

➤ **Vendor: Microsoft**

➤ **Exam Code: DP-201**

➤ **Exam Name: Designing an Azure Data Solution**

➤ **New Updated Questions from [Braindump2go](#) (Updated in [Feb./2021](#))**

[Visit Braindump2go and Download Full Version DP-201 Exam Dumps](#)

**QUESTION 181**

You are planning a solution that combines log data from multiple systems. The log data will be downloaded from an API and stored in a data store.

You plan to keep a copy of the raw data as well as some transformed versions of the data. You expect that there will be at least 2 TB of log files. The data will be used by data scientists and applications.

You need to recommend a solution to store the data in Azure. The solution must minimize costs.

What storage solution should you recommend?

- A. Azure Data Lake Storage Gen2
- B. Azure Synapse Analytics
- C. Azure SQL Database
- D. Azure Cosmos DB

**Answer: A**

**Explanation:**

To land the data in Azure storage, you can move it to Azure Blob storage or Azure Data Lake Store Gen2.

In either location, the data should be stored in text files. PolyBase and the COPY statement can load from either location.

Incorrect Answers:

B: Azure Synapse Analytics, uses distributed query processing architecture that takes advantage of the scalability and flexibility of compute and storage resources. Use Azure Synapse Analytics transform and move the data.

Reference:

<https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/design-elt-data-loading>

**QUESTION 182**

You are designing a serving layer for data. The design must meet the following requirements:

- Authenticate users by using Azure Active Directory (Azure AD).
- Serve as a hot path for data.
- Support query scale out.
- Support SQL queries.

What should you include in the design?

- A. Azure Data Lake Storage
- B. Azure Cosmos DB
- C. Azure Blob storage
- D. Azure Synapse Analytics

**Answer: B**

**Explanation:**

Do you need serving storage that can serve as a hot path for your data? If yes, narrow your options to those that are optimized for a speed serving layer. This would be Cosmos DB among the options given in this question.

[DP-201 Exam Dumps](#) [DP-201 Exam Questions](#) [DP-201 PDF Dumps](#) [DP-201 VCE Dumps](#)

<https://www.braindump2go.com/dp-201.html>

Note: Analytical data stores that support querying of both hot-path and cold-path data are collectively referred to as the serving layer, or data serving storage.

There are several options for data serving storage in Azure, depending on your needs:

- Azure Synapse Analytics
- Azure Cosmos DB
- Azure Data Explorer
- Azure SQL Database
- SQL Server in Azure VM
- HBase/Phoenix on HDInsight
- Hive LLAP on HDInsight
- Azure Analysis Services

Incorrect Answers:

A, C: Azure Data Lake Storage & Azure Blob storage are not data serving storage in Azure.

Reference:

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

### QUESTION 183

You are designing a storage solution for streaming data that is processed by Azure Databricks. The solution must meet the following requirements:

- The data schema must be fluid.
- The source data must have a high throughput.
- The data must be available in multiple Azure regions as quickly as possible.

What should you include in the solution to meet the requirements?

- A. Azure Cosmos DB
- B. Azure Synapse Analytics
- C. Azure SQL Database
- D. Azure Data Lake Storage

**Answer:** A

**Explanation:**

Azure Cosmos DB is Microsoft's globally distributed, multi-model database. Azure Cosmos DB enables you to elastically and independently scale throughput and storage across any number of Azure's geographic regions. It offers throughput, latency, availability, and consistency guarantees with comprehensive service level agreements (SLAs). You can read data from and write data to Azure Cosmos DB using Databricks.

Note on fluid schema:

If you are managing data whose structures are constantly changing at a high rate, particularly if transactions can come from external sources where it is difficult to enforce conformity across the database, you may want to consider a more schema-agnostic approach using a managed NoSQL database service like Azure Cosmos DB.

Reference:

<https://docs.databricks.com/data/data-sources/azure/cosmosdb-connector.html>

<https://docs.microsoft.com/en-us/azure/cosmos-db/relational-nosql>

### QUESTION 184

You are designing a log storage solution that will use Azure Blob storage containers.

CSV log files will be generated by a multi-tenant application. The log files will be generated for each customer at five-minute intervals. There will be more than 5,000 customers. Typically, the customers will query data generated on the day the data was created.

You need to recommend a naming convention for the virtual directories and files. The solution must minimize the time it takes for the customers to query the log files.

What naming convention should you recommend?

- A. {year}/{month}/{day}/{hour}/{minute}/{CustomerID}.csv
- B. {year}/{month}/{day}/{CustomerID}/{hour}/{minute}.csv
- C. {minute}/{hour}/{day}/{month}/{year}/{CustomerID}.csv
- D. {CustomerID}/{year}/{month}/{day}/{hour}/{minute}.csv

**Answer:** B

**Explanation:**

<https://docs.microsoft.com/en-us/azure/cdn/cdn-azure-diagnostic-logs>

**QUESTION 185**

You are designing an anomaly detection solution for streaming data from an Azure IoT hub. The solution must meet the following requirements:

- Send the output to Azure Synapse.
- Identify spikes and dips in time series data.
- Minimize development and configuration effort

Which should you include in the solution?

- A. Azure Databricks
- B. Azure Stream Analytics
- C. Azure SQL Database

**Answer:** B

**Explanation:**

You can identify anomalies by routing data via IoT Hub to a built-in ML model in Azure Stream Analytics.

Reference:

<https://docs.microsoft.com/en-us/learn/modules/data-anomaly-detection-using-azure-iot-hub/>

**QUESTION 186**

You have an Azure Databricks workspace named workspace1 in the Standard pricing tier. Workspace1 contains an all-purpose cluster named cluster1.

You need to reduce the time it takes for cluster1 to start and scale up. The solution must minimize costs.

What should you do first?

- A. Upgrade workspace1 to the Premium pricing tier.
- B. Create a pool in workspace1.
- C. Configure a global init script for workspace1.
- D. Create a cluster policy in workspace1.

**Answer:** B

**Explanation:**

Databricks Pools increase the productivity of both Data Engineers and Data Analysts. With Pools, Databricks customers eliminate slow cluster start and auto-scaling times. Data Engineers can reduce the time it takes to run short jobs in their data pipeline, thereby providing better SLAs to their downstream teams.

Reference:

<https://databricks.com/blog/2019/11/11/databricks-pools-speed-up-data-pipelines.html>

**QUESTION 187**

You have a large amount of sensor data stored in an Azure Data Lake Storage Gen2 account. The files are in the Parquet file format.

New sensor data will be published to Azure Event Hubs.

You need to recommend a solution to add the new sensor data to the existing sensor data in real-time. The solution must support the interactive querying of the entire dataset.

Which type of server should you include in the recommendation?

- A. Azure SQL Database
- B. Azure Cosmos DB
- C. Azure Stream Analytics
- D. Azure Databricks

**Answer:** C

**Explanation:**

[DP-201 Exam Dumps](#) [DP-201 Exam Questions](#) [DP-201 PDF Dumps](#) [DP-201 VCE Dumps](#)

<https://www.braindump2go.com/dp-201.html>

Azure Stream Analytics is a fully managed PaaS offering that enables real-time analytics and complex event processing on fast moving data streams.

By outputting data in parquet format into a blob store or a data lake, you can take advantage of Azure Stream Analytics to power large scale streaming extract, transfer, and load (ETL), to run batch processing, to train machine learning algorithms, or to run interactive queries on your historical data.

Reference:

<https://azure.microsoft.com/en-us/blog/new-capabilities-in-stream-analytics-reduce-development-time-for-big-data-apps/>

#### **QUESTION 188**

You are designing a solution that will copy Parquet files stored in an Azure Blob storage account to an Azure Data Lake Storage Gen2 account.

The data will be loaded daily to the data lake and will use a folder structure of {Year}/{Month}/{Day}/.

You need to design a daily Azure Data Factory data load to minimize the data transfer between the two accounts. Which two configurations should you include in the design? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Delete the files in the destination before loading new data.
- B. Filter by the last modified date of the source files.
- C. Delete the source files after they are copied.
- D. Specify a file naming pattern for the destination.

**Answer:** BC

**Explanation:**

B: To copy a subset of files under a folder, specify folderPath with a folder part and fileName with a wildcard filter.

C: After completion: Choose to do nothing with the source file after the data flow runs, delete the source file, or move the source file. The paths for the move are relative.

Reference:

<https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-data-lake-storage>

#### **QUESTION 189**

You have a C# application that process data from an Azure IoT hub and performs complex transformations.

You need to replace the application with a real-time solution. The solution must reuse as much code as possible from the existing application.

- A. Azure Databricks
- B. Azure Event Grid
- C. Azure Stream Analytics
- D. Azure Data Factory

**Answer:** C

**Explanation:**

Azure Stream Analytics on IoT Edge empowers developers to deploy near-real-time analytical intelligence closer to IoT devices so that they can unlock the full value of device-generated data. UDF are available in C# for IoT Edge jobs Azure Stream Analytics on IoT Edge runs within the Azure IoT Edge framework. Once the job is created in Stream Analytics, you can deploy and manage it using IoT Hub.

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-edge>

#### **QUESTION 190**

A company purchases IoT devices to monitor manufacturing machinery. The company uses an IoT appliance to communicate with the IoT devices.

The company must be able to monitor the devices in real-time.

You need to design the solution.

What should you recommend?

- A. Azure Data Factory instance using Azure Portal
- B. Azure Analysis Services using Microsoft Visual Studio
- C. Azure Stream Analytics cloud job using Azure Portal
- D. Azure Data Factory instance using Azure Portal

**Answer: C**

**Explanation:**

The Stream Analytics query language allows to perform CEP (Complex Event Processing) by offering a wide array of functions for analyzing streaming data. This query language supports simple data manipulation, aggregation and analytics functions, geospatial functions, pattern matching and anomaly detection. You can edit queries in the portal or using our development tools, and test them using sample data that is extracted from a live stream.

Note: Stream Analytics is a cost-effective event processing engine that helps uncover real-time insights from devices, sensors, infrastructure, applications and data quickly and easily.

Monitor and manage Stream Analytics resources with Azure PowerShell cmdlets and powershell scripting that execute basic Stream Analytics tasks.

Reference:

<https://cloudblogs.microsoft.com/sqlserver/2014/10/29/microsoft-adds-iot-streaming-analytics-data-production-and-workflow-services-to-azure/>

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-introduction>

**QUESTION 191**

You are designing a storage solution to store CSV files.

You need to grant a data scientist access to read all the files in a single container of an Azure Storage account. The solution must use the principle of least privilege and provide the highest level of security.

What are two possible ways to achieve the goal? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Provide an access key.
- B. Assign the Storage Blob Data Reader role at the container level.
- C. Assign the Reader role to the storage account.
- D. Provide an account shared access signature (SAS).
- E. Provide a user delegation shared access signature (SAS).

**Answer: BE**

**Explanation:**

B: When an Azure role is assigned to an Azure AD security principal, Azure grants access to those resources for that security principal. Access can be scoped to the level of the subscription, the resource group, the storage account, or an individual container or queue.

The built-in Data Reader roles provide read permissions for the data in a container or queue.

Note: Permissions are scoped to the specified resource.

For example, if you assign the Storage Blob Data Reader role to user Mary at the level of a container named sample-container, then Mary is granted read access to all of the blobs in that container.

E: A user delegation SAS is secured with Azure Active Directory (Azure AD) credentials and also by the permissions specified for the SAS. A user delegation SAS applies to Blob storage only.

Reference:

<https://docs.microsoft.com/en-us/azure/storage/common/storage-auth-aad-rbac-portal>

<https://docs.microsoft.com/en-us/azure/storage/common/storage-sas-overview>

**QUESTION 192**

You are designing an Azure Synapse solution that will provide a query interface for the data stored in an Azure Storage account. The storage account is only accessible from a virtual network.

You need to recommend an authentication mechanism to ensure that the solution can access the source data.

What should you recommend?

- A. a shared key
- B. an Azure Active Directory (Azure AD) service principal

- C. a shared access signature (SAS)
- D. anonymous public read access

**Answer: B**

**Explanation:**

Managed Identity authentication is required when your storage account is attached to a VNet.

Regardless of the type of identity chosen a managed identity is a service principal of a special type that may only be used with Azure resources.

Note: An Azure service principal is an identity created for use with applications, hosted services, and automated tools to access Azure resources.

Reference:

<https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/quickstart-bulk-load-copy-tsql-examples>

<https://docs.microsoft.com/en-us/powershell/azure/create-azure-service-principal-azureps>

### QUESTION 193

You are designing a real-time stream processing solution in Azure Stream Analytics. The solution must read data from a blob container in an Azure Storage account via a service endpoint.

You need to recommend an authentication mechanism for the solution.

What should you recommend?

- A. a managed identity
- B. a storage access signature (SAS)
- C. a user-assigned managed identity
- D. an account key

**Answer: A**

**Explanation:**

Azure Stream Analytics supports Managed Identity authentication for both Azure Event Hubs input and output.

Note: First, you create a managed identity for your Azure Stream Analytics job.

1. In the Azure portal, open your Azure Stream Analytics job.

2. From the left navigation menu, select Managed Identity located under Configure. Then, check the box next to Use System-assigned Managed Identity and select Save.

Reference:

<https://docs.microsoft.com/en-us/azure/stream-analytics/event-hubs-managed-identity>

### QUESTION 194

You plan to implement an Azure Data Lake Gen2 storage account.

You need to ensure that the data lake will remain available if a data center fails in the primary Azure region. The solution must minimize costs.

Which type of replication should you use for the storage account?

- A. geo-redundant storage (GRS)
- B. zone-redundant storage (ZRS)
- C. locally-redundant storage (LRS)
- D. geo-zone-redundant storage (GZRS)

**Answer: A**

**Explanation:**

Geo-redundant storage (GRS) copies your data synchronously three times within a single physical location in the primary region using LRS. It then copies your data asynchronously to a single physical location in the secondary region. Incorrect Answers:

B: Zone-redundant storage (ZRS) copies your data synchronously across three Azure availability zones in the primary region. For applications requiring high availability, Microsoft recommends using ZRS in the primary region, and also replicating to a secondary region.

C: Locally redundant storage (LRS) copies your data synchronously three times within a single physical location in the primary region. LRS is the least expensive replication option, but is not recommended for applications requiring high

availability.

D: GZRS is more expensive compared to GRS.

Reference:

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

**QUESTION 195**

Hotspot Question

You are designing a solution to process data from multiple Azure event hubs in near real-time. Once processed, the data will be written to an Azure SQL database.

The solution must meet the following requirements:

- Support the auditing of resource and data changes.
- Support data versioning and rollback.

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

NOTE: Each correct selection is worth one point.

**Answer Area**

Azure service to use:

	▼
Azure Databricks	
Azure Stream Analytics	
Azure Analysis Services	

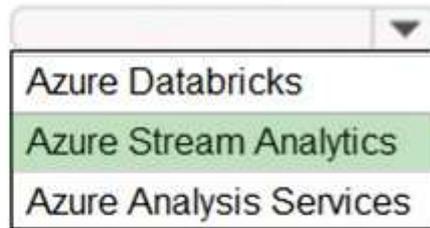
Feature to use:

	▼
Delta	
Replay	
Persistence point	

**Answer:**

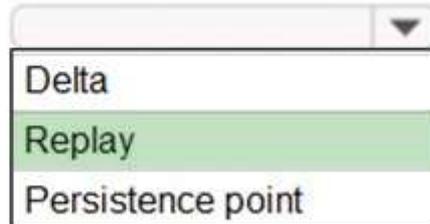
## Answer Area

Azure service to use:



Azure Databricks  
Azure Stream Analytics  
Azure Analysis Services

Feature to use:



Delta  
Replay  
Persistence point

### Explanation:

Box 1: Azure Stream Analytics

Users can now ingest, process, view, and analyze real-time streaming data into a table directly from a database in Azure SQL Database. They do so in the Azure portal using Azure Stream Analytics.

In the Azure portal, you can select an events source (Event Hub/IoT Hub), view incoming real-time events, and select a table to store events.

Stream Analytics leverages versioning of reference data to augment streaming data with the reference data that was valid at the time the event was generated. This ensures repeatability of results.

Box 2: Replay

Reference data is versioned, enabling to always get the same results, even when we "replay" the stream.

Reference:

<https://docs.microsoft.com/en-us/azure/azure-sql/database/stream-data-stream-analytics-integration>

<https://azure.microsoft.com/en-us/updates/additional-support-for-managed-identity-and-new-features-in-azure-stream-analytics/>

### QUESTION 196

Drag and Drop Question

You are designing a real-time processing solution for maintenance work requests that are received via email. The solution will perform the following actions:

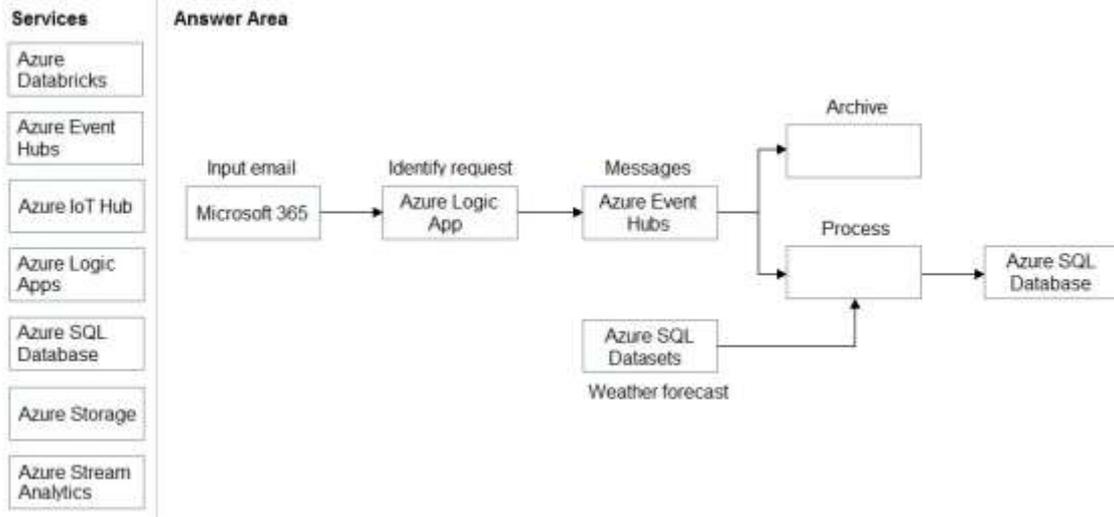
- Store all email messages in an archive.
- Access weather forecast data by using the Python SDK for Azure Open Datasets.
- Identify high priority requests that will be affected by poor weather conditions and store the requests in an Azure SQL database.

The solution must minimize costs.

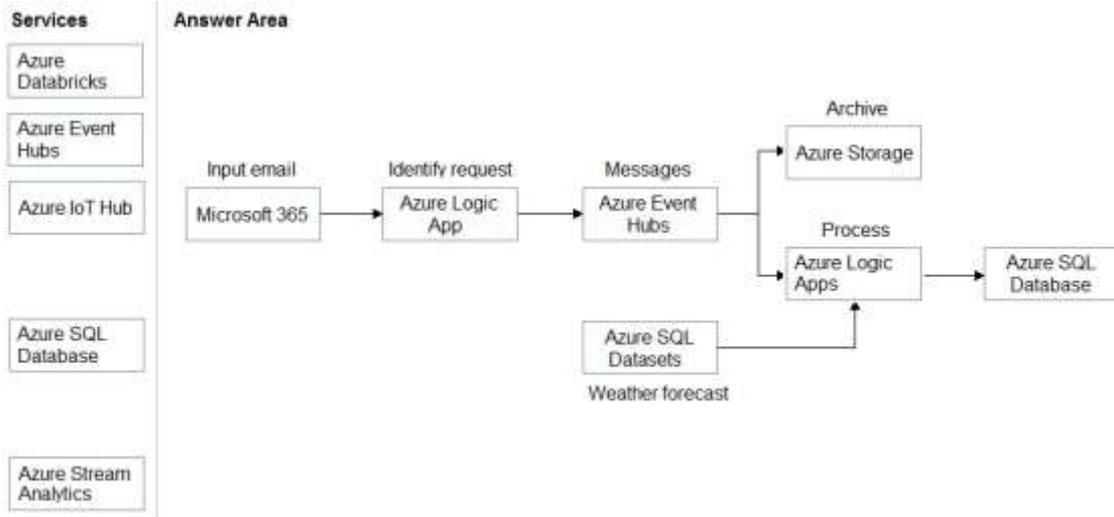
How should you complete the solution? To answer, drag the appropriate services to the correct locations.

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.



**Answer:**



**Explanation:**

Box 1: Azure Storage

Azure Event Hubs enables you to automatically capture the streaming data in Event Hubs in an Azure Blob storage or Azure Data Lake Storage Gen 1 or Gen 2 account of your choice, with the added flexibility of specifying a time or size interval. Setting up Capture is fast, there are no administrative costs to run it, and it scales automatically with Event Hubs throughput units. Event Hubs Capture is the easiest way to load streaming data into Azure, and enables you to focus on data processing rather than on data capture.

Box 2: Azure Logic Apps

You can monitor and manage events sent to Azure Event Hubs from inside a logic app with the Azure Event Hubs connector. That way, you can create logic apps that automate tasks and workflows for checking, sending, and receiving events from your Event Hub.

Reference:

<https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-capture-overview>

<https://docs.microsoft.com/en-us/azure/connectors/connectors-create-api-azure-event-hubs>

**QUESTION 197**

Hotspot Question

You have an Azure Storage account that generates 200,000 new files daily. The file names have a format of {YYYY}/{MM}/{DD}/{HH}/{CustomerID}.csv.

You need to design an Azure Data Factory solution that will load new data from the storage account to an Azure Data Lake once hourly. The solution must minimize load times and costs.

[DP-201 Exam Dumps](#) [DP-201 Exam Questions](#) [DP-201 PDF Dumps](#) [DP-201 VCE Dumps](#)

<https://www.braindump2go.com/dp-201.html>

How should you configure the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point

### Answer Area

Load methodology:

Full Load
Incremental load
Load individual files as they arrive

Trigger:

Fixed schedule
New file
Tumbling window

Answer:

### Answer Area

Load methodology:

Full Load
Incremental load
Load individual files as they arrive

Trigger:

Fixed schedule
New file
Tumbling window

#### Explanation:

Box 1: Incremental load

When you start to build the end to end data integration flow the first challenge is to extract data from different data stores, where incrementally (or delta) loading data after an initial full load is widely used at this stage. Now, ADF provides a new capability for you to incrementally copy new or changed files only by LastModifiedDate from a file-based store. By using this new feature, you do not need to partition the data by time-based folder or file name. The new or changed file will be automatically selected by its metadata LastModifiedDate and copied to the destination store.

Box 2: Tumbling window

[DP-201 Exam Dumps](#) [DP-201 Exam Questions](#) [DP-201 PDF Dumps](#) [DP-201 VCE Dumps](#)

<https://www.braindump2go.com/dp-201.html>

Tumbling window triggers are a type of trigger that fires at a periodic time interval from a specified start time, while retaining state. Tumbling windows are a series of fixed-sized, non-overlapping, and contiguous time intervals. A tumbling window trigger has a one-to-one relationship with a pipeline and can only reference a singular pipeline. Reference:

<https://azure.microsoft.com/en-us/blog/incrementally-copy-new-files-by-lastmodifieddate-with-azure-data-factory/>  
<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-tumbling-window-trigger>

**QUESTION 198**

Hotspot Question

You are designing an Azure Data Factory solution that will download up to 5 TB of data from several REST APIs.

The solution must meet the following staging requirements:

- Ensure that the data can be landed quickly and in parallel to a staging area.
- Minimize the need to return to the API sources to retrieve the data again should a later activity in the pipeline fail.

The solution must meet the following analysis requirements:

- Ensure that the data can be loaded in parallel.
- Ensure that users and applications can query the data without requiring an additional compute engine.

What should you include in the solution to meet the requirements? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Answer Area**

Staging requirements:

<input type="checkbox"/>	Azure Blob storage
<input type="checkbox"/>	Azure SQL Database
<input type="checkbox"/>	Azure Synapse Analytics

Analysis requirements:

<input type="checkbox"/>	Azure Blob storage
<input type="checkbox"/>	Azure SQL Database
<input type="checkbox"/>	Azure Synapse Analytics

**Answer:**

## Answer Area

Staging requirements:

	▼
Azure Blob storage	
Azure SQL Database	
Azure Synapse Analytics	

Analysis requirements:

	▼
Azure Blob storage	
Azure SQL Database	
Azure Synapse Analytics	

### Explanation:

Box 1: Azure Blob storage

When you activate the staging feature, first the data is copied from the source data store to the staging storage (bring your own Azure Blob or Azure Data Lake Storage Gen2).

Box 2: Azure Synapse Analytics

The Azure Synapse Analytics connector in copy activity provides built-in data partitioning to copy data in parallel.

Reference:

<https://docs.microsoft.com/en-us/azure/data-factory/copy-activity-performance-features>

<https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-sql-data-warehouse>

### QUESTION 199

Hotspot Question

You have an Azure subscription that contains an Azure Data Lake Storage account. The storage account contains a data lake named DataLake1.

You plan to use an Azure data factory to ingest data from a folder in DataLake1, transform the data, and land the data in another folder.

You need to ensure that the data factory can read and write data from any folder in the DataLake1 file system. The solution must meet the following requirements:

- Minimize the risk of unauthorized user access.
- Use the principle of least privilege.
- Minimize maintenance effort.

How should you configure access to the storage account for the data factory? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point

**Answer Area**

Use  to authenticate by using

Azure Active Directory (Azure AD)	a managed identity
a shared access signature (SAS)	a stored access policy
a shared key	an Authorization header

**Answer:**

**Answer Area**

Use  to authenticate by using

Azure Active Directory (Azure AD)	a managed identity
a shared access signature (SAS)	a stored access policy
a shared key	an Authorization header

**Explanation:**

Box 1: Azure Active Directory (Azure AD)

On Azure, managed identities eliminate the need for developers having to manage credentials by providing an identity for the Azure resource in Azure AD and using it to obtain Azure Active Directory (Azure AD) tokens.

Box 2: a managed identity

A data factory can be associated with a managed identity for Azure resources, which represents this specific data factory. You can directly use this managed identity for Data Lake Storage Gen2 authentication, similar to using your own service principal. It allows this designated factory to access and copy data to or from your Data Lake Storage Gen2.

Note: The Azure Data Lake Storage Gen2 connector supports the following authentication types.

Account key authentication

Service principal authentication

Managed identities for Azure resources authentication

Reference:

<https://docs.microsoft.com/en-us/azure/active-directory/managed-identities-azure-resources/overview>

<https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-data-lake-storage>

**QUESTION 200**

Hotspot Question

You are designing a solution that uses Azure Cosmos DB to store and serve data.

You need to design the Azure Cosmos DB storage to meet the following requirements:

- Provide high availability.
- Provide a recovery point objective (RPO) of less than 15 minutes.
- Provide a recovery time objective (RTO) of less than two minutes.
- Minimize data loss in the event of a disaster.

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

NOTE: Each correct selection is worth one point.

### Answer Area

Write region:   

Multiple
Single

Consistency level:   

Bounded staleness
Session
Strong

Answer:

### Answer Area

Write region:   

Multiple
Single

Consistency level:   

Bounded staleness
Session
Strong

**Explanation:**

Box 1: Multiple

For higher write availability, configure your Azure Cosmos account to have multiple write regions.

Box 2: Bounded staleness

Region(s)	Replication mode	Consistency level	RPO	RTO
1	Single or Multiple write regions	Any Consistency Level	< 240 Minutes	<1 Week
>1	Single write region	Session, Consistent Prefix, Eventual	< 15 minutes	< 15 minutes
>1	Single write region	Bounded Staleness	$K & T$	< 15 minutes
>1	Single write region	Strong	0	< 15 minutes
>1	Multiple write regions	Session, Consistent Prefix, Eventual	< 15 minutes	0
>1	Multiple write regions	Bounded Staleness	$K & T$	0

$K$  = The number of "K" versions (i.e., updates) of an item.

$T$  = The time interval "T" since the last update.

Reference:

<https://docs.microsoft.com/en-us/azure/cosmos-db/high-availability>

<https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels#consistency-levels-and-throughput>

**QUESTION 201**

Hotspot Question

You have an Azure Data Lake Storage Gen2 account named account1 that stores logs as shown in the following table.

Type	Designated retention period
Application	360 days
Infrastructure	60 days

You do not expect that the logs will be accessed during the retention periods.

You need to recommend a solution for account1 that meets the following requirements:

- Automatically deletes the logs at the end of each retention period
- Minimizes storage costs

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

NOTE: Each correct selection is worth one point.

**Answer Area**

To minimize storage costs:

Store the infrastructure logs and the application logs in the Archive access tier.
Store the infrastructure logs and the application logs in the Cool access tier.
Store the infrastructure in the Cool access tier and the application logs in the Archive access tier.

To delete the logs automatically:

Azure Data Factory pipelines
Azure Blob storage lifecycle management rules
Immutable Azure Blob storage time-based retention policies

**Answer:**

**Answer Area**

To minimize storage costs:

Store the infrastructure logs and the application logs in the Archive access tier.
Store the infrastructure logs and the application logs in the Cool access tier.
Store the infrastructure in the Cool access tier and the application logs in the Archive access tier.

To delete the logs automatically:

Azure Data Factory pipelines
Azure Blob storage lifecycle management rules
Immutable Azure Blob storage time-based retention policies

**Explanation:**

Box 1: Store the infrastructure in the Cool access tier and the application logs in the Archive access tier. Cool - Optimized for storing data that is infrequently accessed and stored for at least 30 days. Archive - Optimized for storing data that is rarely accessed and stored for at least 180 days with flexible latency requirements, on the order of hours.

Box 2: Azure Blob storage lifecycle management rules

Blob storage lifecycle management offers a rich, rule-based policy that you can use to transition your data to the best access tier and to expire data at the end of its lifecycle.

Reference:

<https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-storage-tiers>