



# 70-475<sup>Q&As</sup>

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.pass4itsure.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

References: <https://docs.microsoft.com/en-us/azure/batch/batch-technical-overview>

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 implement DB2.

You need to configure the tables in DB2 to host the data from DB1. The solution must meet the requirements for DB2.

Which type of table and history table storage should you use for the tables? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



## Answer area

Table:

Change Data Capture
Change tracking
Temporal table

History table storage:

Clustered columnstore
In-Memory OLTP
Row store

Correct Answer:

## Answer area

Table:

Change Data Capture
Change tracking
Temporal table

History table storage:

Clustered columnstore
In-Memory OLTP
Row store

From Scenario: Relecloud plans to implement a data warehouse named DB2.

Box 1: Temporal table

From Scenario:

Relecloud identifies the following requirements for DB2:

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

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

A system-versioned temporal table is a new type of user table in SQL Server 2017, designed to keep a full history of data changes and allow easy point in time analysis. A temporal table also contains a reference to another table with a

mirrored schema. The system uses this table to automatically store the previous version of the row each time a row in the temporal table gets updated or deleted. This additional table is referred to as the history table, while the main table that



stores current (actual) row versions is referred to as the current table or simply as the temporal table.

Reference: <https://docs.microsoft.com/en-us/sql/relational-databases/tables/temporal-tables>

Box 2: Clustered Columnstore

From Scenario:

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.

Relecloud identifies the following requirements for DB2:

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

How SQL Database deletes aged rows?

The cleanup process depends on the index layout of the history table. It is important to notice that only history tables with a clustered index (B-tree or columnstore) can have finite retention policy configured.

Reference: <https://docs.microsoft.com/en-us/sql/relational-databases/tables/manage-retention-of-historical-data-in-system-versioned-temporal-tables>

---

## QUESTION 2

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 deploy a Microsoft Azure SQL data warehouse and a web application.

The data warehouse will ingest 5 TB of data from an on-premises Microsoft SQL Server database daily. The web application will query the data warehouse.

You need to design a solution to ingest data into the data warehouse.

Solution: You use SQL Server Integration Services (SSIS) to transfer data from SQL Server to Azure SQL Data Warehouse.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

Integration Services (SSIS) is a powerful and flexible Extract Transform and Load (ETL) tool that supports complex workflows, data transformation, and several data loading options. The main drawback is speed. We should use Polybase instead.



References: <https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services>

---

### QUESTION 3

A company named Fabrikam, Inc. plans to monitor financial markets and social networks, and then to correlate global stock movements to social network activity.

You need to recommend a Microsoft Azure HDInsight cluster solution that meets the following requirements: Provides continuous availability Can process asynchronous feeds

What is the best type of cluster to recommend to achieve the goal? More than one answer choice may achieve the goal. Select the BEST answer.

- A. Apache Hbase
- B. Apache Hadoop
- C. Apache Spark
- D. Apache Storm

Correct Answer: C

---

### QUESTION 4

You are designing a partitioning scheme for ingesting real-time data by using Kafka. Kafka and Apache Storm will be integrated.

You plan to use four event processing servers that each run as a Kafka consumer. Each server will have a two quad-core processor.

You need to identify the minimum number of partitions required to ensure that the load is distributed evenly.

How many should you identify?

- A. 1
- B. 4
- C. 16
- D. 32

Correct Answer: C

---

### QUESTION 5

You plan to deploy a Microsoft Azure Data Factory pipeline to run an end-to-end data processing workflow.





You need to recommend which Azure Data Factory features must be used to meet the following requirements:

Track the run status of the historical activity.

Enable alerts and notifications on events and metrics.

Monitor the creation, updating, and deletion of Azure resources.

Which features should you recommend? To answer, drag the appropriate features to the correct requirements. Each feature 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.

Select and Place:

**Features**

- Azure Data Factory alerts
- Azure Data Factory events
- Azure Data Factory metrics
- Azure HDInsight logs

**Answer Area**

Track the run status of the historical activity: Feature

Enable alerts and notifications on events and metrics: Feature

Monitor the creation, updating, and deletion of Azure resources: Feature

Correct Answer:

**Features**

- Azure HDInsight logs
- Azure Data Factory alerts
- Azure Data Factory metrics
- Azure Data Factory events

**Answer Area**

Track the run status of the historical activity: Azure HDInsight logs

Enable alerts and notifications on events and metrics: Azure Data Factory alerts

Monitor the creation, updating, and deletion of Azure resources: Azure Data Factory events

Box 1: Azure Hdinsight logs Box 2: Azure Data Factory alerts Box 3: Azure Data Factory events

Reference: <https://docs.microsoft.com/en-us/azure/log-analytics/log-analytics-azure-storage>

[Latest 70-475 Dumps](#)

[70-475 VCE Dumps](#)

[70-475 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:



 <b>One Year Free Update</b> <p>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>	 <b>Money Back Guarantee</b> <p>To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <b>Security &amp; Privacy</b> <p>We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; 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 © pass4itsure, All Rights Reserved.