Snowflake Health Check Metadata Download Guide

Product Version	4.8.0
Document Type	Health Check Preparation Guide
Authors	Snowflake Data source Team
Reviewer	Red Team & Architects
Approver	Shivnath
Total Pages	10
Document Status	Released

Table Of Contents

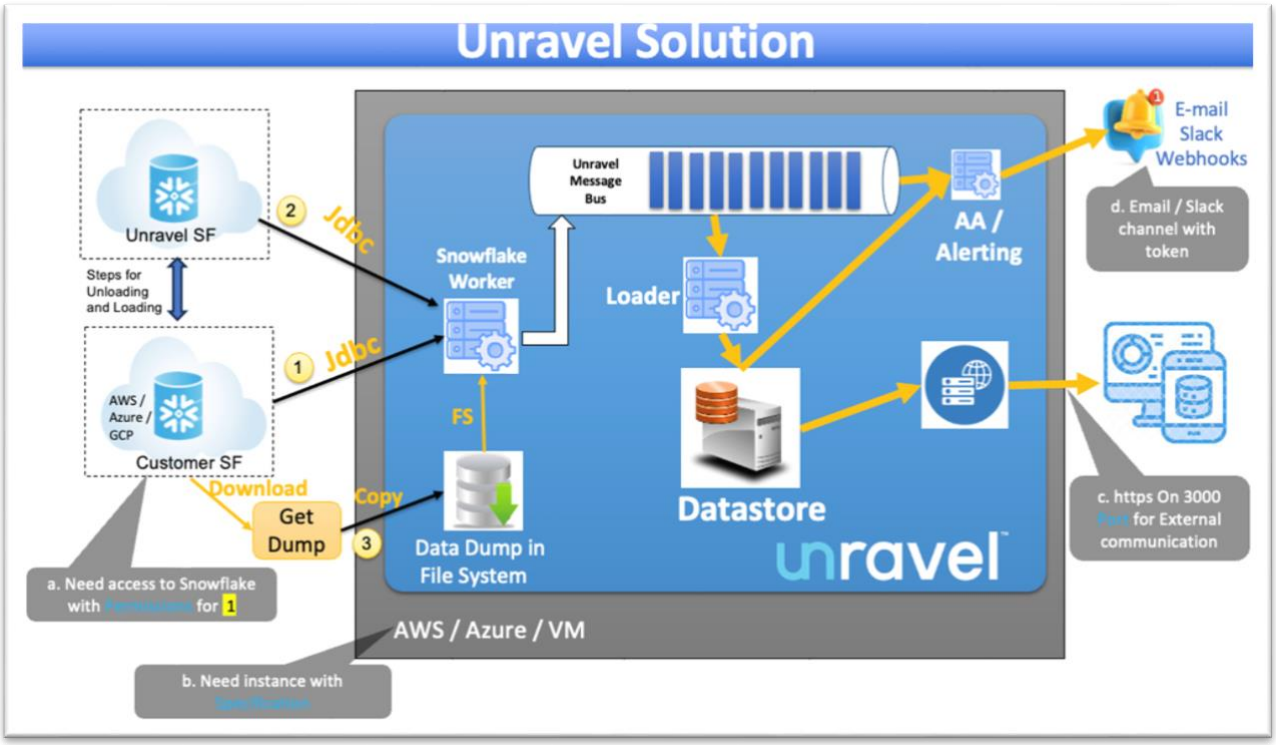
Document Version Record

Date	Version #	Author	Remarks / Reason
02-May-23	1.0	Dev Team	New Document
28-Aug-23	1.1	Dev Team	Updated sessions and access_control

1.1 Objectives

Health check download for snowflake unravel product.

1.2 Architecture



We are going to follow 2 Download from customer snowflake environment and upload to unravel snowflake environment.

1.3 Pre-requisite

A. **Snowflake account Access**

a. Health check data download from snowflake
   i. Create or select a user in snowflake, which has read access to Snowflake#Account_usage & Snowflake#Information_Schema.
   ii. Grant selects for unravel user on schema SNOWFLAKE.ACCOUNT_USAGE & SNOWFLAKE.INFORMATION_SCHEMA.
   iii. User should have permission to create stage and file format.
   iv. Snowsql 1.2.27 should be installed
   v. Python3 should be available
   vi. snowflake-connector-python required to execute snowflake queries in python (Refer the steps here - https://docs.snowflake.com/en/developer-guide/python-connector/python-connector-install).
   vii. Tested on Mac-OS(12.3), Windows 10, Ubuntu(16.04.7 LTS), CentOS-8

1.4 Download Snowflake Metadata for health check.

The following scripts let you download the account usage information from your Snowflake account.

- **snowsql_download_data.sql**: Downloads Snowflake metadata or telemetry data.
- **snowsql_show_wareshouses.sql**: Downloads Snowflake warehouse's data.
- **snowflake_query.py**: Downloads Snowflake warehouse parameters

Do the following to download Snowflake metadata:

1. Download the download/upload scripts from this location:

   https://github.com/unraveldata-org/snowflake-data-loader/tree/main/script

2. Using Snowsql connect to a Snowflake account. This account must have access to creating **stage** and **file** format.
3. Execute the **snowsql_download_data.sql** script with the required arguments to download the metadata from Snowflake **#account_usage** views. Refer to the following list of arguments for more details:

Step-1:

Note : ${variable} -> need to be modified as per site credentials and data

Once you run the script, it will prompt for the password.

```
snowsql -f ${downloaded-path}/snowsql_download_data.sql -d ${db} -s ${schema} -r ${role} -a ${account} -u ${user} -w ${warehouse} -o variable_substitution=true -o log_file=${script output file path/filename} --variable path=$(local/path/to store/downloaded metadata) --variable stage_name=unravel_stage_name --variable file_format=unravel_file_format
```

For example:

Windows :

```
snowsql -f "snowsql_download_data.sql" -d unravel_db -s unravel_schema -r accountadmin -a rtb81672.us-east-1 -u unraveluser -w unraveldata -o variable_substitution=true -o log_file=./snowsql_download_data.log --variable path=C:\opt\unraveldata\ --variable stage_name=unravel_stage --variable file_format=unravel_file_format
```

During copy make sure double quotes are correctly transferred.

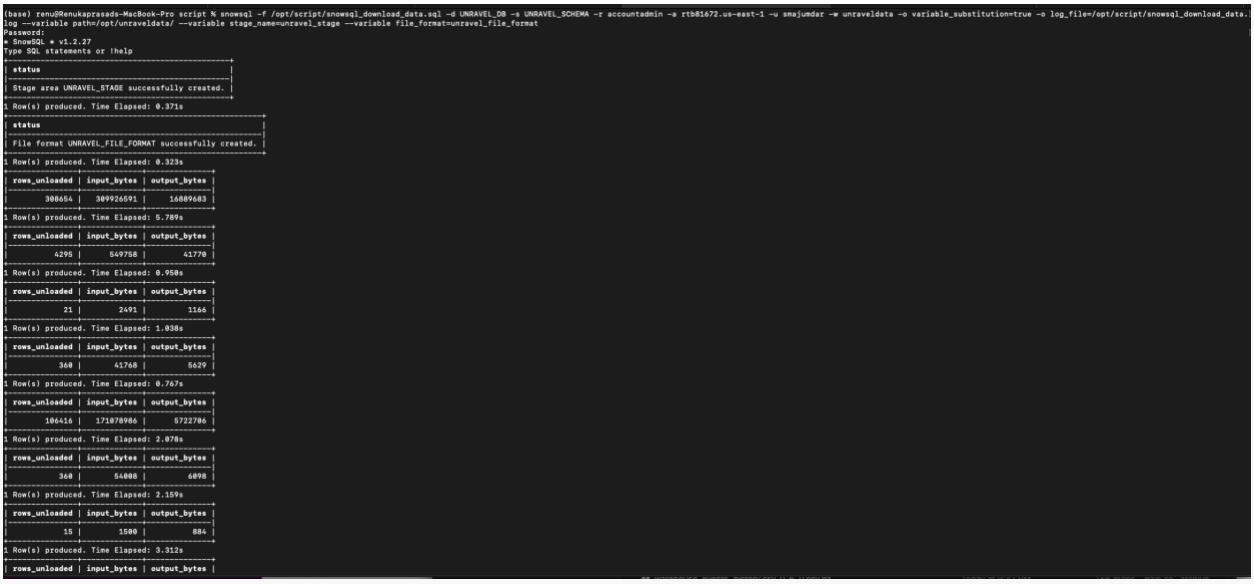
Mac / Linux :

```
snowsql -f snowsql_download_data.sql -d unravel_db -s unravel_schema -r accountadmin -a rtb81672.us-east-1 -u unraveluser -w unraveldata -o variable_substitution=true -o log_file=./snowsql_download_data.log --variable path=/opt/unraveldata/ --variable stage_name=unravel_stage --variable file_format=unravel_file_format
```

Parameter	Description
-f	Specify the file name of the script that is executed. Either specify the fully qualified name or just filename from the current directly. **For Windows:** Please put the parameter value in double quotes ("")
-d	Specify the database used for the script execution.
-s	Specify the name of the schema name used for the script execution. User should have **rights to create the stage and file formats under this** schema.
-r	Specify the role of the user who executes the script.

Parameter	Description
-a	Specify the Snowflake account, which will be used for the script execution.
-u	Specify the username of the Snowflake user who executes the script.
-w	Specify the Snowflake warehouse.
-o log_file	Specify the path to the log file that will be generated when you execute the script. If any issues with the relative path, please provide absolute path.
-o	Provide the output-related arguments to get the logs in the specified path and format.
-o variable_substitution	Enable the variable substitution switch in the script. Some variables are used in the script for which the values must be passed from the CLI. Set this to **true**.
--variable path	Specify the local path to store the downloaded Snowflake system metadata. Path mentioned here should exists already, otherwise create it manually.
--variable stage_name	Specify the stage name, which is used to keep the temporary files for download and upload.
--variable file_format	Specify the file format name, which is used by upload/download scripts.

After the script is executed, the Snowflake metadata gets downloaded to the specified location.



Step-2:

Execute the snowsql_show_wareshouses.sql script with the required arguments to download the warehouse data. Refer to the following list of arguments for more details:

Once you run the script, it will prompt for the password.

```
snowsql -f ${downloaded-path}/snowsql_show_wareshouses.sql -d ${db} -s ${schema} -r ${role} -a ${account} -u ${user} -w ${warehouse}  -o output_format=csv -o output_file=${path}/warehouses.csv -o variable_substitution=true
```

Windows :

```
snowsql -f "snowsql_show_wareshouses.sql" -d unravel_db -s unravel_schema -r accountadmin -a rtb81672.us-east-1 -u unraveluser -w unraveldata  -o output_format=csv -o output_file=C:\opt\unraveldata\warehouses.csv -o variable_substitution=true
```

During copy make sure double quotes are correctly transferred.

Mac / Linux :

```
snowsql -f snowsql_show_wareshouses.sql -d unravel_db -s unravel_schema -r accountadmin -a rtb81672.us-east-1 -u unraveluser -w unraveldata  -o output_format=csv -o output_file=/opt/unraveldata/warehouses.csv -o variable_substitution=true
```

Parameter	Description
-f	Specify the file name of the script that is executed. Either specify the fully qualified name or just filename from the current directly. For Windows: Please put the parameter value in double quotes ("")
-d	Specify the database used for the script execution.
-u	Specify the username of the Snowflake user who executes the script.
-s	Specify the name of the schema name used for the script execution.
-r	Specify the role of the user who executes the script.
-a	Specify the Snowflake account, which will be used for the script execution.
-o	Provide the output-related arguments to get the logs in the specified path and format.
-o variable_substitution	Enable the variable substitution switch in the script. Some variables are used in the script for which the values must be passed from the CLI. Set this to **true**.
-o output_file	Specify the path to the output file generated on script execution. Generated filename should be 'warehouses.csv' as mentioned in the command/example. Don't change the filename here. **Note**: The `output_file` path should be same as the path mentioned in the `snowsql_download_data.sql` command If any issues with the relative path, please provide absolute path.
-o output_format	Specify the output file format on script execution. This is in CSV format.

After the script is executed, the warehouse data gets downloaded in CSV format at the specified output location.

Step-3:

Execute the snowflake_query.py script with the required arguments to download the warehouse parameters data. Refer to the following list of arguments for more details:

```
python3 snowflake_query.py --user '${user}' --password '${password}' --account '${account}' --warehouse '${warehouse}' --database '${db}' --schema '${schema}' --out '/opt/unravel' --role ${role}
```

For example:

Windows :

```
python3 snowflake_query.py --user 'unraveluser' --password 'Password' --account 'rtb81672.us-east-1' --warehouse 'unraveldata' --database 'SNOWFLAKE' --schema 'ACCOUNT_USAGE' --out 'C:\opt\unraveldata' --role 'accountadmin'
```

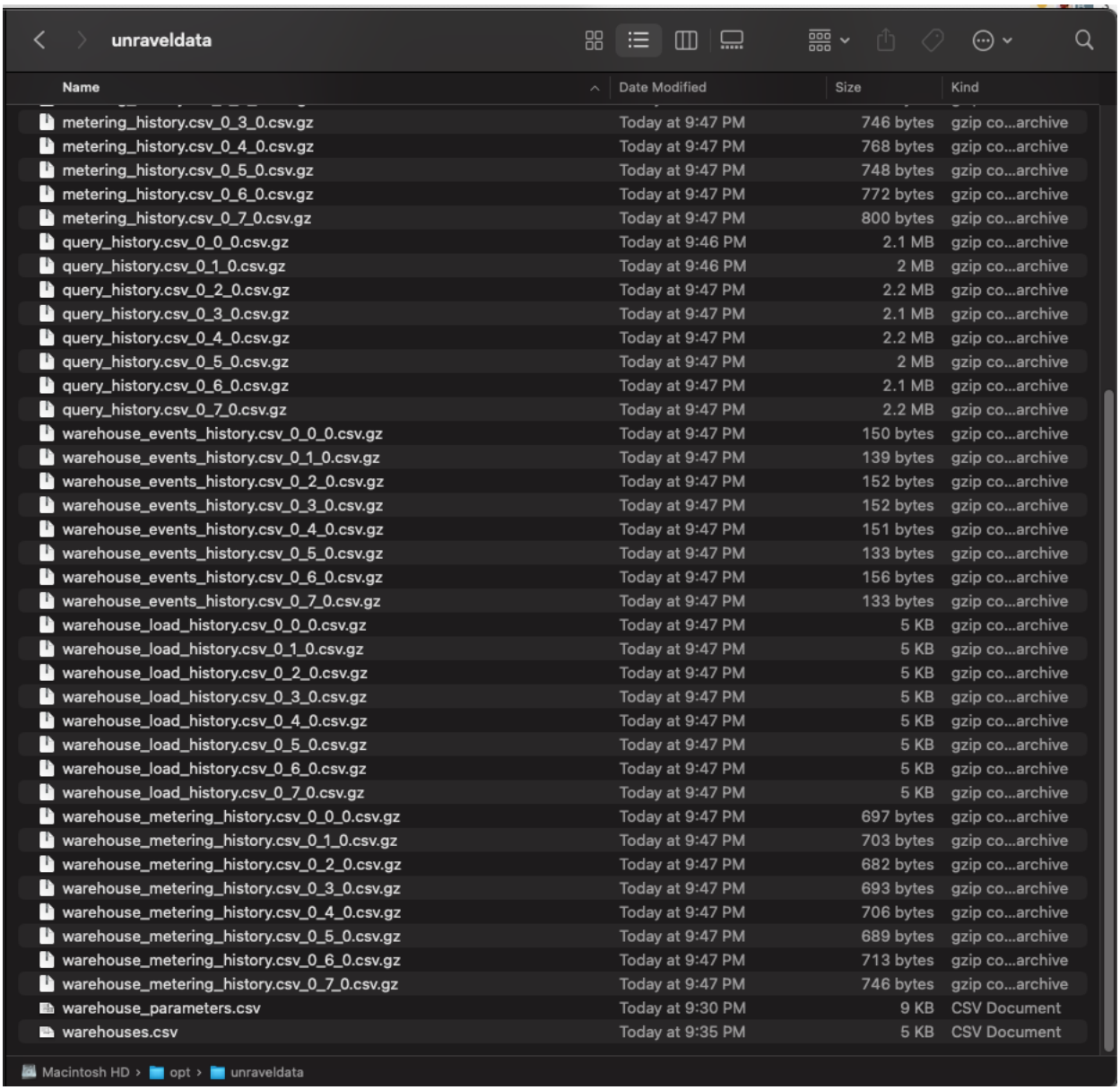
Mac / Linux :

```
python3 snowflake_query.py --user 'unraveluser' --password 'Password' --account 'rtb81672.us-east-1' --warehouse 'unraveldata' --database 'SNOWFLAKE' --schema 'ACCOUNT_USAGE' --out '/opt/unraveldata' --role 'accountadmin'
```

Parameter	Description
`--user`	Specify the name of the Snowflake user for the script execution.
`--password`	Specify the Snowflake user account passcode.
`--database`	Specify the database used for the script execution.

Parameter	Description
`--schema`	Specify the name of the schema that must be used to for the script execution.
`--role`	Specify the role of the user.
`--account`	Specify the Snowflake account, which will be used for the script execution.
`--out`	Specify the path to the output folder.

1.5 Output Files (Sample)



1.6 Video Guide

https://drive.google.com/open?id=13FvL6XTOzvTYtVAI2MZRRzeOWx0Y_s2r&usp=drive_fs