# Practice Test 1

### **Question 1**

Domain :Monitor and optimize data solutions

A company has an Azure SQL database. The database contains tables that have masked columns. The company wants to identify when a user tries to attempt to access the data from any one of the masked columns. Which of the following would you use for this requirement?

]A.

**Azure Advanced Threat Protection**

]B.

**Auditing**

]C.

**Transparent Data Encryption**

]D.

**Azure Monitor Audit Logs**

### **Question 2**

Domain :Monitor and optimize data solutions

A company has an Azure SQL Data Warehouse defined as part of its Azure subscription. The company wants to ensure that its support department gets an alert when the Data Warehouse consumes the maximum allotted resources to it?

Which of the following would they use as the resource type when configuring the alert in Azure Monitor?

]A.

**Resource Group**

]B.

**SQL Server**

]C.

**SQL Data Warehouse**

]D.

**Subscription**

### **Question 3**

Domain :Monitor and optimize data solutions

A company has an Azure SQL Data Warehouse defined as part of its Azure subscription. The company wants to ensure that its support department gets an alert when the Data Warehouse consumes the maximum allotted resources.

Which of the following would you use as the signal type for the alert?

]A.

**CPU used**

]B.

**DWU limit**

]C.

**DWU used**

]D.

**Data IO Percentage**

### **Question 4**

Domain :Monitor and optimize data solutions

A company has an Azure SQL Database defined as part of its Azure subscription. The Automatic tuning settings are configured, as shown below.

|  |  |  |
| --- | --- | --- |
| **Option** | **Server level** | **Database level** |
| **Force Plan** | Inherited | Inherited |
| **Create Index** | Inherited | Inherited |
| **Drop Index** | Inherited | Inherited |

Would the setting of “Force Plan” be ON for the database?

]A.**Yes**

]B.**No**

### **Question 5**

Domain :Monitor and optimize data solutions

A company has an Azure SQL Database defined as part of its Azure subscription. The Automatic tuning settings are configured, as shown below.

|  |  |  |
| --- | --- | --- |
| **Option** | **Server level** | **Database level** |
| **Force Plan** | Inherited | Inherited |
| **Create Index** | Inherited | Inherited |
| **Drop Index** | Inherited | Inherited |

Would the setting of “Create Index” be ON for the database?

]A.**Yes**

]B.**No**

### **Question 6**

Domain :Monitor and optimize data solutions

A company has an Azure SQL Database defined as part of its Azure subscription. The Automatic tuning settings are configured, as shown below.

|  |  |  |
| --- | --- | --- |
| **Option** | **Server level** | **Database level** |
| **Force Plan** | Inherited | Inherited |
| **Create Index** | Inherited | Inherited |
| **Drop Index** | Inherited | Inherited |

Would the setting of “Drop Index” be ON for the database?

]A.**Yes**

]B.**No**

### **Question 7**

Domain :Implement data storage solutions

A company has an application that is storing its data in an Azure Cosmos DB Account. The database currently has around 100 GB worth of data. Each entry in a collection in the database is shown below.

{

               OrderId: number,

               OrderDescriptionId: number,

               ProductName: string

               OrderValue: number

}

The partition key for the collection is set as OrderId.

Users report that queries take a long time to execute when retrieving data using the Product Name attribute.

You have to resolve the issue.

You decide to create a lookup collection that uses ProductName as the partition key.

Would this resolve the issue?

]A.**Yes**

]B.**No**

### **Question 8**

Domain :Implement data storage solutions

A company has an application that is storing its data in an Azure Cosmos DB Account. The database currently has around 100 GB worth of data. Each entry in a collection in the database is shown below.

{

               OrderId: number,

               OrderDescriptionId: number,

               ProductName: string

               OrderValue: number

}

The partition key for the collection is set as OrderId.

Users report that queries take a long time to execute when retrieving data using the Product Name attribute.

You have to resolve the issue.

You decide to create a lookup collection that uses ProductName as the partition key and OrderId as a value.

Would this resolve the issue?

]A.**Yes**

]B.**No**

### **Question 9**

Domain :Implement data storage solutions

A company has an application that is storing its data in an Azure Cosmos DB Account. The database currently has around 100 GB worth of data. Each entry in a collection in the database is shown below.

{

               OrderId: number,

               OrderDescriptionId: number,

               ProductName: string

               OrderValue: number

}

The partition key for the collection is set as OrderId.

Users report that queries take a long time to execute when retrieving data using the Product Name attribute.

You have to resolve the issue.

You decide to change the partition key to include the ProductName

Would this resolve the issue?

]A.**Yes**

]B.**No**

### **Question 10**

Domain :Manage and develop data processing

You need to create a new Azure Databricks cluster. This cluster would connect to Azure Data Lake Storage Gen2 by using Azure Active Directory (Azure AD) integration.

Which of the following would you use as the Cluster Mode?

]A.

**High Concurrency**

]B.

**Low Concurrency**

]C.

**Premium**

]D.

**Standard**

### **Question 11**

Domain :Manage and develop data processing

You need to create a new Azure Databricks cluster. This cluster would connect to Azure Data Lake Storage Gen2 using Azure Active Directory (Azure AD) integration.

Which of the following advanced option would you enable?

]A.

**Blob access control**

]B.

**Table access control**

]C.

**Credential Passthrough**

]D.

**Single Sign-On**

### **Question 12**

Domain :Manage and develop data processing

You currently have an Azure Storage Account and an Azure SQL Database defined as part of your Azure subscription. You need to move data from Azure Storage Account to the SQL database using Azure Data Factory. You have to ensure that the following requirements are met.

* Ensure that the data remains in the same region as the Azure Storage Account and the Azure SQL Database at all times.
* Minimize administrative effort.

Which of the following would you use as the Integration runtime type?

]A.

**Azure**

]B.

**Self-Hosted**

]C.

**Primary**

]D.

**Azure-SSIS**

### **Question 13**

Domain :Manage and develop data processing

You have to implement Azure Stream Analytics Functions as part of your data streaming solution. The solution has the following requirements.

* Segment the data stream into distinct time segments that do not repeat or overlap.
* Segment the data stream into distinct time segments that repeat and can overlap.
* Segment the data stream to produce an output when an event occurs.

Which of the following windowing function would you use for the following requirement?

***“Segment the data stream into distinct time segments that do not repeat or overlap.”***

]A.

**Hopping**

]B.

**Session**

]C.

**Sliding**

]D.

**Tumbling**

### **Question 14**

Domain :Manage and develop data processing

You have to implement Azure Stream Analytics Functions as part of your data streaming solution. The solution has the following requirements.

* Segment the data stream into distinct time segments that do not repeat or overlap.
* Segment the data stream into distinct time segments that repeat and can overlap.
* Segment the data stream to produce an output when an event occurs.

Which of the following windowing function would you use for the following requirement?

***“Segment the data stream into distinct time segments that repeat and can overlap.”***

]A.

**Hopping**

]B.

**Session**

]C.

**Sliding**

]D.

**Tumbling**

### **Question 15**

Domain :Manage and develop data processing

You have to implement Azure Stream Analytics Functions as part of your data streaming solution. The solution has the following requirements.

* Segment the data stream into distinct time segments that do not repeat or overlap.
* Segment the data stream into distinct time segments that repeat and can overlap.
* Segment the data stream to produce an output when an event occurs.

Which of the following windowing function would you use for the following requirement?

***“Segment the data stream to produce an output when an event occurs.”***

]A.

**Hopping**

]B.

**Session**

]C.

**Sliding**

]D.

**Tumbling**

### **Question 16**

Domain :Manage and develop data processing

You have JSON files stored in an Azure Data Lake Storage Gen2 account. The JSON file contains the FirstName and LastName of customers. You need to use Azure Data bricks to copy the data in the JSON files to an Azure data warehouse. A new column must be created which concatenates the FirstName and LastName values. You have the following components in place in Azure.

* A destination table in the SQL Data Warehouse
* An Azure Blob storage container
* A service principal

Which of the following are actions you would perform to transfer the data onto the Azure SQL Data warehouse table? Choose 5 answers from the options given below.

A.

**Write the results onto Azure Data Lake Storage.**

B.

**Drop the data frame.**

C.

**Perform transformations on the data frame.**

D.

**Mount the Data Lake Storage onto DBFS.**

E.

**Perform transformations on the file.**

F.

**Read the file into a data frame.**

G.

**Specify a temporary folder to stage the data.**

H.

**Write the results to a table in SQL Data Warehouse.**

### **Question 17**

Domain :Manage and develop data processing

You have created an instance of Azure Data Bricks. You have gone ahead and created a cluster and a notebook. The notebook will use R as the primary language. But you also need to be able to switch the notebook to support Scala and SQL. Which of the following can be used to switch between languages in the notebook?

]A.

**%**

]B.

**#**

]C.

**@**

]D.

**&**

### **Question 18**

Domain :Manage and develop data processing

You have an Azure Data Lake Storage Gen 2 account. You have several CSV files loaded into the account. Each file has a header row. After the header row is a property formatted by carriage return (/r) and line feed (/n).

You need to load the files daily as a batch into Azure SQL Data warehouse using Polybase.

You have to skip the header row when the files are imported. Which of the following actions would you take to implement this requirement? Choose 3 answers from the options given below.

A.

**Create an external data source and ensure to use the abfs location.**

B.

**Create an external data source and ensure to use the Hadoop location.**

C.

**Create an external file format and set the First\_row option.**

D.

**Create a database scoped credential that uses OAuth2 token and a key.**

E.

**Use the CREATE EXTERNAL TABLE AS SELECT and create a view that removes the empty row.**

### **Question 19**

Domain :Implement data storage solutions

A company is planning to create an Azure Cosmos DB account. This account will contain a database and a collection. Around 10,000 JSON records will be written to the collection every 24 hours. The company wants to set a consistency level for the database that would meet the following requirements.

* Enable monotonic reads and writes within a session.
* Provide fast throughput.
* Provide the lowest latency.

Which of the following should be set as the consistency level for the database?

]A.

**Strong**

]B.

**Bounded Staleness**

]C.

**Eventual**

]D.

**Session**

]E.

**Consistent Prefix**

### **Question 20**

Domain :Implement data storage solutions

A company has an Azure SQL Datawarehouse. They have a table named comp\_salesfact that contains data for the past 12 months. The data is partitioned by month. The table contains around a billion rows. The table has clustered columnstore indexes. At the beginning of each month, you need to remove the data from the table that is older than 12 months. Which of the following actions would you implement for this requirement? Choose 3 answers from the options given below.

A.

**Create a new empty table named comp\_salesfact\_new that has the same schema as comp\_salesfact.**

B.

**Drop the comp\_salesfact\_new table.**

C.

**Copy the data to the new table by using CREATE TABLE AS SELECT (CTAS).**

D.

**Truncate the partition containing the stale data.**

E.

**Switch the partition containing the stale data from comp\_salesfact to comp\_salesfact\_new.**

F.

**Execute the DELETE statement where the value in the Date column is greater than 12 months.**

### **Question 21**

Domain :Implement data storage solutions

You have an Azure SQL data warehouse. You have used Polybase to create a table named [Ext].[compitems] to query Parquet files stored in Azure Data lake Storage Gen 2. The external table has been defined with 3 columns. You have now discovered that the Parquet files contain a fourth column named ItemID. Which of the following command can you use to add the fourth column to the external table?

]A.

**ALTER EXTERNAL TABLE [Ext].[compitems] ADD [ItemID] int;**

]B.**ALTER TABLE [Ext].[compitems] ADD [ItemID] int;**

]C.**DROP TABLE [Ext].[compitems] CREATE EXTERNAL FILE FORAMT parquetfilenew WITH ( FORMAT\_TYPE=PARQUET, DATA\_COMPRESSION = ‘org.apache.hadoop.io.compress.SnappyCodec’ );**

]D.

**DROP TABLE [Ext].[compitems]**

**CREATE EXTERNAL TABLE [Ext].[compitems]**

**( [ItemID] [Int] NULL,**

**[ItemName] nvarchar(50) NULL,**

**[ItemType] nvarchar(20) NULL,**

**[ItemDescription] nvarchar(250))**

**WITH**

**(**

**LOCATION=’/Items/’,**

**DATA\_SOURCE=AzureDataLakeStore,**

**FILE\_FORMAT=PARQUET,**

**REJECT\_TYPE=VALUE,**

**REJECT\_VALUE=0**

**);**

### **Question 22**

Domain :Implement data storage solutions

You are planning to create a dimension table in an Azure SQL Data Warehouse. The data in the table will be less than 1 GB. You need to ensure that the table meets the following requirements.

* Minimize data movement.
* Provide the fastest query time.

Which of the following would you choose as the table type?

]A.

**Hash distributed**

]B.

**Heap**

]C.

**Replicated**

]D.

**Round-Robin**

### **Question 23**

Domain :Implement data storage solutions

You have an Azure SQL Database named compdb. The database contains a table named compcustomer. The table has a column named customerID that is of the type varchar(22). You have to implement masking for the customerID, which would meet the following requirements.

* The first two prefix characters must be exposed.
* The last four prefix characters must be exposed.
* All other characters must be masked.

You decide to implement data masking and use a credit card function mask.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

### **Question 24**

Domain :Implement data storage solutions

You have an Azure SQL Database named compdb. The database contains a table named compcustomer. The table has a column named customerID that is of the type varchar(22). You have to implement masking for the customerID, which would meet the following requirements.

* The first two prefix characters must be exposed.
* The last four prefix characters must be exposed.
* All other characters must be masked.

You decide to implement data masking and use a random number function mask.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

### **Question 25**

Domain :Implement data storage solutions

You have an Azure SQL Database named compdb. The database contains a table named compcustomer. The table has a column named customerID that is of the type varchar(22). You have to implement masking for the customerID, which would meet the following requirements.

* The first two prefix characters must be exposed.
* The last four prefix characters must be exposed.
* All other characters must be masked.

You decide to implement data masking and use an email function mask.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

### **Question 26**

Domain :Implement data storage solutions

You have an Azure Data Lake Storage Gen 2 account. Your user account has contributor access to the storage account. You have the application ID and access key. You need to use PolyBase to load data into the Azure SQL Data warehouse.

You need to configure PolyBase to connect the data warehouse to the storage account. Which of the following would you need to create for this requirement? Choose 3 answers from the options given below.

A.

**A database encryption key**

B.

**An Asymmetric key**

C.

**An external data source**

D.

**An external file format**

E.

**A database scoped credential**

### **Question 27**

Domain :Monitor and optimize data solutions

A company has an application that allows developers to share and compare code. The conversations for the code snippets, the code snippets themselves and the linked shared are all stored in an Azure SQL database instance. The application also allows to search for historical conversations and code snippets. Matches to previous code snippets also take place in the application. This comparison is made via Transact-SQL functions. If the application finds a match, a link to the match is added to the conversation.

Currently, the following issues are occurring within the application.

* Delays occur during live conversations.
* Delay occurs before the matching link appears after the code snippet is added to the conversation.

Which of the following can be used to resolve the below issue?

**“There are delays which occur during live conversations.”**

]A.

**columnstore index**

]B.

**non-durable table**

]C.

**materialized view**

]D.

**memory-optimized table**

### **Question 28**

Domain :Monitor and optimize data solutions

A company has an application that allows developers to share and compare code. The conversations for the code snippets, the code snippets themselves and the linked shared are all stored in an Azure SQL database instance. The application also allows to search for historical conversations and code snippets. Matches to previous code snippets also take place in the application. This comparison is made via Transact-SQL functions. If the application finds a match, a link to the match is added to the conversation.

Currently, the following issues are occurring within the application.

* Some delays occur during live conversations.
* A delay occurs before the matching link appears after the code snippet is added to the conversation.

Which of the following can be used to resolve the below issue?

**“There is a delay which occurs before the matching link appears after the code snippet is added to the conversation.”**

]A.

**columnstore index**

]B.

**non-durable table**

]C.

**materialized view**

]D.

**memory-optimized table**

### **Question 29**

Domain :Monitor and optimize data solutions

You have an Azure SQL Data Warehouse. You plan to use PolyBase to load data from CSV files located in Azure Data Lake Gen 2 by using an external table. You need to monitor files with invalid schema errors. Which of the following is an error you would monitor for?

]A.

**‘Java exception raised on call to HdfsBridge\_Connect:Error [com.microsoft.polybase.client.KerberosSecureLogin] occurred while accessing external files’**

]B.

**Cannot execute the query “Remote Query” against OLE DB provider “SQLNCLI11”: for linked server “(null)”, Query aborted-the maximum reject threshold (0 rows) was reached while reading from external source**

]C.

**‘Java exception raised on call to HdfsBridge\_Connect:Error[Unable to Instantiate LoginClass] occurred while accessing external files.’**

]D.

**‘Java exception raised on call to HdfsBridge\_Connect:Error[No FileSystem for schema:wabs]**

### **Question 30**

Domain :Monitor and optimize data solutions

The security team in your company currently uses Azure Databricks to analyze data emitted from various sources. You have to send the Apache Spark level events, the Spark structured streaming metrics and application metrics to Azure Monitor.

Which of the following would you implement for this requirement? Choose 3 answers from the options given below.

A.

**In Azure Monitor, go ahead and create a new data source.**

B.

**Configure the current Azure Databricks cluster to use the monitoring library.**

C.

**Deploy an instance of Grafana to a new Azure virtual machine.**

D.

**Build the spark-listeners-loganalytics-1.0-SNAPSHOT.jar file.**

E.

**Create the required Dropwizard counters in the application code.**

### **Question 31**

Domain :Implement data storage solutions

You need to enable Transparent Data Encryption for an Azure SQL database. Which of the following steps would you perform for this requirement? Choose 4 answers from the options given below.

A.

**Create a database encryption key using a certificate.**

B.

**Create a certificate protected by the master key.**

C.

**Set the context to the master database.**

D.

**Create a master key using a password.**

E.

**Set the context to the company database.**

F.

**Enable Encryption.**

### **Question 32**

Domain :Implement data storage solutions

An application is currently making use of a database on the Azure platform. Below is a snippet of the code base.

private static readonly string compendpointUrl = ConfigurationManager.AppSettings[“EndpointUrl”];

private static readonly SecureString compkey=ToSecureString(ConfigurationManager.AppSettings[“AuthorizationKey”]);

var comp\_client= new CosmosClient(new Url(compendpointUrl), compkey);

Database database= await comp\_client.CreateDatabaseAsync(new Database { Id=”compdb” );

Which of the following is the type of database the code is connecting to?

]A.

**Azure Cosmos DB**

]B.

**Azure SQL Database**

]C.

**Azure Storage Account – Blob**

]D.

**Azure SQL Datawarehouse**

### **Question 33**

Domain :Implement data storage solutions

An application is currently making use of a database on the Azure platform. Below is a snippet of the code base.

private static readonly string compendpointUrl = ConfigurationManager.AppSettings[“EndpointUrl”];

private static readonly SecureString compkey=ToSecureString(ConfigurationManager.AppSettings[“AuthorizationKey”]);

var comp\_client= new DocumentClient(new Url(compendpointUrl), compkey);

Database database= await comp\_client.CreateDatabaseAsync(new Database { Id=”compdb” );

Which of the following is the key type used in the code?

]A.

**Resource token**

]B.

**Master Key**

]C.

**Certificate**

]D.

**Password**

### **Question 34**

Domain :Implement data storage solutions

A company is planning to set up an Azure SQL database to store sensitive data. The company wants to monitor data usage and data copied from the system to prevent data leakage. The company also wants to configure the Azure SQL database to email a specific user when the data leakage occurs. Which of the following activities would you need to perform? Choose 3 answers from the options given below.

A.

**In Auditing, enable the auditing feature.**

B.

**Configure the service to create alerts for threat detections of the type “Data Exfiltration”.**

C.

**In the Firewalls and virtual networks section, enable “Allow access to Azure services”.**

D.

**Enable Advanced threat protection.**

E.

**Configure the service to send email alerts to the IT security administrator.**

### **Question 35**

Domain :Monitor and optimize data solutions

Your company currently has an enterprise data warehouse in Azure Synapse Analytics.

You have to monitor the solution to see whether the data warehouse needs to be scaled up based on the current workloads.

Which of the following metric would you monitor for this requirement?

]A.

**CPU Percentage**

]B.

**DWU used**

]C.

**DWU percentage**

]D.

**Data IO percentage**

### **Question 36**

Domain :Manage and develop data processing

A company wants to implement a lambda architecture on Microsoft Azure. The following are the key requirements for each architecture layer.

**Data storage**

* The data store should serve as a repository for high volumes of files.
* The files can be large and of different formats.
* It should be optimized for big data analytics workloads.
* The data should be organized using a hierarchical structure.

**Batch processing**

* This layer should provide a managed solution for in-memory computation processing.
* It should provide support for a variety of programming languages.
* It should provide the ability to resize and terminate the cluster automatically.

**Analytical data store**

* This layer must provide support for SQL language.
* It must implement native columnar storage.
* It should support parallel processing.

Which of the following should be used as a technology for the “Data Storage” layer?

]A.

**Azure SQL Database**

]B.

**Azure Blob storage**

]C.

**Azure Cosmos DB**

]D.

**Azure Data Lake Storage**

### **Question 37**

Domain :Manage and develop data processing

A company wants to implement a lambda architecture on Microsoft Azure. The following are the key requirements for each architecture layer.

**Data storage**

* The data store should serve as a repository for high volumes of files.
* The files can be large and of different formats.
* It should be optimized for big data analytics workloads.
* The data should be organized using a hierarchical structure.

**Batch processing**

* This layer should provide a managed solution for in-memory computation processing.
* It should provide support for a variety of programming languages.
* It should provide the ability to resize and terminate the cluster automatically.

**Analytical data store**

* This layer must provide support for SQL language.
* It must implement native columnar storage.

It should support parallel processing.

Which of the following should be used as a technology for the “Batch processing” layer?

]A.

**HDInsight Spark**

]B.

**HDInsight Hadoop**

]C.

**Azure Databricks**

]D.

**HDInsight Interactive Query**

### **Question 38**

Domain :Manage and develop data processing

A company wants to implement a lambda architecture on Microsoft Azure. The following are the key requirements for each architecture layer.

**Data storage**

* The data store should serve as a repository for high volumes of files.
* The files can be large and of different formats.
* It should be optimized for big data analytics workloads.
* The data should be organized using a hierarchical structure.

**Batch processing**

* This layer should provide a managed solution for in-memory computation processing.
* It should provide support for a variety of programming languages.
* It should provide the ability to resize and terminate the cluster automatically.

**Analytical datastore**

* This layer must provide support for SQL language.
* It must implement native columnar storage.

It should support parallel processing.

Which of the following should be used as a technology for the “Analytical data store” layer?

]A.

**HDInsight Base**

]B.

**Azure SQL Data warehouse**

]C.

**Azure Analysis services**

]D.

**Azure Cosmos DB**

### **Question 39**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

Which of the following should be used as the API for the Cosmos DB account?

]A.

**Cassandra**

]B.

**Gremlin**

]C.

**MongoDB**

]D.

**SQL**

]E.

**Table**

### **Question 40**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

Which of the following would you use for the consistency level for the database?

]A.

**Eventual**

]B.

**Session**

]C.

**Strong**

]D.

**Consistent Prefix**

### **Question 41**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

You need to build the Azure SQL Data warehouse data store. Which of the following would you use as the underlying table type?

]A.

**Hash distributed**

]B.

**Replicated**

]C.

**Round-Robin**

]D.

**Primary**

### **Question 42**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

You need to build the Azure SQL Data warehouse data store. Which of the following would you use as the underlying index type?

]A.

**Clustered**

]B.

**Clustered column store**

]C.

**Heap**

]D.

**Nonclustered**

### **Question 43**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

Which of the following masking functions should be used for the “carID” column?

]A.

**Credit Card**

]B.

**Default**

]C.

**Email**

]D.

**Random number**

### **Question 44**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

Which of the following masking functions should be used for the “carWeight” column?

]A.

**Credit Card**

]B.

**Default**

]C.

**Email**

]D.

**Random number**

### **Question 45**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

Which of the following should be included in the Data Factory Pipeline?

]A.

**A copy activity that needs to use a stored procedure as the source**

]B.

**A copy activity that needs to use schema mappings**

]C.

**A delete activity that needs to have logging enabled**

]D.

**A filter activity that needs to have a condition**

### **Question 46**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

The team is monitoring the Data Factory pipeline. They can see that the Cosmos DB to SQL database run time is taking 45 minutes. Which of the following can be carried out to improve the performance of the job?

]A.

**Increase in the number of data integration units.**

]B.

**Ensure that the copy activity uses a staged copy.**

]C.

**Ensure that the copy activity performs compression.**

]D.

**Ensure to decrease the amount pf parallelism in the copy activities.**

### **Question 47**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

Which of the following can be used to satisfy the case study requirement?

**“The query performance for data in the Azure SQL database must be stable without the need for administrative overhead.”**

]A.

**sp\_update stored procedure**

]B.

**Using the Query store**

]C.

**Using the dbcc checkdb command**

]D.

**Using automatic tuning**

### **Question 48**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They also provide a yearly gaming competition for their students. The competition is held every month in different locations.

**Current Environment**

The company currently has the following environment in place

* The racing cars for the competition send their telemetry data to a MongoDB database. The telemetry data has around 100 attributes.
* A custom application is then used to transfer the data from the MongoDB database to a SQL Server 2017 database. The attribute names are changed when they are sent to the SQL Server database.
* Another application named “Comp workflow” is then used to perform analytics on the telemetry data to look for improvements on the racing cars.
* The SQL Server 2017 database has a table named “cardata” which has around 1 TB of data. “Comp workflow” performs the required analytics on the data in this table. Large aggregations are performed on a column of the table.

**Proposed Environment**

The company now wants to move the environment to Azure. Below are the key requirements

* The racing car data will now be moved to Azure Cosmos DB and Azure SQL database. The data must be written to the closest Azure data center and must converge in the least amount of time.
* The query performance for data in the Azure SQL database must be stable without the need of administrative overhead
* The data for analytics will be moved to an Azure SQL Data warehouse
* Transparent data encryption must be enabled for all data stores wherever possible
* An Azure Data Factory pipeline will be used to move data from the Cosmos DB database to the Azure SQL database. If there is a delay of more than 15 minutes for the data transfer, then configuration changes need to be made to the pipeline workflow.
* The telemetry data must be monitored for any sort of performance issues.
* The Request Units for Cosmos DB must be adjusted to maintain the demand while also minimizing costs.
* The data in the Azure SQL Server database must be protected via the following requirements
  + Only the last four digits of the values in the column carID must be shown
  + A zero value must be shown for all values in the column carWeight

You need to monitor the telemetry data being sent to Cosmos DB so that you can decide on the amount of Request Units to provision for Cosmos DB. Which of the following metrics must you monitor? Choose 2 answers from the options given below.

A.

**The number of requests**

B.

**The Session consistency**

C.

**The data and index storage consumed**

D.

**The Average Throughput/second**

### **Question 49**

Domain :Manage and develop data processing

You have the following query defined in Azure Stream Analytics

WITH

step1 AS (SELECT \*

FROM compinput1

PARTITION BY OrderId

INTO 10),

step1 AS (SELECT \*

FROM compinput2

PARTITION BY OrderId

INTO 10)

SELECT \*

INTO compoutput

FROM step1

PARTITION BY OrderId

UNION step2

BY OrderId

Would the above query join two streams of partitioned data?

]A.**Yes**

]B.**No**

### **Question 51**

Domain :Manage and develop data processing

You have the following query defined in Azure Stream Analytics

WITH

step1 AS (SELECT \*

FROM compinput1

PARTITION BY OrderId

INTO 10),

step1 AS (SELECT \*

FROM compinput2

PARTITION BY OrderId

INTO 10)

SELECT \*

INTO compoutput

FROM step1

PARTITION BY OrderId

UNION step2

BY OrderId

Would the supply of 60 streaming units optimize the performance of the query?

]A.**Yes**

]B.**No**

### **Question 52**

Domain :Manage and develop data processing

A company wants to use the Azure Databricks service. There is a need to create clusters based on the following configuration.

* ClusterA – Here the cluster needs to be configured to terminate automatically after 120 minutes.
* ClusterB – Here an environment needs to be created for each notebook.
* ClusterC – Here a group of data engineers will be sharing the same cluster.

Which of the following cluster type would you set for ClusterA?

]A.

**Standard**

]B.

**Basic**

]C.

**Job**

]D.

**High concurrency**

### **Question 53**

Domain :Manage and develop data processing

A company wants to use the Azure Databricks service. There is a need to create clusters based on the following configuration.

* ClusterA – Here the cluster needs to be configured to terminate automatically after 120 minutes.
* ClusterB – Here an environment needs to be created for each notebook.
* ClusterC – Here a group of data engineers will be sharing the same cluster.

Which of the following cluster type would you set for ClusterB?

]A.

**Standard**

]B.

**Basic**

]C.

**Job**

]D.

**High concurrency**

### **Question 54**

Domain :Manage and develop data processing

A company wants to use the Azure Databricks service. There is a need to create clusters based on the following configuration.

* ClusterA – Here the cluster needs to be configured to terminate automatically after 120 minutes.
* ClusterB – Here an environment needs to be created for each notebook.
* ClusterC – Here a group of data engineers will be sharing the same cluster.

Which of the following cluster type would you set for ClusterC?

]A.

**Standard**

]B.

**Basic**

]C.

**Job**

]D.

**High concurrency**

### **Question 55**

Domain :Monitor and optimize data solutions

 A company has an Azure SQL Database. They want to enable diagnostics logging for the database. Which of the following can be used to store the diagnostic logs for the database? Choose 2 answers from the options given below.

A.

**Azure Event Hubs**

B.

**Azure Storage**

C.

**Azure Cosmos DB**

D.

**Azure SQL Data warehouse**