Oracle RDS to Aurora PostgreSQL migration lab using DMS and SCT:

The following instructions help you launch a CloudFormation stack which sets up the resources required for this lab. Once the resources are launched you are guided through the migration of a sample database from Oracle or SQLServer to Aurora PostgreSQL or MySQL using the Database Migration Service (DMS) and the Schema Conversion Tool (SCT). The resources launched by the CloudFormation stack include:

- 1. An Oracle RDS source within a VPC with all networking resources configured with a sample database installed from https://github.com/awslabs/aws-database-migrationsamples
- 2. An Aurora target within same VPC as #1 with all networking resources configured
- 3. A Windows Host with the Schema Conversion Tool pre-installed within same VPC as #1
- 4. A Database Migration Service Instance within same VPC as #1
- 5. The CloudFormation script will provision the necessary security groups with access to database resources from #3 & #4 and allow RDP access to #3

High level steps:

- 1. Log into your AWS account
- 2. Launch the CloudFormation stack
- 3. Log in to the Windows Host and install JDBC Drivers
- 4. Use the SCT to convert source Oracle schema objects to target Aurora PostgreSQL
- 5. Create a source endpoint (the source database is launched using the CloudFormation stack)
- 6. Create a target endpoint (the target database is launched using the CloudFormation stack)
- 7. Create a task to migrate the data from the source database to the target database 8. Verify the data has been migrated to the target

Create an EC2 Key Pair if Needed

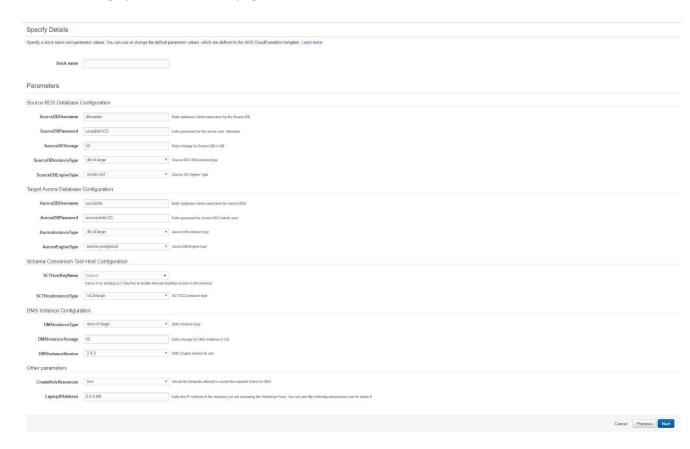
You will need an EC2 key pair in order to log into the SCT instance launched as part of your stack. If you have access to an existing one in the region where your assets are launched, you can skip this step and use it, (however you should check to be sure.) If you don't have access to an existing key pair, you will need to create one, (you can remove it once the lab is complete if you want to.) To create a new key pair:

- 1. Open the EC2 service console
- 2. In the left hand gutter under "Network & Security" click "Key Pairs"
- 3. Supply a name and click "Create"
- 4. Download or otherwise save the .pem file **Lab Instructions**:
- 1. Log in to your AWS console and go to the CloudFormation landing page https://us-east-2.console.aws.amazon.com/cloudformation/home?region=us-east-2#/stacks?filter=active

- Click create stack, select <u>'Specify an Amazon S3 template URL'</u> and launch the CloudFormation stack from this url: <u>https://s3-us-west-2.amazonaws.com/mtt-reinvent-2018/reInvent-2018-</u> DMS-Workshop-v1.4.template
- 3. Here is a screenshot of the first page -

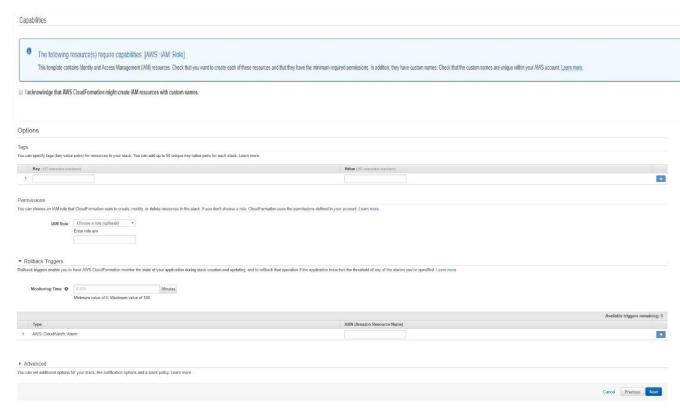


4. Hit next to bring up the stack details page



- 5. Do the following for the entries on this page
 - a. Stack name give a friendly name such as dmslab for the stack

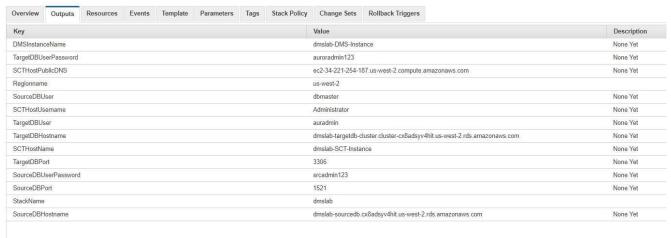
- b. **sourceDBInstanceType** Pick any instance type. Maybe pick something different than your neighbor to spread things around a bit.
- c. SourceDBEngineType choose oracle-se2
- d. **SCTHostKeyName** Select one of your existing EC2 KeyPairs to enable access to the SCT Host. *You select a value from the list!* If you do not have KeyPairs already please see the instructions above on how to create one.
- e. CreateRoleResources This should be true, unless you have run DMS before in this region and the required IAM roles have already been created. Setting this to true when the roles already exist will cause the Stack Creation to fail, and you will need to re-run the stack creation with this set to false. If you aren't sure, try it with "true" first. If it fails, try again with "false."
- f. LaptopIPAddress the IP address you are running the workshop from, this will limit access to the SCT Host only from your laptop. If you want to limit access to your SCT host enter your laptop IP address here.
- g. Finally Hit next
- 6. On the next page, simply hit next



7. On the next page, check the check box to allow CloudFormation to create IAM resources with custom names. Then click next!

^{*}Leave everything else default. Once the CloudFormation stack launch is complete, you should see 2 RDS instances in your account in the region you created the stack. Approximately, it will **take 15 to 20 minutes** for creating the Stack.

8. When the stack creation completes, look at the Outputs section. It will provide you with the names etc. of the resources needed to complete the workshop.



- 9. Once stack is available, Log in to the SCT host using RDP and the information provided in the stack Output section. The username to connect as is "Administrator". You will need to open a new tab and go to the EC2 Console to decrypt the password using the KeyPair you provided during the Stack creation.
 - a. Go to the EC2 console.
 - b. Find and select the ec2 instance you just created. The name should contain your stack name.
 - c. Click the "connect" button at the top of the console.
 - d. Supply the keypair when prompted to do so.
 - e. Launch RDP and connect as Administrator using the password just obtained.
- 10. Once on the host, open the document "JDBC Driver Install Instructions.rtf" located on the desktop and follow the instructions for installing the JDBC drivers for the database engines selected during stack creation.
 - If you are doing the Oracle -> Aurora PostgreSQL lab you will need the Oracle and PostgreSQL drivers
 - If you are doing the SQL Server -> Aurora MySQL lab you will need the SQL Server and MySQL drivers

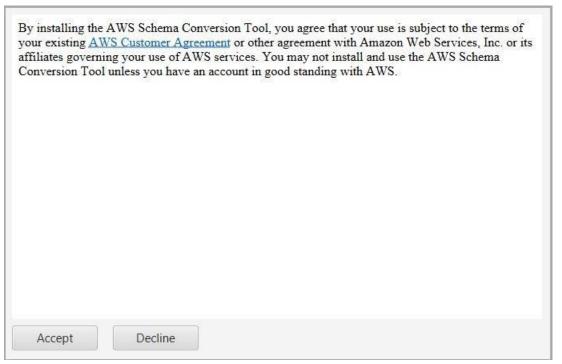
JDBC Driver Install Instructions for AWS **Schema Conversion Tool** Oracle: 1) Download the Oracle JDBC driver from the following location: https://www.oracle.com/technetwork/database/features/jdbc/jdbcucp-122-3110062.html You want the file named "ojdbc8.jar", you will also need to accept the license agreement and sign in using an OTN account. 2) Once you have downloaded this file you should move it to the following location: C:\JDBC Drivers\Oracle 3) When providing connection information for the database endpoint in the Schema Conversion Tool ensure you also specify the path to the location where you placed the JDBC driver. PostgreSQL: 1) Now you need to download the PostgreSQL JDBC Driver from the following location: https://jdbc.postgresql.org/download.html you want the file named "PostgreSQL JDBC 4.2 Driver, 42.2.5". 2) Next the PostgreSQL JDBC driver("postgresql-42.2.5.jar") should be moved to the following

Schema Conversion:

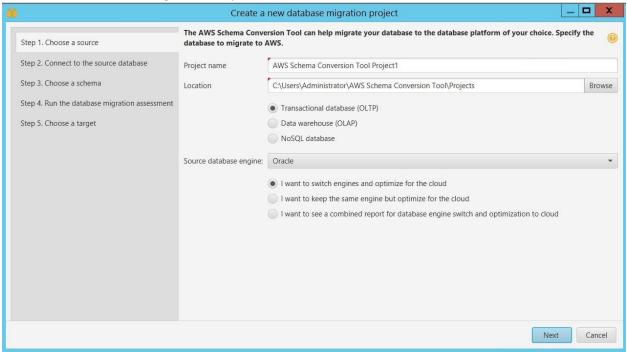
- 1. Launch SCT on your Windows Host there's a shortcut for it on your desktop
- 2. You can choose to update the SCT installation or dismiss update dialog(this is recommended for the Workshop, in all other situations you will want to download the update and install it)



3. Accept the license agreement



- 4. The "Create a new database Migration Project" Wizard will launch. If it doesn't you should select this option from the File Menu. Make sure you have the following selected and then click Next.
 - a. Transactional database (OLTP)
 - b. Source database engine: Oracle
 - c. I want to switch engines and optimize for the cloud.



5. Provide the connection information for your source RDS Database.

This information can be obtained from the Outputs section of the CloudFormation Stack.

Leave "Use SSL" unchecked

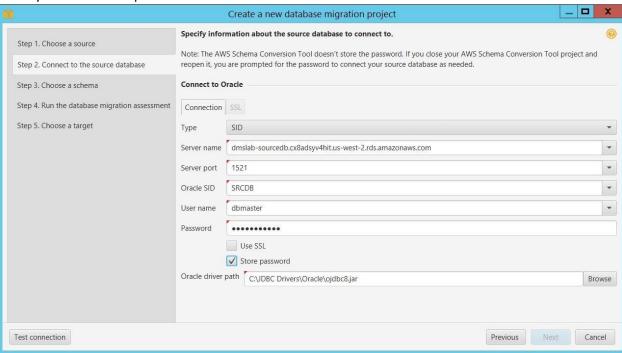
Check "Store password"

Browse for and find the jdbc driver you just installed.

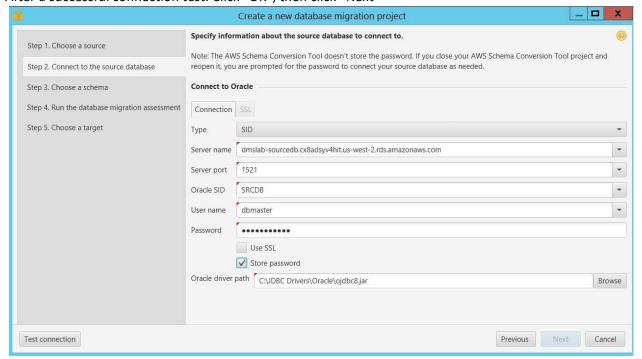
For Oracle sources the SID to use is SRCDB.

For SQL Server sources - leave Instance Name blank

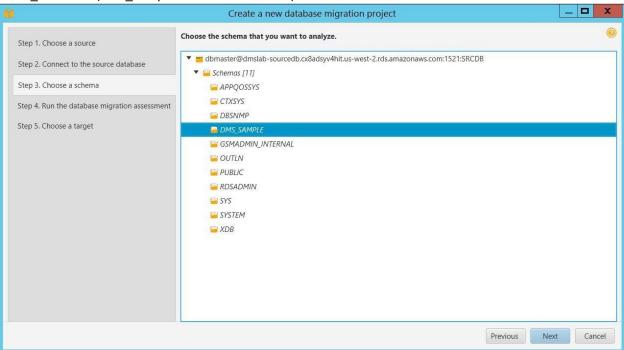
When you are finished press the "Test Connection" button.



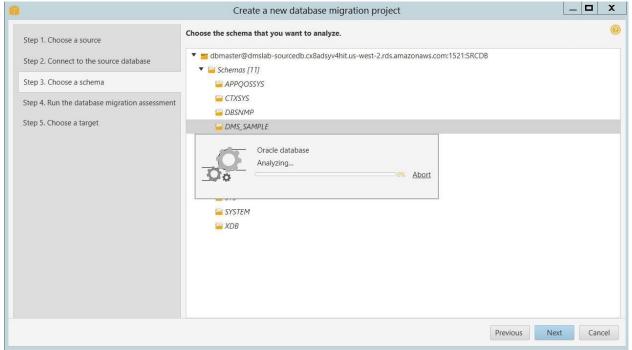
6. After a successful connection test. Click "OK", then click "Next"



7. Now select the schema to convert and click next. For this workshop the schema will be "DMS_SAMPLE" (dms_sample.dbo for SQL Server)



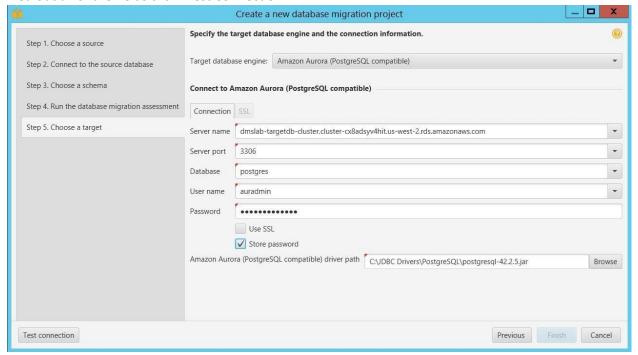
8. The Schema Conversion Tool will now perform an analysis of the DMS_SAMPLE, wait while it completes.



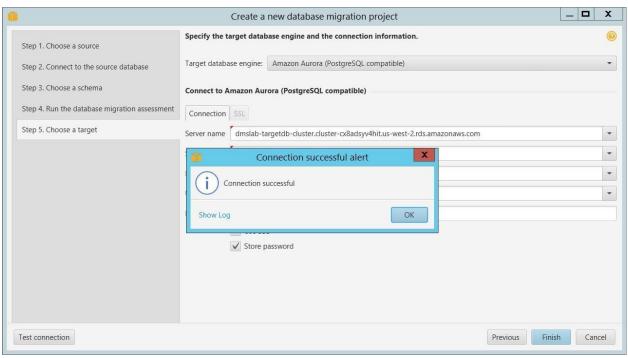
9. When the analysis is complete you will see the "Database migration assessment report". Review it as it will provide information on various target databases and how much can automatically be converted. When done reviewing the report click "Next"



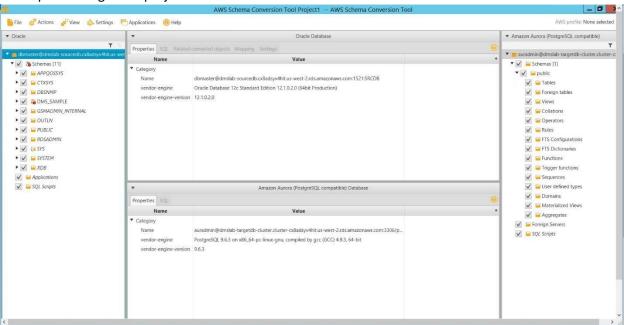
10. Now you will setup your connection to your target database. Select the correct target engine type (Amazon Aurora PostgreSQL Compatible d). For Aurora PostgreSQL compatible use postgres for the Database field, The rest of the information for filling in the option can be obtained from the outputs section of the CloudFormation stack you created. Once you have filled out all of the fields click "Test Connection"



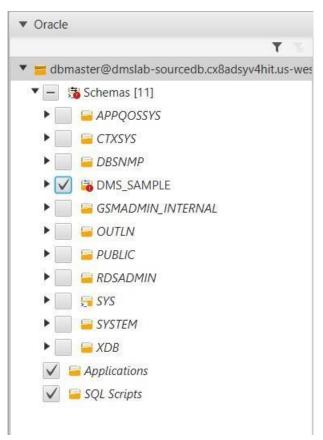
11. Once the connection test completes successfully click "OK" and then click "Finish".



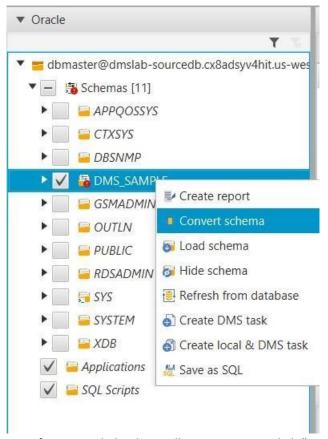
12. At this point a migration project will launch.



13. In the left hand panel deselect all of the source schemas except for DMS_SAMPLE.



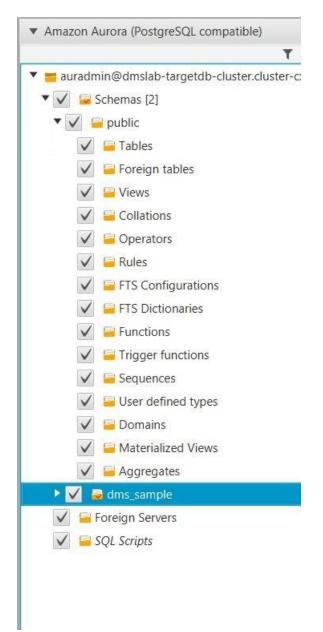
14. Now on the source schema panel. Select the DMS_SAMPLE (or dms_sample -> dbo) and right click to bring up the menu to convert the schema. You can also select convert schema from the Actions menu.



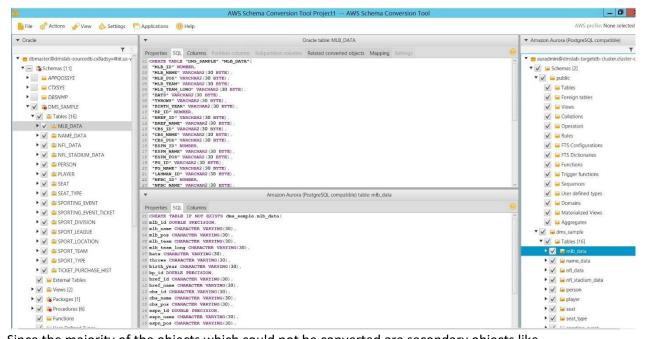
15. A confirmation dialog box will come popup. Click "Yes" to continue as there are no objects in the target database at this point.



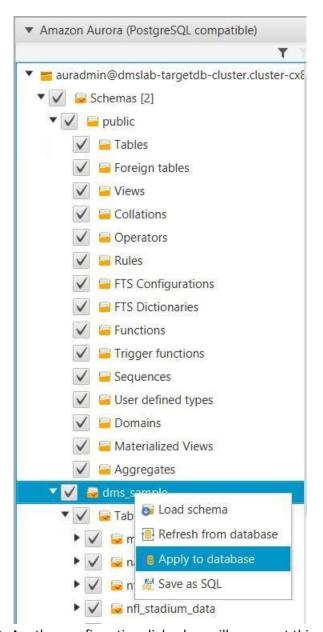
16. SCT will convert your source schema to a target schema that will be displayed in the right hand panel as either dms_sample or dms_sample_dbo.



17. At this point you can select objects in either the source or target panels and see a comparison of original and converted code and objects. You can also make changes to the converted objects as desired. Look through which objects could be automatically converted and which could not be.



18. Since the majority of the objects which could not be converted are secondary objects like functions or procedures, right click on the schema that was created (in the right panel) and select "Apply to database". This will write the converted objects to the target database.



19. Another confirmation dialog box will pop up at this point. Click "Yes" to continue.



20. The schema conversion is now complete. You are ready to move on to the next step and use the Database Migration Service to move the data from your Source to your Target.

Data Migration:

Now we will use the Database Migration Service (DMS) to move the data from the Source to the Target. Go to the AWS Console and bring up the Database Migration Service page. To migrate date from source to target we will:

- 1. Create a source endpoint
- 2. Create a target endpoint
- 3. Create a task to move data from the source endpoint to the target endpoint

Click Endpoints on the left hand side of the screen. Then click "Create Endpoint" and the create endpoint page will appear. You will do this twice, once to create a source endpoint and once to create a target endpoint.

Create endpoint

AWS DMS accesses your data sources and targets using endpoints. A source endpoint allows AWS DMS to read data from a database (on-premise or in the cloud), or from a non-database source such as Amazon S3. A target endpoint allows AWS DMS to write data to a database, or to a non-database target. We recommend that you choose "Run test" on this page, to verify that your endpoint is valid before using it in an AWS DMS task. Endpoint type*
Source Target Select RDS DB Instance 0 Endpoint identifier* e.g. ProdEndpoint - Select One -Source engine* - 0 Server name 0 - Select One -User name* 0 Password* 0 Advanced · Test endpoint connection (optional) Test your endpoint connection by selecting a replication instance within your desired VPC. After clicking "Run test", an endpoint will be created with the details provided and attempt to connect to the instance. If the connection fails, you can edit and test it again. Endpoints that aren't saved will be deleted. - Select One -Replication instance* - Select One -Refresh schemas after successful connection test Run test

Create a source endpoint:

• Select Endpoints from the menu on the left and "click the "create endpoint" button at the top.

Create endpoint

Endpoint Type: source

Note: if you check "select RDS DB Instance" you can select your particular instance from drop down menus and several entries will be populated for you.

· Endpoint identifier: dbsource

· Engine: Oracle

• Server Name: RDS Source endpoint from your account (found in your cloud formation stack outputs)

Port: 1521SSL Mode: noneUser name: dbmaster

- Password: srcadmin123 (Please note that if you changed the default settings when creating the Cloudformation stack, then you need to enter the password that you have provided)
- Database name: dms_sample
- For Oracle SID: SRCDB
- VPC: VPC launched by the CloudFormation template
- Replication Instance: choose the instance that was created as part of the stack creation

 Refresh schemas: Yes check this box

Test the connection and save it if the test is successful.

Create a target endpoint:

- Click the "create endpoint" button at the top
- Endpoint Type: Target

Note: if you check "select RDS DB Instance" you can select your particular instance from drop down menus and several entries will be populated for you.

Endpoint identifier: targetdbEngine: aurora-postgresgl

Server name: endpoint of Aurora target instance

Port: 5432SSL Mode: noneUsername: auradmin

- Password: auroradmin123 (Please note that if you changed the default settings when creating the Cloudformation stack, then you need to enter the password that you have provided)
- For Aurora PostgreSQL Database name: postgres
- VPC: VPC launched by the CloudFormation template
- Replication Instance: choose the instance that was created as part of the workshop stack

Test the connection and save it if the test is successful

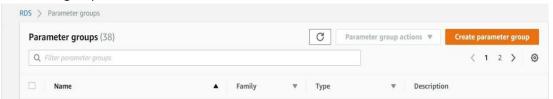
Dealing with Foreign Keys on the target.

The initial load in DMS is done table by table which means that the target tables cannot have active foreign key constraints. As we are using SCT to convert source objects into target objects, all secondary objects were created as part of the process. This means that we would need to disable all foreign key constraints on the target for the initial full load to be successful..

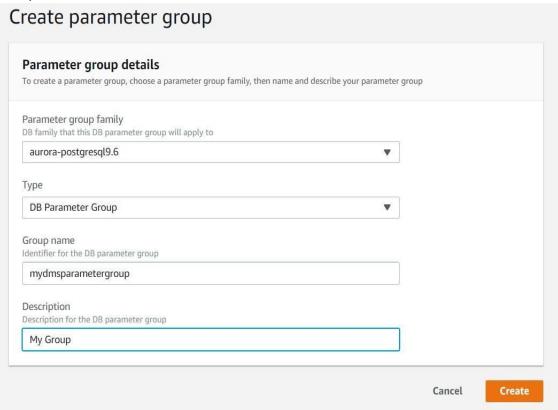
Aurora PostgreSQL

One of the ways to do this is to use the session_replication_role parameter in PostgreSQL. Triggers also have a state in PostgreSQL (Origin, replica, always or disabled). When the session_replication_role parameter is set to replica, only triggers of the state replica will be active and are fired when called. If not, the triggers remain inactive. So you need to change this parameter on your PostgreSQL instance. Here are the steps to do that.

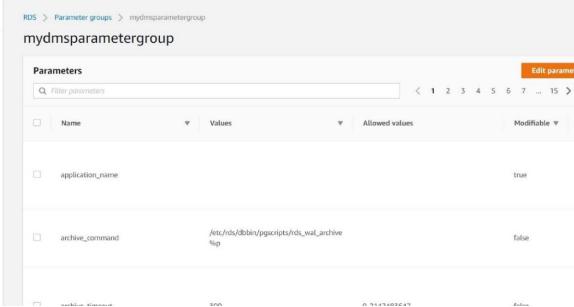
1. Go to the RDS parameter group page, and select the button to create a parameter group



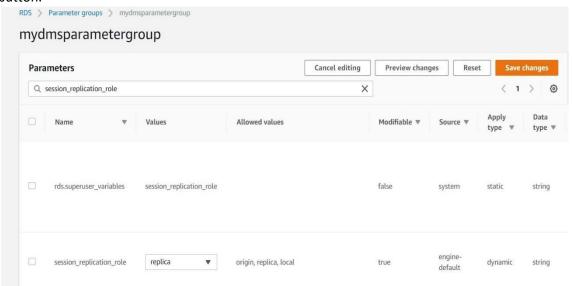
2. Now select aurora-posgresql9.6 and give your parameter group a name and description.



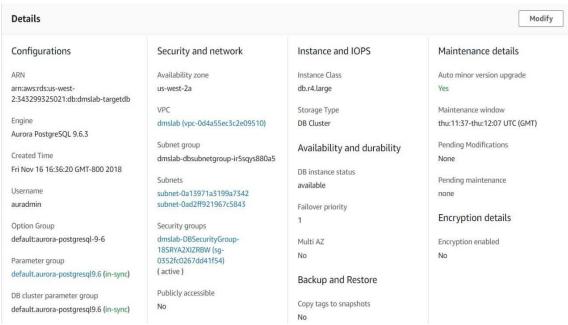
- 3. Now in the parameter group list select your parameter group
- 4. In the filter parameters box enter "session replication role"



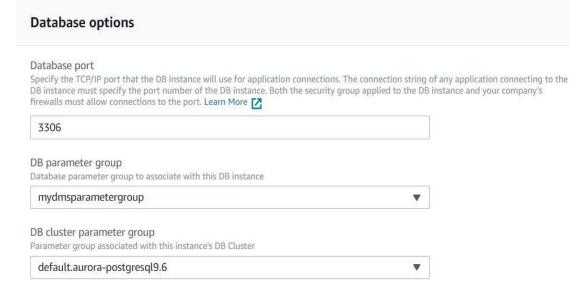
Now click the "Edit Paramters" button, select the "session_replication_role" parameter and change it's value to "replica", then click the "save changes" button.



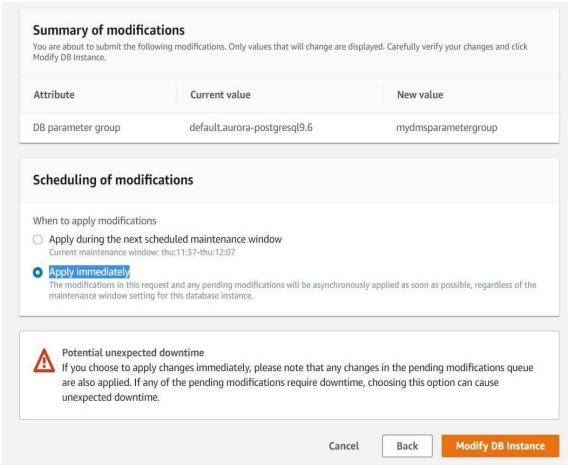
- 5. Find your target DB instance under the RDS instance page.
- 6. Select your target DB instance.
- 7. Scroll down until you see Details for the instance.
- 8. Now click on the Modify button



9. Look for the Database Options section and change the database parameter group from the default to the one you just created. Then click Continue.



10. On the next page select "Apply immediately" and then click the Modify DB instance button.



11. Now you need to select Reboot from the Instance action menu. Once the database reboot completes you can continue with the workshop

CREATE A MIGRATION TASK:

Next, we need to create a migration task. Select "Tasks" from the menu on the left in your DMS console. Click the "Create task" button in the upper left and fill out the form as follows:

- Name: Give your task a relevant name (dms-workshop or something like that)
- Replication Instance: choose the one you created (see your cloud formation stack outputs)
- Source Endpoint: choose the source endpoint you just created
- Target Endpoint: choose the target endpoint you just created
- Migration Type: Migrate existing data
- Start task on create: check this (or don't, up to you). If you don't, you'll need to start your task separately.

^{**} Target table preparation mode: Do Nothing (Important for PostgreSQL targets)

A task can contain one or more table mappings which define what data is moved from the source to the target. If a table does not exist on the ta

Task name*	dms-test-task	0
Replication instance*	test-dms-lab-instance - vpc-7cc65715 ▼	0
Source endpoint*	cf-source ▼	0
Target endpoint*	cf-target ▼	0
Migration type*	Migrate existing data and replicate on ▼	0
	The same of the sa	ANGENTAL STATE OF THE STATE OF

Why a do-nothing task?

Since the initial load in DMS is done table by table which means that the target tables cannot have active foreign key constraints. As we are using SCT to convert source objects into target objects, all secondary objects were created as part of the process. This means that we would need to disable all foreign key constraints on the target for the initial full load to be successful.

One of the ways to do this is to use the session_replication_role parameter in PostgreSQL. Triggers also have a state in PostgreSQL (Origin, replica, always or disabled). When the session_replication_role parameter is set to replica, only triggers of the state replica will be active and are fired when called. If not, the triggers remain inactive. We have already setup the parameter group on the target to set this role to replica which means all foreign key constraints (innately triggers in the origin state) will not be active. However, PostgreSQL has a failsafe mechanism of not letting a table truncate even with this role set. As we are using prepopulated tables on the target and cannot truncate the table, we are using do_nothing for the target table prep mode.

More details in this awesome blog post -

http://blog.endpoint.com/2015/01/postgressessionreplicationrole.html

- Stop task after full load completes: don't stop
- Include LOB columns in replication: Accept the default
- Enable validation: up to you
- Enable logging: YES! (Check this box)

Target table preparation mode*		Do nothing	0	
		Drop tables on target		
		Truncate		
Stop task after full load completes*	•	Don't stop	0	
		Stop Before Applying Cached Changes		
		Stop After Applying Cached Changes		
Include LOB columns in replication*		Don't include LOB columns	0	
		Full LOB mode		
	•	Limited LOB mode		
Max LOB size (kb)*	32	2	0	
Enable logging	~			

<u>Table Mappings – Selection rules(For Oracle to Aurora PostgreSQL Migrations):</u>

- 1. Select DMS_SAMPLE from the drop-down box Click the "Add selection rule" button.
- 2. Click on "add transformation rule" -> Choose schema for target dropdown -> Schema name is DMS_SAMPLE -> Choose "make lower case" for action & click add transformation rule.
- 3. Click on "add transformation rule" -> Choose table for target dropdown -> Schema name is DMS_SAMPLE and table name is % -> Choose "make lower case" for action & click add transformation rule.
- 4. Click on "add transformation rule" -> Choose column for target dropdown -> schema name is DMS_SAMPLE, table name is % -> column name is % -> Choose "make lower case" for action & click add transformation rule.



Note – We are doing 2, 3 and 4 above as the schema has been pre-created on the target with lowercase (PostgreSQL convention). Your task json should look like this –

```
"rules": [
               "rule-type": "selection",
              "rule-id": "1",
              "rule-name": "1",
               "object-locator": {
                      "schema-name": "DMS SAMPLE",
                      "table-name": "%"
               "rule-action": "include"
       },
       {
              "rule-type": "transformation",
              "rule-id": "2",
              "rule-name": "2",
              "rule-target": "schema",
               "object-locator": {
                      "schema-name": "DMS SAMPLE"
               "rule-action": "convert-lowercase"
       },
              "rule-type": "transformation",
              "rule-id": "3",
              "rule-name": "3",
              "rule-target": "table",
               "object-locator": {
                      "schema-name": "DMS_SAMPLE",
                      "table-name": "%"
               "rule-action": "convert-lowercase"
       },
```

Start Moving Data

→ Click create task now and let the task run.

<u>Note:</u> Task may fail for 3 tables while it should go well for other tables in the DMS_SAMPLE (or dbo) schema. As a bonus step you can investigate why this is happening and fix the problem before reloading failed tables.

Checking your results:

Links to downloads for PGAdmin have been added to the windows desktop you used to run the SCT.

Clean up CloudFormation stack:

- 1. After testing, stop the task and delete the task.
- 2. Delete the source/target endpoints.
- 3. Delete CloudFormation stack from the CloudFormation console.