



70-475^{Q&As}

Designing and Implementing Big Data Analytics Solutions

Pass Microsoft 70-475 Exam with 100% Guarantee

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

<https://www.passapply.com/70-475.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

Overview:

Litware, Inc. is a company that manufactures personal devices to track physical activity and other health-related data.

Litware has a health tracking application that sends health-related data from a user's personal device to Microsoft Azure.

Litware has three development and commercial offices. The offices are located in the United States, Luxembourg, and India.

Litware products are sold worldwide. Litware has commercial representatives in more than 80 countries.

Existing Environment:

In addition to using desktop computers in all of the offices, Litware recently started using Microsoft Azure resources and services for both development and operations.

Litware has an Azure Machine Learning solution.

Litware recently extended its platform to provide third-party companies with the ability to upload data from devices to Azure. The data can be aggregated across multiple devices to provide users with a comprehensive view of their global health activity.

While the upload from each device is small, potentially more than 100 million devices will upload data daily by using an Azure event hub. Each health activity has a small amount of data, such as activity type, start date/time, and end date/time.

Each activity is limited to a total of 3 KB and includes a customer identification key.

In addition to the Litware health tracking application, the users' activities can be reported to Azure by using an open API.

The developers at Litware perform Machine Learning experiments to recommend an appropriate health activity based on the past three activities of a user.

The Litware developers train a model to recommend the best activity for a user based on the hour of the day.

Requirements:

Litware plans to extend the existing dashboard features so that health activities can be compared between the users based on age, gender, and geographic region.

Minimize the costs associated with transferring data from the event hub to Azure Storage.

Litware identifies the following technical requirements:

Data from the devices must be stored for three years in a format that enables the fast processing of date fields and filtering.

The third-party companies must be able to use the Litware Machine Learning models to generate recommendations to their users by using a third-party application.



Any changes to the health tracking application must ensure that the Litware developers can run the experiments without interrupting or degrading the performance of the production environment.

Activity tracking data must be available to all of the Litware developers for experimentation. The developers must be prevented from accessing the private information of the users.

When the Litware health tracking application asks users how they feel, their responses must be reported to Azure.

You need to recommend a data handling solution to support the planned changes to the dashboard.

What is the best recommendation to achieve the goal? More than one answer choice may achieve the goal. Select the BEST answer.

- A. anonymization
- B. encryption
- C. obfuscation
- D. compression

Correct Answer: C

From scenario: Litware plans to extend the existing dashboard features so that health activities can be compared between the users based on age, gender, and geographic region. The developers must be prevented from accessing the private information of the users.

Dynamic Data Masking can be used to hide or obfuscate sensitive data, by controlling how the data appears in the output of database queries. Dynamic Data Masking rules can be defined on particular columns, indicating how the data in those columns will appear when queried. There are no physical changes to the data in the database itself; the data remains intact and is fully available to authorized users or applications. Database operations remain unaffected, and the masked data has the same data type as the original data, so DDM can often be applied without making any changes to database procedures or application code.

Reference: <https://blogs.technet.microsoft.com/dataplatforminsider/2016/01/25/use-dynamic-data-masking-to-obfuscate-your-sensitive-data/>

QUESTION 2

You plan to design a solution to gather data from 5,000 sensors that are deployed to multiple machines. The sensors generate events that contain data on the health status of the machines.

You need to create a new Microsoft Azure event hub to collect the event data.

Which command should you run? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Answer Area

| | | | |
|---|---|--|---|
| ▼ | ▼ | ▼ | ▼ |
| Add-Type -Path "C:\temp\Microsoft.ServiceBus.dll" | Get-AzureSBNamespace -Name \$Namespace -NamespaceType Messaging | Get-AzureSBNamespace -Name \$Namespace | Add-Type -TypeName Microsoft.ServiceBus.Messaging.EventHubDescription |
| New-Object -Path "C:\temp\Microsoft.ServiceBus.dll" | New-AzureSBNamespace -Name \$Namespace -NamespaceType Messaging | Set-AzureSBNamespace -Name \$Namespace | New-Object -TypeName Microsoft.ServiceBus.Messaging.EventHubDescription |

Correct Answer:

Answer Area

| | | | |
|---|---|--|---|
| ▼ | ▼ | ▼ | ▼ |
| Add-Type -Path "C:\temp\Microsoft.ServiceBus.dll" | Get-AzureSBNamespace -Name \$Namespace -NamespaceType Messaging | Get-AzureSBNamespace -Name \$Namespace | Add-Type -TypeName Microsoft.ServiceBus.Messaging.EventHubDescription |
| New-Object -Path "C:\temp\Microsoft.ServiceBus.dll" | New-AzureSBNamespace -Name \$Namespace -NamespaceType Messaging | Set-AzureSBNamespace -Name \$Namespace | New-Object -TypeName Microsoft.ServiceBus.Messaging.EventHubDescription |

Set-AzureSBNameSpace

Box 1: Add-Type -Path "C:\temp\Microsoft.ServiceBus.dll"

Make sure to reference the latest version of the \Microsoft.ServiceBus.dll by using the Add-Type -Path command.

Box 2: New-AzureSBNamespace -Name \$Namespace -NameSpaceType Messaging

To create a new Azure Service Bus namespace you can use the New-AzureSBNamespace PowerShell cmdlet

Box 3: Get-AzureSBNamespace -Name \$Namespace -NameSpaceType Messaging

Example:

Write-Output "Creating the [\$Namespace] namespace in the [\$Location] region..."

New-AzureSBNamespace -Name \$Namespace -Location \$Location -CreateACSNamespace \$CreateACSNamespace -NamespaceType Messaging

\$CurrentNamespace = Get-AzureSBNamespace -Name \$Namespace

Write-Host "The [\$Namespace] namespace in the [\$Location] region has been successfully created."

Box 4: New-Object -TypeName Microsoft.ServiceBus.Messaging.EventHubDescription

References:

<https://blogs.msdn.microsoft.com/paolos/2014/12/01/how-to-create-a-service-bus-namespace-and-an-event-hub-using-a-powershell-script/>

QUESTION 3

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while

the others might not have a correct solution.



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

You plan to implement a new data warehouse.

You have the following information regarding the data warehouse:

The first data files for the data warehouse will be available in a few days.

Most queries that will be executed against the data warehouse are ad-hoc.

The schemas of data files that will be loaded to the data warehouse change often.

One month after the planned implementation, the data warehouse will contain 15 TB of data.

You need to recommend a database solution to support the planned implementation.

Solution: You recommend an Apache Hadoop system.

Does this meet the goal?

A. Yes

B. No

Correct Answer: A

QUESTION 4

You have an analytics solution in Microsoft Azure that must be operationalized.

You have the relevant data in Azure Blob storage. You use an Azure HDInsight Cluster to process the data.

You plan to process the raw data files by using Azure HDInsight. Azure Data Factory will operationalize the solution.

You need to create a data factory to orchestrate the data movement. Output data must be written back to Azure Blob storage.

Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:



Actions

Create input and output datasets for the files in Azure.

Rename the data factory hub.

Create a data factory.

Create a client to connect to Azure Blob storage.

Create linked services for Azure Blob storage and the Azure HDInsight cluster.

Create an Azure HDInsight activity in a pipeline to process the data.

Answer Area

1

2

3

4

Correct Answer:

Actions

Rename the data factory hub.

Create a client to connect to Azure Blob storage.

Answer Area

1

2

3

4

QUESTION 5

Overview:

Relecloud is a social media company that processes hundreds of millions of social media posts per day and sells advertisements to several hundred companies. Relecloud has a Microsoft SQL Server database named DB1 that stores

information about the advertisers. DB1 is hosted on a Microsoft Azure virtual machine.



Relecloud has two main offices. The offices are located in San Francisco and New York City.

The offices connect to each other by using a site-to-site VPN. Each office connects directly to the Internet.

Relecloud modifies the pricing of its advertisements based on trending topics. Topics are considered to be trending if they generate many mentions in a specific country during a 15-minute time frame. The highest trending topics generate the

highest advertising revenue.

Relecloud wants to deliver reports to the advertisers by using Microsoft Power BI. The reports will provide real-time data on trending topics, current advertising rates, and advertising costs for a given month. Relecloud will analyze the trending

topics data, and then store the data in a new data warehouse for ad-hoc analysis. The data warehouse is expected to grow at a rate of 1 GB per hour or 8.7 terabytes (TB) per year. The data will be retained for five years for the purpose of long-term trending.

Requirements:

Management at Relecloud must be able to view which topics are trending to adjust advertising rates in near real-time.

Relecloud plans to implement a new streaming analytics platform that will report on trending topics.

Relecloud plans to implement a data warehouse named DB2.

Relecloud identifies the following technical requirements:

Social media data must be analyzed to identify trending topics in real-time.

The use of Infrastructure as a Service (IaaS) platforms must be minimized, whenever possible.

The real-time solution used to analyze the social media data must support scaling up and down without service interruption.

Relecloud identifies the following technical requirements for the advertisers:

The advertisers must be able to see only their own data in the Power BI reports.

The advertisers must authenticate to Power BI by using Azure Active Directory (Azure AD) credentials.

The advertisers must be able to leverage existing Transact-SQL language knowledge when developing the real-time streaming solution.

Members of the internal advertising sales team at Relecloud must be able to see only the sales date of the advertisers to which they are assigned.

The internal Relecloud advertising sales team must be prevented from inserting, updating, and deleting rows for the advertisers to which they are not assigned.

The internal Relecloud advertising sales team must be able to use a text file to update the list of advertisers, and then to upload the file to Azure Blob storage.

Relecloud identifies the following requirements for DB1:

Data generated by the streaming analytics platform must be stored in DB1.



The user names of the advertisers must be mapped to CustomerID in a table named Table2.

The advertisers in DB1 must be stored in a table named Table1 and must be refreshed nightly.

The user names of the employees at Relecloud must be mapped to EmployeeID in a table named Table3.

Relecloud identifies the following requirements for DB2:

DB2 must have minimal storage costs.

DB2 must run load processes in parallel.

DB2 must support massive parallel processing.

DB2 must be able to store more than 40 TB of data.

DB2 must support scaling up and down, as required.

Data from DB1 must be archived in DB2 for long-term storage.

All of the reports that are executed from DB2 must use aggregation.

Users must be able to pause DB2 when the data warehouse is not in use.

Users must be able to view previous versions of the data in DB2 by using aggregates.

Relecloud identifies the following requirements for extract, transformation, and load (ETL):

Data movement between DB1 and DB2 must occur each hour.

An email alert must be generated when a failure of any type occurs during ETL processing.

Sample code and data:

You execute the following code for a table named rls_table1.

```
create function rls_table1 (@CustomerId int, @SalesPersonId int)
    returns table
    with schemabinding
as
return
select 1 as result
from dbo.table1
join dbo.table2 on table1.customerid = Table2.CustomerId
where table2.UserName = suser_sname()
    and table1.customerid = @CustomerId
union all
select 1 as result
from dbo.table1
join dbo.table3 on table1.salespersonid = table3.EmployeeId
where table3.UserName = suser_sname()
    and table1.salespersonid = @SalesPersonId
go
```




You use the following code to create Table1.

create table table1 (customerid int, salespersonid int ...) Go

The following is a sample of the streaming data.

| User | Country | Topic | Time |
|-------|---------|--------|------------------------------|
| user1 | USA | Topic1 | 2017-01-01T00:00:01.0000000Z |
| user1 | USA | Topic3 | 2017-01-01T00:02:01.0000000Z |
| user2 | Canada | Topic2 | 2017-01-01T00:01:11.0000000Z |
| user3 | India | Topic1 | 2017-01-01T00:03:14.0000000Z |

You need to implement a solution that meets the data refresh requirement for DB1.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:

Actions

- In DB1, create external objects.
- From the Azure portal, export the storage account key.
- In DB1, create a stored procedure that imports data from an external table to Table1.
- From the Azure portal, create and schedule an Azure Automation job that executes the stored procedure.
- In DB1, create a staging table.

Answer Area

>
<

^
v

Correct Answer:



| Actions | Answer Area |
|----------------------------------|--|
| In DB1, create external objects. | In DB1, create a stored procedure that imports data from an external table to Table1. |
| | From the Azure portal, export the storage account key. |
| | From the Azure portal, create and schedule an Azure Automation job that executes the stored procedure. |
| | |
| In DB1, create a staging table. | |

Azure Data Factory can be used to orchestrate the execution of stored procedures. This allows more complex pipelines to be created and extends Azure Data Factory's ability to leverage the computational power of SQL Data Warehouse.

From scenario:

Relecloud has a Microsoft SQL Server database named DB1 that stores information about the advertisers. DB1 is hosted on a Microsoft Azure virtual machine.

Relecloud identifies the following requirements for DB1:

Data generated by the streaming analytics platform must be stored in DB1.

The advertisers in DB1 must be stored in a table named Table1 and must be refreshed nightly.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/machine-learning-data-science-move-sql-server-virtual-machine>

[Latest 70-475 Dumps](#)

[70-475 VCE Dumps](#)

[70-475 Exam Questions](#)



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

Try our product !

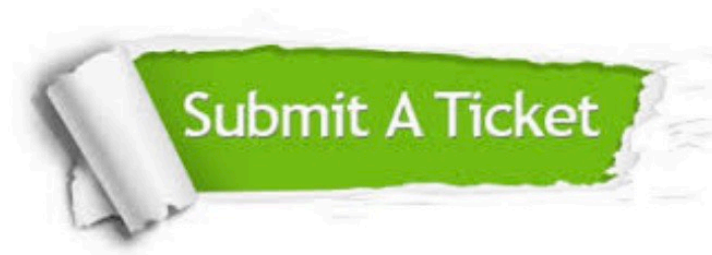
100% Guaranteed Success
100% Money Back Guarantee
365 Days Free Update
Instant Download After Purchase
24x7 Customer Support
Average 99.9% Success Rate
More than 800,000 Satisfied Customers Worldwide
Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [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.passapply.com/allproducts>

Need Help

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



| | | |
|---|---|--|
|  <p>One Year Free Update Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p> |  <p>Money Back Guarantee To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p> |  <p>Security & Privacy We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.</p> |
|---|---|--|

Any charges made through this site will appear as Global Simulators Limited.
All trademarks are the property of their respective owners.
Copyright © passapply, All Rights Reserved.