

Pre-requisites:

- RangerMSP SQL edition (this will not work because the JDBC data loader is not available for the non-SQL local database)
I'm all ears if someone has ideas on how to overcome this
- Java 11 (make sure the JAVA_HOME environment variable is set)
- A Neo4j Database - several options:
 - Cloud hosted (AuraDB): Free edition is limited to 200k nodes, 400k relationships: <https://neo4j.com/cloud/platform/aura-graph-database/>
 - Self hosted: (community Edition is free) <https://neo4j.com/download-center/#community>
 - If you are installing a self-hosted Neo4j database I have instructions here: <https://blog.graphcommit.com/2021/10/configuring-neo4j-4x-server-on-windows.html>
- Neo4j requires the APOC libraries (included in the auraDB)
- Advantage JDBC Driver <https://devzone.advantagedatabase.com/dz/content.aspx?Key=20&Release=19&Product=12>
- Apache Hop <https://blog.graphcommit.com/2022/06/bare-bones-configure-apache-hop.html>

You must SHARE your RangerMSP (CommitCRM) folder. You will need to grant both FILE and SHARE permissions to whichever account you run the data ingestion from.

Install Apache HOP: <https://blog.graphcommit.com/2022/06/bare-bones-configure-apache-hop.html>

Copy the installed JDBC Advantage adsjdbc.jar file from C:\Program Files (x86)\Advantage 11.10\JDBC into the /lib folder of your HOP installation.

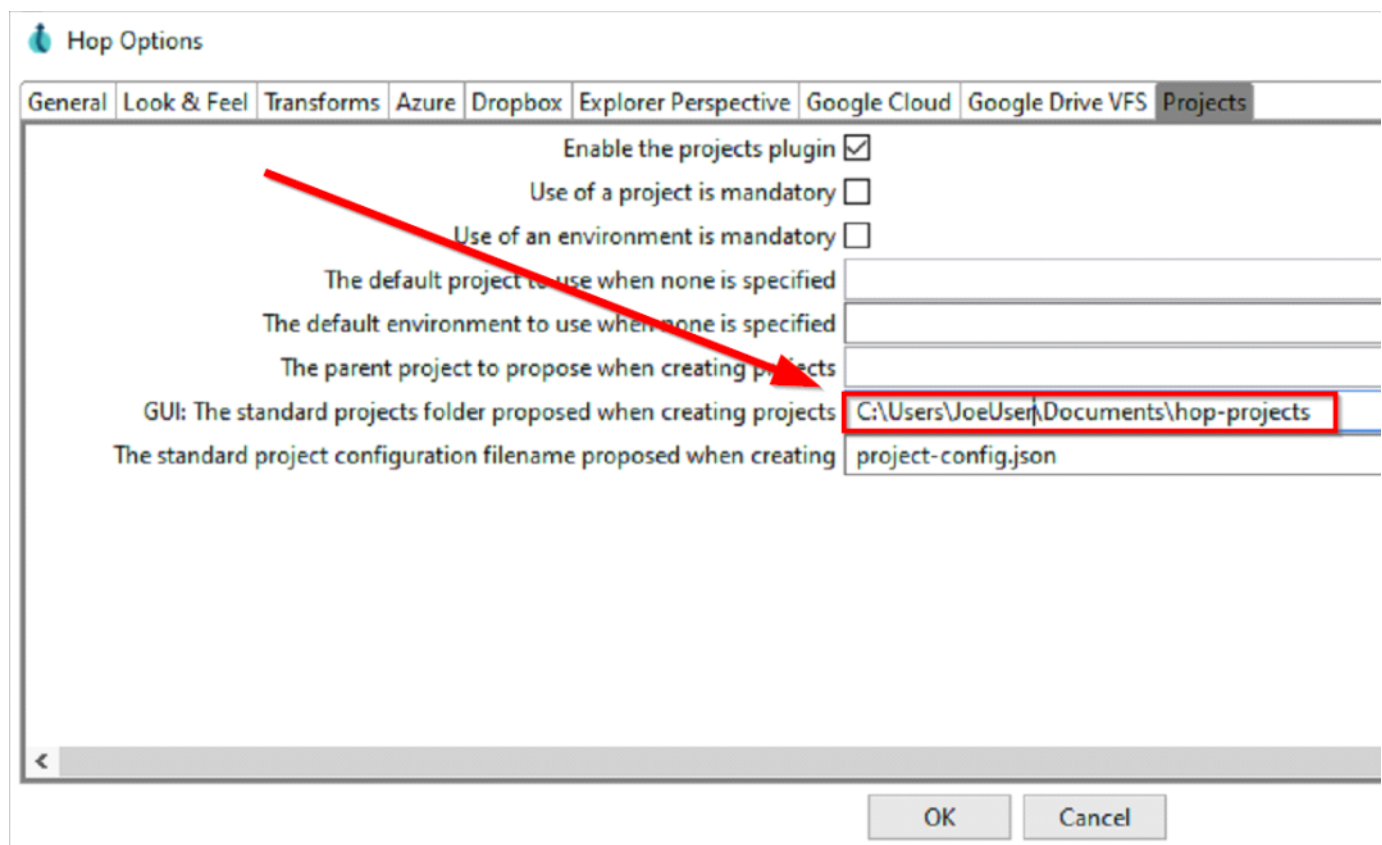
Launch HOP

Tools → Options

Select the Projects tab

Select the field "GUI: The standard project folder proposed when creating projects"

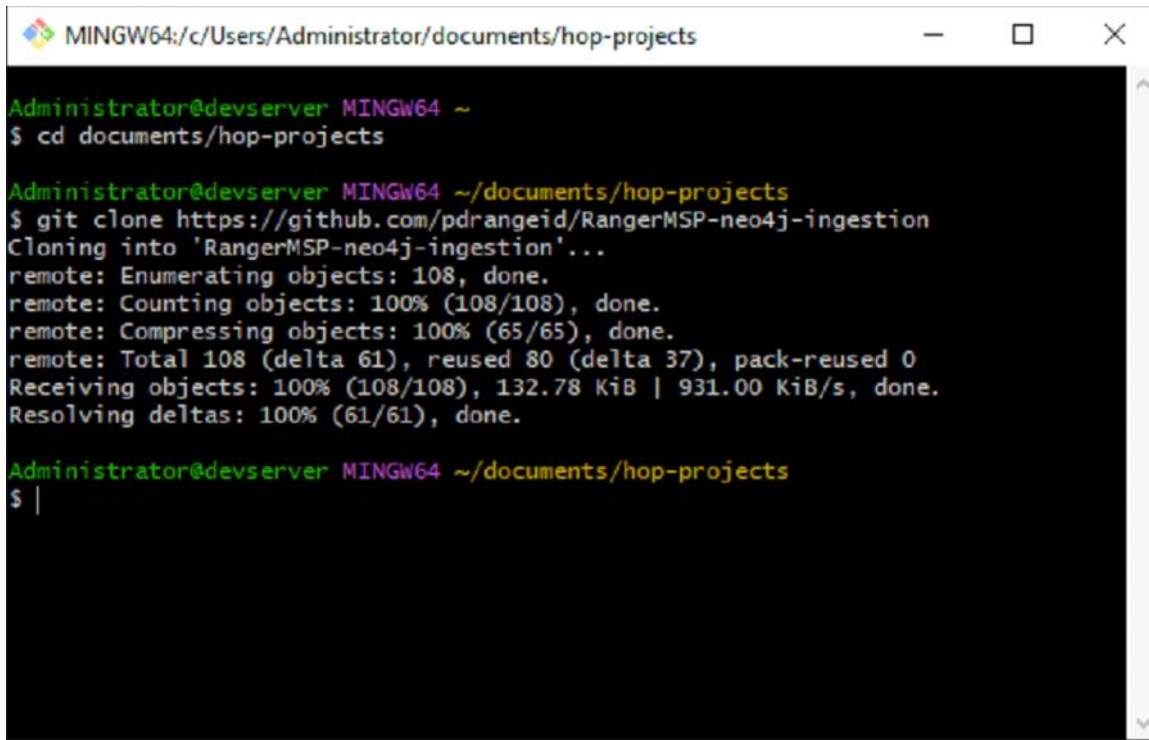
Choose a top level folder for your projects (recommend this NOT be in the folder where you installed HOP - also recommend a folder that is backed up)



Navigate to your PROJECTS folder,

git clone <https://github.com/pdrangeid/RangerMSP-neo4j-ingestion>

this will create a folder in hop-projects named 'RangerMSP-neo4j-ingestion'



```
MINGW64:/c/Users/Administrator/documents/hop-projects
Administrator@devserver MINGW64 ~
$ cd documents/hop-projects

Administrator@devserver MINGW64 ~/documents/hop-projects
$ git clone https://github.com/pdrangeid/RangerMSP-neo4j-ingestion
Cloning into 'RangerMSP-neo4j-ingestion'...
remote: Enumerating objects: 108, done.
remote: Counting objects: 100% (108/108), done.
remote: Compressing objects: 100% (65/65), done.
remote: Total 108 (delta 61), reused 80 (delta 37), pack-reused 0
Receiving objects: 100% (108/108), 132.78 KiB | 931.00 KiB/s, done.
Resolving deltas: 100% (61/61), done.

Administrator@devserver MINGW64 ~/documents/hop-projects
$ |
```

There are 3 configuration files needed for this project I have provided samples that you will need to modify for your environment:

- Environment Configuration
- (2) Datasource configs (1 for the Neo4j database, 1 for your RangerMSP Advantage SQL Database)

Download the sample environment configuration file <https://drive.google.com/file/d/1-JucoF5WemeOlkSluiWqLorU2wpJF-t/view?usp=sharing> and save it in your environments folder for this new project.

I store my environment configurations in their own folder structure SEPARATE from my projects (this segments configurations from the actual pipeline/workflow code to allow for easier upgrades in the future).

ie: /documents/hop-environments/rangerm-sp-neo4j/

Here are sample metadata config files:

Neo4j - Save this within the /metadata/neo4j-connection folder

Neo4j cloud database connection <https://drive.google.com/file/d/1-OkGlg29ArCFXk4FgqoT9vpqnFUX9UYF/view?usp=sharing>

or

Neo4j local database connection <https://drive.google.com/file/d/1-Owk7W4lpsGOa7BaZfpIMKSgqRTcs0BH/view?usp=sharing>

and

RangerMSP server connection <https://drive.google.com/file/d/1-NgLa6xXKI0pZV6-szKidrRHuCMsY18g/view?usp=sharing>

Save this within the /metadata/rdbms folder

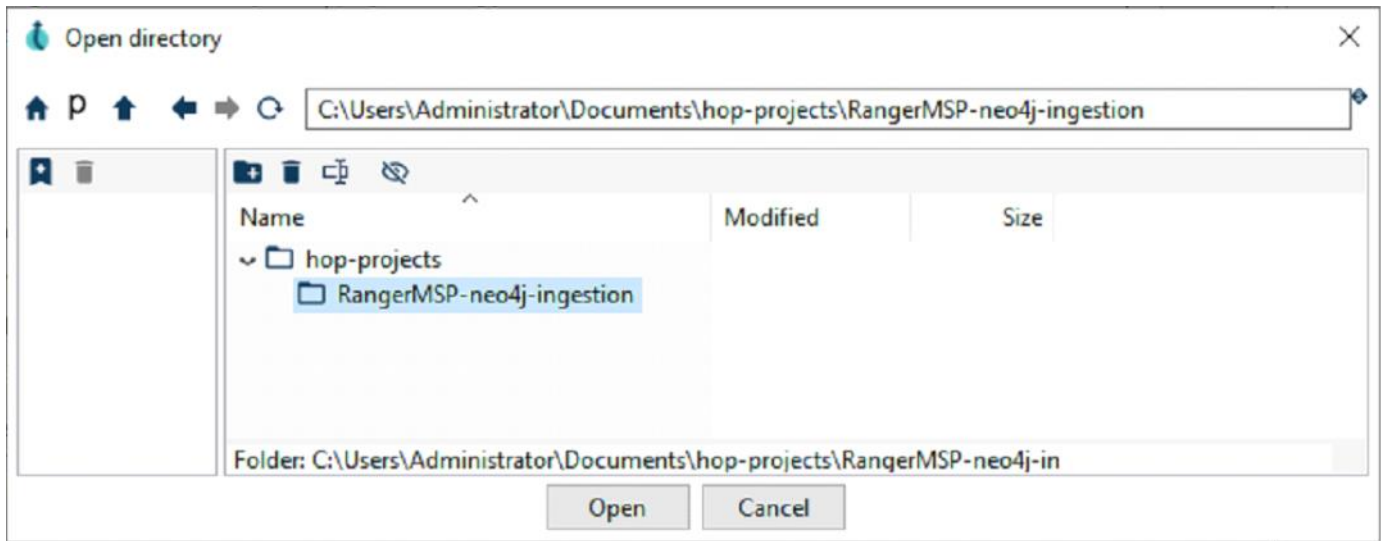
Launch HOP

Under Projects →

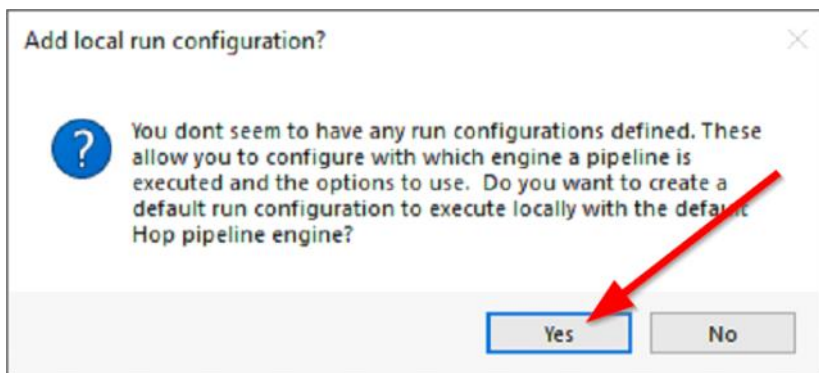
Click the p+ icon to ADD a project

Give it a name: RangerMSP-neo4j

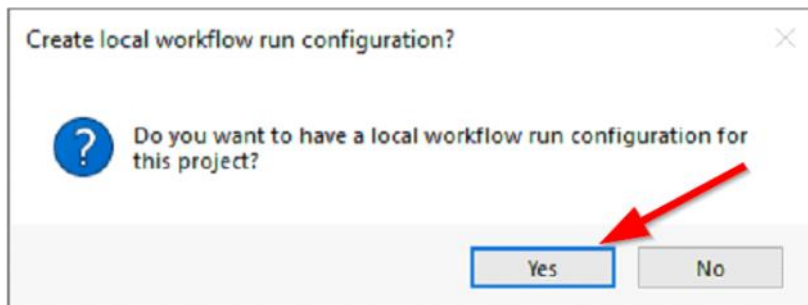
for HOME folder browse to the GIT repository you previously cloned:



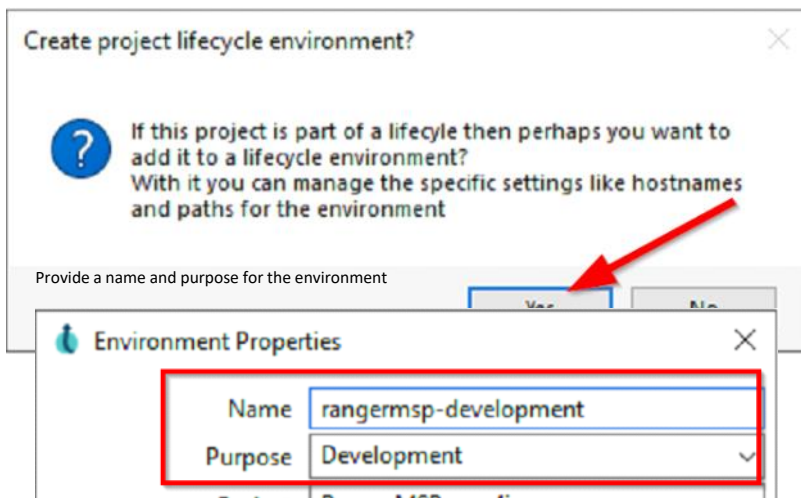
After OPENING the project - Allow HOP to Add a default run configuration

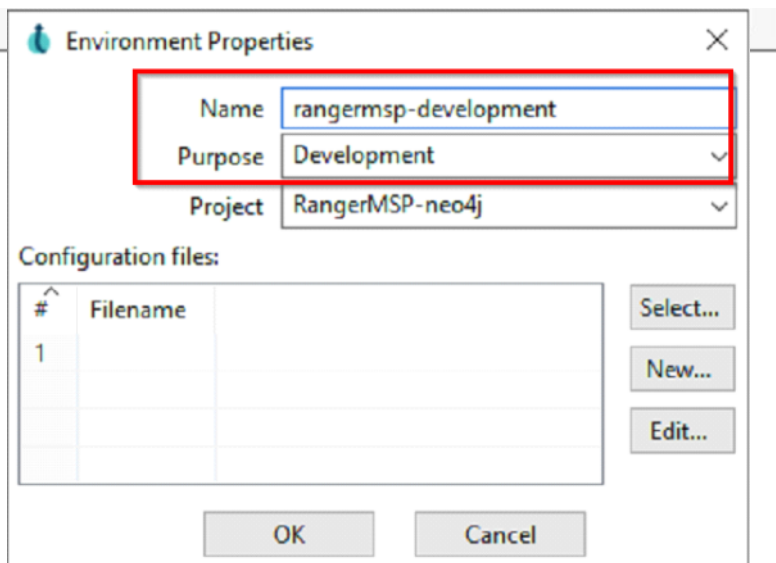


Also allow it to create a local workflow run configuration



Also allow it to Create a lifecycle environment. You will need to provide some parameters (explained below)





Environment Properties

Name: rangermisp-development

Purpose: Development

Project: RangerMSP-neo4j

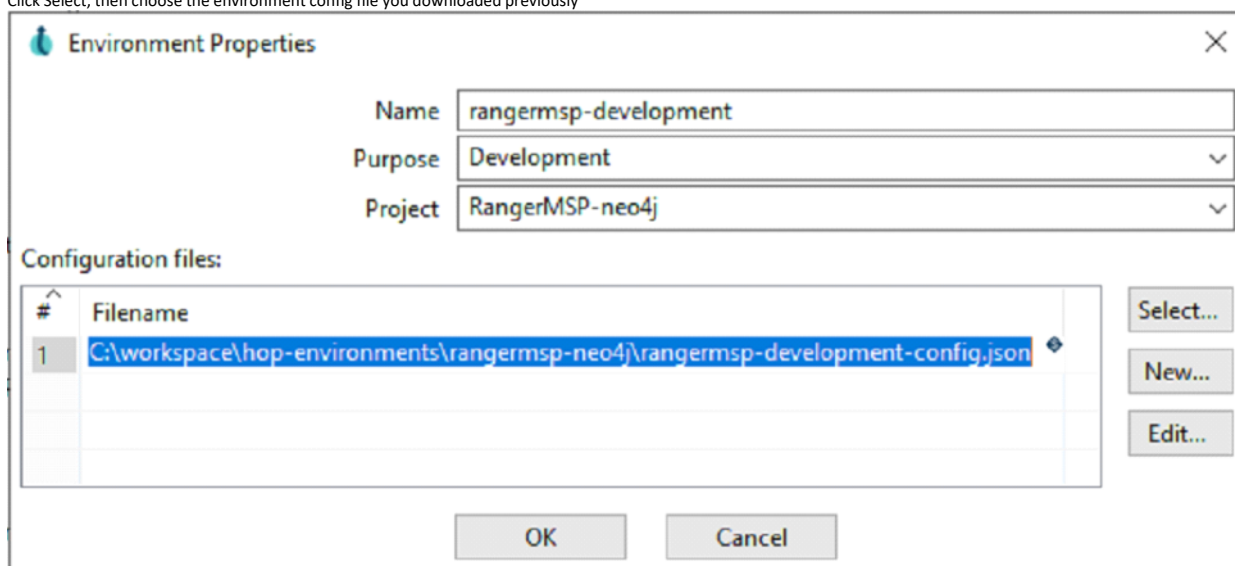
Configuration files:

#	Filename
1	

Select... New... Edit...

OK Cancel

Click Select, then choose the environment config file you downloaded previously



Environment Properties

Name: rangermisp-development

Purpose: Development

Project: RangerMSP-neo4j

Configuration files:

#	Filename
1	C:\workspace\hop-environments\rangermisp-neo4j\rangermisp-development-config.json

Select... New... Edit...

OK Cancel

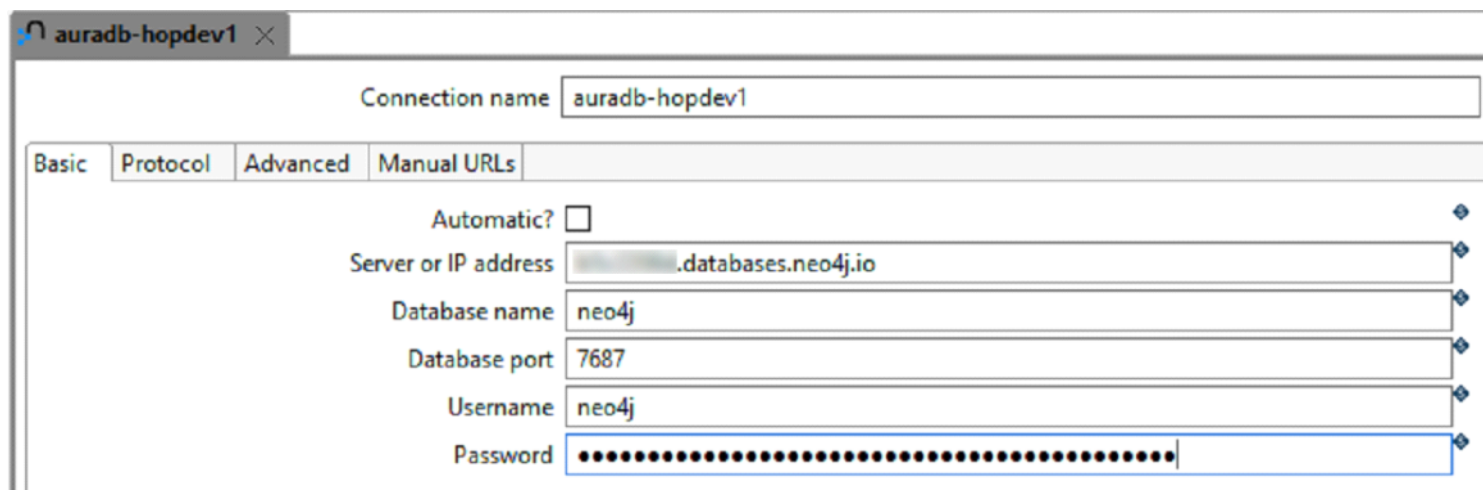
Adjust Configuration Parameters. Click "Edit". Make adjustments to any of the import parameters for your environment. This is particularly important if you are using the free cloud hosted version of the neo4j AuraDB, as it is limited to 200k nodes, and 400k relationships.

Next you will need to modify the datasource properties for both Neo4j and RangerMSP:

On the left side of HOP click the "Metadata" explorer.

Select auradb-hopdev1 by doubleclicking (this example configuration is for Neo4j Cloud hosted auradb)

Edit the server hostname, and change the user/password with credentials you were provided when signing up for your Neo4j Aura DB cloud instance.



auradb-hopdev1

Connection name: auradb-hopdev1

Basic Protocol Advanced Manual URLs

Automatic? ☐

Server or IP address: .databases.neo4j.io

Database name: neo4j

Database port: 7687

Username: neo4j

Password:

Click TEST and ensure you can connect to the neo4j database.