VCE & PDF
Pass4itSure.com

DP-201^{Q&As}

Designing an Azure Data Solution

Pass Microsoft DP-201 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass4itsure.com/dp-201.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

A company is designing a solution that uses Azure Databricks.

The solution must be resilient to regional Azure datacenter outages.

You need to recommend the redundancy type for the solution. What should you recommend?

A. Read-access geo-redundant storage

B. Locally-redundant storage

C. Geo-redundant storage

D. Zone-redundant storage

Correct Answer: C

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.

References: https://medium.com/microsoftazure/data-durability-fault-tolerance-resilience-in-azure-databricks-95392982bac7

QUESTION 2

You need to design a real-time stream solution that uses Azure Functions to process data uploaded to Azure Blob Storage. The solution must meet the following requirements:

1.

 Support up to 1 million blobs.

2.

 Scaling must occur automatically.

3.

 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.

Correct Answer: C

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

QUESTION 3

A company is evaluating data storage solutions.

You need to recommend a data storage solution that meets the following requirements:

1.

 Minimize costs for storing blob objects.

2.

 Optimize access for data that is infrequently accessed.

3.

 Data must be stored for at least 30 days.

4.

 Data availability must be at least 99 percent. What should you recommend?

A. Premium

B. Cold

C. Hot

D. Archive

Correct Answer: B

Azure\'s cool storage tier, also known as Azure cool Blob storage, is for infrequently-accessed data that needs to be stored for a minimum of 30 days. Typical use cases include backing up data before tiering to archival systems, legal data, media files, system audit information, datasets used for big data analysis and more.

The storage cost for this Azure cold storage tier is lower than that of hot storage tier. Since it is expected that the data stored in this tier will be accessed less frequently, the data access charges are high when compared to hot tier. There are no additional changes required in your applications as these tiers can be accessed using APIs in the same manner that you access Azure storage.

References: https://cloud.netapp.com/blog/low-cost-storage-options-on-azure

QUESTION 4

HOTSPOT

```
{
    "num_workers": null,
    "autoscale": {
        "min_workers": 2,
        "max_workers": 8
    },
    "cluster_name": "MyCluster",
    "spark_version": "latest-stable-scala2.11",
    "spark_conf": {
        "spark.databricks.cluster.profile": "serverless",
        "spark.databricks.repl.allowedLanguages": "sql,python,r"
    },
    "node_type_id": "Standard_DS13_v2",
    "ssh_public_keys": [],
    "custom_tags": {
        "ResourceClass" "Serverless"
    },
    "spark_env_vars": {
        "PYSPARK_PYTHON": "/databricks/python3/bin/python3"
    },
    "autotermination_minutes": 90,
    "enable_elastic_disk": true,
    "init_scripts": []
}
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Statements	Yes	No
The Databricks cluster supports multiple concurrent users.	○	○
The Databricks cluster minimizes costs when running scheduled jobs that execute notebooks.	○	○
The Databricks cluster supports the creation of a Delta Lake table.	○	○

Correct Answer:

Answer Area

Statements	Yes	No
The Databricks cluster supports multiple concurrent users.	◉	○
The Databricks cluster minimizes costs when running scheduled jobs that execute notebooks.	○	◉
The Databricks cluster supports the creation of a Delta Lake table.	◉	○

Box 1: Yes Box 2: No autotermination_minutes: Automatically terminates the cluster after it is inactive for this time in minutes. If not set, this cluster will not be automatically terminated. If specified, the threshold must be between 10 and 10000 minutes. You can also set this value to 0 to explicitly disable automatic termination. Box 3: Yes References: https://docs.databricks.com/dev-tools/api/latest/clusters.html

QUESTION 5

HOTSPOT

You need to design the storage for the Health Insights data platform.

Which types of tables should you include in the design? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

High performance of joins
on large fact tables:

| Hash distributed tables |
| Merge Joins |
| Replicated tables |
| Round-robin distributed tables |

Joins to the date
dimension table:

| Hash distributed tables |
| Merge Joins |
| Replicated tables |
| Round-robin distributed tables |

Correct Answer:

Answer Area

High performance of joins on large fact tables:

Hash distributed tables
Merge Joins
Replicated tables
Round-robin distributed tables

Joins to the date dimension table:

Hash distributed tables
Merge Joins
Replicated tables
Round-robin distributed tables

Box 1: Hash-distributed tables

The new Health Insights application must be built on a massively parallel processing (MPP) architecture that will support the high performance of joins on large fact tables.

Hash-distributed tables improve query performance on large fact tables.

Box 2: Round-robin distributed tables

A round-robin distributed table distributes table rows evenly across all distributions. The assignment of rows to distributions is random.

Scenario:

ADatum identifies the following requirements for the Health Insights application:

The new Health Insights application must be built on a massively parallel processing (MPP) architecture that will support the high performance of joins on large fact tables.

References:

https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-distribute

DP-201 Study Guide   DP-201 Exam Questions   DP-201 Braindumps

To Read the Whole Q&As, please purchase the Complete Version from Our website.

Try our product !

100% Guaranteed Success
100% Money Back Guarantee
365 Days Free Update
Instant Download After Purchase
24x7 Customer Support
Average 99.9% Success Rate
More than 800,000 Satisfied Customers Worldwide
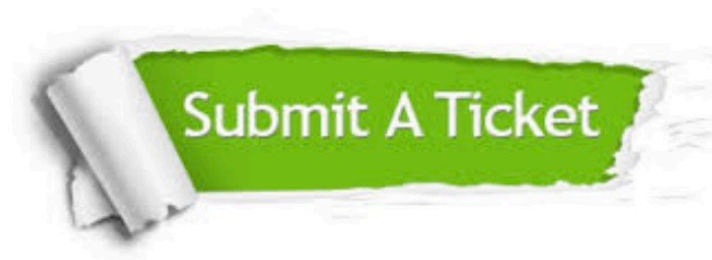Multi-Platform capabilities - Windows, Mac, Android, iPhone, iPod, iPad, Kindle

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications.
You can view Vendor list of All Certification Exams offered:

https://www.pass4itsure.com/allproducts

Need Help

Please provide as much detail as possible so we can best assist you.
To update a previously submitted ticket: