
- C. Azure Data Factory
- D. Azure Database Migration Service
- E. Data Migration Assistant

**Answer:** C

#### Explanation:

<https://docs.microsoft.com/en-us/azure/data-factory/introduction>

<https://azure.microsoft.com/en-us/blog/operationalize-azure-databricks-notebooks-using-data-factory/> <https://azure.microsoft.com/en-us/blog/data-ingestion-into-azure-at-scale-made-easier-with-latest-enhancements>

### NEW QUESTION 6

- (Exam Topic 4)

You design data engineering solutions for a company.

You must integrate on-premises SQL Server data into an Azure solution that performs Extract-Transform-Load (ETL) operations have the following requirements:

- ▶ Develop a pipeline that can integrate data and run notebooks.
- ▶ Develop notebooks to transform the data.
- ▶ Load the data into a massively parallel processing database for later analysis. You need to recommend a solution.

What should you recommend? To answer, select the appropriate options in the answer area.  
 NOTE: Each correct selection is worth one point.

| Requirement                                    | Service   |
|--|---|
| Integrate the on-premises data into the cloud. | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Develop notebooks to transform the data.       | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Run notebooks.                                 | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Load the data.                                 | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Store the transformed data.                    | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

| Requirement                                    | Service   |
|--|---|
| Integrate the on-premises data into the cloud. | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Develop notebooks to transform the data.       | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Run notebooks.                                 | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Load the data.                                 | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |
| Store the transformed data.                    | <div>▼</div> <div>                     Azure Databricks<br/>                     Azure Data Factory<br/>                     Azure SQL Data Warehouse<br/>                     Azure Batch                 </div> |

## NEW QUESTION 7

- (Exam Topic 4)

You are designing a real-time stream solution based on Azure Functions. The solution will process data uploaded to Azure Blob Storage.

The solution requirements are as follows:

New blobs must be processed with a little delay as possible. Scaling must occur automatically.

Costs must be minimized. What should you recommend?

- A. Deploy the Azure Function in an App Service plan and use a Blob trigger.
- B. Deploy the Azure Function in a Consumption plan and use an Event Grid trigger.
- C. Deploy the Azure Function in a Consumption plan and use a Blob trigger.
- D. Deploy the Azure Function in an App Service plan and use an Event Grid trigger.

**Answer: C**

**Explanation:**

Create a function, with the help of a blob trigger template, which is triggered when files are uploaded to or updated in Azure Blob storage.

You use a consumption plan, which is a hosting plan that defines how resources are allocated to your function app. In the default Consumption Plan, resources are added dynamically as required by your functions. In this serverless hosting, you only pay for the time your functions run. When you run in an App Service plan, you must manage the scaling of your function app.

References:

<https://docs.microsoft.com/en-us/azure/azure-functions/functions-create-storage-blob-triggered-function>

**NEW QUESTION 8**

- (Exam Topic 4)

You need to design the unauthorized data usage detection system. What Azure service should you include in the design?

- A. Azure Databricks
- B. Azure SQL Data Warehouse
- C. Azure Analysis Services
- D. Azure Data Factory

**Answer: B**

**NEW QUESTION 9**

- (Exam Topic 4)

A company has many applications. Each application is supported by separate on-premises databases. You must migrate the databases to Azure SQL Database.

You have the following requirements: Organize databases into groups based on database usage.

Define the maximum resource limit available for each group of databases.

You need to recommend technologies to scale the databases to support expected increases in demand. What should you recommend?

- A. Read scale-out
- B. Managed instances
- C. Elastic pools
- D. Database sharding

**Answer: C**

**Explanation:**

SQL Database elastic pools are a simple, cost-effective solution for managing and scaling multiple databases that have varying and unpredictable usage demands. The databases in an elastic pool are on a single Azure SQL Database server and share a set number of resources at a set price.

You can configure resources for the pool based either on the DTU-based purchasing model or the vCorebased purchasing model.

**NEW QUESTION 10**

- (Exam Topic 4)

You are designing a solution for a company. You plan to use Azure Databricks. You need to recommend workloads and tiers to meet the following requirements:

- ▶ Provide managed clusters for running production jobs.
- ▶ Provide persistent clusters that support auto-scaling for analytics processes.
- ▶ Provide role-based access control (RBAC) support for Notebooks.

What should you recommend? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

| Requirement  | Workload   | Tier   |
|--|--|--|
| Provide managed clusters for running production jobs.                          | <div> <div>▼</div> <div> Data Engineering only<br/> Data Analytics only<br/> Data Engineering and Data Analytics </div> </div> | <div> <div>▼</div> <div>Standard</div> </div>                |
| Provide persistent clusters that support auto-scaling for analytics processes. | <div> <div>▼</div> <div> Data Engineering only<br/> Data Analytics only<br/> Data Engineering and Data Analytics </div> </div> | <div> <div>▼</div> <div> Standard<br/> Premium </div> </div> |
| Provide role-based access control (RBAC) support for Notebooks.                | <div> <div>▼</div> <div> Data Engineering only<br/> Data Analytics only<br/> Data Engineering and Data Analytics </div> </div> | <div> <div>▼</div> <div> Standard<br/> Premium </div> </div> |

A. Mastered





NOTE: Each correct selection is worth one point.

- A. Use the Premium tier and the default backup retention policy.
- B. Use the Basic tier and the default backup retention policy.
- C. Use the Standard tier and the default backup retention policy.
- D. Use the Standard tier and configure a long-term backup retention policy.
- E. Use the Premium tier and configure a long-term backup retention policy.

**Answer:** DE

**Explanation:**

The default retention period for a database created using the DTU-based purchasing model depends on the service tier:

- ▶ Basic service tier is 1 week.
- ▶ Standard service tier is 5 weeks.
- ▶ Premium service tier is 5 weeks.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-long-term-retention>

**NEW QUESTION 19**

- (Exam Topic 4)

You are designing an Azure SQL Data Warehouse. You plan to load millions of rows of data into the data warehouse each day.

You must ensure that staging tables are optimized for data loading. You need to design the staging tables.

What type of tables should you recommend?

- A. Round-robin distributed table
- B. Hash-distributed table
- C. Replicated table
- D. External table

**Answer:** A

**Explanation:**

To achieve the fastest loading speed for moving data into a data warehouse table, load data into a staging table. Define the staging table as a heap and use round-robin for the distribution option.

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/guidance-for-loading-data>

**NEW QUESTION 24**

- (Exam Topic 4)

You are designing a solution for a company. The solution will use model training for objective classification. You need to design the solution.

What should you recommend?

- A. an Azure Cognitive Services application
- B. a Spark Streaming job
- C. interactive Spark queries
- D. Power BI models
- E. a Spark application that uses Spark MLlib.

**Answer:** E

**Explanation:**

Spark in SQL Server big data cluster enables AI and machine learning.

You can use Apache Spark MLlib to create a machine learning application to do simple predictive analysis on an open dataset.

MLlib is a core Spark library that provides many utilities useful for machine learning tasks, including utilities that are suitable for:

- ▶ Classification
- ▶ Regression
- ▶ Clustering
- ▶ Topic modeling
- ▶ Singular value decomposition (SVD) and principal component analysis (PCA)
- ▶ Hypothesis testing and calculating sample statistics

References:

<https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-machine-learning-mllib-ipynb>

**NEW QUESTION 25**

- (Exam Topic 4)

You have an on-premises MySQL database that is 800 GB in size.

You need to migrate a MySQL database to Azure Database for MySQL. You must minimize service interruption to live sites or applications that use the database.

What should you recommend?

- A. Azure Database Migration Service
- B. Dump and restore
- C. Import and export
- D. MySQL Workbench

**Answer:** A

**Explanation:**

You can perform MySQL migrations to Azure Database for MySQL with minimal downtime by using the newly introduced continuous sync capability for the Azure

Database Migration Service (DMS). This functionality limits the amount of downtime that is incurred by the application. References:  
<https://docs.microsoft.com/en-us/azure/mysql/howto-migrate-online>

#### NEW QUESTION 28

- (Exam Topic 4)

You are designing an Azure Databricks cluster that runs user-defined local processes. You need to recommend a cluster configuration that meets the following requirements:

- Minimize query latency.
- Reduce overall costs.
- Maximize the number of users that can run queries on the cluster at the same time. Which cluster type should you recommend?

- A. Standard with Autoscaling
- B. High Concurrency with Auto Termination
- C. High Concurrency with Autoscaling
- D. Standard with Auto Termination

**Answer:** C

#### Explanation:

High Concurrency clusters allow multiple users to run queries on the cluster at the same time, while minimizing query latency. Autoscaling clusters can reduce overall costs compared to a statically-sized cluster.

References:

<https://docs.azuredatabricks.net/user-guide/clusters/create.html> <https://docs.azuredatabricks.net/user-guide/clusters/high-concurrency.html#high-concurrency>  
<https://docs.azuredatabricks.net/user-guide/clusters/terminate.html> <https://docs.azuredatabricks.net/user-guide/clusters/sizing.html#enable-and-configure-autoscaling>

#### NEW QUESTION 29

- (Exam Topic 4)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are designing an HDInsight/Hadoop cluster solution that uses Azure Data Lake Gen1 Storage. The solution requires POSIX permissions and enables diagnostics logging for auditing.

You need to recommend solutions that optimize storage.

Proposed Solution: Implement compaction jobs to combine small files into larger files. Does the solution meet the goal?

- A. Yes
- B. No

**Answer:** A

#### Explanation:

Depending on what services and workloads are using the data, a good size to consider for files is 256 MB or greater. If the file sizes cannot be batched when landing in Data Lake Storage Gen1, you can have a separate compaction job that combines these files into larger ones.

Note: POSIX permissions and auditing in Data Lake Storage Gen1 comes with an overhead that becomes apparent when working with numerous small files. As a best practice, you must batch your data into larger files versus writing thousands or millions of small files to Data Lake Storage Gen1. Avoiding small file sizes can have multiple benefits, such as:

Lowering the authentication checks across multiple files Reduced open file connections

Faster copying/replication

Fewer files to process when updating Data Lake Storage Gen1 POSIX permissions References:

<https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-best-practices>

#### NEW QUESTION 30

- (Exam Topic 4)

You are evaluating data storage solutions to support a new application.

You need to recommend a data storage solution that represents data by using nodes and relationships in graph structures.

Which data storage solution should you recommend?

- A. Blob Storage
- B. Cosmos DB
- C. Data Lake Store
- D. HDInsight

**Answer:** B

#### Explanation:

For large graphs with lots of entities and relationships, you can perform very complex analyses very quickly. Many graph databases provide a query language that you can use to traverse a network of relationships efficiently.

Relevant Azure service: Cosmos DB

References:

<https://docs.microsoft.com/en-us/azure/architecture/guide/technology-choices/data-store-overview>

#### NEW QUESTION 31

- (Exam Topic 4)

A company has locations in North America and Europe. The company uses Azure SQL Database to support business apps.

Employees must be able to access the app data in case of a region-wide outage. A multi-region availability solution is needed with the following requirements:

- Read-access to data in a secondary region must be available only in case of an outage of the primary region.
- The Azure SQL Database compute and storage layers must be integrated and replicated together.



You need to design the multi-region high availability solution.  
 What should you recommend? To answer, select the appropriate values in the answer area.  
 NOTE: Each correct selection is worth one point.

| Option          | Value  |
|-----------------|--|
| Service tier    | <div>▼</div> <div>Basic</div> <div>Standard</div> <div>General</div> <div>Premium</div>      |
| Redundancy type | <div>▼</div> <div>SQL Sync</div> <div>Zone-redundancy</div> <div>Geo-redundant storage</div> |

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Box 1: Standard

The following table describes the types of storage accounts and their capabilities:

| Storage account type | Supported services                 | Supported performance tiers    | Supported access tiers          | Replication options                 | Deployment model <sup>1</sup> |
|----------------------|------------------------------------|--------------------------------|---------------------------------|-------------------------------------|-------------------------------|
| General-purpose V2   | Blob, File, Queue, Table, and Disk | Standard, Premium <sup>5</sup> | Hot, Cool, Archive <sup>3</sup> | LRS, ZRS <sup>4</sup> , GRS, RA-GRS | Resource Manager              |
| General-purpose V1   | Blob, File, Queue, Table, and Disk | Standard, Premium <sup>5</sup> | N/A                             | LRS, GRS, RA-GRS                    | Resource Manager, Classic     |

Box 2: Geo-redundant storage

If your storage account has GRS enabled, then your data is durable even in the case of a complete regional outage or a disaster in which the primary region isn't recoverable.

Note: If you opt for GRS, you have two related options to choose from:

GRS replicates your data to another data center in a secondary region, but that data is available to be read only if Microsoft initiates a failover from the primary to secondary region.

Read-access geo-redundant storage (RA-GRS) is based on GRS. RA-GRS replicates your data to another data center in a secondary region, and also provides you with the option to read from the secondary region. With RA-GRS, you can read from the secondary region regardless of whether Microsoft initiates a failover from the primary to secondary region.

| Scenario   | LRS | ZRS | GRS | RA-GRS |
|--|-----|-----|-----|--------|
| Node unavailability within a data center                       | Yes | Yes | Yes | Yes    |
| An entire data center (zonal or non-zonal) becomes unavailable | No  | Yes | Yes | Yes    |
| A region-wide outage   | No  | No  | Yes | Yes    |

**References:**

<https://docs.microsoft.com/en-us/azure/storage/common/storage-introduction> <https://docs.microsoft.com/en-us/azure/storage/common/storage-redundancy-grs>

**NEW QUESTION 35**

- (Exam Topic 4)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are designing an Azure SQL Database that will use elastic pools. You plan to store data about customers in a table. Each record uses a value for CustomerID.

You need to recommend a strategy to partition data based on values in CustomerID. Proposed Solution: Separate data into customer regions by using horizontal partitioning. Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

**Explanation:**

We should use Horizontal Partitioning through Sharding, not divide through regions.  
 Note: Horizontal Partitioning - Sharding: Data is partitioned horizontally to distribute rows across a scaled out data tier. With this approach, the schema is identical on all participating databases. This approach is also called “sharding”. Sharding can be performed and managed using (1) the elastic database tools libraries or (2) self-sharding. An elastic query is used to query or compile reports across many shards.  
 References:  
<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-query-overview>

**NEW QUESTION 36**

- (Exam Topic 4)  
 You are designing a Spark job that performs batch processing of daily web log traffic.  
 When you deploy the job in the production environment, it must meet the following requirements:

- ▶ Run once a day.
  - ▶ Display status information on the company intranet as the job runs. You need to recommend technologies for triggering and monitoring jobs.
- Which technologies should you recommend? To answer, drag the appropriate technologies to the correct locations. Each technology may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.  
 NOTE: Each correct selection is worth one point.

**Technologies**

Livy

Beeline

Azure Logic App

Azure API App

| Requirement        | Technology |
|--------------------|------------|
| Triggering of jobs |            |
| Monitoring of jobs |            |

- A. Mastered
- B. Not Mastered

Answer: A

**Explanation:**

Box 1: Livy  
 You can use Livy to run interactive Spark shells or submit batch jobs to be run on Spark. Box 2: Beeline  
 Apache Beeline can be used to run Apache Hive queries on HDInsight. You can use Beeline with Apache Spark.  
 Note: Beeline is a Hive client that is included on the head nodes of your HDInsight cluster. Beeline uses JDBC to connect to HiveServer2, a service hosted on your HDInsight cluster. You can also use Beeline to access Hive on HDInsight remotely over the internet.  
 References:  
<https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-livy-rest-interface> <https://docs.microsoft.com/en-us/azure/hdinsight/hadoop/apache-hadoop-use-hive-beeline>

**NEW QUESTION 41**

- (Exam Topic 4)  
 You need to design the system for notifying law enforcement officers about speeding vehicles.  
 How should you design the pipeline? To answer, drag the appropriate services to the correct locations. Each service may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.  
 NOTE: Each correct selection is worth one point.

**Services**

kafka
 Storage queue
 API Management

Azure Databricks
 Traffic Manager
 Azure SQL database

Telemetry Data

service

Telemetry

service

Speeding Data

service

- A. Mastered
- B. Not Mastered

Answer: A

**Explanation:**



NEW QUESTION 45  
.....

## Thank You for Trying Our Product

### We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

### DP-201 Practice Exam Features:

- \* DP-201 Questions and Answers Updated Frequently
- \* DP-201 Practice Questions Verified by Expert Senior Certified Staff
- \* DP-201 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- \* DP-201 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

**100% Actual & Verified — Instant Download, Please Click**  
**[Order The DP-201 Practice Test Here](#)**