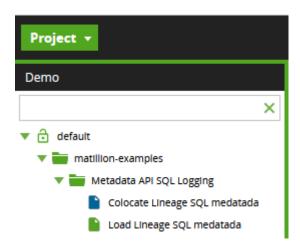


**Source**: <a href="https://exchange.matillion.com/metl/job/metadata-api-sql-logging/">https://exchange.matillion.com/metl/job/metadata-api-sql-logging/</a>

**Project Files**: <a href="https://github.com/anandjha90/MASTER-THE-ART-OF-EXTRACT-TRANSFORM-LOAD-WITH-MATILLION/tree/main/End-To-End%20Project/Matillion\_API\_Logger">https://github.com/anandjha90/MASTER-THE-ART-OF-EXTRACT-TRANSFORM-LOAD-WITH-MATILLION/tree/main/End-To-End%20Project/Matillion\_API\_Logger</a>

A pair of Matillion ETL jobs that maintain a table of the SQL executed by Matillion ETL.

The SQL information can be used to construct data lineage.



The orchestration job uses the <u>Metadata API</u> to acquire the data. Afterwards it calls the transformation job to copy the newly loaded data into a permanent table.

Features of these jobs include:

- High water mark loading with Matillion ETL
- Demonstration of bronze-to-silver layer data transformation in a medallion data architecture
- How to un-nest and relationalize semi-structured data
- Using an API Extract

# **Prerequisites**

You must first install the Metadata API Extract connector.

To run the jobs in your own environment you must supply values for the four Job Variables



Job Variable	Description
prvt_api_username	A Matillion ETL user that has permission to use the API
prvt_api_password	That user's password
prvt_hwm_timestamp	You can leave this blank and the job will still run OK. Set it to 1000000 if you wish to validate the orchestration job
prvt_staging_bucket_name	The name of an S3 bucket to be used for staging. Do not include the $s3://$ prefix nor any path

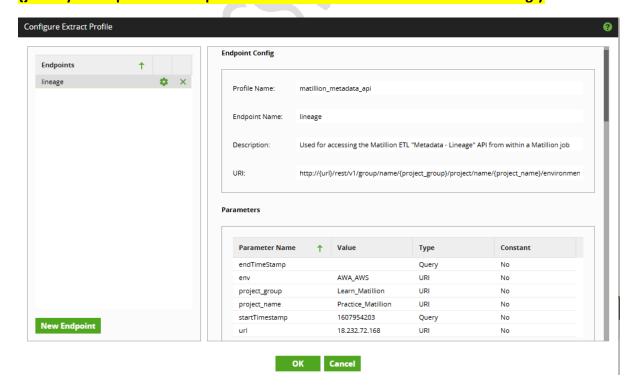
#### Steps:

Download the API Profile from the above link and export it successfully. Once done, configure Configure the **Matillion Extract Profile** as shown below by passing matillion below parameters from your **matillion url**:

https://18.232.72.168/#Learn Matillion/Practice Matillion/dev anand/

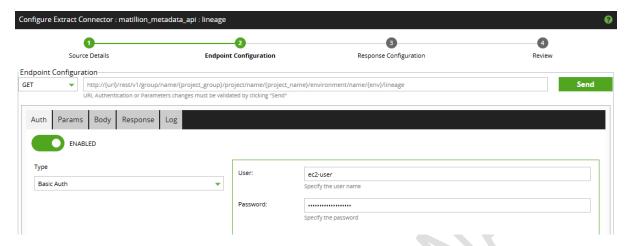
http://{url}/rest/v1/group/name/{project\_group}/project/name/{project\_name}/environ ment/name/{env}/lineage

{} -> any value placed in the parameters section will be susbstituted accordingly

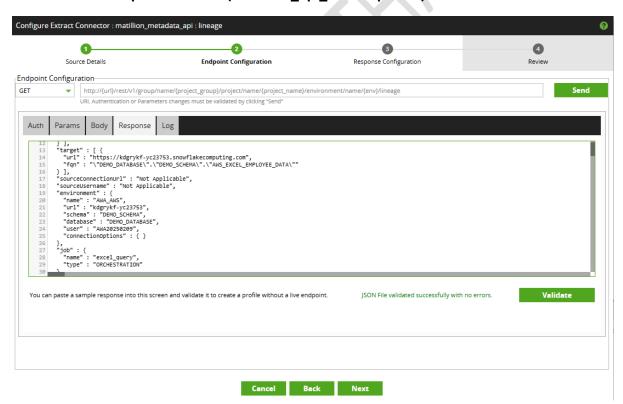




In the Auth section feed your credentials from AWS EC2 page as below where user -> ec2user(matillion\_username) and Password -> instance\_id

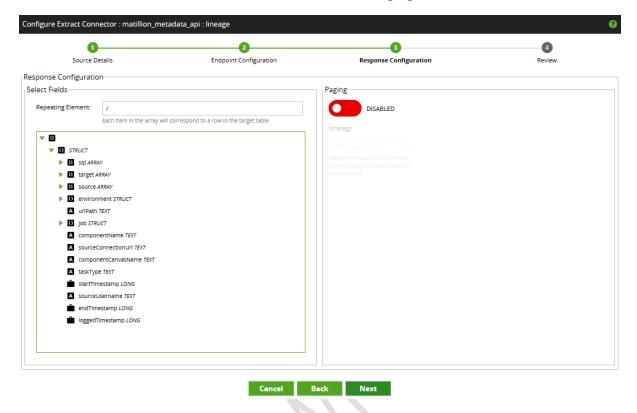


Then click on **Send** and validate the created JSON file from different users working in the same project across different version. **All such details are configured in the form of API** which we have imported before (matillion\_api\_extract profile)

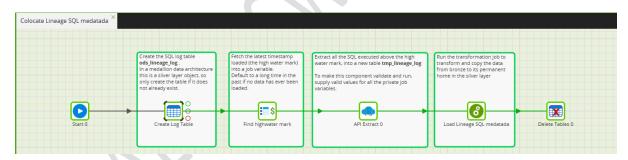




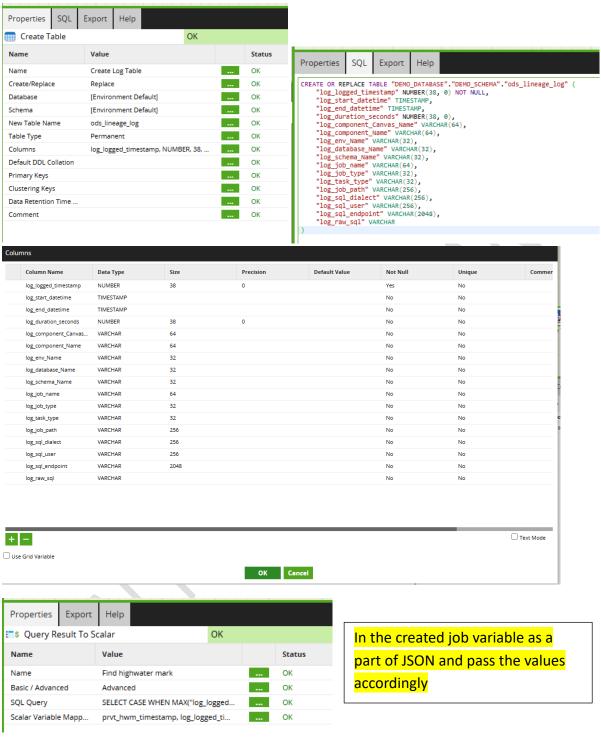
## Once Validated Click on Next and Close it. Disable the Paging



# Then export the attached JSON file and configure the parameters as shown below:

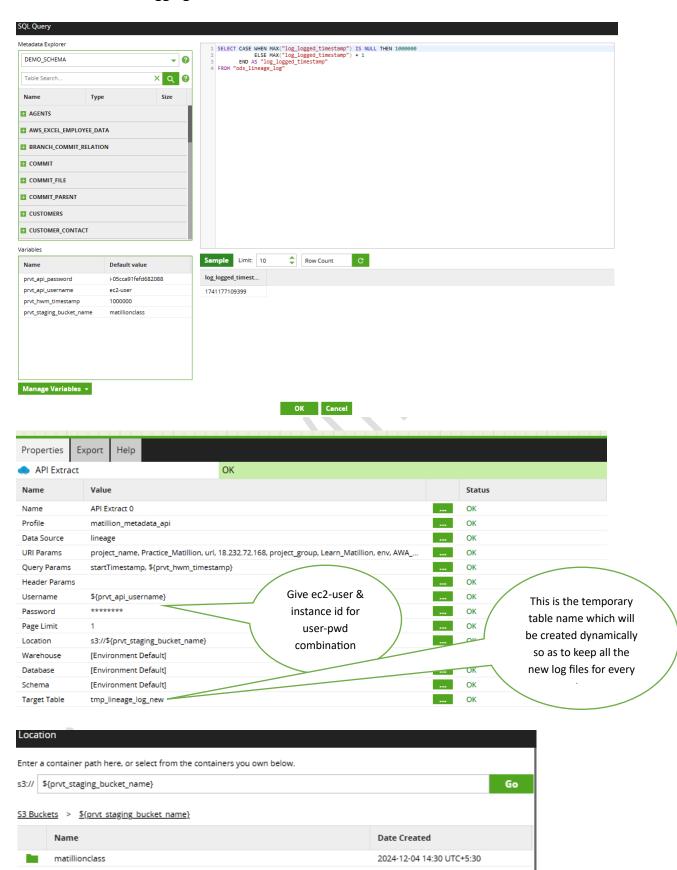






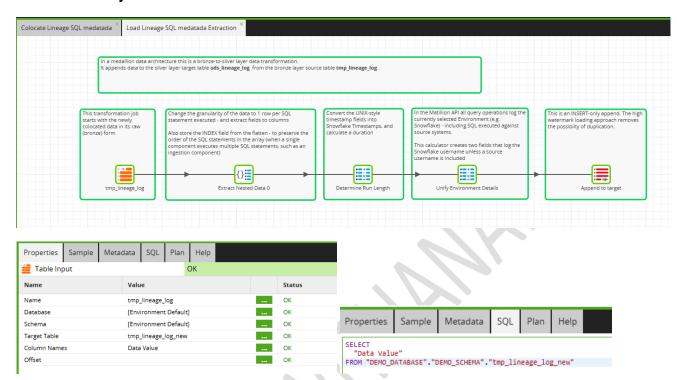








# Once orchestration job have been configured, goa head and configure below transformation job

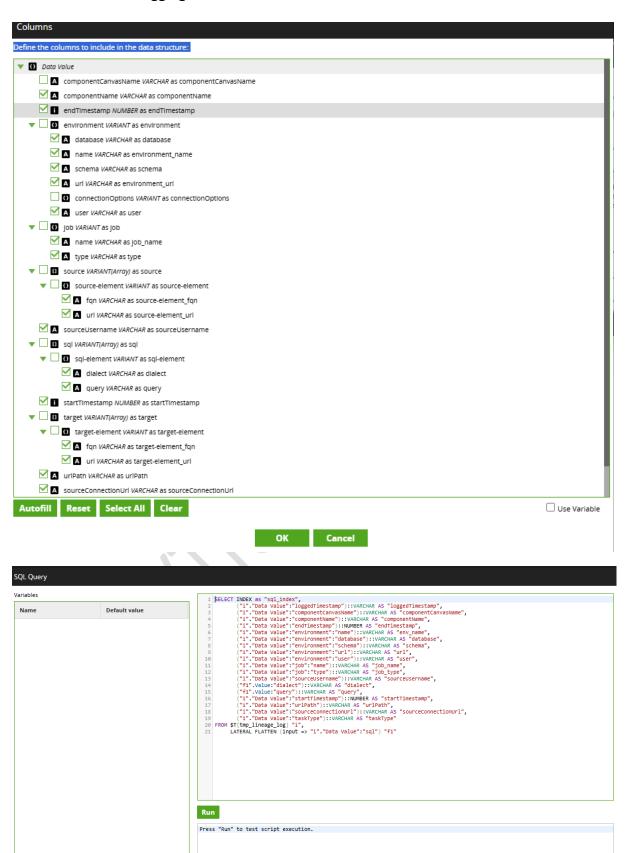


Now we will extract the parameters from the log files generated in the form of JSON through API Extract Profile as shown below using Extract Nested Data component or one can use SQL Script component and can extract the components via SCRIPT.

If you are not able to understand the SQL script refer my Youtube Lecture no 39: https://www.youtube.com/live/TbuH6SOD7uw?si=wYRrjSmyCyKGPg4K

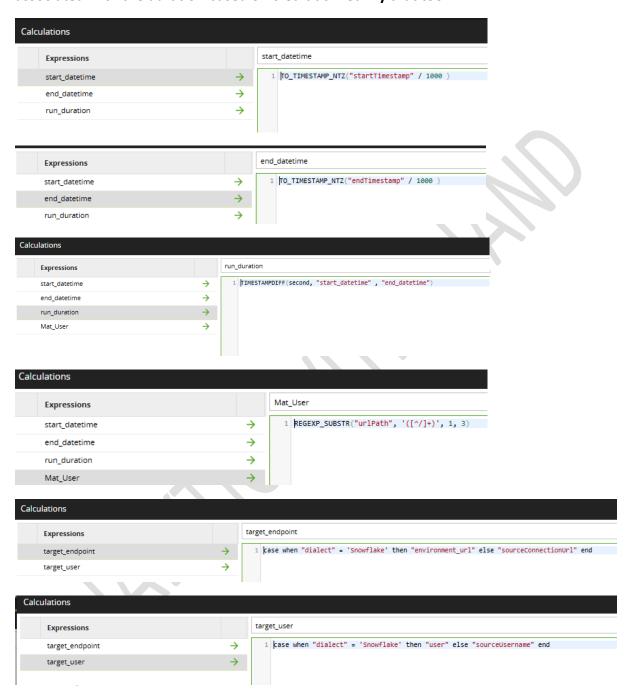








Using calculator component, extract the below attributes. You can also extract cost associated with the duration based on credit utilized. Try that too



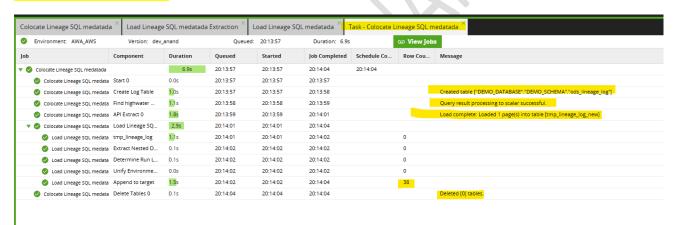


# **Table Output Properties:**



Run the orchestration job in different version by switching into it and along with that other jobs also you can run and and verify the result in snowflake

# **Final Matillion Output:**





## Final Snowflake Output:

