Understanding compute cost

Compute costs represent credits used for:

- Virtual Warehouse compute — Virtual warehouses consume credits as they execute queries, load data and perform other DML operations. Virtual Warehouses are ***user-managed***, which means you can directly control credit consumption of these resources.
- Serverless compute — Serverless features use compute resources that are managed by Snowflake instead of using virtual warehouses.
- Cloud Services compute — Cloud Services is the layer of the Snowflake architecture that performs services that tie together all the different components of Snowflake to process user requests, login, query display, and more. Cloud Services compute resources are managed by Snowflake.

Virtual warehouse credit usage

A virtual warehouse is one or more clusters of compute resources that enable executing queries, loading data, and performing other DML operations. The web interface and other features use warehouses, such as Cross-Cloud Auto-Fulfillment or display information in dashboards.

Snowflake credits are used to pay for the processing time used by each virtual warehouse. Snowflake credits are charged based on the number of virtual warehouses you use, how long they run, and their size.

Warehouses come in many sizes. In this table, the size specifies the compute resources per cluster available to the warehouse. Each increase in size to the next larger warehouse approximately doubles the computing power and the number of credits billed per full hour that the warehouse runs.

Type of Virtual Warehouse	X-Small	Small	Medium	Large	X-Large	2X-Large	3X-Large	4X-Large	5X-Large	6X-Large
Standard	1	2	4	8	16	32	64	128	256	512

Snowpark-optimized	n/a	n/a	6	12	24	48	96	192	384	768

Important
Warehouses are only billed for credit usage while running. When a warehouse is suspended, it does not use any credits.

The credit numbers shown above are for a full hour of usage; however, credits are billed per-second, with a 60-second (i.e. 1-minute) minimum:

- Each time a warehouse is ***started*** or ***resumed***, the warehouse is billed for 1 minute's worth of usage based on the hourly rate shown above.
- Each time a warehouse is ***resized*** to a larger size, the warehouse is billed for 1 minute's worth of usage; however, the number of credits billed are only for the ***additional*** compute resources that are provisioned. For example, resizing from Small (2 credits/hour) to Medium (4 credits/hour) results in billing charges for 1 minute's worth of 2 additional credits.
- After 1 minute, all subsequent billing is per-second as long as the warehouse runs continuously.
- Suspending and then resuming a warehouse within the first minute results in ***multiple*** charges because the 1-minute minimum starts over each time a warehouse is resumed.
- Resizing a warehouse from 5X-Large or 6X-Large to 4X-Large (or smaller) results in a brief period during which the warehouse is billed for both the new compute resources and the old resources while the old resources are quiesced.

For more information on warehouses in general, see Overview of warehouses and Warehouse considerations.
To learn how to view the historical cost of consuming compute resources with virtual warehouses, see Exploring compute cost.

Serverless credit usage

Serverless credit usage is the result of features relying on compute resources provided by Snowflake rather than user-managed virtual warehouses. These compute resources are automatically resized and scaled up or down by Snowflake as required for each workload.

For these serverless features, which usually require continuous and/or maintenance operations, this model is more efficient, allowing Snowflake to charge based on the time spent using the resources. In contrast, user-managed virtual warehouses consume credits while running, regardless of whether they are performing any work, which may cause them to be overutilized or sit idle.

Charges for serverless features are calculated based on total usage of snowflake-managed compute resources measured in *compute-hours*. Compute-Hours are calculated on a per second basis, rounded up to the nearest whole second. The number of credits consumed per compute hour varies depending on the serverless feature.

To learn how many credits are consumed by a serverless feature, refer to the “Serverless Feature Credit Table” in the Snowflake service consumption table.

Charges for the use of a serverless feature appear on your bill as an individual line item. Charges for both Snowflake-managed compute resources and Cloud Services appear as a single line item for that serverless feature.

Serverless features

Feature	Required Compute	Additional Information
Automatic Clustering	Automated background maintenance of each clustered table, including initial clustering and reclustering as needed.	Automatic Clustering Cost
COPY FILES	Copying files.	COPY FILES cost
Data Quality and Data Metric Functions (DMFs)	Calling a scheduled DMF to measure the state and integrity of your data and then ingesting the results into an event table.	Data Quality Cost
External Tables	Automated refreshing of the external table metadata with the latest set of associated files in the external stage and path.	External Tables Cost

Hybrid tables	Hybrid Table Requests credits consumed based on the amount of data that is read from or written to hybrid tables. Additional resources are used in the creation and use of indexes.	Understand cost for hybrid tables
Logging and tracing	Collecting log and trace messages in batches, then ingesting those batches into an event table.	Costs of Logging
Materialized Views	Automated background synchronization of each materialized view with changes in the base table for the view.	Materialized Views Cost
Query Acceleration Service	Execution of portions of eligible queries.	Query Acceleration Service Cost
Replication	Automated copying of data between accounts, including initial replication and maintenance as needed.	Replication Cost
Search Optimization Service	Automated background maintenance of the search access paths used by the search optimization service.	Search Optimization Service Cost
Snowpipe	Automated processing of file loading requests for each pipe object.	Snowpipe Cost
Snowpipe Streaming	Automated loading of streaming data rows.	Snowpipe Streaming Cost

Serverless alerts	The execution of the SQL code for the alert.	Understanding the costs of alerts
Serverless tasks	The execution of the SQL code for the task.	Tasks Cost

To learn how to view the historical cost of using serverless compute resources, see Exploring compute cost.

Cloud service credit usage

The cloud services layer of the Snowflake architecture is a collection of services that coordinate activities across Snowflake. This layer authenticates users, enforces security, performs query compilation and optimization, handles request query caching, and more. Cloud services tie together all of the different components of Snowflake, including supporting the use of virtual warehouses.

The cloud services layer is constructed of stateless compute resources, running across multiple availability zones and using a highly available, distributed metadata store for global state management. The cloud services layer runs on compute instances provisioned by Snowflake from the cloud provider.

Similar to virtual warehouse usage, Snowflake credits are used to pay for the usage of the cloud services.

Understanding billing for cloud services usage

Usage for cloud services is charged only if the daily consumption of cloud services exceeds 10% of the daily usage of virtual warehouses. The charge is calculated daily (in the UTC time zone). This ensures that the 10% adjustment is accurately applied each day, at the credit price for that day.

Keep the following in mind:

- Serverless compute does not factor into the 10% adjustment for cloud services.
- The 10% adjustment for cloud services is calculated daily (in the UTC time zone) by multiplying daily warehouse usage by 10%.
- The adjustment on the monthly usage statement is equal to the sum of these daily calculations.
- If cloud services consumption is less than 10% of warehouse compute credits on a given day, then the adjustment for that day is equal to the cloud services used by your

account. The daily adjustment never exceeds actual cloud services usage for that day. Thus, the total monthly adjustment may be significantly less than 10%.

For example:

Date	Compute Credits Used (Warehouses only)	Cloud Services Credits Used	Credit Adjustment for Cloud Services (Lesser of 10% of Compute or Cloud Services)	Credits Billed (Sum of Compute, Cloud Services, and Adjustment)
Nov 1	100	20	-10	110
Nov 2	120	10	-10	120
Nov 3	80	5	-5	80
Nov 4	100	13	-10	103
Total	**400**	**48**	**-35**	**413**

More about cloud services

- To learn how to view the historical cost of consuming cloud services resources, see Exploring compute cost, which includes sample queries you can run to see how much of cloud services consumption was actually billed and which queries and warehouses have the highest cloud services usage.
- To learn about patterns that drive cloud services consumption and ways that you might be able to reduce that consumption, see Optimizing cloud services for cost.

What are credits?

Snowflake credits are used to pay for the consumption of resources on Snowflake. A Snowflake credit is a unit of measure, and it is consumed only when a customer is using resources, such

as when a virtual warehouse is running, the cloud services layer is performing work, or serverless features are used.

Next Topic

- Exploring compute cost