WhereScape Enablement Pack for Snowflake - SQL Server

This is a guide to installing the WhereScape Enablement Pack for Snowflake for WhereScape RED 8.6.6.1 or higher

Table of Contents

Prerequisites For SQL Server Metadata

Before you begin the following prerequisites must be met:

- Create Database and ODBC DSN :
  - Supported* version of SQL Server or Azure SQL
    - A database to house the RED Metadata Repository.
      Note : This needs to be Empty Database with optional permissions SELECT,INSERT,UPDATE,EXECUTE
    - A database for the Range Table DB (Optional)
- Software Installations
  - WhereScape RED version 8.6.6.1 or higher with valid license key entered and EULA accepted
  - WhereScape Enablement Pack for target database version 8.6.6.1 or higher

- Windows Powershell (64 bit) version 4 or higher
  - To check Windows Powershell Version:
    - Run below command in Windows Powershell

      ```
      Get-Host|Select-Object Version
      ```

    - Run below command in Command Prompt

      ```
      powershell $psversiontable
      ```

* : *RED supports the following versions for the metadata repository: MS SQL SERVER 2012 to 2019 and Azure SQL DB*

Prerequisites For Snowflake

Before you begin the following prerequisites must be met:

- Create Database and ODBC DSN :
  - Snowflake (ODBC driver version 2.22.01 or higher(64-bit))
    - At least one schema available to use as a RED Data Warehouse Target
    - Snowflake driver Log Level set to 0 in the Windows Registry and the Snowflake DSN Tracing set to 0

- Software Installations
  - Snowflake SnowSQL (CLI Client)

- Python 3.8
  - Select "Add Python 3.8 to PATH" from installation Window
  - Pip Manager Install with command : *python -m pip install --upgrade pip*

Enablement Pack Setup Scripts

The Enablement Pack Install process is entirely driven by scripts. The below table outlines these scripts, their purpose and if "Run as Administrator" is required.

#	Enablement Pack Setup Scripts	Script Purpose	Run as Admin	Intended Application
1	install_New_RED_Repository.ps1	Setup and configure a RED Metadata Repository for Snowflake If RED repository exists then updates the repository with 1.Templates 2.Scripts 3.Extended Properties 4.Datatype Mappings 5.UI Configurations	Yes	New and Existing installations
2	install_WslPowershell_Modules.bat	Installs or updates WslPowershell Modules on this machine	Yes	New and Existing installations
3	install_WslPython_Modules.bat	Installs or updates WslPython Modules and required Python libraries on this machine	Yes	New and Existing installations
4	import_powershell_templates.ps1	Imports or updates the Powershell Templates to a RED Metadata Repository. Also includes any Script or Procedure Imports	No*	Existing installations
5	import_python_templates.ps1	Imports or updates the Python Templates to a RED Metadata Repository. Also includes any Script Imports	No*	Existing installations
6	set_default_templates.ps1	Applies the RED Connection defaults in a RED Metadata Repository for Python or Powershell templates	No*	Existing installations

** Note that on some systems executing Windows Powershell scripts is disabled by default, see troubleshooting for workarounds*

Each Powershell script in the list above provides some help at the command line, this can be output by passing the "-help" parameter to the script. For Example > .\install_New_RED_Repository.ps1 -help

Step-By-Step Guide

Setup and configure RED Metadata Repository for Snowflake

Run Powershell as Administrator:

```
        Script 1 > Powershell -ExecutionPolicy Bypass -File .\install_New_RED_Repository.ps1
```

Important Upgrade Notes

If RED repository exists,it will prompt to upgrade the repository.
This enablement pack will overwrite any existing Source Enablement Pack UI Configs:

Connection UI Config	Load UI Config
Amazon S3	Load From Amazon S3
Azure Data Lake Storage Gen2	Load From Azure Data Lake Storage Gen2
Google Cloud	Load From Google Cloud

To ensure existing Source Enablement Pack connections and associated Load Tables continue to browse and load:

Go into UI Configuration Maintenance in RED prior to installing this Enablement Pack and rename the affected UI Configurations. While the updated Load Template will work with previous Source Enablement Pack's we recommend moving these previous versions of Load Tables to newly created Parser based connections following this install. The earlier versions of the Source Enablement Pack will be deprecated following this release.

Install or Update WhereScape Powershell Modules

Run Script As Administrator

```
        Script 2 > install_WslPowershell_Modules.bat
```

Install or Update WhereScape Python Modules

Run Script As Administrator

```
        Script 3 > install_WslPython_Modules.bat
```

There are two steps in this script:

1. Installs the WhereScape WslPython modules to C:\Program Data\WhereScape\Modules\

2. PIP to download/update required Python libraries - for offline install please see the required library list for Python in the Troubleshooting section.

Install or Update WhereScape Powershell Templates (For Existing Installations)

Run Script as Administrator

```
         Script 2 > install_WslPowershell_Modules.bat
         Script 4 > . .\import_powershell_templates.ps1
         Script 6 > . .\set_default_templates.ps1
```

Note : Skip this step for new installations.

Install or Update WhereScape Python Templates (For Existing Installations)

Run Script as Administrator

```
         Script 3 > install_WslPython_Modules.bat
         Script 5 > . .\import_python_templates.ps1
         Script 6 > . .\set_default_templates.ps1
```

Note : Skip this step for new installations.

Set Connection defaults for a Template Set (For Existing Installations)

```
         Script 6 > . .\set_default_templates.ps1
```

Choose either "Powershell" or "Python" when prompted.

Post Install Steps - Optional

If you used the script 'install_New_RED_Repository.ps1' then the following optional post install steps are available

Configure Connections

There were Three connections added that will optionally require your attention:

1. Connection: 'Snowflake' - This connection was setup as per parameters provided in the installation script 1
    a. open it's properties and check extended properties tab, set it up for **SNOWSQL_ACCOUNT** and **SNOWSQL_WAREHOUSE**
2. Connection: 'Database Source System' - this connection was setup as an example source connection,
    - open it's properties and set it up for a source DB in your environment
    - or you can remove it if not required
3. Connection: 'Range Table Location' - this connection is an example target connection for your Range Table DB on SQL Server,
    - if you do not intend to use the "Ranged Loading" templates then this can be removed
    - otherwise open it's properties and set it up to point to a suitable SQL DB target to store your control tables for Ranged Loading

Note:

Snowflake RED21 supports Range Table load only for Snowflake Work Range Table.

Enable Script Launcher Toolbar

There are a number of stand-alone scripts which provide some features such as "Ranged Loading", these scripts have been added to the Script Launcher menu but you will need to enable the menu toolbar item to see them.

To enable the Script Launcher menu in RED: Select menu item 'View->Toolbars->Script Launcher'

Source Enablement Pack Support

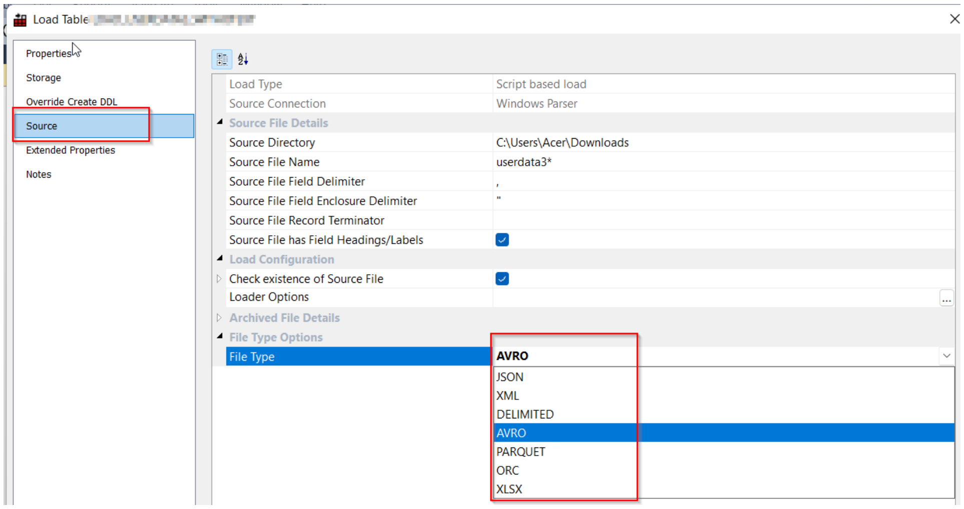
Source Pack Name	Supported By Snowflake	Supported Features	Prerequisites

Amazon S3	Yes	Bulk load to Snowflake	For Powershell Load :
			`[Net.ServicePointManager]::SecurityProtocol = [Net.SecurityProtocolType]::Tls12`
			`Install-Module -Name AWSPowerShell`

Azure DataLake Storage Gen2	Yes	Bulk load to Snowflake	Generate and add SAS token to Azure DataLake Storage Gen2 connection properties: https://docs.snowflake.com/en/user-guide/data-load-azure-config.html#option-2-generating-a-sas-token

Google Cloud Storage	Yes	Bulk load to Snowflake	1.Create Storage Integration (Snowflake)
			`CREATE OR REPLACE STORAGE INTEGRATION <STORAGE_INTEGRATION_NAME>`
			`TYPE = EXTERNAL_STAGE`
			`STORAGE_PROVIDER = GCS`
			`ENABLED = TRUE`
			`STORAGE_ALLOWED_LOCATIONS = (‘<BUCKET_PATH>’);`
			2.Create Staging Area (Snowflake)
			`USE DATABASE <DATABASE_NAME>;`
			`CREATE OR REPLACE STAGE <STAGE_NAME>`
			`url = ‘<BUCKET_PATH>’`
			`storage_integration = <STORAGE_INTEGRATION_NAME>`
			3.Get Storage Integration Desc (Snowflake)
			`DESC STORAGE INTEGRATION <STORAGE_INTEGRATION_NAME>`
			4.Copy *STORAGE_GCP_SERVICE_ACCOUNT* and create IAM on Google Cloud Console with following permissions
			`storage.buckets.get`
			`storage.objects.createstorage.objects.delete`
			`storage.objects.get`
			`storage.objects.list`
			5.Add extended properties to Google Cloud Storage Connection in RED
			`GCS_STORAGE_INTEGRATIONGCS_STAGE_AREA_NAME`

Google Drive	Yes	Downloads file From Google Drive and uploads to Snowflake Table	None

Windows Parser	Yes	1.CSV 2.Parquet 3.Avro 4.Orc 5.Excel 6.JSON 7.XML	Refer to Windows Parser Guide. Load Template ,Source Properties will have option to select parser type to load the files.  Load Table [redacted] Properties Storage Override Create DDL Source Extended Properties Notes Load Type: Script based load Source Connection: Windows Parser Source File Details Source Directory: C:\Users\Acer\Downloads Source File Name: userdata3* Source File Field Delimiter: , Source File Field Enclosure Delimiter: " Source File Record Terminator: Source File has Field Headings/Labels: ☑ Load Configuration Check existence of Source File: ☑ Loader Options: ... Archived File Details File Type Options File Type: AVRO

Troubleshooting and Tips

Run As Administrator

Press the Windows Key on your keyboard and start typing cmd.exe, when the cmd.exe icon shows up in the search list right click it to bring up the context menu, select "Run As Administrator"

Now you have an admin prompt navigate to to the folder where you have unpacked your WhereScape Red Enablement Pack to using the 'cd' command:

```
C:\Windows\system32> cd <full path to the unpacked folder>
```

Run batch (.bat) scripts from the administrator prompt by simply typing the name at the prompt and hit enter, for example:

```
C:\temp\EnablementPack>install_WslPowershell_Modules.bat
```

Run Powershell (.ps1) scripts from the administrator prompt by typing the Powershell run script command, for example:

```
C:\temp\EnablementPack>Powershell -ExecutionPolicy Bypass -File .\install_New_RED_Repository.ps1
```

Notes: In the event you can not bypass the Powershell execution policy due to group policies you can instead try "-ExecutionPolicy RemoteSigned" which should allow unsigned local scripts.

Windows Powershell Script Execution

On some systems Windows Powershell script execution is disabled by default. There are a number of workarounds for this which can be found by searching the term "Powershell Execution Policy".

Here is the most common workaround which WhereScape suggests, which does not permanently change the execution rights:

Start a Windows CMD prompt as Administrator, change directory to your script directory and run the WhereScape Powershell scripts with this command:

- cmd:>Powershell -ExecutionPolicy Bypass -File .\<script_file_name.ps1>

Restarting failed scripts

Some of the setup scripts will track each step and output the step number when there is a failure. To restart from the failed step (or to skip the step) provide the parameter "-startAtStep <step number>" to the script.

Example:

```
Powershell -ExecutionPolicy Bypass -File .\<script_file_name.ps1> -startAtStep 123
```

Tip: to avoid having to provide all the parameters again you can copy the full command line with parameters from the first "INFO" message from the beginning of the console output.

Setting the Snowflake driver LogLevel setting in the Windows Registry

The default LogLevel for Snowflake driver is set too high and can produce undesired info messages on the error stream. This in turn can cause some tasks look like failures in RED even though they were successful.

To disable the diver logging you can run the registry file provided in this Enablement Pack, or alternatively open it in a text editor to see the required changes.

> Registry file: ".\Auto Execute\Disable Snowflake ODBC Driver Logging.reg"

Note: The Snowflake ODBC DSN level setting for Tracing must also be set to 0

Python requirements for offline install

Additionally to the base Python installation being required, the WhereScape Python Template set also requires certain additional Python libraries. The install scripts uses the PIP (package manager) to download these libraries, however for offline installs you will need to install the required libraries yourself.

Required Python libraries/add-ons:

- pywin32-ctypes
- python-tds
- pywin32
- glob2
- gzip-reader
- regex
- pyodbc

If a valid RED installation can not be found

If you have Red 8.5.1.x or higher installed but the script (install_New_RED_Repository.ps1) fails to find it on you system then you are most likely running PowerShell (x86) version which does not show installed 64 bit apps by default. Please open a 64 bit version of Powershell instead and re-run the script.