1) Which statements are true of micro-partitions? (Select all that apply)
They are approximately 16MB in size
They are stored compressed only if COMPRESS=TRUE on Table
They are immutable
They are only encrypted in the Enterprise edition and above

2) Micro-partition metadata enables some operations to be completed without requiring Compute

3)***True*** or False: When data is staged to a Snowflake internal staging area using the PUT command, the data is encrypted on the client's machine'.

4) The longer the data retention period, the higher the resulting storage costs.
True or false Ans -***true***

5) warehouse- ***compute layer***?
Yes

6) GRANT SELECT, INSERT, UPDATE, DELETE ON MYTABLE TO USER JOHN; will this fail ---?***TRUE***

8) What is the minimum Snowflake edition that provides data sharing?
Standard also has the feature?

9) A single database can exist in more than one Snowflake account.
T/f ans -***false***

10) You can query the files in an External Stage directly without having to load the data into a table.
T/F
Ans -***True***

11) True or False: A single schema can exist in more than one database. Ans ***True***

12) A Virtual Warehouse can only be resized while suspended. -- T/F ans **False**

13) **True** or False: It is possible for a user to run a query against the query result cache without a requiring an active Warehouse.
14) True or **False**: Fail-safe can be disabled within a Snowflake account.
T? F ans -**false**

15) It is possible to load data into Snowflake without creating a named File Format object. ????**True**

16) **True** or False: You can define multiple columns within a clustering key on a table.,,yes

17) Snowpipe's underlying architecture utilizes the COPY command to ingest data into tables.
true

18) Which scenarios will benefit from the Snowflake cross-region replication feature?
Mostly: High availability

19) Snowflake Virtual Warehouses are part of which layer of the Snowflake architecture? --
Global or compute Ans -**compute**

20) Snowflake data can only be accessed by the Compute instances that own its micro-partitions. ? **False**

21) Snow pipe's underlying architecture utilizes the COPY command to ingest data into tables. ? **True**

22) True or **False**: Multi-Factor Authentication (MFA) in Snowflake is only supported in conjunction with Single Sign-On (SSO).

23) Snow pipe''s underlying architecture utilizes the COPY command to ingest data into tables. ?**yes**

24) True or False: It is possible to load data into Snowflake without creating a named File Format object.
Yes ? ans -**true**

25) True or **False**: A single database can exist in more than one Snowflake account.

26) It is possible for a user to run a query against the query result cache without an requiring an active Warehouse. ???**TRUE**

27) Increasing the maximum number of clusters in a Multi-Cluster Warehouse is an example of:?**scale-Out**

28) What command is used to load files into an Internal Stage within Snowflake?**Put**

29) The default Time Travel retention period is**: 1 day yes**

30) Multi-Factor Authentication (MFA) is available only to customers with Enterprise edition and above. ????? **False**

31) True or **False**: Snowpipe via REST API can only reference External Stages as source.

32)True or **False**: Loading data via the COPY command only allows error handling on a per-file basis--all files are loaded in their entirety or marked as error, there are no partial loads.

33) **True** or False: It is possible to unload structured data to semi-structured formats such as JSON and Parquet.
(ORC, XML, and AVRO don’t supporting for unloading)

34) True or **False**: Every object is owned by one and only one role; therefore a user must be using that role to access the object. ??

35) **True** or False: It is possible to set a user's default role to a role they have not been granted access to.

36) True or **False**: A single schema can exist in more than one database.

37) True or **False**: The user has to specify which cluster a query will run on in a multi-cluster Warehouse.

38) When specifying a table''s clustering key, it is recommended you use a column with high cardinality. ??? T/F ??? **False**

39) True or **False**: The user has to specify which cluster a query will run on in a multi-cluster Warehouse.

40) True or **False**: Snowflake charges additional fees to Data Providers for each Share they create.

41) A single schema can exist in more than one database. ??? T/**False** ???

43) **True** or False: Micro-partition metadata enables some operations to be completed without requiring Compute.

45) Snowflake's Global Services Layer gathers and maintains statistics on all columns in all micro-partitions. ??? **True** /F ??

46)True or **False**: Loading data via the COPY command only allows error handling on a per-file basis--all files are loaded in their entirety or marked as error, there are no partial loads.

48) True or **False**: Users are able to see the result sets of queries executed by other users that share their same role.

49) AWS Private Link provides a secure connection from the Customer's on-premise data center to the Snowflake Virtual Private Cloud ??? **True** /F ??

50) When Snowflake is configured to use Single Sign-On (SSO), Snowflake receives the usernames and credentials from the SSO service and loads them into the customer's Snowflake account. ?????
True /F?

51) Every query should be expected to execute faster on a larger Virtual Warehouse than on a smaller Virtual Warehouse. ???? T/**False** ???

52) It is possible for a user to run a query against the query result cache without an requiring an active Warehouse. ??? **True** /F

53) Which statement best describes Snowflake tables?

It is possible to load data into Snowflake without creating a named File Format object. ???? **True**
Snowflake data can only be accessed by the Compute instances that own its micro- partitions? **false**
It is possible for a user to run a query against the query result cache without an requiring an active Warehouse. ??? **True**

54) When specifying a table's clustering key, it is recommended you use a column with high cardinality. ??? T/**False** ???

57) **True** or False: Micro-partition metadata enables some operations to be completed without requiring Compute.

--========== Screen shot

1) Which of the following terms best describes snowflake's database architecture?
columnar shared nothing
shared disk
multi-cluster,shared data
cloud-native shared memory

2) T or **False** :- it is best practice to define a clustering key on every table

3) **True** or F :- you can define multiple columns within a clustering key on a table

4) which of the following statements are true?
a**.the copy command must refernce a stage or cloud storage location**
b.a named file format is optinal when using the copy command
c.a file format option used in the copy command will override the one in stage
d.a file format object must be defined when creating a stage

5) snowflake collects which of the following metadata on each of its micro partitions?
a)range of values
b)numberof distince values
c)MIN/MAX values
d)NULL Count

6) which of the following commands sets the Virtual warehouse for a session?
a)COPY WAREHOUSE FROM <<config file>>;
b)SET WAREHOUSE FROM =<<warehouse name>>;
c)USE WAREHOUSE<<warehouse name>>;
d)USE VIRTUAL_WAREHOUSE<<warehouse name>>;

7) which of the following are valid approaches to loading data into a snowflake table?
a)bulk copy from a external storage
b)continous load using snowpipe REST API

c**)the snowflake web interface (UI)data loading wizard**
d) bulk copy from intenal storage

8) The PUT command
a)Automatically creates a file format object
b)Automatically uses the last storage created
c**)automatically compresses files using GZIP**
d)Automatically encrypts files

9) **True** or F:- pipes can be suspended and resumed

10) **True** or F:- An Active warehouse is required to run a COPY INTO Statement

11)when should you consider disabling auto suspend for a virtual warehouse?
a)when users will be using compute at different time throughout a 24/7 period
b)when managing a steady workload
c)when the compute must be available with no delay or lag time
d)when you don't want to have to manually turn on the warehouse each time as user needs it

12)which of the following statements are true of VALIDATION_MODE in snowflake?
a)the validation_mode option is used when creating the internal storage
b)validation_mode = return_all_errors is a parameter of the copy command
c)the Validation_mode option will validate data to be loaded by the copy statement while completing the load and will retrun the rows that could not be loaded without error
d)the validation_mode option will validate data to be loaded byt the copy statementwithout completing the load and will retrun possible errors

14)t or **false** the copy command must specify a file format in order to execute

15)the information schema and account usage share provides storage information for which of the following objects?
a)users
b)tables
c)database
d)internal stages

16)A client uses a PUT command to upload a new file to an internal stage.at what point is the data encrypted?
a)when it reaches the virutal warehouse
b)when it gets micro-partitioned and stored in cloud storage
c)on the client's machine before being sent
d)clients must encrypts files before sending to snowflake

17)**true** ot f the follwing command will fail:-

GRANT SELECT,INSERT,UPDATE,DELETE, ON MYTABLE TO USERJOHN;

18)T Or **False** when active, a pipe requires a decdicated virtual warehouse to execute.

19) which of the following are options when creating a virtual warehouse?
a)auto drop
b)auto resize
c)auto resume
d)auto suspend

23)which of the following are use cases that fit well in snowflake?
a)data lake
b)enterprise warehouse
c)online transaction processing
d)datamart consolidation

26)T or **False**:- multi factor authentication (MFA) in snowflake is only supported in conjunction with single sign-on(SSO).

27)which of the following commands are not blocking operations?
a)UPDATE
b)INSERT
c)MERGE
d)COPY

28)T or **False** snowflake data warehouse was built from the ground up for the cloud in lieu of using on the existing database or a platform like hadoop as a base.

29)the default time travel rentention period is :
a)1 day
b)7 days
c)45 days
d)90 days

30)t or **false** :- a virtual warehouse can only be resized while suspended

31)**true** or f :- reader accounts incur no additional storage costs to data provider since they are simply reading the shared data without making changes

32)t or **false** :- loading data via the copy command only allows error handling on a per file basis -- all files are loaded in their entirety or marked as error, there are no partial loads.

33)**true** or f:- snowpipe's underlying architecture utilizes the copy command to ingest data into tables.

34)what is the most performant file format for loading?
a)CSV(unzipped)
b)parquet
c)CSV(gZipped)
d)ORC

35)what is the minimum snowflake edition that provides the multi cluster warehouses and up to 90 days of timetravel?
a)standard
b)premier
c)enterprise
d)enterprise for sensitive data

36)t or **false** :- loading data into snowflake requires that source data files be no larger than 16 MB

37)which of the following are options for the ON_ERROR property when using the copy command?
a)continue
b)abort_statement
c)skip_file
d)stop_statement

38)which of the following options when creating a virutal warehouse?
a)auto suspend
b)auto resume
c)Local SSD Size
d)User Count

39)t or **false** :- snowflake enforces unique, primary key and foreign key constraints during DML operations.

40)which of the following languages can be used to implement snowflake user defined functions(UDFS)
a)Java
b)Javascript
c)SQL
d)python

41)storage is calculated based on data in which of the following statuses?
a)Active
b)time travel
c)fail safe
d)purged

42) which of the following are examples of operations that require a virtual warehouse to complete, assuming no queries have been executed previously?
a)MIN(<<column value>>)

b)COPY
c)SUM(<<column value>>)
d)UPDATE

43) t or **false** :- A snowflake account is charged for data stored in both internal and external stages

44) which of the following are main sections of the top navigation of the snowflake web interface(UI)?
a)Databases
b)tables
c)warehouses
d)worksheets

45)t or **false**:- multi factor authentication (MFA) is available only to customers with enterprise edition and above.

46)what are the three things customers want most from their enterprise data warehouse solution?
a)on premise availability
b)simplicity
c)open source based
d)concurrency
e)performance

47)when scaling up virtual warehouses by increasing virtual warehouse
t-shirt size, you are primarily scaling for improved.
a)concurrency
b)Performance

48)which of the following items does the global services layer manage?
a)user authentication
b)metadata
c)query compilation and optimization
d)external blob storage
e)data security

49)**True** or f:- you can query the files in an external stage directly without having to load the data into the table?
50)snowsql can unload query results to a local file system in the following file formats:-
a)JSON
b)parquet
c)orc
d)CSV and any other delimted formats

51)which of the following statements are true of transient tables? Transient tables:-
a)Are tied to a snowflake session

b)have a fail safe period of 7 days
c)have a maximum time travel data retention period of 1 day
d)are visible to all users with roles that have permissions to access that table

52)query results are stored in the result cache for how long after they are last accessed, assuming no data changes have occurred?
a)1 hour
b)3 hours
c)12 hours
d)24 hours

53)**true** or f:- once created, a micro partition will never be changed

54)t or **false** :- accountadmins are able to see the results sets of queries executed by all users within a snowflake account

55)which snowflake object enables loading data from files as soon as they are available in a cloud storage location?
a)pipe
b)external storage
c)file format
d)VARIANT

56)when a pipe is recreated using the CREATE OR REPLACE PIPE command?
a)the pipe load history is reset
b)the Refresh parameter is set to true
c)previoulsy loaded files will be ignored
d)All of the above

57)which transformations are available when using the COPY into command to load data files into snowflake from a stage?
a)filters
b)aggregates
c)column data type conversion
d)column concatenation

59)snowflake provides a mechanism for its customers to override its natural clustering algorithms. this method is?
a)Micro Partitions
b)clustering keys
c)Key partitions
d)clustered partitions

60)what is the most granular object that the time travel retention period can be defined on?

a)Account
b)database
c)schema
d)table

61) fail-safe is unavailable on which table types?
a)temporary
b)transient
c)provisional
d)permanent

62)**true** or f :- although a user can be granted multiple roles, each session has a single current role which dertermines current privileges?

63)in which layer of its architecture does snowflake store its metdata statistics?
a)storage layer
b)compute layer
c)database layer
d)global service layer

63) which interface can be used to create and/or manage virtual warehouse?
a)the snowflake web interface(UI)
b)SQL commands
c)data integartion tools
d)all of the above

64)what parameter controls if the virtual warehouse starts immediately after the create warehouse statement?
A)INITIALLY_SUSPENDED = TRUE/FALSE
B)START_AFTER_CREATE = TRUE/FALSE
C)START_TIME = 60//(Seconds from now)
D)START_TIME = CURRENT_DATE()

65)which formats are supported for unloading data from snowflake?
a)Delimited(CSV,TSV,etc).
b)Avro
c)JSON
d)ORC

66)true or **false** :-when active, a pipe requires a dedicated virtual warehouse to execute

67)to run a multi-cluster warehouse in auto scale mode, a user would:-
a)configure the maximum clusters setting to "Auto Scale"
b)Set the warehouse type to Auto

c)Set the minimum clusters and Maximum clusters settings to the same value
d)set the minimum cluster and maximum clusters settings to the different value

68)what is share?
a)a named, first class snowflake object that encapsulates all the information required to share objects within a database
b)the name of the database created by the data provider and shared with data consumers
c)a virtual warehouse that is used by data consumers to query a data provider's shared database
d)the name of the snowflake account that has shared objects with data consumers

70) which of the following statement is true of snowflake micro-partitioning? Micro-Partitioning?
a)has been known to introduce data skew
b)requires a partitioning schema to be defined up front
c)is transparently completed using the ordering that occurs when the data is inserted/loaded
d)can be disabled within a snowflake account

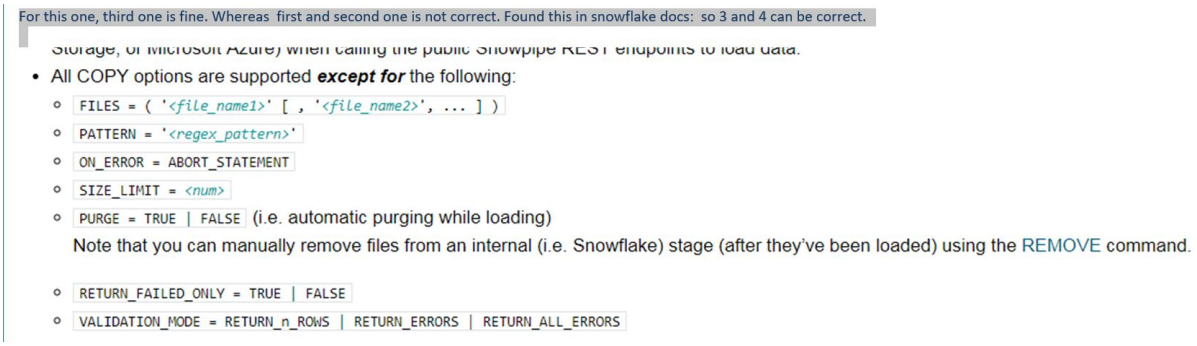
71)t or **False**:- snowflake requires that a single run its ETL , reporting, and data science workloads on the same virtual warehouse

72)what happens when a data providers revokes privileges to a share on a object in their source database?
a)the object immediately becomes unavailable for all data consumers
b)any additional data arriving after this point in time will not be visible to data consumer
c)the data consumer stop seeing the data updates and become responsible for storage charges in a object
d)a static copy of the object at the time the privilege was revoked is created is created in the data consumers accounts

73)t or **false** :- a single database can exist in more than one snowflake account?

74)what command is used to delete files from a stage?
a)Truncate
b)delete
c)remove
d)purge

75)which of the following is true of snowpipe via REST API ?
a)you can only use it on internal stages
b)All copy into options are available during pipe creation
c)Snowflake automatically manages that compute required to execute the pipe's copy into commands
d)Snowpipe keeps track of which files it has loaded


For this one, third one is fine. Whereas first and second one is not correct. Found this in snowflake docs: so 3 and 4 can be correct.

Storage, or Microsoft Azure) when calling the public Snowpipe REST endpoints to load data.

- All COPY options are supported ***except for*** the following:
  - `FILES = ( '<file_name1>' [ , '<file_name2>', ... ] )`
  - `PATTERN = '<regex_pattern>'`
  - `ON_ERROR = ABORT_STATEMENT`
  - `SIZE_LIMIT = <num>`
  - `PURGE = TRUE | FALSE` (i.e. automatic purging while loading)
    Note that you can manually remove files from an internal (i.e. Snowflake) stage (after they've been loaded) using the REMOVE command.
  - `RETURN_FAILED_ONLY = TRUE | FALSE`
  - `VALIDATION_MODE = RETURN_n_ROWS | RETURN_ERRORS | RETURN_ALL_ERRORS`

76)which of the following commands are not blocking operations?
a)update
b)**insert**
c)merge
d)copy

77)the query history in the snowflake web interface (UI) is kept for approximately?
a)60 mins
b)24 hrs
c)14 days
d)30 days
e)1 year

79)The PUT command is used to place local files into:-
a)an external stage
b)a stage zone
c)a transient stage
d)an internal stage

80) **true** or f :- pipes can be suspended and resumed

81)when loading data into snowflake, the copy commands supports:-
a)Joins
b)filters
c)data type conversions
d)column reordering
e)aggregates

84)the following factors affect the data load rates?
a)**physical location on the stage**
b)RAM on the Virtual warehouse
c)GZip compression efficiency
d)Thread Size

85)which of the following roles is recommended to be used to create and manage users and roles?
a)SYSADMIN
b)SECURITYADMIN
c)PUBLIC
D)ACCOUNTADMIN

-==========

1.True or False: IN snowflake, Data unloading can be done only to CSV, JSON in compressed format
a. **False**

2.Minimum snowflake edition that supports data sharing
a. Standard
b. Premium
c. Enterprise
d. ESD
e. VPS

3.Which of the following are valid data sharing accounts
a. Consumer account
b. Reader account
c. Provider account
d. None

4.SSO can be done in how many ways. Select all applicable:
a. Web UI
b. SnowSQL
c. Python
d. ODBC
e. JDBC
f. Etc..

5A third party tool which can support JDBC Or ODBC but doesn't have snowflake drivers. Can it still connect to Snowflake
a. **True**
b. False

6.Min level of account required for MFA
a. **Standard**
b. Premium
c. Enterprise
d. ESD
e. VPS

7.**True** or False: When a user creates a role, The user owns that role until he transfers the role

8.When a user creates an object with a role, And that role is dropped, What happens?

a. **Answer: Role which dropped the original role will own the object**

9. True or **False**: Data share can happen between any edition without Snowflake support (b/w E & ESD editions)

10.True or False: When a reader account is created, No extra charges will be applied to provider

a. True
b. **False (Answer)**

11.Best practice: When a role creates an object, Best practice to give grants on the object to

a. ACCOUNT ADMIN
b. **SYSADMIN (Answer)**
c. SECURITYADMIN

12.What are the activities to be done by the customer, when an on-premise installation is done

a. **Physical security**
b. Applying releases and patches
c. Managing partitions
d. Etc.. (pl check) Answer is to check all options

13.Regarding Snowflake releases, Select all that apply

a. Weekly once
b. During the release transparently user queries will be redirected to new version

14.Clustering information can be checked with below commands

a. **SYSTEM$CLUSTERING_INFORMATION**
b. SYSTEM$CLUSTERING_DEPTH

15.FLATTEN is used when loading

a. Structured data
b. **Semi structured data (Answer)**
c. None
d. Both

16.While unloading, file format can be

a**. CSV – yes**
b. XML – No
c. Parquet – Yes
d. ORC – No
e. Avro – No

f. **Json – Yes**

17.While uploading , to convert semi structured null values to structured null values, What option will be used

a. **STRIP_TO_NULL** is the answer

18.While loading data what are the factors that impacts rate : below are the answers

a. Location of the stage

b. Thread size

19.Ideal file size should be 16MB because micro partition size is 16MB

a. True

b. **False** (Answer)

20.What are all the RECOMMENDATIONS when new user is created, Below are all answers

a. **Password reset**

b. Giving default role

21.Data replication use cases

a. Catastrophy

b. For Migration activities

c. Data sharing

d. Schema sharing (wrong)

22.True or False: provisioning a 4XL warehouse can take more time than XS WH ? **True**

23.Can a reader account incurs charges to data providers : **Storage no and compute Yes.**

24.Can a reader account extract data to use outside snowflake : **True**

25.Default time-period for auto-suspend for WH through WEB UI ? **10 mins**

26.Each worksheet in WEB UI can have multiple warehouses, schemas, databases : **True**

27. Does a snow-pipe requires an active warehouse? **False**

28. Which maintains the credits for snowflake account : **Resource Monitor**

32. From where we can see account billing ? Account : **Billing and Usage**

29. When we can resize the WH ? ? **at any time**

30. Which DML is not supported in Snowflake ? **Upsert**

31. Facts about Multi clusters :

a. Can scale up and down

b. Can auto suspend and auto resume

33.What are the factors to count credits will be accounted on WH Usage

a. Warehouse Size and Num. of Clusters used to process

34.Small WH will have 2 clusters. Then how many clusters will M have ? ? **4**

35.How billing is done in pipe ? **per second/per core granularity**

36. At the client local area only file will get encrypted. -- **TRUE**

37.Which of these DML doesn't require a WH

a. Insert

b. Update

c. Merge

d. **Drop** ? because it's an metadata operation. (Delete also)

38.WH will fit into which layer of architecture ? **Compute**

39.When a query is executed a query result cache, will it require an active WH ? ? **False**

40.Account admin can see all the query results executed by the other roles ? **False**

41.Result cache can be accessed by other users of same role ? **False**

42.How can we view the storage of a tables :

a. SHOW TABLES and INFORMATION_SCHEMA : TABLE_STORAGE_METRICS

43.While migrating from different databases to snowflake, what will be migrated > -? **Tables, schemas and pipes. (No indexes can be migrated)**

44.Minimum Edition to maintain secure data ? **ESD**

45.Default time travel for transient tables ? **1 day.**

46.Fail safe can be disabled -? **False**

47.Does stages have time travel and failsafe ? **False**

48.Snowflake Billing will be done for both internal and external stages ? **False**.

49.What objects can be shared ? **Tables, Schemas , databases, file formats.**

50.Which operations can be done while loading data from stage ? ? **column ordering, Cast(data type conversion), column omission**

51.Can regular expression be used in copy command ? (explore)

52.User has Account_admin/sysadmin roles , recommended practices

a. Default role is sys admin (which is min)
b. Only login thru account_admin for some critical options.
c. Other two options needed to be confirmed

--==========

Factors that impact Unit Cost
Differences in unit costs for credits and data storage are calculated by Snowflake Region and not by cloud platform.
another factor that impacts unit costs is whether your Snowflake account is On Demand or Capacity.

Pay for compute and query.

snowflake supports loading data from files staged in any of the following locations, regardless of the cloud platform for your Snowflake account:
Internal (i.e. Snowflake) stages
Amazon S3
Microsoft Azure Blob storage

Limitations of Azure network
Virtual Private Snowflake (VPS) is not currently offered for Snowflake accounts hosted on the Azure cloud platform.
No support for secure connectivity to customer-owned virtual networks (similar to AWS PrivateLink).
No support for accessing an external AWS S3 stage using policies attached to an IAM role.

Regions :

Note that regions do not limit user access to Snowflake; they only dictate the geographic location where data is stored and compute resources are provisioned.
Each Snowflake account is located in a single region (i.e. multi-region accounts are not supported).
In addition, Snowflake does not yet support accessing or sharing data between regions. If you wish to use Snowflake across multiple regions, you must maintain a separate Snowflake account in each region.

If your account is hosted on AWS and latency is a concern, you should choose the available region with the closest geographic proximity to your end users.

Disaster recovery of modified/deleted data (for 7 days beyond Time Travel

Snowflake Time Travel (1 day)
Access Control in Snowflake
Snowflake provides granular control over access to objects — who can access what objects, what operations can be performed on those objects, and who can create or alter access control policies.

Discretionary Access Control (DAC): Each object has an owner, who can in turn grant access to that object.
Role-based Access Control (RBAC):

Company History :
Founded 2012

SNOWFLAKE VS. TRADITIONAL ARCHITECTURES
Snowflake uses the cloud to enable elasticity Varies Traditional databases are inflexible

- Complete sql database
- Zero management
- All of your data
- All of your users
- Pay for what you use
- Instant live data sharing.
- Structured and semi
- Pay for only
- what you use with no
- overprovisioning
- Eliminate
- overbuy
- Scale compute
- up and down, transparently
- and automatically
- No need for capacity
- planning, make capacity
- decisions on the fly
- -

structured data
(JSON, XML, Avro)
Centralized storage of data, accessible
by any user and application
Multi
-
petabyte scale

SCALE & CONCURRENCY

Up to 200x faster than solutions
not built for the cloud
Maintains a consistent SLA
–
resources grow and shrink automatically
Loading data does not impact
query performance
Multiple groups access data
at the same time with no
performance degradation
Supports an unlimited number
of simultaneous users
Diverse applications

ZERO MANAGEMENT

- Load data and run queries,
- we do all the rest
- Zero infrastructure and admin costs
- Secure and highly available
- Fully managed with no knobs
- or tuning required
- No indexes, distribution keys,
- partitioning, or vacuuming

INSTANT, LIVE DATA SHARING

- Share with unlimited number
- of consumers, without duplicating storage.
- Data consumers immediately see all updates
- Consumers can immediately start querying
- Reader Accounts enable sharing with non

Snowflake customers

What customers want
Perf
Concurr
Simplicity
200 X faster
1/10th cost

Avro,parquet,xml

Terms & Conditions
THE IMPACT OF DATABASE PROBLEMS AT ?
Slow queries
With massive concurrency
across dozens of business
units, database performance
was strained to the maximum,
slowing analysis times and
limiting data to a few.

Couldn't meet the analytics needs of the business System
Data scientists were unable to
scale the service to larger
queries, significantly inhibiting
their ability to find new ways
to deliver personalized
experiences

High Availability

Without a way to failover
between regions, the existing
solution lacked a critical
capability in delivering and
protecting data

Comprehensive Security
Without the built
-in security
they needed, ? had to
painstakingly secure their
data warehouse from scratch.
Changes or updates had
significant implication

Unlimited storage because of s3

It is not for high frequency changing data but analytics warehouse.

Services :

- Management
- Optimization
- Security
- Transaction

- Metadata

Transaction recommitted.

Table level lock and who starts first gets privilege.
Within snowflake https transaction.
AWS is more secure.

Questions :
How many editions of snowflake are supported : 5

How many regions is Snowflake deployed in today? 11

Where will you download snowsql from (help ---downloads)

Which tools are supported All of the above
Biggest connector : jdbc
Micropartitions are immutable
Micro partition column level (min/max, distinct,null) --- for any query using this snowfalke does not use warehouse.
Metadata only operations (deleting all rows from a table)
Truncate
Query pruning :
At a micro partition level we store the min max value of every column.
Clustered by determines the distribution of data
Data Clustering :
Clustering is sorting and ordering
Typically done naturally by date.(time series analysis)
Depending on the query you will need clustering
As data is loaded, Snowflake co-locates column data with the same values in the same micro-partition, if possible.
If you know clustering is needed use order by while loading . This reduces the cost.
How to choose clustering keys : Order of clustered columns is imp
Cluster by datetime is bad
When to turn on clustering
Table over 1 TB
More partitions
With Alter new virtual cluster gets created.
Hoe to choose cluster key
Low cardinality
Time dimention
Join columns
More keys in cluster key is bad (5)
Clustering on semi structured data : sub object can be used
One table can have only one cluster.

foundation database --- snowflake used it internally)
How to get 90 days of time travel (enterprise edition)
Only way to eliminate fail safe is to use transient tables.
When the retention period ends for an object, the historical data is purged from Snowflake
Purged data is not available for querying.
Purged objects can no longer be cloned.
Purged objects can no longer be restore

How to reduce CDP
Create these tables as transient with zero Time Travel retention
Copy these tables on a periodic basis into a permanent table (full backup)
Once the new backup has been created, delete the old one

Time Travel cannot be cloned.

1 TB database can be cloned in 1 min.
When a Clone of a Table is created, the original Table's data is physically copied.

From
<https://mediaplayer.mindtickle.com/pdfViewer/?width=547&height=401&showPreviousViewOnLoad=false&reset_to_page=1>

Data is never stored in cluster.it gets stored in global storage.

Compute :
Each virtual machine has 8 V cpu/threads.

The access to a compute cluster is controlled by roles.

8 concurrent users
60 secs and then per second billing.

If query runs more than 60 secs let it scale and once it executes turn it off.
Scale out -- concurrency , Scale up - performance
Cost and performance :
Warehouse size, manual vs automated
True or False: When a Resource Monitor shuts down a Warehouse, the Customer can issue an ALTER WAREHOUSE command and it will manually resume.

Result set cache sits on global storage and lives for 24 hours and if the cache accessed in the 23 rd hour it lives for another 24 hours.
Tableau optimization .

Query history stays for 14 days / 1 year. The cache is always the result set of the role.Set of features that help protect data stored in Snowflake against human error, malicious acts, and software or hardware failure

Is the query id,timetravel accessible to other users ? Only if the role is the same.

The database snowflake has query history .
Every DB has a information schema (metadata)
Time travel is sliding window on micro partitions .each micro partition Is timestamped on date of creation.

Cloning cannot be done across accounts. Sharing can be done across accounts.

Timetravel cannot be disabled at account level.
Time travel can be at table or db or schema
The longer the data retention period more is the cost.

Encryption :
AES -256

SOC 2 Type II certified with support for PHI (HIPAA) and PCI DSS

Only enterprise has PHI (HIPAA) and PCI DSS every one has SOC2 Type2

Tri Secret secure : only available for enterprise sensitive data.
CMK, KMS
Third key : password to snowflake account.
If cmk is unavailable for 10 mins then snowflake looses access.

Copy command source, destination, fileformat
Default file format is .csv
File format is in the schema.

Snow pipe : through rest API
Small files , infrequent less than 100 MB --- snowpipe
Stage does not have failsafe and timetravel.
It is possible to load directly from a Cloud Storage location without creating a named Stage(Y)
File that loads faster csv gzip (3 times faster)
10 MB to 100 MB (File Size)
File locAtions (Many locations , each with few files)
Designed for Bulk Loading not for large transaction.
Concurrency --- Partition lock, Table lock.
Single updfste statement is better.

Copy Command :

NOT SUPPORTED: Joins, filters, aggregations
Can include SEQUENCE columns, current_timestamp(), or other column functions during data load
Execute copy in validation mode.
Best practices for JSON
Validation
Null handling
Staging
Extraction
Skip the file or number of rows.
Strip_null_value

Talking to metadata does not need compute.
Snow pipe si near real time.
Pipe is wrapper for copy command.
Rest Api , Auto ingest.
Snowflake has
Recommended size for snowpipe 1 to 10 mb
If something is in Public preview do not put it in production.
Auto ingest needs external storage.
Notification channel is on auto ingest.

Copy command looks for hash name and pipe looks for a name.
Max file size snowflake can output to is 5 GB
Variant datatype is composed of Variant , Array , Object
Normalization best practices:
Variant (flexible access, no dupliczte ,changes to str imply changes to query and view)
What are integers stored as in json documents (strings)
Material views better pruning.
Keep variant and MV in the same table

Data sharing :
Pay for compute share storage.

3 types of account data providers
Data consumers
Reader account
There is no limit on share.
Access is provided to selected objects.

Consumers can share unlimited number of shares.
Consumers can make only 1 DB per share.
No access to time travel
Cannot clone.

Reader Account :
Data provider pays for it .
Cost effective way to share data
Consumer cannot write data to that account
Cost of reader account will be same as the main provider account.

When editing a share
Add or remove tables and secure views
Add/ revoke access for data consumers.

Shares cannot span multiple data bases (F)
If a table is dropped and recreated it needs to be shared again.

Secure view :
Secure UDF's.
ESD

1 ec2 x small

Aws customer can share data with another AWS cutomer but not with Azure in that across regions.
Pod is effectively in a region.

Fail over is through replication.

Standard cannot scale out . Multi cluster can
Auto scale ---- scale out

Secure View --- definition is visible to all the roles.
Materialized view ---- only chance when the data changes

Even if we use result set cache
NOTE: compilation is STILL REQUIRED
How to find the best warehouse size (use query profiler and go from large to small & small to large)
Can there be a resource monitor on user … NO

History of queries for more than 14 days snowflake (account usage)
Accessing Account usage uses a warehouse.

Information schema is not case sensitive.
Can u get the query history lookinmg at query profile no

Practice questions :

1.when does a virtual warehouse start consuming (active)
2.why good to keep it on (cache will be gone)
3. TP (N)
4. UI or Script
5. Best arch (Multicluster shared data)
6. Concurrency(work load separation,elasticity) , performance,
7. Always load JSON into varient dataype.
8. Read (JSON, PARQUET, AVRO, ORC) Unload (ORC not possible)
9. Data sharing -- compute cost (provider)
- With snowflake customer (compute consumer pays , storage provider)
- With reader account compute provider ,storage provider
10. See data instantly
11. If access is taken away instantly the changes reflect
12. I have shared data with you and I have deleted data then it will let the query run
13. If a query is running and WH is suspended can I do it (Y)
14. How often snowflake updates (weekly on fridays)
15. Is there any downtime for users while snowfalke updates (NO)
16. When snowflake updates which reason gets it forst (all)
17. Created a WH in the UI and want to suspend it (initially suspend)
18. Created a user(assign to role) , change password, get them with multifactor authentication
19. MFA through script (yes)
20. Enable MFA through account preferences .
21. If MFA is lost we can ask admin to temporarily disable MFA.
22. How to create a Dataware house that is initially suspended (add intially suspeneded to the sql script)
23. Where does table reside (N)
24. Never update statistics
25. Micro partition gets updated (N)
26. When there is a change we pay for 2 micro partition .
27. Can range of a micro part over lap yes
28. Data 50
29. Snow pipe (1 to 10)
30. Optimum for copy (10 to 100)
31. Query min max use compute (N)
32. Get more data added to snowflake
33. Copy command use a file format (no)
34. Is defalut compressed in copy command (n)
35. Copy command ---- stage name , destination , file format
36. Source
37. Destination
38. Pipe ---- wrapper for copy command, continuous load
39. Admin can see the query not the query results.
40. Stages in snowflake (table , named, user)
41. Table stage (internal)

42. Named (internal/external)
43. User(internal)
44. Variant contain variant, array, object
45. Array contain variant
46. How to set the context (use command)
47. What users, roles, DW,resource monitor, database.
48. Every object must be owned by a role.
49. Security admin(users, roles)
50. Sysadmin(objects)
51. Insert gives access to copy.
52. A user not having access to Ware house can access metadata.
53. Role owns the object.
54. Sys admin does not have managed access
55. HIPPA (ESD, Enterprise, VPC)
56. Encryption AES 256
57. Function and sequence exist inside schema.
58. Sequence one schema accessible from another schema yes if the role permits it
59. Case sensitive and if double quote is added then it retains the name with double quote.
60. Metadata is case insensitive.
61. 3 types of views standard, secure, materialized view
62. Secure views can bypass optimizations
63. MV can speak only to one view ,Cannot use window functions.
64. Functions(Table, Secure), can be shared only if secure.
65. Pipe Wrapper over copy command for continuous loading
66. Stored procedures (Java script) can execute a sql .
67. Stored procedure runs with the privilege of the role that owns it.

Stored procedure Transaction control statements are not permitted in a stored procedure
Stored procedure is executed within a single transaction (explicitly or implicitly)
Stored procedure Can run out of memory
Stored procedure Executed through API objects

68. Stored procedure May return a value , UDF must return a value.
69. Stored procedure (DDL , DML) , UDF (NO DDL , DML)
70. UDF's (Java script, sql)

UDF 's can Return a singular scalar value or, if defined as
a table function, a set of rows

71. Sql variables can be a varient cannot be a table, only valid within a session.
72. Constraints (only enforces not null constraints)
73. String : Max legth of string by default is 16 mb
74. String : case sensitive.
75. Max row size 16mb
76. Date format is controlled by session parameters.
77. Timestamp (timestamp with local timestamp
78. timestamp with no timezone
79. timestamp)

80. Show parameters.
81. Boolean (only 0 is false rest of the numbers are 1)
82. Session context using set command

Role is an Entity to which we can give access.
At any point of time there can be only 1 role against the session.
Users
Roles can inherit other roles.
Privileges can be un granted.
User assumes role and roles are given privileges.

Snowflake DB
Account usage
Reader account usage.
SNOWFLAKE TOPOLOGY
A deployment is a "multi
-
tenant" system
?
Some resources are SHARED across multiple accounts
?
Increases peak capacity to customers (performance)
?
Achieves better overall resource utilization (cost)
?
Leads to higher bang for the buck for our customers (value)

Security is through IM role

Encryption
?
Web UI, command line client, and drivers
communicate solely over HTTPS
?
Connections encrypted using TLS 1.2 from
client through to Snowflake Service
All access controlled
?
Authentication required for all connections
?
IP whitelisting available to restrict client
communication to specific IP addresses
All access are controlled
Authentication required for all connections

?
IP whitelisting available to restrict client
communication to specific IP addresses

Compression : 5 times
Data gets saved to micro partitions 16 MB each

Comprehensive data protection

COMPREHENSIVE DATA PROTECTION
44
Protection against infrastructure failures

Protection against corruption & user errors
Long
-
term data protection

ADAPTIVE CACHING
- Metadata cached for fast access during query planning
- Active working set transparently cached on virtual warehouse SSD
- Results sets cached for reuse without requiring compute

Interfaces and connectivity :

Shares : Inbound , Outbound

Reader Account : To share data with users not part of snowflake.

Cloud Storage Layer :

Staging (S3, AWS)
Maximum size of micropartition is 16 MB'

FDN -- Flick on the edge

No indexes.
Key differentiators(timetravel,zero copy cloning)

TABLE & MICRO
-

PARTITION METADATA
13
•
Snowflake automatically collects and maintains metadata
about tables and their underlying micro
-
partitions, including:
?
Table
¦
Row count
¦
Table size (in bytes)
¦
File references and table versions
?
Micro
-
Partition Column Level:
¦
MIN/MAX values (range of values)
¦
Number of distinct values
¦
NULL count
•
Essentially a zone map
USE ROLE ANALYTICS_USER;
USE DATABASE SNOWFLAKE_SAMPLE_DATA;
USE SCHEMA TPCH_SF100;
SELECT
MIN(L_SHIPDATE),
MAX(L_SHIPDATE)
FROM LINEITEM;
Snowflake's metadata allows
some queries to be serviced
as metadata only operations.
No compute required!

METADATA TYPES & STORAGE
14
Statistical
All Other

•
Stored in the Global Services Layer
•
Entry is an “Expression Property” (or “EP”)
•
This type of metadata includes:
?
FDN
-
level
¦
Row Count
?
FDN
-
column
-
level
¦
MIN/MAX values
¦
Number of distinct values
¦
Number of NULLs
•
Stored in FDB
•
This type of metadata includes:
?
Reference to physical file/objects
(FDNs) in Cloud Storage
?
Table versions (supports Time Travel
and Zero
-
Copy Cloning)

IMPACT OF METADATA
15
DML
Query Pruning
•
All DML operations take advantage micro

-
partition metadata for table maintenance
•
Some operations are metadata
-
only
operations and require no Compute to
complete
--
ex: deleting all rows from a table
•
Micro
-
partition metadata enables precise
pruning of columns in micro
-
partitions at query
run
-
time, including columns containing semi
-
structured data
•
Snowflake’s pruning algorithm first identifies
the micro
-
partitions required to answer a
query, and then scans only the portion of
these micro
-
partitions that contain the required
columns
•
Snowflake uses columnar scanning of
partitions so that an entire partition is not
scanned if a query only filters by one column

Storage Management:

Data Encryption:

Network Policy :
Account level, IP Address access control
Supports whitelists and blacklists using explicit IP addresses

Management available through the Snowflake Web UI under
Account
-
> Policies
User level support in private preview

Create the policy and set the policy .

Private link is only for AWS customers.
C:\Users\ni74075\Downloads\data_loading_lab

PUT file://C:\Users\ni74075\Downloads\data_loading_lab\region.tbl@%REGION auto_compress=false;
TrainMe1

PUT "file://C:\\Users\ni74075\Downloads\data_loading_lab.zip\data_loading_lab\region.tbl"
@%REGION auto_compress=false;
TrainMe1
28

PUT file:///C:/Users/ni74075/Downloads/data_loading_lab/data_loading_lab/region.tbl
@%REGION auto_compress=false;

```
COPY INTO @dbhol.schol.aws_unload1/INANDA_TECHMAHINDRA_parquet_region
FROM (select * from INANDA_TECHMAHINDRA.public.region)
FILE_FORMAT = ( TYPE = parquet );


COPY INTO @dbhol.schol.aws_unload1/INANDA_TECHMAHINDRA_parquet_region
FROM (select object_construct(*) from
INANDA_TECHMAHINDRA_DB.public.region)
FILE_FORMAT = ( TYPE = PARQUET );


COPY INTO @dbhol.schol.aws_unload1/[login]_json_region
FROM (select object_construct(*) from
[login]_DB.public.region)
FILE_FORMAT = ( TYPE = JSON );


COPY INTO @dbhol.schol.aws_unload1/
INANDA_TECHMAHINDRA_DB_parquet_region
FROM (select * from [login_DB].public.region)
FILE_FORMAT = ( TYPE = parquet );
```

--==========

For PUT command:
1. AUTO_COMPRESS option is by default TRUE and compresses in GZIP format
2.
Uploads (i.e. stages) data files from a local directory/folder on a client machine to one of the following Snowflake stages:

- Named internal stage.
- Internal stage for a specified table.
- Internal stage for the current user.

Recreate the pipe (using the CREATE OR REPLACE PIPE syntax). Internally, the pipe is dropped and created. The file loading metadata is associated with the pipe object rather than the table. Recreating the pipe removes the history of files loaded. Ensure that files already loaded by Snowpipe are not accidentally resubmitted to the pipe and loaded into the target table again

For snowflake with rest api (Select all that apply)
1. you can only use it on internal stages -> Definitely false. we can use for both internal and external stages
2. ALL COPY INTO OPTIONS ARE AVAILABLE DURING PIPE CREATION -> Couldn't confirm from documentation, But usually it allows
3. Snowflake automatically manages the compute required to execute the pipe's COPY INTO command -> True
4. Snowpipe keeps track of which files it has loaded -> True
--==========

. ALL COPY INTO OPTIONS ARE AVAILABLE DURING PIPE CREATION -> False.

Found in the documentation regarding it.

- All COPY options are supported ***except for*** the following:
  - `FILES = ( '<file_name1>' [ , '<file_name2>', ... ] )`
  - `PATTERN = '<regex_pattern>'`
  - `ON_ERROR = ABORT_STATEMENT`
  - `SIZE_LIMIT = <num>`
  - `PURGE = TRUE | FALSE` (i.e. automatic purging while loading)
    Note that you can manually remove files from an internal (i.e. Snowflake) stage (after they've be
  - `RETURN_FAILED_ONLY = TRUE | FALSE`
  - `VALIDATION_MODE = RETURN_n_ROWS | RETURN_ERRORS | RETURN_ALL_ERRORS`

1. Which of the following statements is true of Virtual Warehouse resizing?
   (i) A resize requires the warehouse to be in suspended status
   (ii) A resize can be completed at any time
   (iii) A resize will affect running, queued, and new queries
   (iv) A resize can only be completed once per day
2. Once created, a micro-partition will never be changed.
   Ans: **TRUE**/FALSE. (bcz micro partitions are immutable)
3. It is best practice to define a clustering key on every table:
   Ans: TRUE/**FALSE**
4. Snowflake provides standard and powerful features that ensure the highest levels of security for your account and users if used properly. Which aare the true statements about Snowflake security?
   (i) Tri-secret requires that customer manage their own keys
   (ii) Federated authentication in Snowflake is compliant with SAML 2.0
   (iii) Snowflake support user-based access control.
5. Query statement encryption is supported on ____ accounts.
   (i) Standard
   (ii) Enterprise
   (iii) Enterprise for Sensitive Data (ESD)
   (iv) Virtual Private Snowflake (VPS).
6. All security information is stored in the ____ layer in the Snowflake architecture.
   (i) Storage
   (ii) Compute
   (iii) Service
   (iv) All the above.
7. The benefits of client-side encryption are:
   (i) It provides a secure system for managing data in cloud storage(**True**)
   (ii) Data is encrypted before loading into storage layer**(False)**
   (iii) The storage service layer only contains encrypted version of the data**(True)**
   (iv) Queries can be encrypted on the client side**(False)**
8. Snowflake supports multi-factor authentication (MFA) to provide increased login security for users connecting to Snowflake. Which statements are true about MFA security?

1. MFA is automatically enabled for your account and available for all users to self-enroll	◉ True	○ False ✔
2. MFA is an integrated feature powered by the Duo Security Service	◉ True	○ False ✔
3. MFA can be used for connecting to Snowflake via the Snowflake JDBC driver	◉ True	○ False ✔
4. MFA login is designed only for connecting to Snowflake through the web interface	○ True	◉ False ✔

9. With an IdP configured for your account, (e.g. Okta, ADFS, or any of the other supported SAML 2.0-compliant service/applications) Snowflake supports using SSO to connect and authenticate with the following clients:
   (i) Python Connector
   (ii) JDBC Driver
   (iii) SnowSQL

(iv) ODBC Driver

10. Snowflake includes Role-Based Access Control to enable administrators to:
    (i) Linit access to data and privileges
    (ii) Manage secure access to the Snowflake account and data
    (iii) Establish role hierarchy and privilege inheritance to align access
    (iv) All of the above.
11. The snowflake user interface can execute many tasks that can be performed using SQL and the command line, including:

1. Creating and managing users and other account-level objects	◉ True	○ False ✔
2. Creating and using virtual warehouses	◉ True	○ False ✔
3. Creating and modifying databases	◉ True	○ False ✔
4. Loading data into tables	◉ True	○ False ✔

12. You can manage user preferences in the following areas through the Snowflake user interface except:
    (i) Adding users and permissions
    (ii) Enrolling in multi-factor authentication
    (iii) Executing Fail-Safe recovery
    (iv) Monitoring queries using the History page.
13. Mark the processes that are used to load data into the Snowflake system as True:

1. Bulk loading using local file system using the COPY command	◉ True	○ False ✔
2. Load continuously using Snowpipe function	◉ True	○ False ✔
3. Using the Loading Wizard in Snowflake	◉ True	○ False ✔

14. Select the features that are true about the Snowflake Data Warehouse Architecture:

1. Data storage is independent from compute	◉ True	○ False ✔
2. All virtual warehouses have access to all data	◉ True	○ False ✔
3. All interactions with data are initiated through the services layer	◉ True	○ False ✔
4. Multi-Cluster warehouses support high concurrency	◉ True	○ False ✔
5. Warehouses can be dynamically expanded to adjust to workloads	◉ True	○ False ✔

15. The Snowflake Architecture consists of 3 layers. Match the labels of the layers that deliver the functionality described:

Which layer provides centralized services for the data warehouse?	Cloud Services
Which layer contains virtual warehouses for data access?	Query Processing
Which layer contains the data in compressed, columnar format?	Database Storage

16. The Snowflake platform prioritizes security and authentication and includes the following key features:
    (i) Multi-factor authentication (MFA) -- TRUE
    (ii) Snowflake failure alerts -- FALSE
    (iii) Data encryption using Snowflake-managed keys -- FLASE
    (iv) Object-level access control -- TRUE
17. The Snowflake architecture includes advanced capabilities in the cloud services layer that deliver:
    (i) Transaction management for consistent operation on the same data at the same time.
    (ii) Matadata service
    (iii) Security and authentication control
    (iv) Query optimization
18. What are the layers of the Snowflake architecture?
    (i) Storage
    (ii) Severs
    (iii) Nodes
    (iv) Compute
    (v) Services
    (vi) Metadata
19. What can you use the worksheet for in the Snowflake Web Interface?
    (i) Crate and execute SQL queries
    (ii) Load data
    (iii) View query history
    (iv) Access Snowflake Tutorials
    (v) Manage Metadata
20. A virtual warehouse is part of what laye in the snowflake architecture.
    (i) Servers
    (ii) Compute
    (iii) Storage
    (iv) Other

21. A zero-copy clone uses additional storage when?
    (i) When create
    (ii) When accessed
    (iii) When data is added
    (iv) When scanned
    (v) When data is modified
22. Snowflake uses what command to move data into Snowflake?
    (i) Create
    (ii) Clone
    (iii) Copy
    (iv) Drop
    (v) Insert
23. A database transaction is said to have ACID properties if it is:
    (i) Atomic
    (ii) Consistent
    (iii) Isolated
    (iv) Durable
    (v) All of the above
24. Which statement is accurate when discussing the difference between the commands CREATE OR REPLACE table vs. CREATE table if none exists?
    (i) If a table already exists, create or replace will simply remove that table and replace it with new definition
    (ii) User CREATE TABLE if none exists; if the table already exists, then the statement will just be skipped
    (iii) Both statements above are accurate
    (iv) Neither statement above is accurate.
25. Stored procedures enable procedural logic (branching and looping) and error handling, which straight SQL may not support. It is capable of dynamically creating a SQL statement and executing it.
    Review the example below and confirm which statements are true:

STORED PROCEDURE EXAMPLE

```
create or replace procedure get_row_count(table_name VARCHAR)
  returns float not null
  language javascript
  as
  $$
  var row_count = 0;
  // Dynamically compose the SQL statement to execute.
  var sql_command = "select count(*) from " + TABLE_NAME;
  // Run the statement.
  var stmt = snowflake.createStatement(
        {
        sqlText: sql_command
        }
      );
  var res = stmt.execute();
  // Get back the row count. Specifically,
  // ... get the first (and in this case only) row from the result set ...
  res.next();
  // ... and then get the returned value, which in this case is the number of
  // rows in the table.
  row_count = res.getColumnValue(1);
  return row_count;
  $$
  ;
call get_row_count('stproc_test_table1');
-- Double-check that we got the correct result.
select count(*) from stproc_test_table1;
```

1. This procedure is written in JavaScript	◉ True	○ False ✔
2. It will execute a SQL statement	◉ True	○ False ✔
3. It can retrieve result of a query or information about the result set	◉ True	○ False ✔

26. Which of the following methods for traversing data or paths is effective in Snowflake?
    (i) Insert a colon: between VARIANT column name and first-level element
    (ii) Dot notation to traverse a path in a JSON object: <column>:<level1_element>.>level2_element>.
    (iii) All of the above.
27. Which of these items are necessary when preparing to migrate a data warehouse to the cloud?
    (i) List of databases to migrate
    (ii) List of database objects to migrate
    (iii) List of processes and tools that populate and access the data warehouse
    (iv) List of security roles, users and permissions
    (v) All of the above.
28. When preparing for migration, what's the objective of examining current processes and data structures? (mark TRUE for those that apply)

1. Identify processes to migrate "as-is"	◉ True	○ False ✔
2. Prioritize processes that require re-engineering	◉ True	○ False ✔
3. Discard processes that are no longer necessary	◉ True	○ False ✔
4. Figure out how to eliminate testing phase	○ True	◉ False ✔

29. A data warehouse migration plan requires a project team to cover all aspects of requirements, preparation and execution. Match the role with the description:

Evaluates current processes and re-engineers systems to support new environment	Developer
Responsible for test planning and implementation; from design to performance	Quality Assurance
Maintains schedules, priorities of planning and implementation	Project Manager
Sets priority timeline for databases tools and processes to be migrated	Database Infrastructure Manager

30. Which of the following can not be migrated or administered using the user interface in Snowflake.
    (i) Stored procedures.
    (ii) SQL scripts and commands
    (iii) Security / user roles settings
    (iv) Resource or credit usage limits.
31. The longer the data retention period, the higher the resulting storage costs
    Ans: **TRUE**/FALSE
32. Warehouse – compute layer?
    Ans: **TRUE**/FALSE
33. Fail-safe can be disables within a Snowflake account
    Ans: TRUE/**FALSE**
34. Match the key feature of Snowflake's Data Warehouse architecture to data sharing benefits:

Storage separated from Compute	Allows multiple users to concurrently share live data
Elastic Scaling	Supports any size of data; removes concurrency limitations
Centralized management of services	Connect consumers directly to live data rapidly and securely

35. Data provides can manage access to Snowflake data shares using:
    (i) Role-based account privileges
    (ii) Account mapping tables
    (iii) Manual updates to user tables.
36. Which of the following conditions are required for sharing data in Snowflake?
    (i) Data providers with ACCOUNTADMIN role can set up shares
    (ii) Consumer accounts must be in same Snowflake region as the provider account
    (iii) Secure views are not required when query performance is priority
    (iv) Each data share must contain a single database
    (v) Any object in a share can be from different databases
37. What are the 3 data sharing account profiles provided by Snowflake?

Data Provider	Creates and administers access to data shares
Data Consumer	Creates a database from shares
Reader Account	Access to account created by provider to preview data share

38. Zero copy cloning allows users to have multiple copies of your data without the additional cost of storage usually associated with replication data. Which other statements about the cloning features in Snowflake are True?
    (i) Clone is a "point in time version" of the table data as of the time the clone was made

(ii) The clone is a pointer to the original table data

(iii) Cloning is an efficient and cost effective approach for code migration for Agile Release Management.

39. The cloning feature in Snowflake requires less storage because:

(i) Only metadata is copied; no physical data is copied

(ii) The cloned data is constantly updated to synchronize with original table data

(iii) Data is replicated to ensure integrity of data in the original table data.

40. Which of the following database objects can be cloned in Snowflake?

(i) Tables

(ii) Schemas

(iii) Databases

(iv) All of the above

41. Which SQL statements can work on a cloned table?

(i) SHOW command

(ii) DROP command

(iii) SELECT command

(iv) All of the above

42. Which of the following is a common case for cloning in Snowflake? (Select all that apply)

(i) "point in time" snapshot

(ii) Agile release and development

(iii) Data life cycle management

(iv) Data encryption protection.

43. Choosing separate accounts in Snowflake enables users to have:

(i) Different editions of Snowflake and different regions

(ii) Billing at the account level

(iii) Simpler database object deployment between environments

(iv) All of the above

44. Account structure is a strategic decision that impacts access to different capabilities in Snowflake because:

(i) Different editions of Snowflake instances require separate accounts

(ii) Snowflake instances in different regions require separate accounts

(iii) Data can be shared READ ONLY across accounts, and can be cloned across accounts.

45. Compare data sharing vs. data cloning in snowflake:

Defined data set can be shared outside the Snowflake account	Data Sharing
Defined data set is a point in time snapshot and can be updated by customers	Cloning

46. An enterprise view of data is useful because:

(i) A data set can be stored once and shared multiple times

(ii) Data set provisioning is assigned to the owner of the data set
(iii) Data sets can be shared securely
(iv) All of the above.

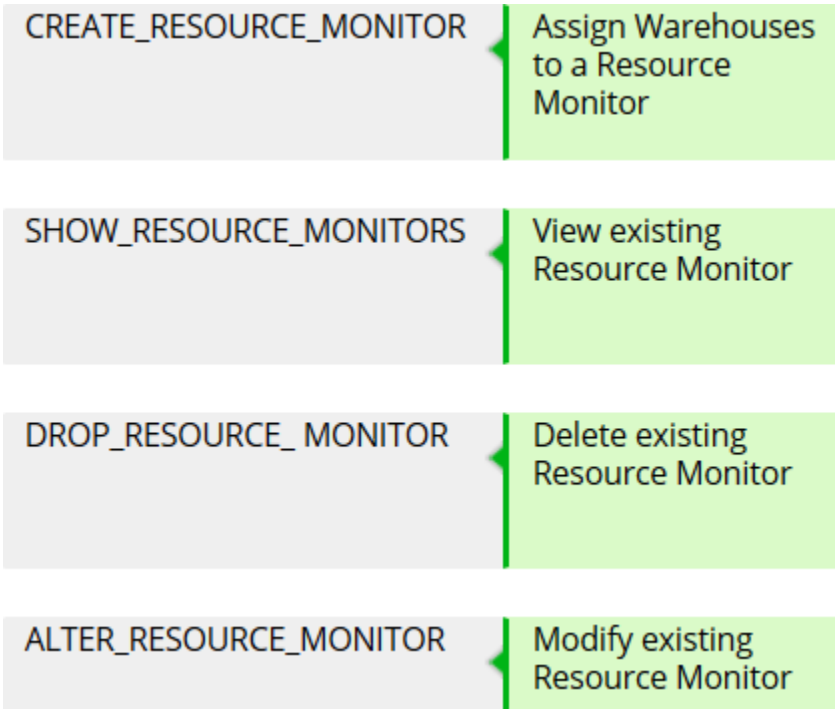
47. The ability to clone and share data between accounts factors into the decision to set up separate accounts in Snowflake because:
(i) Data sharing is only supported between accounts in the same Snowflake region
(ii) A share can't be cloned by a consumer account but the shared data can be copied into a table
(iii) Data can be shared READ-ONLY across Snowflake accounts and can also be cloned
(iv) There are cases where separate accounts are required such as different editions or regions
48. A micro-partition will be updated as updates are made to the rows of data it contains
Ans: TRUE/**FALES**
49. Snowflake clients include connection and drivers for:
(i) Data integration tools
(ii) Business intelligence tools
(iii) Advanced analytics tools
(iv) All of above.
50. Connectors and drivers available to Snowflake users include:
(i) Native ones built by Snowflake
(ii) Partner provided solutions
(iii) Standardized general purpose ones
(iv) All of above.
51. The Snowflake connector for Spark support:
(i) Exporting Snowflake data into external stage for Spark consumption
(ii) Exporting Snowflake data to on premise data center for Spark consumption
(iii) Importing Spark data in internal stage into Snowflake table
(iv) Importing Spark data in external stage into Snowflake table
52. SQLAlchemy is a database toolkit for developers of the following programming interface:
(i) ODBC
(ii) Node.js
(iii) Python
(iv) C/C++
53. Which of the following application tools can work with Snowflake?
(i) Querying tools using JDBC interface such as DBeaver
(ii) Modern data science notebooks such as Jupyter Notebook
(iii) Cobol application.
54. Map the basic backup and recovery capability in Snowflake with their use case:

All data and data objects are fully recoverable by users during retention period	Time Travel
Cost effective alternative to back up enabled by Snowflake Architecture	Fail Safe
Creating an exact copy or mirror of another application program or object	Cloning

55. Snowflake Time Travel enables access to data that has been changes or deleted at any point within a defined period. With Time Travel, a Snowflake administrator can:
    (i) Query data in the past that has since been updated or deleted
    (ii) Create clones of entire tables, schemas, and databases at or before specific points in the past
    (iii) Restore tables, schemas, and databases that have been dropped after the retention period lapses
    (iv) Analyzing data usage/manipulation over specified periods of time
56. Fail-Safe ensures historical data is protected in the event of a system failure or other catastrophic event, e.g. a hardware failure or security breach. Under what circumstances would Fail-Safe NOT be an effective method for data recovery?
    (i) As a means for accessing historical data after the Time Travel retention period has ended
    (ii) To recover data that may have been lost or damaged due to extreme operational failures
    (iii) To protect data in the event of a system failure or other catastrophic event, e.g. a hardware failure or security breach.
57. What is the maximum data retention period that an enterprise account can have?
    (i) 90 days
    (ii) 120 days
    (iii) 30 days
58. What is the maximum data retention for Fail-Safe access?
    (i) 30 days
    (ii) 7 days
    (iii) 60 days
59. Snowflake Resource Monitor allows administrators to set thresholds and triggers to track credit usage. What method are most effective for resource monitoring? (Check all that apply)
    (i) Tracking consumption on an hourly basis
    (ii) Tracking and controlling credit consumption on a monthly basis
    (iii) Setting quota thresholds for trigger actions and notifications

(iv) Control global monthly credit usage for an account

60. The Snowflake resource monitor feature enables administrators to execute the following actions based on thresholds and notifications:
    (i) Limiting the credit for a group warehouse
    (ii) Suspending a specific warehouse
    (iii) Resuming suspended warehouse when the monitor is dropped
    (iv) Changing credit quotas after a resource monitor is created.
61. Snowflake resource monitors can be managed using DDL commands. Match the command with the action:



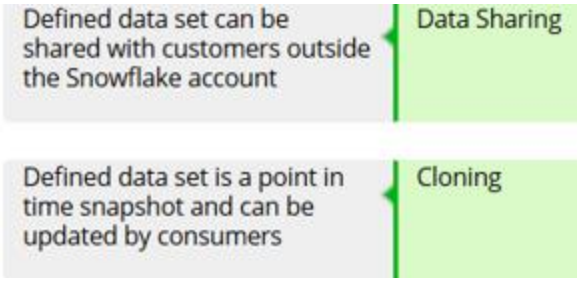
62. Suspended warehouses cannot be resumed until one of the following conditions is met:
    (i) New monthly billing cycle starts
    (ii) Credit quota for the monitor is increased
    (iii) Credit threshold for the trigger is increased
    (iv) The monitor is dropped
    (v) Anyone of the above will meet the required condition.
63. Snowflake notification are disabled by default. How do administrators enable resource monitor notifications? (check all that apply)
    (i) Enable notification through the preferences in the Snowflake web interface
    (ii) Provide and verify valid email address if email notification is preferred
    (iii) Include notifications in the resource monitoring reports.
64. Which of the following are valid approaches to loading data into a snowflake table? (select all that applu)
    (i) Bulk copy from an external stage.
    (ii) Continuous load using Snowpipe REST API
    (iii) The snowflake web interface(UI) data loading wizard
    (iv) Bulk copy from an internal stage.

65. Which approach would result in improved performance through linear scaling of data ingestion workload?
    (i) Resize virtual warehouse.
    (ii) Consider the practice of organizing data by granular path
    (iii) Consider the practice of splitting input file batch the recommended range of 10 MB to 100MB
    (iv) All of the above.
66. Snowflake support service addresses customer issues covering:
    (i) Product usage questions
    (ii) Troubleshooting failed queries
    (iii) Individual query syntax improvement
    (iv) 3rd party application configuration support.
67. Snowflake users with support contracts that have a Severity-1 issues should contact Snowflake in these ways EXCEPT:
    (i) Snowflake lodge – set sppropriate Severity(1-4)
    (ii) Send email
    (iii) 844-SNOWFLAKE
68. The snowflake lodge is a community site that
    (i) Has technical information for support customers only
    (ii) Is the recommended place to submit support cases
    (iii) Contains the most up to date security alerts and product release information
    (iv) Does not allow members to post questions
69. Map the severity levels to their definition:

An Error that has a medium-to-low impact on the service.	Severity Level 3 (Medium Severity)
An Error that (a) renders the Snowflake service completely inoperative.	Severity Level 1 (Critical Severity)
An Error that has low-to-no impact on Customer's access to / use of the service.	Severity Level 4 (Low Severity)
An Error that (a) has high impact to key portions of service or (b) serious.	Severity Level 2 (High Severity)

70. Snowflake includes parameters that allow administrators to control the behavior of individual user sessions and objects within a Snowflake account including:
    (i) DATA_RETENTION_TIME_IN_DAYS
    (ii) MAX_CONCURRENCY_LEVEL
    (iii) STATEMENT_QUEUED_TIMEOUT_IN_SECONDS
    (iv) All the above.

71. Which of the following are options when creating a virtual warehouse? (select all that apply)
    - **(i) Auti-suspend**
    - **(ii) Auto-resume**
    - (iii) Local SSD size
    - (iv) User count.
72. Snowflake included administration settings for resource consumption in order to:
    - **(i) Help control costs associated with unexpected credit usage of warehouses.**
    - (ii) Manage access to data warehouses for specified users
    - (iii) Maintain data availability
73. Which roles and permissions can be granted within a Snowflake account by the administrator?
    - **(i) Create role**
    - **(ii) Drop role**
    - **(iii) Create user**
    - **(iv) Drop user**
    - **(v) Grant privilege**
74. Snowflake provided specific administration features and capabilities to support the following activities except:
    - (i) Managing databases and warehouses within a Snowflake account
    - (ii) Managing roles and users within a Snowflake account
    - (iii) Monitoring usages and manage resources to control costs in a Snowflake Account
    - **(iv) Manage 3rd party applications providing data to a Snowflake account.**
75. Compare Data Sharing vs. Data Cloning in Snowflake:



76. The following performance optimizing query methods are support by Snowflake:
    - **(i) Caching techniques**
    - (ii) B-tree type indexes
    - **(iii) Retrieving results if previous query from cache**
77. Which of the concepts should be considered when loading data into snowflake?
    - (i) Stage objects
    - (ii) File format
    - (iii) Transformation and error validation
    - **(iv) All of the above.**

78. Which of the following are unique database objects in snowflake?
    - **(i) Stage**
    - **(ii) Pipe**
    - **(iii) Table**
79. The compute resource used by Snowflake for data loading jobs can be provided by

(i) User managed virtual warehouse
(ii) Snowflake managed services
(iii) Hardware provisioned by user directly from cloud providers

80. Snowflake supports loading data into
(i) Internal stage on cloud storage platform
(ii) External stage on cloud storage platform
(iii) Bring your own device

81. Which of the following objects is not covered by rest
(i) Tables
(ii) Schemas
(iii) Databases
(iv) Stages

---

1. Snowflake notifications are disabled by default. How do administrators enable resource monitor notifications? (check all that apply)
(i) Enable notifications through the preferences in the Snowflake web interface
(ii) Provide and verify valid email address if email notification is preferred
(iii) **Include notifications in the Resource Monitoring reports.**
2. Snowflake administrators should utilize resource monitors to help control costs and avoid unexpected credit usage. Which of the following actions can Snowflake's resource monitor triggers initiate automatically? (Check all that apply)
(i) Impose limits on the number of credits that warehouses consume each month
(ii) Trigger alert notifications for high usage
(iii) Roll over query executions to under-utilized warehouses
(iv) Trigger warehouse suspension for high usage.
3. The snowflake architecture consists of:
(i) A monolithic, integrate stack
(ii) Tightly coupled storage and compute layers
(iii) Three layers: Cloud Services, Virtual Warehouse, Hybrid Columnar Storage.
(iv) None of the above.
4. Which statement accurately describes the snowflake cloud services layer?
(i) A set of cloud provider's general services made available to users via the internet
(ii) A collection of independent, scalable, and stateless services providing crucial data management capabilities
(iii) A collection of tightly coupled database management feature
(iv) Managed and scaled by the user
5. Which statement is the most accurate regarding the snowflake metadata service?
(i) Snowflake MDD service is a separate product offering from snowflake

(ii) It provides crucial, built-in management services for all metadata from capabilities such as query

(iii) Snowflake MDD services can be managed by the user

(iv) Snowflake MDD services capabilities do not impact data sharing, time travel or cloning capabilities

6. Which statement is most accurate regarding Snowflake's transaction support?

(i) Snowflake's transaction support is ACID-compliant

(ii) It does not require updates

(iii) Snowflake transaction support requires special user configuration

(iv) Snowflake transaction support allows only statement

7. How does Snowflake's support for High Availability work?

(i) The feature requires users to manage provisioning & replication of the cloud provider's virtual instances

(ii) Snowflake's support for High Availability provides built-in, resilient capabilities in all three architectural layers

(iii) It does not support fault tolerance

(iv) High availability capability is limited to the storage layer.

8. Which security feature is supported in Snowflake?

(i) Role-Based Access Control

(ii) Multi-Factor Authentication

(iii) Tri-Secret Secure Encryption

(iv) All of the above.

9.