Topic 1 - Exam A

Question #1   *Topic 1*

What built-in Snowflake features make use of the change tracking metadata for a table? (Choose two.)

A. The MERGE command

B. The UPSERT command

C. The CHANGES clause

D. A STREAM object

E. Thee CHANGE_DATA_CAPTURE command

Question #2   *Topic 1*

When using the Snowflake Connector for Kafka, what data formats are supported for the messages? (Choose two.)

A. CSV

B. XML

C. Avro

D. JSON

E. Parquet

Question #3   *Topic 1*

At which object type level can the APPLY MASKING POLICY, APPLY ROW ACCESS POLICY and APPLY SESSION POLICY privileges be granted?

A. Global

B. Database

C. Schema

D. Table

An Architect uses COPY INTO with the ON_ERROR=SKIP_FILE option to bulk load CSV files into a table called TABLEA, using its table stage. One file named file5.csv fails to load. The Architect fixes the file and re-loads it to the stage with the exact same file name it had previously. Which commands should the Architect use to load only file5.csv file from the stage? (Choose two.)

A. COPY INTO tablea FROM @%tablea RETURN_FAILED_ONLY = TRUE;

B. COPY INTO tablea FROM @%tablea;

C. COPY INTO tablea FROM @%tablea FILES = ('file5.csv');

D. COPY INTO tablea FROM @%tablea FORCE = TRUE;

E. COPY INTO tablea FROM @%tablea NEW_FILES_ONLY = TRUE;

F. COPY INTO tablea FROM @%tablea MERGE = TRUE;

A large manufacturing company runs a dozen individual Snowflake accounts across its business divisions. The company wants to increase the level of data sharing to support supply chain optimizations and increase its purchasing leverage with multiple vendors.
The company's Snowflake Architects need to design a solution that would allow the business divisions to decide what to share, while minimizing the level of effort spent on configuration and management. Most of the company divisions use Snowflake accounts in the same cloud deployments with a few exceptions for European-based divisions.
According to Snowflake recommended best practice, how should these requirements be met?

A. Migrate the European accounts in the global region and manage shares in a connected graph architecture. Deploy a Data Exchange.

B. Deploy a Private Data Exchange in combination with data shares for the European accounts.

C. Deploy to the Snowflake Marketplace making sure that invoker_share() is used in all secure views.

D. Deploy a Private Data Exchange and use replication to allow European data shares in the Exchange.

A user has the appropriate privilege to see unmasked data in a column.
If the user loads this column data into another column that does not have a masking policy, what will occur?

A. Unmasked data will be loaded in the new column.

B. Masked data will be loaded into the new column.

C. Unmasked data will be loaded into the new column but only users with the appropriate privileges will be able to see the unmasked data.

D. Unmasked data will be loaded into the new column and no users will be able to see the unmasked data.

How can an Architect enable optimal clustering to enhance performance for different access paths on a given table?

A. Create multiple clustering keys for a table.

B. Create multiple materialized views with different cluster keys.

C. Create super projections that will automatically create clustering.

D. Create a clustering key that contains all columns used in the access paths.

Company A would like to share data in Snowflake with Company B. Company B is not on the same cloud platform as Company A.
What is required to allow data sharing between these two companies?

A. Create a pipeline to write shared data to a cloud storage location in the target cloud provider.

B. Ensure that all views are persisted, as views cannot be shared across cloud platforms.

C. Setup data replication to the region and cloud platform where the consumer resides.

D. Company A and Company B must agree to use a single cloud platform: Data sharing is only possible if the companies share the same cloud provider.

What are some of the characteristics of result set caches? (Choose three.)

A. Time Travel queries can be executed against the result set cache.

B. Snowflake persists the data results for 24 hours.

C. Each time persisted results for a query are used, a 24-hour retention period is reset.

D. The data stored in the result cache will contribute to storage costs.

E. The retention period can be reset for a maximum of 31 days.

F. The result set cache is not shared between warehouses.

Which organization-related tasks can be performed by the ORGADMIN role? (Choose three.)

A. Changing the name of the organization

B. Creating an account

C. Viewing a list of organization accounts

D. Changing the name of an account

E. Deleting an account

F. Enabling the replication of a database

A Data Engineer is designing a near real-time ingestion pipeline for a retail company to ingest event logs into Snowflake to derive insights. A Snowflake Architect is asked to define security best practices to configure access control privileges for the data load for auto-ingest to Snowpipe. What are the MINIMUM object privileges required for the Snowpipe user to execute Snowpipe?

A. OWNERSHIP on the named pipe, USAGE on the named stage, target database, and schema, and INSERT and SELECT on the target table

B. OWNERSHIP on the named pipe, USAGE and READ on the named stage, USAGE on the target database and schema, and INSERT end SELECT on the target table

C. CREATE on the named pipe, USAGE and READ on the named stage, USAGE on the target database and schema, and INSERT end SELECT on the target table

D. USAGE on the named pipe, named stage, target database, and schema, and INSERT and SELECT on the target table

The IT Security team has identified that there is an ongoing credential stuffing attack on many of their organization's system.
What is the BEST way to find recent and ongoing login attempts to Snowflake?

A. Call the LOGIN_HISTORY Information Schema table function.

B. Query the LOGIN_HISTORY view in the ACCOUNT_USAGE schema in the SNOWFLAKE database.

C. View the History tab in the Snowflake UI and set up a filter for SQL text that contains the text "LOGIN".

D. View the Users section in the Account tab in the Snowflake UI and review the last login column.

An Architect has a VPN_ACCESS_LOGS table in the SECURITY_LOGS schema containing timestamps of the connection and disconnection, username of the user, and summary statistics.
What should the Architect do to enable the Snowflake search optimization service on this table?

A. Assume role with OWNERSHIP on future tables and ADD SEARCH OPTIMIZATION on the SECURITY_LOGS schema.

B. Assume role with ALL PRIVILEGES including ADD SEARCH OPTIMIZATION in the SECURITY LOGS schema.

C. Assume role with OWNERSHIP on VPN_ACCESS_LOGS and ADD SEARCH OPTIMIZATION in the SECURITY_LOGS schema.

D. Assume role with ALL PRIVILEGES on VPN_ACCESS_LOGS and ADD SEARCH OPTIMIZATION in the SECURITY_LOGS schema.

The table contains five columns and it has millions of records. The cardinality distribution of the columns is shown below:

Column	Number of Distinct Values
C1	10,790
C2	108
C3	302,605
C4	1.117,736
C5	2.205,400

Column C4 and C5 are mostly used by SELECT queries in the GROUP BY and ORDER BY clauses. Whereas columns C1, C2 and C3 are heavily used in filter and join conditions of SELECT queries.
The Architect must design a clustering key for this table to improve the query performance.
Based on Snowflake recommendations, how should the clustering key columns be ordered while defining the multi-column clustering key?

A. C5, C4, C2

B. C3, C4, C5

C. C1, C3, C2

D. C2, C1, C3

Which security, governance, and data protection features require, at a MINIMUM, the Business Critical edition of Snowflake? (Choose two.)

A. Extended Time Travel (up to 90 days)

B. Customer-managed encryption keys through Tri-Secret Secure

C. Periodic rekeying of encrypted data

D. AWS, Azure, or Google Cloud private connectivity to Snowflake

E. Federated authentication and SSO

A company wants to deploy its Snowflake accounts inside its corporate network with no visibility on the internet. The company is using a VPN infrastructure and Virtual Desktop Infrastructure (VDI) for its Snowflake users. The company also wants to re-use the login credentials set up for the VDI to eliminate redundancy when managing logins.
What Snowflake functionality should be used to meet these requirements? (Choose two.)

A. Set up replication to allow users to connect from outside the company VPN.

B. Provision a unique company Tri-Secret Secure key.

C. Use private connectivity from a cloud provider.

D. Set up SSO for federated authentication.

E. Use a proxy Snowflake account outside the VPN, enabling client redirect for user logins.

How do Snowflake databases that are created from shares differ from standard databases that are not created from shares? (Choose three.)

A. Shared databases are read-only.

B. Shared databases must be refreshed in order for new data to be visible.

C. Shared databases cannot be cloned.

D. Shared databases are not supported by Time Travel.

E. Shared databases will have the PUBLIC or INFORMATION_SCHEMA schemas without explicitly granting these schemas to the share.

F. Shared databases can also be created as transient databases.

What integration object should be used to place restrictions on where data may be exported?

A. Stage integration

B. Security integration

C. Storage integration

D. API integration

The following DDL command was used to create a task based on a stream:

```
CREATE TASK ts_insert_new_customers
  WAREHOUSE = MY_WH
  Schedule = '5 minute'
WHEN
  System$STREAM_HAS_DATA('MYSTREAM')
AS
  INSERT INTO new customers(id, name) SELECT id, name
FROM mystream WHERE METADATA$ACTION = 'INSERT';
```

Assuming MY_WH is set to auto_suspend – 60 and used exclusively for this task, which statement is true?

A. The warehouse MY_WH will be made active every five minutes to check the stream.

B. The warehouse MY_WH will only be active when there are results in the stream.

C. The warehouse MY_WH will never suspend.

D. The warehouse MY_WH will automatically resize to accommodate the size of the stream.

What is a characteristic of loading data into Snowflake using the Snowflake Connector for Kafka?

A. The Connector only works in Snowflake regions that use AWS infrastructure.

B. The Connector works with all file formats, including text, JSON, Avro, Ore, Parquet, and XML.

C. The Connector creates and manages its own stage, file format, and pipe objects.

D. Loads using the Connector will have lower latency than Snowpipe and will ingest data in real time.

A healthcare company wants to share data with a medical institute. The institute is running a Standard edition of Snowflake; the healthcare company is running a Business Critical edition.
How can this data be shared?

A. The healthcare company will need to change the institute's Snowflake edition in the accounts panel.

B. By default, sharing is supported from a Business Critical Snowflake edition to a Standard edition.

C. Contact Snowflake and they will execute the share request for the healthcare company.

D. Set the share_restriction parameter on the shared object to false.

An Architect is designing a pipeline to stream event data into Snowflake using the Snowflake Kafka connector. The Architect's highest priority is to configure the connector to stream data in the MOST cost-effective manner.
Which of the following is recommended for optimizing the cost associated with the Snowflake Kafka connector?

A. Utilize a higher Buffer.flush.time in the connector configuration.

B. Utilize a higher Buffer.size.bytes in the connector configuration.

C. Utilize a lower Buffer.size.bytes in the connector configuration.

D. Utilize a lower Buffer.count.records in the connector configuration.

An Architect has chosen to separate their Snowflake Production and QA environments using two separate Snowflake accounts.

The QA account is intended to run and test changes on data and database objects before pushing those changes to the Production account. It is a requirement that all database objects and data in the QA account need to be an exact copy of the database objects, including privileges and data in the Production account on at least a nightly basis.

Which is the LEAST complex approach to use to populate the QA account with the Production account's data and database objects on a nightly basis?

A. 1. Create a share in the Production account for each database
2. Share access to the QA account as a Consumer
3. The QA account creates a database directly from each share
4. Create clones of those databases on a nightly basis
5. Run tests directly on those cloned databases

B. 1. Create a stage in the Production account
2. Create a stage in the QA account that points to the same external object-storage location
3. Create a task that runs nightly to unload each table in the Production account into the stage
4. Use Snowpipe to populate the QA account

C. 1. Enable replication for each database in the Production account
2. Create replica databases in the QA account
3. Create clones of the replica databases on a nightly basis
4. Run tests directly on those cloned databases

D. 1. In the Production account, create an external function that connects into the QA account and returns all the data for one specific table
2. Run the external function as part of a stored procedure that loops through each table in the Production account and populates each table in the QA account

A user can change object parameters using which of the following roles?

A. ACCOUNTADMIN, SECURITYADMIN

B. SYSADMIN, SECURITYADMIN

C. ACCOUNTADMIN, USER with PRIVILEGE

D. SECURITYADMIN, USER with PRIVILEGE

A media company needs a data pipeline that will ingest customer review data into a Snowflake table, and apply some transformations. The company also needs to use Amazon Comprehend to do sentiment analysis and make the de-identified final data set available publicly for advertising companies who use different cloud providers in different regions.

The data pipeline needs to run continuously ang efficiently as new records arrive in the object storage leveraging event notifications. Also, the operational complexity, maintenance of the infrastructure, including platform upgrades and security, and the development effort should be minimal.

Which design will meet these requirements?

A. Ingest the data using COPY INTO and use streams and tasks to orchestrate transformations. Export the data into Amazon S3 to do model inference with Amazon Comprehend and ingest the data back into a Snowflake table. Then create a listing in the Snowflake Marketplace to make the data available to other companies.

B. Ingest the data using Snowpipe and use streams and tasks to orchestrate transformations. Create an external function to do model inference with Amazon Comprehend and write the final records to a Snowflake table. Then create a listing in the Snowflake Marketplace to make the data available to other companies.

C. Ingest the data into Snowflake using Amazon EMR and PySpark using the Snowflake Spark connector. Apply transformations using another Spark job. Develop a python program to do model inference by leveraging the Amazon Comprehend text analysis API. Then write the results to a Snowflake table and create a listing in the Snowflake Marketplace to make the data available to other companies.

D. Ingest the data using Snowpipe and use streams and tasks to orchestrate transformations. Export the data into Amazon S3 to do model inference with Amazon Comprehend and ingest the data back into a Snowflake table. Then create a listing in the Snowflake Marketplace to make the data available to other companies.

A Snowflake Architect is designing an application and tenancy strategy for an organization where strong legal isolation rules as well as multi-tenancy are requirements.

Which approach will meet these requirements if Role-Based Access Policies (RBAC) is a viable option for isolating tenants?

A. Create accounts for each tenant in the Snowflake organization.

B. Create an object for each tenant strategy if row level security is viable for isolating tenants.

C. Create an object for each tenant strategy if row level security is not viable for isolating tenants.

D. Create a multi-tenant table strategy if row level security is not viable for isolating tenants.

Which statements describe characteristics of the use of materialized views in Snowflake? (Choose two.)

A. They can include ORDER BY clauses.

B. They cannot include nested subqueries.

C. They can include context functions, such as CURRENT_TIME().

D. They can support MIN and MAX aggregates.

E. They can support inner joins, but not outer joins.

The Data Engineering team at a large manufacturing company needs to engineer data coming from many sources to support a wide variety of use cases and data consumer requirements which include:

1. Finance and Vendor Management team members who require reporting and visualization
2. Data Science team members who require access to raw data for ML model development
3. Sales team members who require engineered and protected data for data monetization

What Snowflake data modeling approaches will meet these requirements? (Choose two.)

    A. Consolidate data in the company's data lake and use EXTERNAL TABLES.

    B. Create a raw database for landing and persisting raw data entering the data pipelines.

    C. Create a set of profile-specific databases that aligns data with usage patterns.

    D. Create a single star schema in a single database to support all consumers' requirements.

    E. Create a Data Vault as the sole data pipeline endpoint and have all consumers directly access the Vault.

An Architect on a new project has been asked to design an architecture that meets Snowflake security, compliance, and governance requirements as follows:

1. Use Tri-Secret Secure in Snowflake
2. Share some information stored in a view with another Snowflake customer
3. Hide portions of sensitive information from some columns
4. Use zero-copy cloning to refresh the non-production environment from the production environment

To meet these requirements, which design elements must be implemented? (Choose three.)

    A. Define row access policies.

    B. Use the Business Critical edition of Snowflake.

    C. Create a secure view.

    D. Use the Enterprise edition of Snowflake.

    E. Use Dynamic Data Masking.

    F. Create a materialized view.

Which of the following are characteristics of how row access policies can be applied to external tables? (Choose three.)

    A. An external table can be created with a row access policy, and the policy can be applied to the VALUE column.

    B. A row access policy can be applied to the VALUE column of an existing external table.

    C. A row access policy cannot be directly added to a virtual column of an external table.

    D. External tables are supported as mapping tables in a row access policy.

    E. While cloning a database, both the row access policy and the external table will be cloned.

    F. A row access policy cannot be applied to a view created on top of an external table.

An Architect needs to allow a user to create a database from an inbound share.
To meet this requirement, the user's role must have which privileges? (Choose two.)

    A. IMPORT SHARE;

    B. IMPORT PRIVILEGES;

    C. CREATE DATABASE;

    D. CREATE SHARE;

    E. IMPORT DATABASE;

Files arrive in an external stage every 10 seconds from a proprietary system. The files range in size from 500 K to 3 MB. The data must be accessible by dashboards as soon as it arrives.
How can a Snowflake Architect meet this requirement with the LEAST amount of coding? (Choose two.)

    A. Use Snowpipe with auto-ingest.

    B. Use a COPY command with a task.

    C. Use a materialized view on an external table.

    D. Use the COPY INTO command.

    E. Use a combination of a task and a stream.

When loading data into a table that captures the load time in a column with a default value of either CURRENT_TIME() or CURRENT_TIMESTAMP() what will occur?

    A. All rows loaded using a specific COPY statement will have varying timestamps based on when the rows were inserted.

    B. Any rows loaded using a specific COPY statement will have varying timestamps based on when the rows were read from the source.

    C. Any rows loaded using a specific COPY statement will have varying timestamps based on when the rows were created in the source.

    D. All rows loaded using a specific COPY statement will have the same timestamp value.

How does a standard virtual warehouse policy work in Snowflake?

    A. It conserves credits by keeping running clusters fully loaded rather than starting additional clusters.

    B. It starts only if the system estimates that there is a query load that will keep the cluster busy for at least 6 minutes.

    C. It starts only f the system estimates that there is a query load that will keep the cluster busy for at least 2 minutes.

    D. It prevents or minimizes queuing by starting additional clusters instead of conserving credits.

Which feature provides the capability to define an alternate cluster key for a table with an existing cluster key?

A. External table

B. Materialized view

C. Search optimization

D. Result cache

An Architect would like to save quarter-end financial results for the previous six years.
Which Snowflake feature can the Architect use to accomplish this?

A. Search optimization service

B. Materialized view

C. Time Travel

D. Zero-copy cloning

E. Secure views

A company is using a Snowflake account in Azure. The account has SAML SSO set up using ADFS as a SCIM identity provider. To validate Private Link connectivity, an Architect performed the following steps:
Confirmed Private Link URLs are working by logging in with a username/password account
Verified DNS resolution by running nslookups against Private Link URLs
Validated connectivity using SnowCD
Disabled public access using a network policy set to use the company's IP address range
However, the following error message is received when using SSO to log into the company account:
IP XX.XXX.XX.XX is not allowed to access snowflake. Contact your local security administrator.
What steps should the Architect take to resolve this error and ensure that the account is accessed using only Private Link? (Choose two.)

A. Alter the Azure security integration to use the Private Link URLs.

B. Add the IP address in the error message to the allowed list in the network policy.

C. Generate a new SCIM access token using system$generate_scim_access_token and save it to Azure AD.

D. Update the configuration of the Azure AD SSO to use the Private Link URLs.

E. Open a case with Snowflake Support to authorize the Private Link URLs' access to the account.

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.

Which Snowflake data modeling approach is designed for BI queries?

A. 3 NF

B. Star schema

C. Data Vault

D. Snowflake schema

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.

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.

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.

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

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);

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.

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.

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

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

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

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.

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.

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.

What are purposes for creating a storage integration? (Choose three.)

A. Control access to Snowflake data using a master encryption key that is maintained in the cloud provider's key management service.

B. Store a generated identity and access management (IAM) entity for an external cloud provider regardless of the cloud provider that hosts the Snowflake account.

C. Support multiple external stages using one single Snowflake object.

D. Avoid supplying credentials when creating a stage or when loading or unloading data.

E. Create private VPC endpoints that allow direct, secure connectivity between VPCs without traversing the public internet.

F. Manage credentials from multiple cloud providers in one single Snowflake object.

A healthcare company is deploying a Snowflake account that may include Personal Health Information (PHI). The company must ensure compliance with all relevant privacy standards.
Which best practice recommendations will meet data protection and compliance requirements? (Choose three.)

A. Use, at minimum, the Business Critical edition of Snowflake.

B. Create Dynamic Data Masking policies and apply them to columns that contain PHI.

C. Use the Internal Tokenization feature to obfuscate sensitive data.

D. Use the External Tokenization feature to obfuscate sensitive data.

E. Rewrite SQL queries to eliminate projections of PHI data based on current_role().

F. Avoid sharing data with partner organizations.

There are two databases in an account, named fin_db and hr_db which contain payroll and employee data, respectively. Accountants and Analysts in the company require different permissions on the objects in these databases to perform their jobs. Accountants need read-write access to fin_db but only require read-only access to hr_db because the database is maintained by human resources personnel.
An Architect needs to create a read-only role for certain employees working in the human resources department.
Which permission sets must be granted to this role?

    A. USAGE on database hr_db, USAGE on all schemas in database hr_db, SELECT on all tables in database hr_db

    B. USAGE on database hr_db, SELECT on all schemas in database hr_db, SELECT on all tables in database hr_db

    C. MODIFY on database hr_db, USAGE on all schemas in database hr_db, USAGE on all tables in database hr_db

    D. USAGE on database hr_db, USAGE on all schemas in database hr_db, REFERENCES on all tables in database hr_db

An Architect runs the following SQL query:

```
SELECT
    METADATA$FILENAME,
    METADATA$FILE_ROW_NUMBER
FROM @FILEROWS/Food_Reviews.csv
        (file_format=CSV_N)
```

How can this query be interpreted?

    A. FILEROWS is a stage. FILE_ROW_NUMBER is line number in file.

    B. FILEROWS is the table. FILE_ROW_NUMBER is the line number in the table.

    C. FILEROWS is a file. FILE_ROW_NUMBER is the file format location.

    D. FILERONS is the file format location. FILE_ROW_NUMBER is a stage.

An Architect entered the following commands in sequence:

```
CREATE DATABASE SANDBOX;
CREATE ROLE INTERN;
CREATE TABLE SANDBOX.PUBLIC.AGENDA (ID INT, ITEMS STRING);
GRANT SELECT ON ALL TABLES IN SCHEMA SANDBOX.PUBLIC TO ROLE INTERN;
GRANT ROLE INTERN TO USER USER1;
```

USER1 cannot find the table.
Which of the following commands does the Architect need to run for USER1 to find the tables using the Principle of Least Privilege? (Choose two.)

    A. GRANT ROLE PUBLIC TO ROLE INTERN;

    B. GRANT USAGE ON DATABASE SANDBOX TO ROLE INTERN;

    C. GRANT USAGE ON SCHEMA SANDBOX.PUBLIC TO ROLE INTERN;

    D. GRANT OWNERSHIP ON DATABASE SANDBOX TO USER INTERN;

    E. GRANT ALL PRIVILEGES ON DATABASE SANDBOX TO ROLE INTERN;

A DevOps team has a requirement for recovery of staging tables used in a complex set of data pipelines. The staging tables are all located in the same staging schema. One of the requirements is to have online recovery of data on a rolling 7-day basis.
After setting up the DATA_RETENTION_TIME_IN_DAYS at the database level, certain tables remain unrecoverable past 1 day.
What would cause this to occur? (Choose two.)

A. The staging schema has not been setup for MANAGED ACCESS.

B. The DATA_RETENTION_TIME_IN_DAYS for the staging schema has been set to 1 day.

C. The tables exceed the 1 TB limit for data recovery.

D. The staging tables are of the TRANSIENT type.

E. The DevOps role should be granted ALLOW_RECOVERY privilege on the staging schema.

Consider the following COPY command which is loading data with CSV format into a Snowflake table from an internal stage through a data transformation query.

```
copy into home_sales(city, zip, sale_date, price)
from (select t.$1, t.$2, t.$6, t.$7 from @mystage/sales.csv.qz t)
file_format =
(
format_name = mycsvformat
empty_field_as_null = true
field_optionally_enclosed_by = '"'
)
validation_mode - return_all_errors
;
```

This command results in the following error:
SQL compilation error: invalid parameter 'validation_mode'
Assuming the syntax is correct, what is the cause of this error?

A. The VALIDATION_MODE parameter supports COPY statements that load data from external stages only.

B. The VALIDATION_MODE parameter does not support COPY statements with CSV file formats.

C. The VALIDATION_MODE parameter does not support COPY statements that transform data during a load.

D. The value return_all_errors of the option VALIDATION_MODE is causing a compilation error.

A company is storing large numbers of small JSON files (ranging from 1-4 bytes) that are received from IoT devices and sent to a cloud provider.
In any given hour, 100,000 files are added to the cloud provider.
What is the MOST cost-effective way to bring this data into a Snowflake table?

A. An external table

B. A pipe

C. A stream

D. A copy command at regular intervals

A company has a Snowflake account named ACCOUNTA in AWS us-east-1 region. The company stores its marketing data in a Snowflake database named MARKET_DB. One of the company's business partners has an account named PARTNERB in Azure East US 2 region. For marketing purposes the company has agreed to share the database MARKET_DB with the partner account.

Which of the following steps MUST be performed for the account PARTNERB to consume data from the MARKET_DB database?

A. Create a new account (called AZABC123) in Azure East US 2 region. From account ACCOUNTA create a share of database MARKET_DB, create a new database out of this share locally in AWS us-east-1 region, and replicate this new database to AZABC123 account. Then set up data sharing to the PARTNERB account.

B. From account ACCOUNTA create a share of database MARKET_DB, and create a new database out of this share locally in AWS us-east-1 region. Then make this database the provider and share it with the PARTNERB account.

C. Create a new account (called AZABC123) in Azure East US 2 region. From account ACCOUNTA replicate the database MARKET_DB to AZABC123 and from this account set up the data sharing to the PARTNERB account.

D. Create a share of database MARKET_DB, and create a new database out of this share locally in AWS us-east-1 region. Then replicate this database to the partner's account PARTNERB.

What does a Snowflake Architect need to consider when implementing a Snowflake Connector for Kafka?

A. Every Kafka message is in JSON or Avro format.

B. The default retention time for Kafka topics is 14 days.

C. The Kafka connector supports key pair authentication, OAUTH, and basic authentication (for example, username and password).

D. The Kafka connector will create one table and one pipe to ingest data for each topic. If the connector cannot create the table or the pipe it will result in an exception.

A group of Data Analysts have been granted the role ANALYST_ROLE. They need a Snowflake database where they can create and modify tables, views, and other objects to load with their own data. The Analysts should not have the ability to give other Snowflake users outside of their role access to this data.

How should these requirements be met?

A. Grant ANALYST_ROLE OWNERSHIP on the database, but make sure that ANALYST_ROLE does not have the MANAGE GRANTS privilege on the account.

B. Grant SYSADMIN OWNERSHIP of the database, but grant the create schema privilege on the database to the ANALYST_ROLE.

C. Make every schema in the database a MANAGED ACCESS schema, owned by SYSADMIN, and grant create privileges on each schema to the ANALYST_ROLE for each type of object that needs to be created.

D. Grant ANALYST_ROLE OWNERSHIP on the database, but grant the OWNERSHIP ON FUTURE [object type]s in database privilege to SYSADMIN.

What considerations need to be taken when using database cloning as a tool for data lifecycle management in a development environment? (Choose two.)

A. Any pipes in the source are not cloned.

B. Any pipes in the source referring to internal stages are not cloned.

C. Any pipes in the source referring to external stages are not cloned.

D. The clone inherits all granted privileges of all child objects in the source object, including the database.

E. The clone inherits all granted privileges of all child objects in the source object, excluding the database.

Which columns can be included in an external table schema? (Choose three.)

A. VALUE

B. METADATA$ROW_ID

C. METADATA$ISUPDATE

D. METADATA$FILENAME

E. METADATA$FILE_ROW_NUMBER

F. METADATA$EXTERNAL_TABLE_PARTITION

Which SQL ALTER command will MAXIMIZE memory and compute resources for a Snowpark stored procedure when executed on the snowpark_opt_wh warehouse?

A. alter warehouse snowpark_opt_wh set max_concurrency_level = 1;

B. alter warehouse snowpark_opt_wh set max_concurrency_level = 2;

C. alter warehouse snowpark_opt_wh set max_concurrency_level = 8;

D. alter warehouse snowpark_opt_wh set max_concurrency_level = 16;

An Architect clones a database and all of its objects, including tasks. After the cloning, the tasks stop running.

Why is this occurring?

A. Tasks cannot be cloned.

B. The objects that the tasks reference are not fully qualified.

C. Cloned tasks are suspended by default and must be manually resumed.

D. The Architect has insufficient privileges to alter tasks on the cloned database.

Question #68
Topic 1

Which technique will efficiently ingest and consume semi-structured data for Snowflake data lake workloads?

A. IDEF1X

B. Schema-on-write

C. Schema-on-read

D. Information schema

Question #69
Topic 1

Is it possible for a data provider account with a Snowflake Business Critical edition to share data with an Enterprise edition data consumer account?

A. A Business Critical account cannot be a data sharing provider to an Enterprise consumer. Any consumer accounts must also be Business Critical.

B. If a user in the provider account with role authority to CREATE or ALTER SHARE adds an Enterprise account as a consumer, it can import the share.

C. If a user in the provider account with a share owning role sets SHARE_RESTRICTIONS to False when adding an Enterprise consumer account, it can import the share.

D. If a user in the provider account with a share owning role which also has OVERRIDE SHARE RESTRICTIONS privilege SHARE_RESTRICTIONS set to False when adding an Enterprise consumer account, it can import the share.

Question #70
Topic 1

Which of the following ingestion methods can be used to load near real-time data by using the messaging services provided by a cloud provider?

A. Snowflake Connector for Kafka

B. Snowflake streams

C. Snowpipe

D. Spark

Question #71
Topic 1

An Architect is designing a file ingestion recovery solution. The project will use an internal named stage for file storage. Currently, in the case of an ingestion failure, the Operations team must manually download the failed file and check for errors.

Which downloading method should the Architect recommend that requires the LEAST amount of operational overhead?

A. Use the Snowflake Connector for Python, connect to remote storage and download the file.

B. Use the GET command in SnowSQL to retrieve the file.

C. Use the GET command in Snowsight to retrieve the file.

D. Use the Snowflake API endpoint and download the file.

A table for IOT devices that measures water usage is created. The table quickly becomes large and contains more than 2 billion rows.

```
create table water_iot (
UniqueId number,
DeviceId varchar(20),
DeviceManufacturer varchar(50)
CustomerId varchar(20),
IOT_timestamp timestamp_ntz,
City varchar(80),
Location varchar(50)
)
```

The general query patterns for the table are:

1. DeviceId, IOT_timestamp and CustomerId are frequently used in the filter predicate for the select statement
2. The columns City and DeviceManufacturer are often retrieved
3. There is often a count on UniqueId

Which field(s) should be used for the clustering key?

    A. IOT_timestamp

    B. City and DeviceManufacturer

    C. DeviceId and CustomerId

    D. UniqueId

Which Snowflake objects can be used in a data share? (Choose two.)

    A. Standard view

    B. Secure view

    C. Stored procedure

    D. External table

    E. Stream

A company has an external vendor who puts data into Google Cloud Storage. The company's Snowflake account is set up in Azure.

What would be the MOST efficient way to load data from the vendor into Snowflake?

    A. Ask the vendor to create a Snowflake account, load the data into Snowflake and create a data share.

    B. Create an external stage on Google Cloud Storage and use the external table to load the data into Snowflake.

    C. Copy the data from Google Cloud Storage to Azure Blob storage using external tools and load data from Blob storage to Snowflake.

    D. Create a Snowflake Account in the Google Cloud Platform (GCP), ingest data into this account and use data replication to move the data from GCP to Azure.

How can the Snowpipe REST API be used to keep a log of data load history?

    A. Call insertReport every 20 minutes, fetching the last 10,000 entries.

    B. Call loadHistoryScan every minute for the maximum time range.

    C. Call insertReport every 8 minutes for a 10-minute time range.

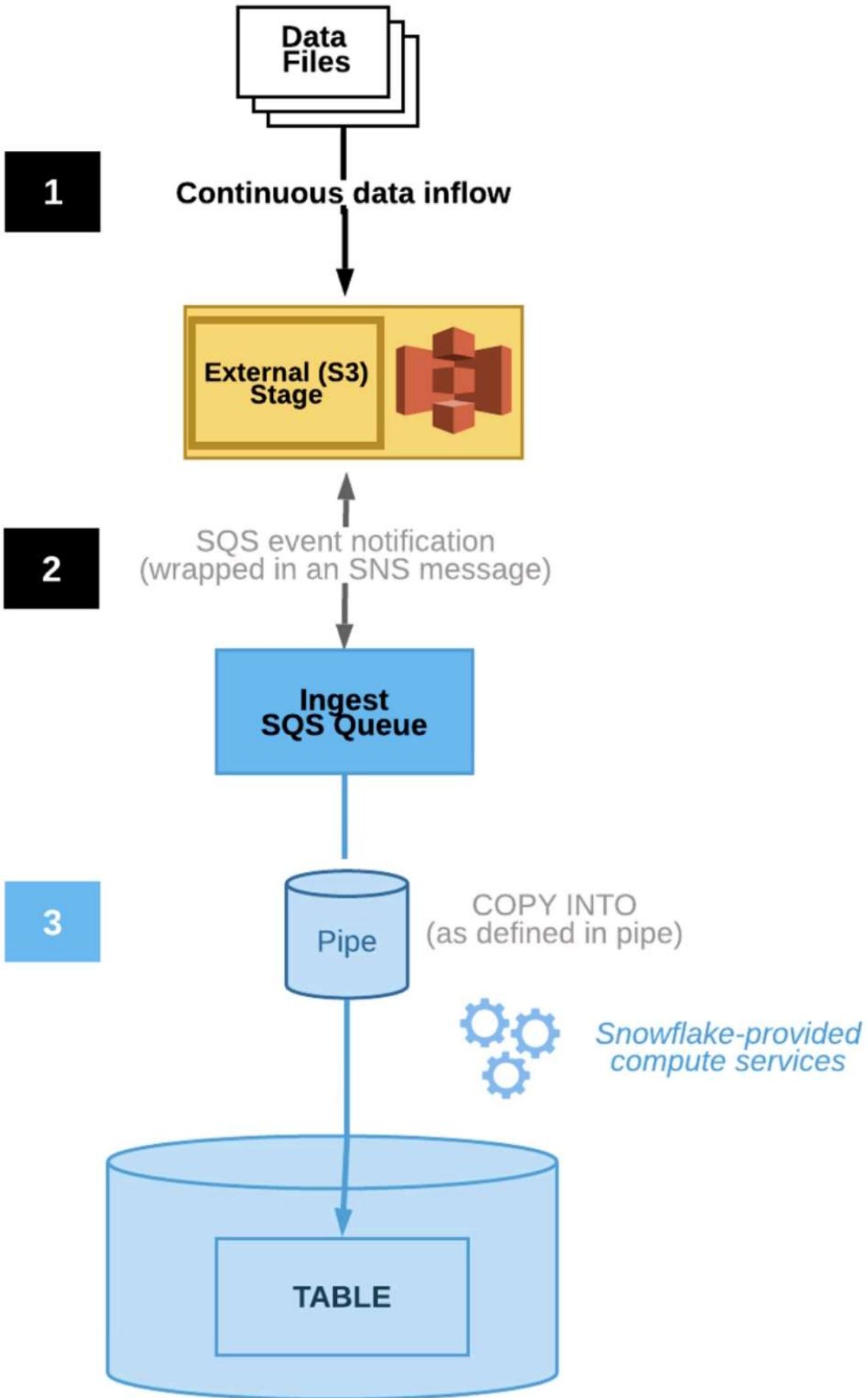   D. Call loadHistoryScan every 10 minutes for a 15-minute time range.

The diagram shows the process flow for Snowpipe auto-ingest with Amazon Simple Notification Service (SNS) with the following steps:

Step 1: Data files are loaded in a stage.

Step 2: An Amazon S3 event notification, published by SNS, informs Snowpipe — by way of Amazon Simple Queue Service (SQS) - that files are ready to load. Snowpipe copies the files into a queue.

Step 3: A Snowflake-provided virtual warehouse loads data from the queued files into the target table based on parameters defined in the specified pipe.

If an AWS Administrator accidentally deletes the SQS subscription to the SNS topic in Step 2, what will happen to the pipe that references the topic to receive event messages from Amazon

S3?

Data Files

1 Continuous data inflow

External (S3) Stage

2 SQS event notification
(wrapped in an SNS message)

Ingest SQS Queue

3 Pipe

COPY INTO
(as defined in pipe)

Snowflake-provided
compute services

TABLE

A. The pipe will continue to receive the messages as Snowflake will automatically restore the subscription to the same SNS topic and will recreate the pipe by specifying the same SNS topic name in the pipe definition.

B. The pipe will no longer be able to receive the messages and the user must wait for 24 hours from the time when the SNS topic subscription was deleted. Pipe recreation is not required as the pipe will reuse the same subscription to the existing SNS topic after 24 hours.

C. The pipe will continue to receive the messages as Snowflake will automatically restore the subscription by creating a new SNS topic. Snowflake will then recreate the pipe by specifying the new SNS topic name in the pipe definition.

D. The pipe will no longer be able to receive the messages. To restore the system immediately, the user needs to manually create a new SNS topic with a different name and then recreate the pipe by specifying the new SNS topic name in the pipe definition.

An Architect needs to meet a company requirement to ingest files from the company's AWS storage accounts into the company's Snowflake Google Cloud Platform (GCP) account.

How can the ingestion of these files into the company's Snowflake account be initiated? (Choose two.)

A. Configure the client application to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 storage.

B. Configure the client application to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 Glacier storage.

C. Create an AWS Lambda function to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 storage.

D. Configure AWS Simple Notification Service (SNS) to notify Snowpipe when new files have arrived in Amazon S3 storage.

E. Configure the client application to issue a COPY INTO command to Snowflake when new files have arrived in Amazon S3 Glacier storage.

When activating Tri-Secret Secure in a hierarchical encryption model in a Snowflake account, at what level is the customer-managed key used?
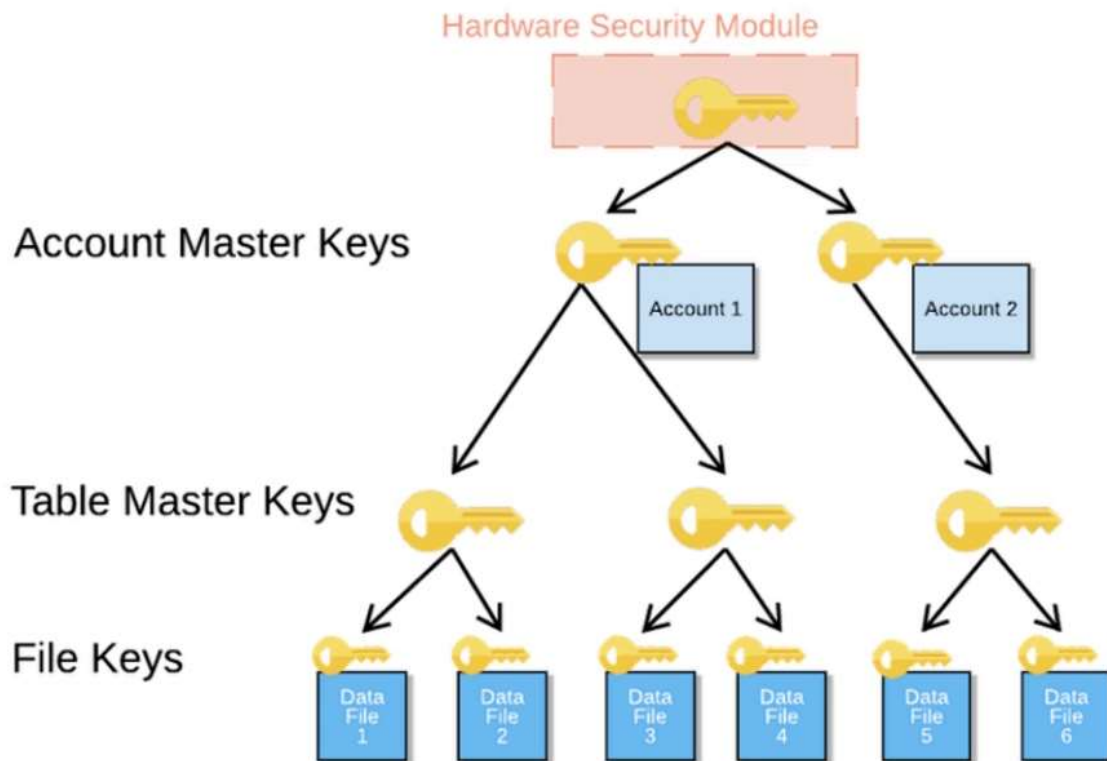
A. At the root level (HSM)

B. At the account level (AMK)

C. At the table level (TMK)

D. At the micro-partition level