



Free Questions for [ARA-C01](#) by [vceexamstest](#)

Shared by [Preston](#) on [15-04-2024](#)

For More Free Questions and Preparation Resources

[Check the Links on Last Page](#)

Question 1

Question Type: MultipleChoice

A company has built a data pipeline using Snowpipe to ingest files from an Amazon S3 bucket. Snowpipe is configured to load data into staging database tables. Then a task runs to load the data from the staging database tables into the reporting database tables.

The company is satisfied with the availability of the data in the reporting database tables, but the reporting tables are not pruning effectively. Currently, a size 4X-Large virtual warehouse is being used to query all of the tables in the reporting database.

What step can be taken to improve the pruning of the reporting tables?

Options:

- A-** Eliminate the use of Snowpipe and load the files into internal stages using PUT commands.
- B-** Increase the size of the virtual warehouse to a size 5X-Large.
- C-** Use an ORDER BY <cluster_key (s) > command to load the reporting tables.
- D-** Create larger files for Snowpipe to ingest and ensure the staging frequency does not exceed 1 minute.

Answer:

C

Explanation:

Effective pruning in Snowflake relies on the organization of data within micro-partitions. By using an ORDER BY clause with clustering keys when loading data into the reporting tables, Snowflake can better organize the data within micro-partitions. This organization allows Snowflake to skip over irrelevant micro-partitions during a query, thus improving query performance and reducing the amount of data scanned¹².

- * Snowflake Documentation on micro-partitions and data clustering²
- * Community article on recognizing unsatisfactory pruning and improving it¹

Question 2

Question Type: MultipleChoice

An Architect needs to design a data unloading strategy for Snowflake, that will be used with the COPY INTO command.

Which configuration is valid?

Options:

A- Location of files: Snowflake internal location

. File formats: CSV, XML

. File encoding: UTF-8

. Encryption: 128-bit

B- Location of files: Amazon S3

. File formats: CSV, JSON

. File encoding: Latin-1 (ISO-8859)

. Encryption: 128-bit

C- Location of files: Google Cloud Storage

. File formats: Parquet

. File encoding: UTF-8

* Compression: gzip

D- Location of files: Azure ADLS

. File formats: JSON, XML, Avro, Parquet, ORC

. Compression: bzip2

. Encryption: User-supplied key

Answer:

C

Question 3

Question Type: MultipleChoice

Which command will create a schema without Fail-safe and will restrict object owners from passing on access to other users?

Options:

- A- create schema EDW.ACCOUNTING WITH MANAGED ACCESS;
- B- create schema EDW.ACCOUNTING WITH MANAGED ACCESS DATA_RETENTION_TIME_IN_DAYS - 7;
- C- create TRANSIENT schema EDW.ACCOUNTING WITH MANAGED ACCESS DATA_RETENTION_TIME_IN_DAYS = 1;
- D- create TRANSIENT schema EDW.ACCOUNTING WITH MANAGED ACCESS DATA_RETENTION_TIME_IN_DAYS = 7;

Answer:

D

Explanation:

A transient schema in Snowflake is designed without a Fail-safe period, meaning it does not incur additional storage costs once it leaves Time Travel, and it is not protected by Fail-safe in the event of a data loss. The WITH MANAGED ACCESS option ensures that all privilege grants, including future grants on objects within the schema, are managed by the schema owner, thus restricting object owners from passing on access to other users¹.

* Snowflake Documentation on creating schemas¹

* Snowflake Documentation on configuring access control²

* Snowflake Documentation on understanding and viewing Fail-safe³

Question 4

Question Type: MultipleChoice

A company is designing a process for importing a large amount of IoT JSON data from cloud storage into Snowflake. New sets of IoT data get generated and uploaded approximately every 5 minutes.

Once the IoT data is in Snowflake, the company needs up-to-date information from an external vendor to join to the data.

a. This data is then presented to users through a dashboard that shows different levels of aggregation. The external vendor is a Snowflake customer.

What solution will MINIMIZE complexity and MAXIMIZE performance?

Options:

- A-** 1. Create an external table over the JSON data in cloud storage.
- 2. Create a task that runs every 5 minutes to run a transformation procedure on new data, based on a saved timestamp.
- 3. Ask the vendor to expose an API so an external function can be used to generate a call to join the data back to the IoT data in the

transformation procedure.

4. Give the transformed table access to the dashboard tool.
5. Perform the aggregations on the dashboard tool.

B- 1. Create an external table over the JSON data in cloud storage.

2. Create a task that runs every 5 minutes to run a transformation procedure on new data based on a saved timestamp.
3. Ask the vendor to create a data share with the required data that can be imported into the company's Snowflake account.
4. Join the vendor's data back to the IoT data using a transformation procedure.
5. Create views over the larger dataset to perform the aggregations required by the dashboard.
6. Give the views access to the dashboard tool.

C- 1. Create a Snowpipe to bring the JSON data into Snowflake.

2. Use streams and tasks to trigger a transformation procedure when new JSON data arrives.
3. Ask the vendor to expose an API so an external function call can be made to join the vendor's data back to the IoT data in a transformation procedure.
4. Create materialized views over the larger dataset to perform the aggregations required by the dashboard.
5. Give the materialized views access to the dashboard tool.

D- 1. Create a Snowpipe to bring the JSON data into Snowflake.

2. Use streams and tasks to trigger a transformation procedure when new JSON data arrives.
3. Ask the vendor to create a data share with the required data that is then imported into the Snowflake account.
4. Join the vendor's data back to the IoT data in a transformation procedure
5. Create materialized views over the larger dataset to perform the aggregations required by the dashboard.
6. Give the materialized views access to the dashboard tool.

Answer:

D

Explanation:

Using Snowpipe for continuous, automated data ingestion minimizes the need for manual intervention and ensures that data is available in Snowflake promptly after it is generated. Leveraging Snowflake's data sharing capabilities allows for efficient and secure access to the vendor's data without the need for complex API integrations. Materialized views provide pre-aggregated data for fast access, which is ideal for dashboards that require high performance¹²³⁴.

- * [Snowflake Documentation on Snowpipe](#)⁴
- * [Snowflake Documentation on Secure Data Sharing](#)²
- * [Best Practices for Data Ingestion with Snowflake](#)¹

Question 5

Question Type: MultipleChoice

A retailer's enterprise data organization is exploring the use of Data Vault 2.0 to model its data lake solution. A Snowflake Architect has been asked to provide recommendations for using Data Vault 2.0 on Snowflake.

What should the Architect tell the data organization? (Select TWO).

Options:

- A-** Change data capture can be performed using the Data Vault 2.0 HASH_DIFF concept.
- B-** Change data capture can be performed using the Data Vault 2.0 HASH_DELTA concept.
- C-** Using the multi-table insert feature in Snowflake, multiple Point-in-Time (PIT) tables can be loaded in parallel from a single join query from the data vault.
- D-** Using the multi-table insert feature, multiple Point-in-Time (PIT) tables can be loaded sequentially from a single join query from the data vault.
- E-** There are performance challenges when using Snowflake to load multiple Point-in-Time (PIT) tables in parallel from a single join query from the data vault.

Answer:

A, C

Explanation:

Data Vault 2.0 on Snowflake supports the HASH_DIFF concept for change data capture, which is a method to detect changes in the data by comparing the hash values of the records. Additionally, Snowflake's multi-table insert feature allows for the loading of multiple PIT tables in parallel from a single join query, which can significantly streamline the data loading process and improve performance¹.

* Snowflake's documentation on multi-table inserts1

* Blog post on optimizing Data Vault architecture on Snowflake2

Question 6

Question Type: MultipleChoice

What is the MOST efficient way to design an environment where data retention is not considered critical, and customization needs are to be kept to a minimum?

Options:

A- Use a transient database.

B- Use a transient schema.

C- Use a transient table.

D- Use a temporary table.

Answer:

A

Question 7

Question Type: MultipleChoice

What actions are permitted when using the Snowflake SQL REST API? (Select TWO).

Options:

- A- The use of a GET command
- B- The use of a PUT command
- C- The use of a ROLLBACK command
- D- The use of a CALL command to a stored procedure which returns a table
- E- Submitting multiple SQL statements in a single call

Answer:

D, E

Question 8

Question Type: MultipleChoice

Role A has the following permissions:

- . USAGE on db1
- . USAGE and CREATE VIEW on schemal in db1
- . SELECT on tablel in schemal

Role B has the following permissions:

- . USAGE on db2
- . USAGE and CREATE VIEW on schema2 in db2
- . SELECT on table2 in schema2

A user has Role A set as the primary role and Role B as a secondary role.

What command will fail for this user?

Options:

- A-** use database db1;
use schema schemal;
create view v1 as select * from db2.schema2.table2;

B- use database db2;
use schema schema2;
create view v2 as select * from db1.schema1. table1;

C- use database db2;
use schema schema2;
select * from db1.schema1.table1 union select * from table2;

D- use database db1;
use schema schema1;
select * from db2.schema2.table2;

Answer:

B

Question 9

Question Type: MultipleChoice

A Developer is having a performance issue with a Snowflake query. The query receives up to 10 different values for one parameter and then performs an aggregation over the majority of a fact table. It then

joins against a smaller dimension table. This parameter value is selected by the different query users when they execute it during business hours. Both the fact and dimension tables are loaded with new data in an overnight import process.

On a Small or Medium-sized virtual warehouse, the query performs slowly. Performance is acceptable on a size Large or bigger warehouse. However, there is no budget to increase costs. The Developer

needs a recommendation that does not increase compute costs to run this query.

What should the Architect recommend?

Options:

- A-** Create a task that will run the 10 different variations of the query corresponding to the 10 different parameters before the users come in to work. The query results will then be cached and ready to respond quickly when the users re-issue the query.
- B-** Create a task that will run the 10 different variations of the query corresponding to the 10 different parameters before the users come in to work. The task will be scheduled to align with the users' working hours in order to allow the warehouse cache to be used.
- C-** Enable the search optimization service on the table. When the users execute the query, the search optimization service will automatically adjust the query execution plan based on the frequently-used parameters.
- D-** Create a dedicated size Large warehouse for this particular set of queries. Create a new role that has USAGE permission on this warehouse and has the appropriate read permissions over the fact and dimension tables. Have users switch to this role and use this warehouse when they want to access this data.

Answer:

C

Explanation:

Enabling the search optimization service on the table can improve the performance of queries that have selective filtering criteria, which seems to be the case here. This service optimizes the execution of queries by creating a persistent data structure called a search access path, which allows some micro-partitions to be skipped during the scanning process. This can significantly speed up query performance without increasing compute costs¹.

* Snowflake Documentation on Search Optimization Service¹.

Question 10

Question Type: MultipleChoice

A company is designing its serving layer for data that is in cloud storage. Multiple terabytes of the data will be used for reporting. Some data does not have a clear use case but could be useful for experimental analysis. This experimentation data changes frequently and is sometimes wiped out and replaced completely in a few days.

The company wants to centralize access control, provide a single point of connection for the end-users, and maintain data governance.

What solution meets these requirements while MINIMIZING costs, administrative effort, and development overhead?

Options:

- A-** Import the data used for reporting into a Snowflake schema with native tables. Then create external tables pointing to the cloud storage folders used for the experimentation data. Then create two different roles with grants to the different datasets to match the different user personas, and grant these roles to the corresponding users.
- B-** Import all the data in cloud storage to be used for reporting into a Snowflake schema with native tables. Then create a role that has access to this schema and manage access to the data through that role.
- C-** Import all the data in cloud storage to be used for reporting into a Snowflake schema with native tables. Then create two different roles with grants to the different datasets to match the different user personas, and grant these roles to the corresponding users.
- D-** Import the data used for reporting into a Snowflake schema with native tables. Then create views that have SELECT commands pointing to the cloud storage files for the experimentation data. Then create two different roles to match the different user personas, and grant these roles to the corresponding users.

Answer:

A

Explanation:

The most cost-effective and administratively efficient solution is to use a combination of native and external tables. Native tables for reporting data ensure performance and governance, while external tables allow for flexibility with frequently changing experimental data. Creating roles with specific grants to datasets aligns with the principle of least privilege, centralizing access control and simplifying user management¹².

* Snowflake Documentation on Optimizing Cost1.

* Snowflake Documentation on Controlling Cost2.

Question 11

Question Type: MultipleChoice

Which Snowflake architecture recommendation needs multiple Snowflake accounts for implementation?

Options:

- A-** Enable a disaster recovery strategy across multiple cloud providers.
- B-** Create external stages pointing to cloud providers and regions other than the region hosting the Snowflake account.
- C-** Enable zero-copy cloning among the development, test, and production environments.
- D-** Enable separation of the development, test, and production environments.

Answer:

D

Explanation:

The Snowflake architecture recommendation that necessitates multiple Snowflake accounts for implementation is the separation of development, test, and production environments. This approach, known as Account per Tenant (APT), isolates tenants into separate Snowflake accounts, ensuring dedicated resources and security isolation¹².

* Snowflake's white paper on "Design Patterns for Building Multi-Tenant Applications on Snowflake" discusses the APT model and its requirement for separate Snowflake accounts for each tenant¹.

* Snowflake Documentation on Secure Data Sharing, which mentions the possibility of sharing data across multiple accounts³.

Question 12

Question Type: MultipleChoice

An Architect needs to design a solution for building environments for development, test, and pre-production, all located in a single Snowflake account. The environments should be based on production data.

Which solution would be MOST cost-effective and performant?

Options:

- A- Use zero-copy cloning into transient tables.
- B- Use zero-copy cloning into permanent tables.
- C- Use CREATE TABLE ... AS SELECT (CTAS) statements.
- D- Use a Snowflake task to trigger a stored procedure to copy data.

Answer:

A

Explanation:

Zero-copy cloning is a feature in Snowflake that allows for the creation of a clone of a database, schema, or table without duplicating any data, which is cost-effective as it saves on storage costs. Transient tables are temporary and do not incur storage costs for the time they are not accessed, making them a cost-effective option for development, test, and pre-production environments that do not require the durability of permanent tables¹²³.

* Snowflake Documentation on Zero-Copy Cloning³.

* Articles discussing the cost-effectiveness and performance benefits of zero-copy cloning¹².

To Get Premium Files for ARA-C01 Visit

<https://www.p2pexams.com/products/ara-c01>

For More Free Questions Visit

<https://www.p2pexams.com/snowflake/pdf/ara-c01>

