Vendor: Snowflake

Exam Code: ARA-C01

Exam Name: Snowflake Certified SnowPro Advanced - Architect

Version: DEMO

QUESTION 1
Which steps are recommended best practices for prioritizing cluster keys in Snowflake? (Choose two.)

A. Choose columns that are frequently used in join predicates.
B. Choose lower cardinality columns to support clustering keys and cost effectiveness.
C. Choose TIMESTAMP columns with nanoseconds for the highest number of unique rows.
D. Choose cluster columns that are most actively used in selective filters.
E. Choose cluster columns that are actively used in the GROUP BY clauses.

Answer: AD
Explanation:
According to the Snowflake documentation, the best practices for choosing clustering keys are:
Choose columns that are frequently used in join predicates. This can improve the join performance by reducing the number of micro-partitions that need to be scanned and joined.
Choose columns that are most actively used in selective filters. This can improve the scan efficiency by skipping micro-partitions that do not match the filter predicates. Avoid using low cardinality columns, such as gender or country, as clustering keys. This can result in poor clustering and high maintenance costs.
Avoid using TIMESTAMP columns with nanoseconds, as they tend to have very high cardinality and low correlation with other columns. This can also result in poor clustering and high maintenance costs.
Avoid using columns with duplicate values or NULLs, as they can cause skew in the clustering and reduce the benefits of pruning.
Cluster on multiple columns if the queries use multiple filters or join predicates. This can increase the chances of pruning more micro-partitions and improve the compression ratio. Clustering is not always useful, especially for small or medium-sized tables, or tables that are not frequently queried or updated. Clustering can incur additional costs for initially clustering the data and maintaining the clustering over time.

QUESTION 2
Which Snowflake data modeling approach is designed for BI queries?

A. 3 NF
B. Star schema
C. Data Vault
D. Snowflake schema

Answer: B
Explanation:
In the context of business intelligence (BI) queries, which are typically focused on data analysis and reporting, the star schema is the most suitable data modeling approach.
Option B: Star Schema - The star schema is a type of relational database schema that is widely used for developing data warehouses and data marts for BI purposes. It consists of a central fact table surrounded by dimension tables. The fact table contains the core data metrics, and the dimension tables contain descriptive attributes related to the fact data. The simplicity of the star schema allows for efficient querying and aggregation, which are common operations in BI reporting.

QUESTION 3
How is the change of local time due to daylight savings time handled in Snowflake tasks? (Choose two.)

A. A task scheduled in a UTC-based schedule will have no issues with the time changes.
B. Task schedules can be designed to follow specified or local time zones to accommodate the time changes.
C. A task will move to a suspended state during the daylight savings time change.
D. A frequent task execution schedule like minutes may not cause a problem, but will affect the task history.
E. A task schedule will follow only the specified time and will fail to handle lost or duplicated hours.

Answer: AB
Explanation:
According to the Snowflake documentation and the web search results, these two statements are true about how the change of local time due to daylight savings time is handled in Snowflake tasks. A task is a feature that allows scheduling and executing SQL statements or stored procedures in Snowflake. A task can be scheduled using a cron expression that specifies the frequency and time zone of the task execution.
A task scheduled in a UTC-based schedule will have no issues with the time changes. UTC is a universal time standard that does not observe daylight savings time. Therefore, a task that uses UTC as the time zone will run at the same time throughout the year, regardless of the local time changes.
Task schedules can be designed to follow specified or local time zones to accommodate the time changes. Snowflake supports using any valid IANA time zone identifier in the cron expression for a task. This allows the task to run according to the local time of the specified time zone, which may include daylight savings time adjustments. For example, a task that uses Europe/London as the time zone will run one hour earlier or later when the local time switches between GMT and BST.

QUESTION 4
An Architect needs to grant a group of ORDER_ADMIN users the ability to clean old data in an ORDERS table (deleting all records older than 5 years), without granting any privileges on the table. The group's manager (ORDER_MANAGER) has full DELETE privileges on the table. How can the ORDER_ADMIN role be enabled to perform this data cleanup, without needing the DELETE privilege held by the ORDER_MANAGER role?

A. Create a stored procedure that runs with caller's rights, including the appropriate "> 5 years" business logic, and grant USAGE on this procedure to ORDER_ADMIN. The ORDER_MANAGER role owns the procedure.
B. Create a stored procedure that can be run using both caller's and owner's rights (allowing the user to specify which rights are used during execution), and grant USAGE on this procedure to ORDER_ADMIN. The ORDER_MANAGER role owns the procedure.
C. Create a stored procedure that runs with owner's rights, including the appropriate "> 5 years" business logic, and grant USAGE on this procedure to ORDER_ADMIN. The ORDER_MANAGER role owns the procedure.
D. This scenario would actually not be possible in Snowflake ?any user performing a DELETE on a table requires the DELETE privilege to be granted to the role they are using.

Answer: C
Explanation:
This is the correct answer because it allows the ORDER_ADMIN role to perform the data cleanup without needing the DELETE privilege on the ORDERS table. A stored procedure is a feature that allows scheduling and executing SQL statements or stored procedures in Snowflake. A stored procedure can run with either the caller's rights or the owner's rights. A caller's rights stored procedure runs with the privileges of the role that called the stored procedure, while an owner's

rights stored procedure runs with the privileges of the role that created the stored procedure. By creating a stored procedure that runs with owner's rights, the ORDER_MANAGER role can delegate the specific task of deleting old data to the ORDER_ADMIN role, without granting the ORDER_ADMIN role more general privileges on the ORDERS table. The stored procedure must include the appropriate business logic to delete only the records older than 5 years, and the ORDER_MANAGER role must grant the USAGE privilege on the stored procedure to the ORDER_ADMIN role. The ORDER_ADMIN role can then execute the stored procedure to perform the data cleanup.

QUESTION 5
A company's daily Snowflake workload consists of a huge number of concurrent queries triggered between 9pm and 11pm. At the individual level, these queries are smaller statements that get completed within a short time period.
What configuration can the company's Architect implement to enhance the performance of this workload? (Choose two.)

A. Enable a multi-clustered virtual warehouse in maximized mode during the workload duration.
B. Set the MAX_CONCURRENCY_LEVEL to a higher value than its default value of 8 at the virtual warehouse level.
C. Increase the size of the virtual warehouse to size X-Large.
D. Reduce the amount of data that is being processed through this workload.
E. Set the connection timeout to a higher value than its default.

Answer: AB
Explanation:
These two configuration options can enhance the performance of the workload that consists of a huge number of concurrent queries that are smaller and faster. Enabling a multi-clustered virtual warehouse in maximized mode allows the warehouse to scale out automatically by adding more clusters as soon as the current cluster is fully loaded, regardless of the number of queries in the queue. This can improve the concurrency and throughput of the workload by minimizing or preventing queuing. The maximized mode is suitable for workloads that require high performance and low latency, and are less sensitive to credit consumption. Setting the MAX_CONCURRENCY_LEVEL to a higher value than its default value of 8 at the virtual warehouse level allows the warehouse to run more queries concurrently on each cluster. This can improve the utilization and efficiency of the warehouse resources, especially for smaller and faster queries that do not require a lot of processing power. The MAX_CONCURRENCY_LEVEL parameter can be set when creating or modifying a warehouse, and it can be changed at any time.

QUESTION 6
Which system functions does Snowflake provide to monitor clustering information within a table (Choose two.)

A. SYSTEM$CLUSTERING_INFORMATION
B. SYSTEM$CLUSTERING_USAGE
C. SYSTEM$CLUSTERING_DEPTH
D. SYSTEM$CLUSTERING_KEYS
E. SYSTEM$CLUSTERING_PERCENT

Answer: AC
Explanation:
According to the Snowflake documentation, these two system functions are provided by

Snowflake to monitor clustering information within a table. A system function is a type of function that allows executing actions or returning information about the system. A clustering key is a feature that allows organizing data across micro-partitions based on one or more columns in the table. Clustering can improve query performance by reducing the number of files to scan.
SYSTEM$CLUSTERING_INFORMATION is a system function that returns clustering information, including average clustering depth, for a table based on one or more columns in the table. The function takes a table name and an optional column name or expression as arguments, and returns a JSON string with the clustering information. The clustering information includes the cluster by keys, the total partition count, the total constant partition count, the average overlaps, and the average depth.
SYSTEM$CLUSTERING_DEPTH is a system function that returns the clustering depth for a table based on one or more columns in the table. The function takes a table name and an optional column name or expression as arguments, and returns an integer value with the clustering depth. The clustering depth is the maximum number of overlapping micro-partitions for any micro-partition in the table. A lower clustering depth indicates a better clustering.

QUESTION 7
A company has a table with that has corrupted data, named Data. The company wants to recover the data as it was 5 minutes ago using cloning and Time Travel.
What command will accomplish this?

A. CREATE CLONE TABLE Recover_Data FROM Data AT(OFFSET => -60*5);
B. CREATE CLONE Recover_Data FROM Data AT(OFFSET => -60*5);
C. CREATE TABLE Recover_Data CLONE Data AT(OFFSET => -60*5);
D. CREATE TABLE Recover Data CLONE Data AT(TIME => -60*5);

Answer: C
Explanation:
This is the correct command to create a clone of the table Data as it was 5 minutes ago using cloning and Time Travel. Cloning is a feature that allows creating a copy of a database, schema, table, or view without duplicating the data or metadata. Time Travel is a feature that enables accessing historical data (i.e. data that has been changed or deleted) at any point within a defined period. To create a clone of a table at a point in time in the past, the syntax is:
CREATE TABLE <clone_name> CLONE <source_table> AT (OFFSET => <offset_in_seconds>);
The OFFSET parameter specifies the time difference in seconds from the present time. A negative value indicates a point in the past. For example, -60*5 means 5 minutes ago.
Alternatively, the TIMESTAMP parameter can be used to specify an exact timestamp in the past. The clone will contain the data as it existed in the source table at the specified point in time.

QUESTION 8
A company has an inbound share set up with eight tables and five secure views. The company plans to make the share part of its production data pipelines. Which actions can the company take with the inbound share? (Choose two.)

A. Clone a table from a share.
B. Grant modify permissions on the share.
C. Create a table from the shared database.
D. Create additional views inside the shared database.
E. Create a table stream on the shared table.

Answer: AD
Explanation:

These two actions are possible with an inbound share, according to the Snowflake documentation and the web search results. An inbound share is a share that is created by another Snowflake account (the provider) and imported into your account (the consumer). An inbound share allows you to access the data shared by the provider, but not to modify or delete it. However, you can perform some actions with the inbound share, such as:
Clone a table from a share. You can create a copy of a table from an inbound share using the CREATE TABLE ... CLONE statement. The clone will contain the same data and metadata as the original table, but it will be independent of the share. You can modify or delete the clone as you wish, but it will not reflect any changes made to the original table by the provider. Create additional views inside the shared database. You can create views on the tables or views from an inbound share using the CREATE VIEW statement. The views will be stored in the shared database, but they will be owned by your account. You can query the views as you would query any other view in your account, but you cannot modify or delete the underlying objects from the share.

QUESTION 9
A company has several sites in different regions from which the company wants to ingest data. Which of the following will enable this type of data ingestion?

A. The company must have a Snowflake account in each cloud region to be able to ingest data to that account.
B. The company must replicate data between Snowflake accounts.
C. The company should provision a reader account to each site and ingest the data through the reader accounts.
D. The company should use a storage integration for the external stage.

Answer: D
Explanation:
This is the correct answer because it allows the company to ingest data from different regions using a storage integration for the external stage. A storage integration is a feature that enables secure and easy access to files in external cloud storage from Snowflake. A storage integration can be used to create an external stage, which is a named location that references the files in the external storage. An external stage can be used to load data into Snowflake tables using the COPY INTO command, or to unload data from Snowflake tables using the COPY INTO LOCATION command. A storage integration can support multiple regions and cloud platforms, as long as the external storage service is compatible with Snowflake.

QUESTION 10
What Snowflake features should be leveraged when modeling using Data Vault?

A. Snowflake's support of multi-table inserts into the data model's Data Vault tables
B. Data needs to be pre-partitioned to obtain a superior data access performance
C. Scaling up the virtual warehouses will support parallel processing of new source loads
D. Snowflake's ability to hash keys so that hash key joins can run faster than integer joins

Answer: A
Explanation:
These two features are relevant for modeling using Data Vault on Snowflake. Data Vault is a data modeling approach that organizes data into hubs, links, and satellites. Data Vault is designed to enable high scalability, flexibility, and performance for data integration and analytics. Snowflake is a cloud data platform that supports various data modeling techniques, including Data Vault. Snowflake provides some features that can enhance the Data Vault modeling, such as:
Snowflake's support of multi-table inserts into the data model's Data Vault tables. Multi-table

inserts (MTI) are a feature that allows inserting data from a single query into multiple tables in a single DML statement. MTI can improve the performance and efficiency of loading data into Data Vault tables, especially for real-time or near-real-time data integration. MTI can also reduce the complexity and maintenance of the loading code, as well as the data duplication and latency. Scaling up the virtual warehouses will support parallel processing of new source loads. Virtual warehouses are a feature that allows provisioning compute resources on demand for data processing. Virtual warehouses can be scaled up or down by changing the size of the warehouse, which determines the number of servers in the warehouse. Scaling up the virtual warehouses can improve the performance and concurrency of processing new source loads into Data Vault tables, especially for large or complex data sets. Scaling up the virtual warehouses can also leverage the parallelism and distribution of Snowflake's architecture, which can optimize the data loading and querying.

QUESTION 11
A company's client application supports multiple authentication methods, and is using Okta. What is the best practice recommendation for the order of priority when applications authenticate to Snowflake?

A. 1) OAuth (either Snowflake OAuth or External OAuth)
   2) External browser
   3) Okta native authentication
   4) Key Pair Authentication, mostly used for service account users
   5) Password
B. 1) External browser, SSO
   2) Key Pair Authentication, mostly used for development environment users
   3) Okta native authentication
   4) OAuth (ether Snowflake OAuth or External OAuth)
   5) Password
C. 1) Okta native authentication
   2) Key Pair Authentication, mostly used for production environment users
   3) Password
   4) OAuth (either Snowflake OAuth or External OAuth)
   5) External browser, SSO
D. 1) Password
   2) Key Pair Authentication, mostly used for production environment users
   3) Okta native authentication
   4) OAuth (either Snowflake OAuth or External OAuth)
   5) External browser, SSO

Answer: A
Explanation:
This is the best practice recommendation for the order of priority when applications authenticate to Snowflake, according to the Snowflake documentation and the web search results. Authentication is the process of verifying the identity of a user or application that connects to Snowflake. Snowflake supports multiple authentication methods, each with different advantages and disadvantages. The recommended order of priority is based on the following factors:
Security: The authentication method should provide a high level of security and protection against unauthorized access or data breaches. The authentication method should also support multi-factor authentication (MFA) or single sign-on (SSO) for additional security.
Convenience: The authentication method should provide a smooth and easy user experience, without requiring complex or manual steps. The authentication method should also support seamless integration with external identity providers or applications. Flexibility: The authentication method should provide a range of options and features to suit different use cases and scenarios. The authentication method should also support customization and configuration to meet specific

requirements.

QUESTION 12
What is a valid object hierarchy when building a Snowflake environment?

A. Account --> Database --> Schema --> Warehouse
B. Organization --> Account --> Database --> Schema --> Stage
C. Account --> Schema > Table --> Stage
D. Organization --> Account --> Stage --> Table --> View

Answer: B
Explanation:
This is the valid object hierarchy when building a Snowflake environment, according to the Snowflake documentation and the web search results. Snowflake is a cloud data platform that supports various types of objects, such as databases, schemas, tables, views, stages, warehouses, and more. These objects are organized in a hierarchical structure, as follows: Organization: An organization is the top-level entity that represents a group of Snowflake accounts that are related by business needs or ownership. An organization can have one or more accounts, and can enable features such as cross-account data sharing, billing and usage reporting, and single sign-on across accounts.

QUESTION 13
Which of the following are characteristics of Snowflake's parameter hierarchy?

A. Session parameters override virtual warehouse parameters.
B. Virtual warehouse parameters override user parameters.
C. Table parameters override virtual warehouse parameters.
D. Schema parameters override account parameters.

Answer: B
Explanation:
In Snowflake's parameter hierarchy, virtual warehouse parameters take precedence over user parameters. This hierarchy is designed to ensure that settings at the virtual warehouse level, which typically reflect the requirements of a specific workload or set of queries, override the preferences set at the individual user level. This helps maintain consistent performance and resource utilization as specified by the administrators managing the virtual warehouses.

QUESTION 14
A Snowflake Architect is designing a multi-tenant application strategy for an organization in the Snowflake Data Cloud and is considering using an Account Per Tenant strategy. Which requirements will be addressed with this approach? (Choose two.)

A. There needs to be fewer objects per tenant.
B. Security and Role-Based Access Control (RBAC) policies must be simple to configure.
C. Compute costs must be optimized.
D. Tenant data shape may be unique per tenant.
E. Storage costs must be optimized.

Answer: BD
Explanation:
The Account Per Tenant strategy involves creating separate Snowflake accounts for each tenant

within the multi-tenant application. This approach offers a number of advantages.
Option B: With separate accounts, each tenant's environment is isolated, making security and RBAC policies simpler to configure and maintain. This is because each account can have its own set of roles and privileges without the risk of cross-tenant access or the complexity of maintaining a highly granular permission model within a shared environment.
Option D: This approach also allows for each tenant to have a unique data shape, meaning that the database schema can be tailored to the specific needs of each tenant without affecting others. This can be essential when tenants have different data models, usage patterns, or application customizations.

QUESTION 15
An Architect has been asked to clone schema STAGING as it looked one week ago, Tuesday June 1st at 8:00 AM, to recover some objects.
The STAGING schema has 50 days of retention.
The Architect runs the following statement:

CREATE SCHEMA STAGING_CLONE CLONE STAGING at (timestamp => '2021-06-01 08:00:00');

The Architect receives the following error: Time travel data is not available for schema STAGING. The requested time is either beyond the allowed time travel period or before the object creation time. The Architect then checks the schema history and sees the following:

CREATED_ON|NAME|DROPPED_ON
2021-06-02 23:00:00 | STAGING | NULL
2021-05-01 10:00:00 | STAGING | 2021-06-02 23:00:00

How can cloning the STAGING schema be achieved?

A. Undrop the STAGING schema and then rerun the CLONE statement.
B. Modify the statement: CREATE SCHEMA STAGING_CLONE CLONE STAGING at (timestamp => '2021- 05-01 10:00:00');
C. Rename the STAGING schema and perform an UNDROP to retrieve the previous STAGING schema version, then run the CLONE statement.
D. Cloning cannot be accomplished because the STAGING schema version was not active during the proposed Time Travel time period.

Answer: D
Explanation:
The error encountered during the cloning attempt arises because the schema STAGING as it existed on June 1st, 2021, is not within the Time Travel retention period. According to the schema history, STAGING was recreated on June 2nd, 2021, after being dropped on the same day. The requested timestamp of '2021-06-01 08:00:00' is prior to this recreation, hence not available. The STAGING schema from before June 2nd was dropped and exceeded the Time Travel period for retrieval by the time of the cloning attempt. Therefore, cloning STAGING as it looked on June 1st, 2021, cannot be achieved because the data from that time is no longer available within the allowed Time Travel window.

Thank You for Trying Our Product

Passleader Certification Exam Features:

★ More than **99,900** Satisfied Customers Worldwide.

★ Average **99.9%** Success Rate.

★ **Free Update** to match latest and real exam scenarios.

★ **Instant Download** Access! No Setup required.

★ Questions & Answers are downloadable in **PDF** format and **VCE** test engine format.

★ Multi-Platform capabilities - **Windows, Laptop, Mac, Android, iPhone, iPod, iPad**.

★ **100%** Guaranteed Success or **100%** Money Back Guarantee.

★ **Fast**, helpful support **24x7**.

View list of all certification exams: http://www.passleader.com/all-products.html

10% Discount Coupon Code: **ASTR14**