



Amazon Web Services Data Engineering Immersion Day

Database Migration Services Lab
March 2020

Table of Contents

<i>Introduction</i>	<i>2</i>
<i>Create the Subnet Group</i>	<i>3</i>
<i>Create the Replication Instance</i>	<i>4</i>
<i>Create the DMS Source Endpoint</i>	<i>6</i>
<i>Create the Target Endpoint</i>	<i>8</i>
<i>Create a task to perform the initial full copy</i>	<i>10</i>
<i>Create the CDC endpoint to replicate ongoing changes (Optional)</i>	<i>15</i>
<i>Create a task to perform the ongoing replication (Optional)</i>	<i>16</i>

Introduction

This lab will give you an understanding of the AWS Database Migration Service (AWS DMS). You will migrate data from an existing Amazon Relational Database Service (Amazon RDS) Postgres database to an Amazon Simple Storage Service (Amazon S3) bucket that you create.



In this lab you will complete the following tasks:

1. Create a subnet group within the DMS Lab VPC
2. Create a DMS replication instance
3. Create a source endpoint
4. Create a target endpoint
5. Create a task to perform the initial migration of the data.

Optionally, you can add ongoing replication of data changes on the source: ***(Only one of the DMS replication instances will enable this feature.)***

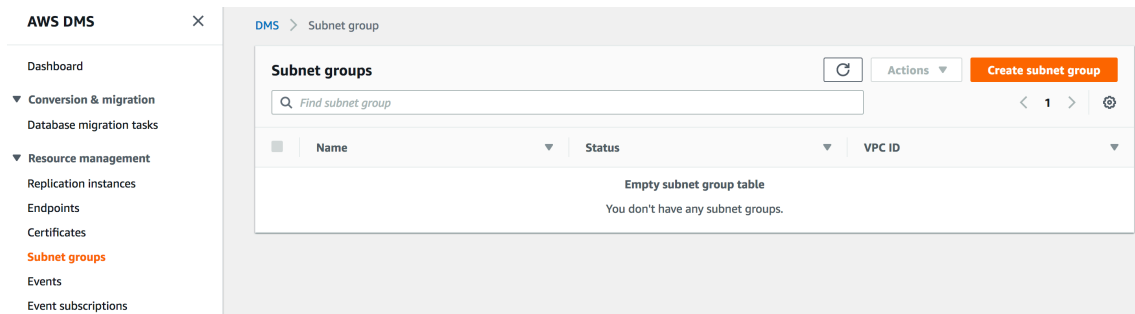
6. Create target endpoint for CDC files to place these files in a separate location than the initial load files
7. Create a task to perform the ongoing replication of data changes

Your instructor has already created and populated the RDS Postgres database that you will use as your source endpoint in this lab. If you have deployed instructor lab, then get RDS endpoint from there.

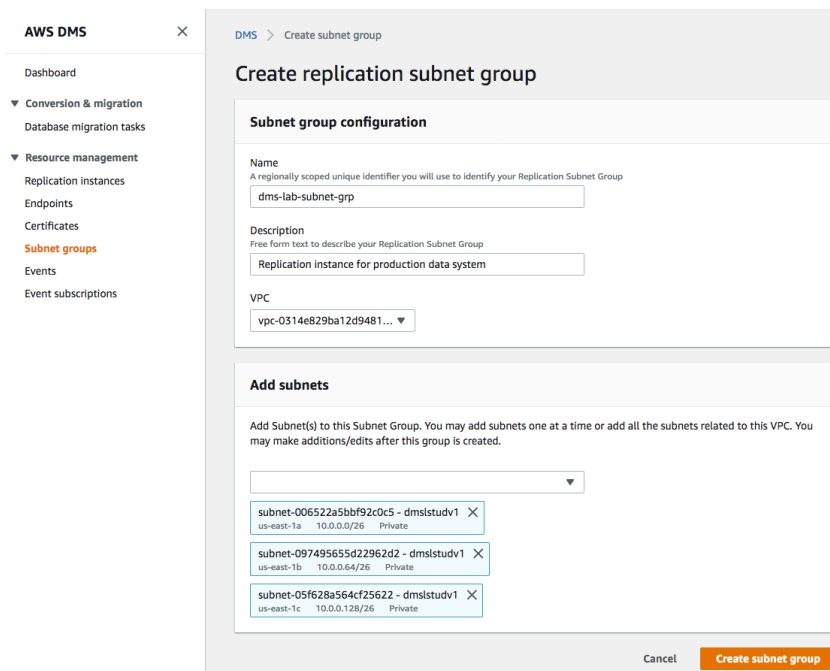
Labs are also available in GitHub - <https://github.com/aws-samples/data-engineering-for-aws-immersion-day>

Create the Subnet Group

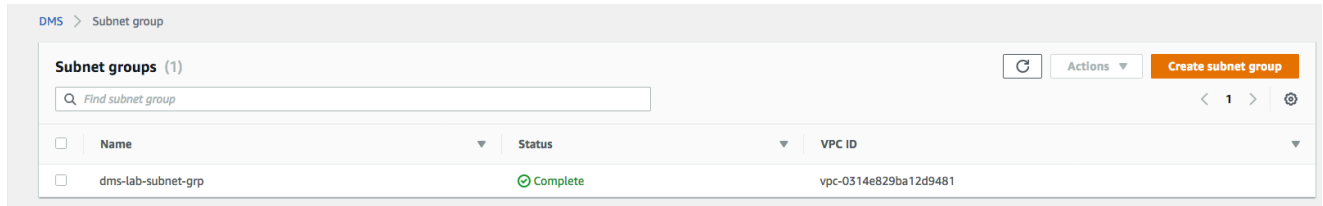
1. On the DMS console, select **Subnet Groups**.



2. Click **Create subnet group**.
 - a. In the Identifier box, type a descriptive name that you will easily recognize (e.g., "dms-lab-subnet-grp").
 - b. In the Description box, type an easily recognizable description (e.g., "Replication instance for production data system").
 - c. For VPC, select the name of the VPC that you created earlier with AWS CloudFormation template. VPC name ending with "dmslstudv1". The subnet list populates in the Available Subnets pane.
 - d. Select as many subnets as you want and click Add. The selected subnets move to the Subnet Group pane. Note: DMS requires at least two separate availability zones to be selected.

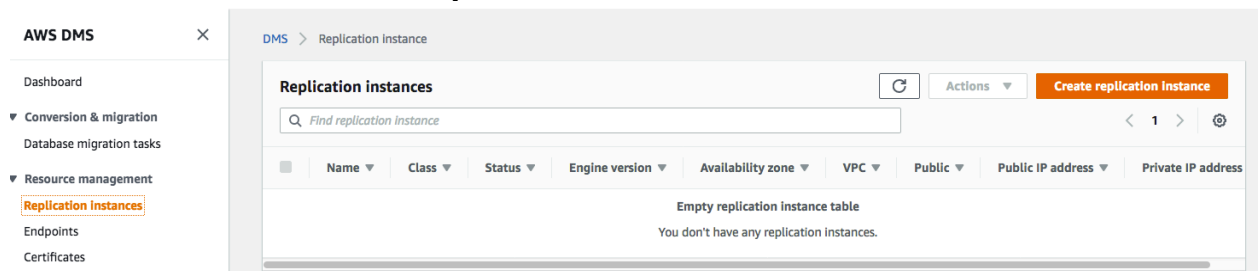


3. Click Create subnet group
4. On the DMS console, the subnet group status displays Complete.



Create the Replication Instance

1. On the DMS console, select **Replication instances**.



2. Click **Create replication instance**.
 - a. For **Name**, type a name for the replication instance that you will easily recognize. (e.g., "DMS-Replication-Instance").
 - b. For **Description**, type a description you will easily recognize. (e.g., "DMS Replication Instance").
 - c. For **Instance class**, choose **dms.t2.medium**
 - d. Select **Engine version** as "3.3.1"
 - e. For VPC, select the name of the VPC that you created earlier with AWS CloudFormation template. VPC name ending with "dmslstudv1".

Database Migration Services Lab

AWS DMS

×

Dashboard

▼ Conversion & migration

Database migration tasks

▼ Resource management

Replication instances

Endpoints

Certificates

Subnet groups

Events

Event subscriptions

What's new

Notifications

Replication instance configuration

Name

The name must be unique among all of your replication instances in the current AWS region.

Replication-Instance

Replication instance name must not start with a numeric value

Description

DMS-Replication-Instance

The description must only have unicode letters, digits, whitespace, or one of these symbols: _:/=+-@. 1000 maximum character.

Instance class

Choose an appropriate instance class for your replication needs. Each instance class provides differing levels of compute, network and memory capacity.

dms.t2.medium

Billing is based on [DMS pricing](#).

Engine version

Choose an AWS DMS version to run on your replication instance.

3.3.1

Allocated storage (GiB)

Choose the amount of storage space you want for your replication instance. AWS DMS uses this storage for log files and cached transactions while replication tasks are in progress.

50

VPC

Choose an Amazon Virtual Private Cloud (VPC) where your replication instance should run.

vpc-0537f7268d522baf3 - dmslstudv1

☐ Multi AZ

If you choose this option, AWS DMS will perform a multi-AZ deployment, with a primary instance in one availability zone (AZ) and a standby instance in another AZ. This configuration provides a highly available, fault-tolerant replication environment.

- f. Click **Advanced** to expand the section.
- g. Select the security group with **student-sgdefault** in the name.

AWS DMS

×

Dashboard

▼ Conversion & migration

Database migration tasks

▼ Resource management

Replication instances

Endpoints

Certificates

Subnet groups

Events

Event subscriptions

☒ Publicly accessible

If you choose this option, AWS DMS will assign a public IP address to your replication instance, and you'll be able to connect to databases outside of your Amazon VPC.

▼ Advanced security and network configuration

Replication subnet group

Choose a subnet group for your replication instance. The subnet group defines the IP ranges and subnets that your replication instance can use within the Amazon VPC you've chosen.

dms-lab-subnet-grp

Availability zone

Choose an availability zone (AZ) where you want your replication instance to run. The default is "No preference", meaning that AWS DMS will determine which AZ to use.

No Preference

VPC security group(s)

Choose one or more security groups for your replication instances. The security group(s) specify inbound and outbound rules to control network access to your replication instance.

Use default

dmslab-student-sgdefault-G2VY06TNTMNZ

KMS master key

[Info](#)

(Default) aws/dms

Account

Description

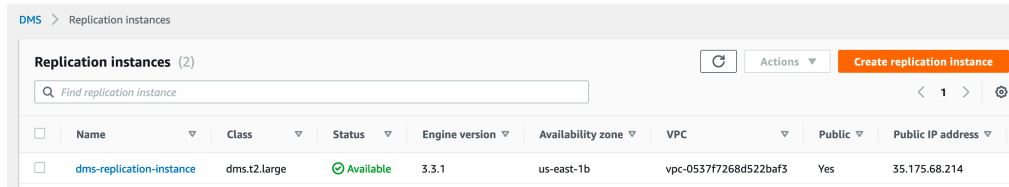
Key ARN

► Maintenance

Cancel Create

3. Click **Create**.

- The DMS console displays **creating** for the instance status. When the replication instance is ready, the status changes to **available**. While replication instance is spinning up, you can proceed to next step for DMS endpoint creation.



DMS > Replication instances

Replication instances (2)

Find replication instance

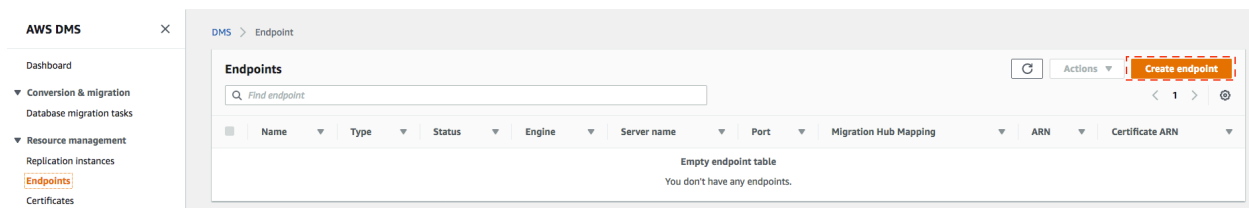
Actions Create replication instance

	Name	Class	Status	Engine version	Availability zone	VPC	Public	Public IP address
<input type="checkbox"/>	dms-replication-instance	dms.t2.large	Available	3.3.1	us-east-1b	vpc-0537f7268d522ba3	Yes	35.175.68.214

Create the DMS Source Endpoint

Please proceed to create your endpoints, without waiting for the step above.

- On the DMS console, select **Endpoints**



AWS DMS

Dashboard

Conversion & migration

Database migration tasks

Resource management

Replication instances

Endpoints

Certificates

DMS > Endpoint

Endpoints

Find endpoint

Actions Create endpoint

	Name	Type	Status	Engine	Server name	Port	Migration Hub Mapping	ARN	Certificate ARN
Empty endpoint table You don't have any endpoints.									

- Click **Create endpoint**.
 - select **Source endpoint** type.
 - For **Endpoint identifier**, select your easily recognized name (e.g. "rds-source-endpoint")
 - For Source engine, select **"postgres"**.
 - Enter the **Server name** provided by your instructor, or if you ran instructor lab then take recorded endpoint from the instructor pre-lab. (e.g. "dmslabinstance.ccla1oozkrry.us-east-1.rds.amazonaws.com")
 - For Port, enter **"5432"**.
 - For SSL mode, choose **"none"**.
 - For User name, type **"master"**.
 - For Password, type **"master123"**.
 - For Database name, type **"sportstickets"**.

Database Migration Services Lab

The screenshot shows the 'Create endpoint' page in the AWS DMS console. The left sidebar lists navigation options: Dashboard, Conversion & migration, Resource management, Endpoints (highlighted), Certificates, Subnet groups, Events, and Event subscriptions. The main content area is titled 'Create endpoint' and contains two sections: 'Endpoint type' and 'Endpoint configuration'. In the 'Endpoint type' section, 'Source endpoint' is selected. The 'Endpoint configuration' section includes fields for 'Endpoint identifier' (prodendpoint-postgre), 'Source engine' (postgres), 'Server name' (dmslabinstance.c1ny3gywsvdz.us-east-1.rds.amazonaws.com), 'Port' (5432), 'Secure Socket Layer (SSL) mode' (none), 'User name' (master), 'Password' (masked), and 'Database name' (sportstickets).

3. Click **Create endpoint** to create the endpoint. When available, the endpoint status changes to **active**.
4. Check the **replication instance** created previously. Make sure the status is **available**.

The screenshot shows the 'Replication instances' page in the AWS DMS console. The left sidebar is the same as the previous screenshot. The main content area is titled 'Replication instances (1)' and contains a table with one instance. The table has columns: Name, Class, Status, Engine version, Availability zone, and VPC. The instance 'dms-replication-instance' is shown with status 'Available'.

Name	Class	Status	Engine version	Availability zone	VPC
dms-replication-instance	dms.t2.medium	Available	3.3.1	us-east-1b	vpc-0f4679

5. Select your newly created source **endpoint**, and choose **Test connection** on the **Actions** drop-down list.

The screenshot shows the 'Endpoints' page in the AWS DMS console. The left sidebar is the same as the previous screenshots. The main content area is titled 'Endpoints (1)' and contains a table with one endpoint. The table has columns: Name, Type, Status, Engine, and Server name. The endpoint 'srcdb' is shown with status 'Active'. The 'Actions' dropdown menu is open, showing options: Modify, Test connection, and Delete. The 'Test connection' option is highlighted.

Name	Type	Status	Engine	Server name
srcdb	Source	Active	PostgreSQL	dmslabinstance.ctmbri3fwuo4.us-east-1.rds.amazonaws.com

- Click **Run test**. This step tests connectivity to the source database system. If successful, the message “Connection tested successfully” appears.

DMS > Endpoints > rds-source-endpoint > Test endpoint connection

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.

Replication instance
A replication instance performs the database migration

dms-replication-instance

Run test

Endpoint identifier	Replication instance	Status	Message
rds-source-endpoint	dms-replication-instance	successful	

Back

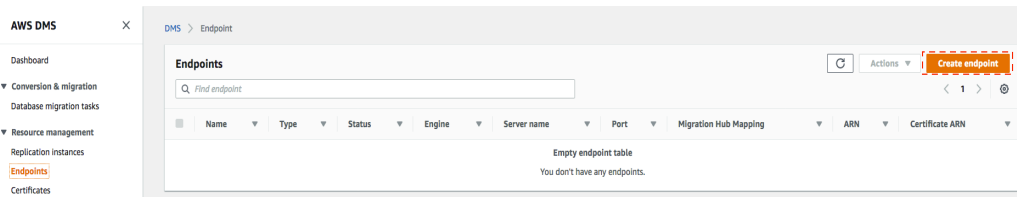
Note for instructor: In case the test fail, two possible reasons: a) RDS security group restriction. To resolve this, revisit instructor lab and follow section “Changing RDS security Group” to troubleshoot. b) RDS is not accessible because its disk is full. The solution is VACUUM FULL on the DB, or free up some disk space by deleting unnecessary auto-snapshots.

Create the Target Endpoint

Before start, make sure you have the following information in-hand:

- DMSLabRoleS3 ARN – It looks like “arn:aws:iam::<Account number>:role/xxx-DMSLabRoleS3-xxxx”
- S3 Bucket Name - It looks like “xxx-dmslabs3bucket-xxxx”

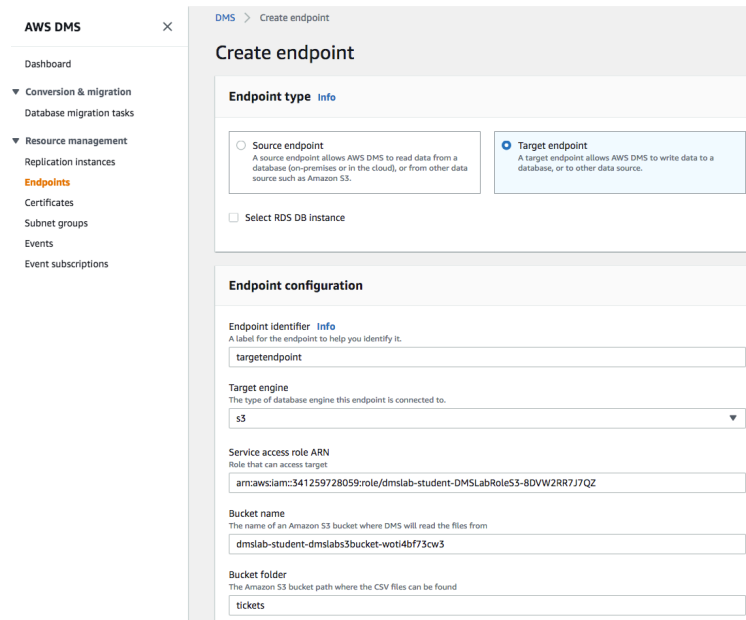
- On the DMS console, select **Endpoints**.



- Click **Create endpoint**.
 - For Endpoint type, select “**Target endpoint**”.
 - For Endpoint identifier, type an easily recognized name such as “s3-target-endpoint”.
 - For Target engine, choose **s3**.

Database Migration Services Lab

- d. For Service access role ARN, paste the **DMSLabRoleS3 ARN** number noted earlier
- e. For Bucket name, paste the **S3 Bucket Name** noted earlier
- f. For Bucket folder, type **"tickets"**.



The screenshot shows the AWS DMS console's 'Create endpoint' page. On the left is a navigation menu with 'AWS DMS' at the top, followed by 'Dashboard', 'Conversion & migration' (with 'Database migration tasks' as a sub-item), and 'Resource management' (with 'Replication instances', 'Endpoints' (highlighted in orange), 'Certificates', 'Subnet groups', 'Events', and 'Event subscriptions' as sub-items). The main panel is titled 'DMS > Create endpoint'. It has a 'Create endpoint' header. Below is the 'Endpoint type' section with 'Info' link, containing two options: 'Source endpoint' (unselected) and 'Target endpoint' (selected). Below these is a checkbox for 'Select RDS DB instance'. The 'Endpoint configuration' section follows, with an 'Endpoint identifier' field containing 'targetendpoint'. The 'Target engine' dropdown is set to 's3'. The 'Service access role ARN' field contains 'arn:awsiam::341259728059:role/dmslab-student-DMSLabRoleS3-8DVW2RR7J7QZ'. The 'Bucket name' field contains 'dmslab-student-dmslabs3bucket-wot4bf73cw3'. The 'Bucket folder' field contains 'tickets'.

- g. Click **Endpoint-specific settings** to expand the section.
- h. In the **Extra connection attributes** box, type **addColumnNames=true**. This attribute includes the column names in the files in the S3 bucket.
- i. Expand the **Test endpoint connection (optional)** section, and choose your "VPC name with dmslstudv1" on the VPC drop-down list.
- j. Click **Run test**. This step tests connectivity to the source database system. If successful, the message "Connection tested successfully" appears.

Database Migration Services Lab

Endpoint-specific settings

Extra connection attributes
Type any additional connection parameters here. See the documentation for more information.

addColumnName=true

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.

VPC
vpc-0314e829ba12d9481 - dmslstudv1

Replication instance
A replication instance performs the database migration
dms-replication-instance

Run test

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.

Endpoint identifier	Replication instance	Status	Message
targetendpoint	dms-replication-instance	successful	

Cancel Create endpoint

3. Click **Create Endpoint**. When available, the endpoint status changes to **active**.

Endpoints (2)

Find endpoint

<input type="checkbox"/>	Name	Type	Status	Engine	Server name	Port	Migration Hub Mapping	ARN
<input type="checkbox"/>	prodendpoint-postgre	Source	Active	PostgreSQL	dmslabinstance.c1ny5gywsvdz.us-east-1.rds.amazonaws.com	5432		arn:aws:dms:us-east-1:341259728059:endpoint-
<input type="checkbox"/>	targetendpoint	Target	Active	Amazon S3	-	-		arn:aws:dms:us-east-1:341259728059:endpoint-

Create a task to perform the initial full copy

1. On the DMS console, select **Database Migration Tasks**.

Database migration tasks

Find task

<input type="checkbox"/>	Name	Status	Source	Target	Type	Progress	Elapsed time	Tables loaded	Tables loading	Tables queued	Tables errored
Empty replication task table You don't have any replication tasks.											

Create task

2. Click **Create Task**.

a. Type an easily recognized **Task name** e.g. "dms-full-dump-task".

- b. Select your **Replication instance** from drop down.
- c. Select your **Source endpoint** from drop down.
- d. Select your **Target endpoint** from drop down.
- e. For, **Migration type** choose "Migrate existing data".
- f. Select "Start task on create"

AWS DMS ×

Dashboard

▼ Conversion & migration

Database migration tasks

▼ Resource management

Replication instances

Endpoints

Certificates

Subnet groups

Events

Event subscriptions

What's new

Notifications

DMS > Database migration tasks > Create database migration task

Create database migration task

Task configuration

Task identifier

dms-full-dump-task

Replication instance

dms-replication-instance - vpc-0537f7268d522baf3 ▼

Source database endpoint

rds-source-endpoint ▼

Target database endpoint

s3-target-endpoint ▼

Migration type [Info](#)

Migrate existing data ▼

☒ Start task on create

- g. Expand **Task Settings**.
- h. Select the **Enable CloudWatch logs** check box.

Database Migration Services Lab

- i. Go to **Table Mappings**.
- j. Click on **Add new selection rule** and select “Enter a Schema” in **Schema** field.
- k. For Schema name, select **dms_sample** . Keep the settings for the remaining fields

3. Click **Create task**. Your task is created and starts automatically. (Note: The complete creation and data extraction process takes 5 to 15 minutes.)
4. Once complete, the console displays 100% complete.

Database Migration Services Lab

DMS > Database migration tasks

Database migration tasks (1)

Find task

1

	Name	Status	Source	Target	Type	Progress	Elapsed time	Tables loaded	Tables loading	Tables queue
<input type="checkbox"/>	dms-task	Load complete	src-rds	targets3	Full load	100%	5 m	16	0	

- Select your task and explore the summary. Scroll down and you can observe all table information loaded in S3 from RDS by DMS

DMS > Database migration tasks > dms-task

dms-task

Summary

Status Load complete	Type Full load	Source prodendpoint-postgre	Target targetendpoint
-------------------------	-------------------	--------------------------------	--------------------------

Overview details

Basic configuration

Task ARN arn:aws:dms:us-east-1:341259728059:task:MUYIRCLBYT45EZESVNFNGAUL4	Status Load complete
Type Full load	Replication instance dms-replication-instance
Source prodendpoint-postgre	Target targetendpoint
Last failure message -	Created 5/29/2019, 10:55:15 AM GMT-0700
Started 5/29/2019, 10:55:51 AM GMT-0700	Migration task logs View logs

Task settings (JSON)

Table statistics (16)

Find schema

	Schema name	Table	Load state	Inserts	Deletes	Updates	DDLs	Full load rows	Total	Validation state	Validation pending
<input type="checkbox"/>	dms_sample	seat_type	Table completed	0	0	0	0	6	6	Not enabled	0
<input type="checkbox"/>	dms_sample	seat	Table completed	0	0	0	0	603,631	603,631	Not enabled	0
<input type="checkbox"/>	dms_sample	mlb_data	Table completed	0	0	0	0	2,230	2,230	Not enabled	0
<input type="checkbox"/>	dms_sample	player	Table completed	0	0	0	0	5,157	5,157	Not enabled	0
<input type="checkbox"/>	dms_sample	ticket_purchase_hist	Table completed	0	0	0	0	6,038,756	6,038,756	Not enabled	0
<input type="checkbox"/>	dms_sample	person	Table completed	0	0	0	0	7,025,584	7,025,584	Not enabled	0
<input type="checkbox"/>	dms_sample	name_data	Table completed	0	0	0	0	5,360	5,360	Not enabled	0
<input type="checkbox"/>	dms_sample	sport_team	Table completed	0	0	0	0	62	62	Not enabled	0
<input type="checkbox"/>	dms_sample	sport_league	Table completed	0	0	0	0	2	2	Not enabled	0
<input type="checkbox"/>	dms_sample	sporting_event	Table completed	0	0	0	0	1,158	1,158	Not enabled	0
<input type="checkbox"/>	dms_sample	sporting_event_ticket	Table completed	0	0	0	0	15,212,460	15,212,460	Not enabled	0
<input type="checkbox"/>	dms_sample	sport_division	Table completed	0	0	0	0	14	14	Not enabled	0
<input type="checkbox"/>	dms_sample	sport_location	Table completed	0	0	0	0	62	62	Not enabled	0
<input type="checkbox"/>	dms_sample	sport_type	Table completed	0	0	0	0	0	0	Not enabled	0
<input type="checkbox"/>	dms_sample	nfl_stadium_data	Table completed	0	0	0	0	32	32	Not enabled	0
<input type="checkbox"/>	dms_sample	nfl_data	Table completed	0	0	0	0	2,928	2,928	Not enabled	0

- Open the S3 console and view the data that was copied by DMS.

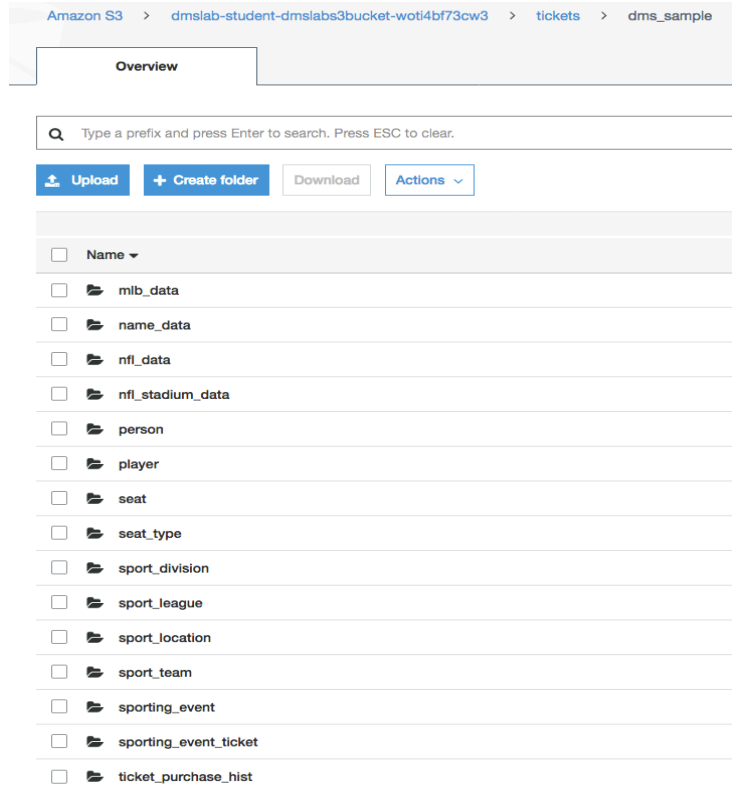
Your S3 bucket name will look like below :

BucketName/bucket_folder_name/schema_name/table_name/objects/

In our lab example this becomes:

"/dmslab-student-dmslabs3bucket-woti4bf73cw3/tickets/dms_sample" with a separate path for each table_name)

Database Migration Services Lab



7. Download one of the files:

- Select the check box next to the file name and click **Download** in the pop-up window.
- Click **Save File**.
- Open the file.

You will notice that the file contains the column headers in the first row as requested by the "addColumnNames=true" connection attribute we included when we created the s3 target endpoint. Note that column names are included in the file in the first row.

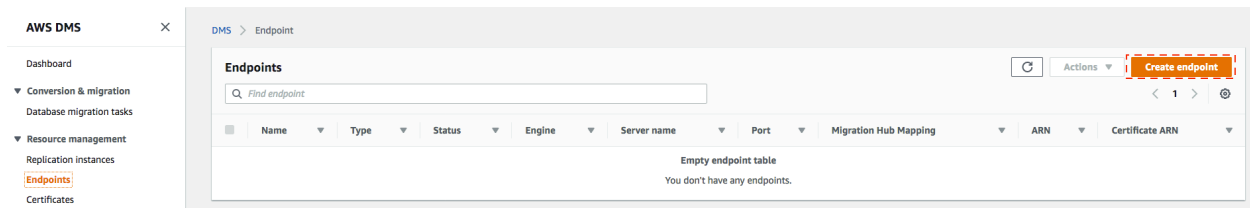
	A	B	C	D	E
1	id	sport_team_id	last_name	first_name	full_name
2	1	131	Adam Loewen	Adam	Loewen
3	11	131	A.J. Pollock	A.J.	Pollock
4	21	131	Alex Sanabia	Alex	Sanabia
5	31	131	Andrew Chafin	Andrew	Chafin
6	41	131	Andy Marte	Andy	Marte
7	51	131	Archie Bradley	Archie	Bradley
8	61	131	Ben Francisco	Ben	Francisco
9	71	131	Braden Shipley	Braden	Shipley
10	81	131	Bradin Hagens	Bradin	Hagens
11	91	131	Brandon Drury	Brandon	Drury
12	101	131	Brett Jackson	Brett	Jackson

Explore the objects in the S3 directory further.

Create the CDC endpoint to replicate ongoing changes (Optional)

As of now we are enabling only one schema replication for CDC

1. On the DMS console, select **Endpoints**.



2. Click **Create endpoint**.
 - a. For **Endpoint type**, select "Target".
 - b. For **Endpoint identifier**, type an easily recognized name that includes "rds-cdc-endpoint" in
 - c. For **Target engine**, choose "s3".
 - d. For **Service Access Role ARN**, paste the ARN value that you noted earlier from lab "prelab - Batch Data Ingestion with DMS - Student Setup"
 - e. For **Bucket name**, type the name of the s3 bucket you noted down from from lab "prelab - Batch Data Ingestion with DMS - Student Setup"
 - f. For **Bucket folder**, type "cdc".

Create endpoint

Endpoint type [Info](#)

☐ Source endpoint
 A source endpoint allows AWS DMS to read data from a database (on-premises or in the cloud), or from other data source such as Amazon S3.

☒ Target endpoint
 A target endpoint allows AWS DMS to write data to a database, or to other data source.

☐ Select RDS DB Instance

Endpoint configuration

Endpoint identifier [Info](#)
 A label for the endpoint to help you identify it.

Target engine
 The type of database engine this endpoint is connected to.

Service access role ARN
 Role that can access target

Bucket name
 The name of an Amazon S3 bucket where DMS will read the files from

Bucket folder
 The Amazon S3 bucket path where the CSV files can be found

- g. Click **Endpoint-specific settings** to expand the section.

Database Migration Services Lab

- h. In the **Extra connection attributes** box, type "addColumnNames=true". This attribute includes the column names in the files in the S3 bucket.
- i. Expand the **Test endpoint connection (optional)** section, and choose your **dmslstudv1** name on the VPC drop-down list.
- j. Click Run test. This step tests connectivity to the source database system. If successful, the message "Connection tested successfully" appears.

▼ Endpoint-specific settings

Extra connection attributes

Type any additional connection parameters here. See the documentation for more information.

addColumnNames=true

▼ 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.

VPC

vpc-0314e829ba12d9481 - dmslstudv1

Replication instance

A replication instance performs the database migration

dms-replication-instance

Run test

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.

Endpoint identifier	Replication instance	Status	Message
cdcendpoint	dms-replication-instance	successful	

Cancel Create endpoint

3. Click **Create endpoint**.

4. When available, the endpoint status changes to active.

DMS > Endpoint

Endpoints (3)

Find endpoint

	Name	Type	Status	Engine	Server name	Port	Migration Hub Mapping	ARN
<input checked="" type="checkbox"/>	cdcendpoint	Target	Active	Amazon S3	-	-		arn:aws:dms:us-east-1:341259728059:endp
<input type="checkbox"/>	prodendpoint-postgre	Source	Active	PostgreSQL	dmslabinstance.c1ny3gywsvdz.us-east-1.rds.amazonaws.com	5432		arn:aws:dms:us-east-1:341259728059:endp
<input type="checkbox"/>	targetendpoint	Target	Active	Amazon S3	-	-		arn:aws:dms:us-east-1:341259728059:endp

Create a task to perform the ongoing replication (Optional)

1. On the DMS console, select **Database Migration Tasks**.

AWS DMS

DMS > Database migration tasks

Database migration tasks

Find task

	Name	Status	Source	Target	Type	Progress	Elapsed time	Tables loaded	Tables loading	Tables queued	Tables errored
Empty replication task table You don't have any replication tasks.											

2. Click **Create Task**.

- Type an easily recognized **Task Identifier**. For example "cdctask".
- Select your **Replication instance**.
- Select your **Source endpoint**.
- Select your **Target endpoint** as **cdc** endpoint created in the previous section.
- For **Migration type**, choose **Replicate data changes only**.
- Select the Start task on create check box.

The screenshot shows the 'Create database migration task' form in the AWS DMS console. The breadcrumb trail is 'DMS > Database migration tasks > Create database migration task'. The form title is 'Create database migration task'. Under the 'Task configuration' section, the following fields are visible: 'Task identifier' with the value 'cdctask'; 'Replication instance' with a dropdown showing 'dms-replication-instance - vpc-0537f7268d522baf3'; 'Source database endpoint' with a dropdown showing 'rds-source-endpoint'; 'Target database endpoint' with a dropdown showing 'rds-cdc-endpoint'; 'Migration type' with a dropdown showing 'Migrate existing data and replicate ongoing changes' and an 'Info' link; and a checked checkbox for 'Start task on create'.

- In **Task Settings**, Select the **Enable CloudWatch logs** check box. Do not enable the validation.

The screenshot shows the 'Task settings' panel in the AWS DMS console. The left sidebar shows the 'AWS DMS' navigation menu with 'Database migration tasks' selected. The 'Task settings' panel includes: 'Target table preparation mode' with radio buttons for 'Do nothing', 'Drop tables on target' (selected), and 'Truncate'; 'Include LOB columns in replication' with radio buttons for 'Don't include LOB columns', 'Full LOB mode', and 'Limited LOB mode' (selected); 'Maximum LOB size (KB)' with a text input field containing '32'; an unchecked checkbox for 'Enable validation' with a descriptive note; and a checked checkbox for 'Enable CloudWatch logs' with an 'Info' link. A blue information box at the bottom states: 'CloudWatch logs usage will be charged at standard rates. See here for more details.'

- h. Go to **Table Mappings**.
- i. Click on **Add new selection rule** and select "Enter a Schema" in Schema field.
- j. For **Schema name**, select "dms_sample" . Keep the settings for the remaining fields

Table mappings

Editing mode

☒ **Guided UI**
Set up your table mapping rules using a step-by-step guided interface.

☐ **JSON editor** [Learn more](#)
Enter your table mapping rules directly, in JSON format.

Specify at least one selection rule with an include action. After you do this, you can add one or more transformation rules.

Selection rules

Choose the schema and/or tables you want to include with, or exclude from, your migration task. [Info](#) Add new selection rule

▼ where schema name is like '%' and table name is like %, include

Schema
Enter a schema

Schema name
Use the % character as a wildcard
dms_sample

Table name
Use the % character as a wildcard
%

Action
Choose "Include" to migrate your selected objects, or "Exclude" to ignore them during the migration.
Include

3. Click Create task. Your task is created and starts automatically. You can see status as **ongoing replication**, after couple of minutes.

DMS > Database migration tasks

Database migration tasks (2)

Actions Create task

	Name	Status	Source	Target	Type	Progress	Elapsed time	Tables loaded	Tables loading	Tables queued
<input type="checkbox"/>	dms-task	Load complete	prodendpoint-postgre	targetendpoint	Full load	100 %	9 m	16	0	0
<input checked="" type="checkbox"/>	newcdc	Replication ongoing	prodendpoint-postgre	cdcendpoint	Ongoing replication	100 %	0 m	16	0	0

Once complete, the console displays 100% complete.

4. Your instructor will generate CDC activity which above migration task will capture, if you ran the instructor setup on your own, then make sure to follow **"Generate the CDC Data"** section from instructor lab.

You may need to wait 5 to 10 minutes for CDC data to first reflect in your RDS postgres database and then picked up by DMS CDC migration task.

- Once the CDC Data gets replicated, you can navigate to CDC task details, scroll down to the "Table statistics" section and verify it, as shown below:

Note: In case you see DMS CDC task in fail/error status. Make sure your replication instance version is 3.3.1 and it is large enough (dms.t2.medium or above) to run CDC replication task

	Schema name ▾	Table ▾	Load state ▾	Inserts ▾	Deletes ▾	Updates ▾	DDLs ▾	Full load rows ▾	Total ▾	Validation state ▾	Validation pending ▾	V
<input type="checkbox"/>	dms_sample	seat_type	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	seat	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	mlb_data	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	player	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	ticket_purchase_hist	Table completed	11,002	0	0	0	0	11,002	Not enabled	0	0
<input type="checkbox"/>	dms_sample	person	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	name_data	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sport_team	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sport_league	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sporting_event	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sporting_event_ticket	Table completed	0	0	11,002	0	0	11,002	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sport_division	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sport_location	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	sport_type	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	nfl_stadium_data	Table completed	0	0	0	0	0	0	Not enabled	0	0
<input type="checkbox"/>	dms_sample	nfl_data	Table completed	0	0	0	0	0	0	Not enabled	0	0

- Open the S3 console and view the CDC data that was copied by DMS.

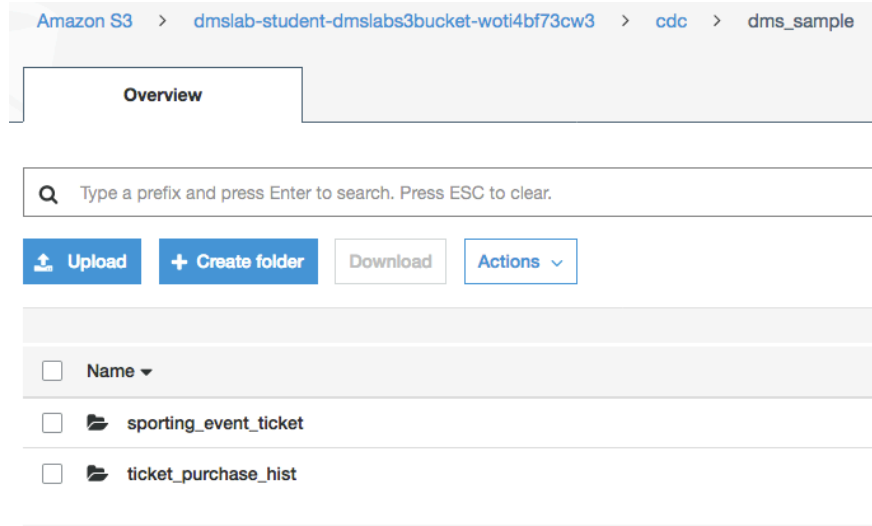
Your S3 bucket name will look like below :

BucketName/bucket_folder_name/schema_name/table_name/objects/

In our lab example this becomes:

"/dmslab-student-dmslabs3bucket-woti4bf73cw3/cdc/dms_sample" with a separate path for each table_name)

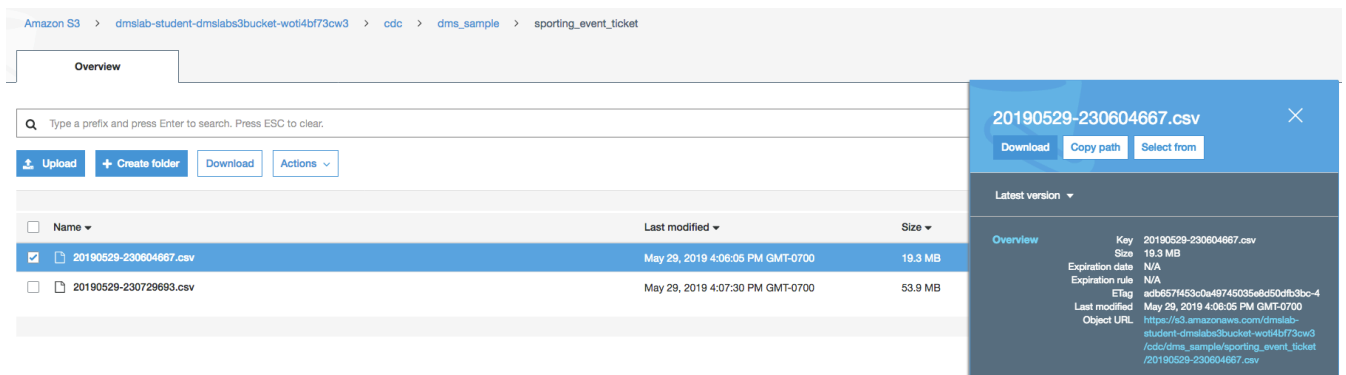
Database Migration Services Lab



7. Download one of the files:

- Select the check box next to the object name and click Download in the pop-up window.
- Click **Save File**.
- Open the file.

You will notice that the file contains the column headers in the first row as requested by the “**addColumnNames=true**” connection attribute we included when we created the s3 target endpoint.



Note that file name has date time - 20190529-230604667.csv

You can see the header is included and the operation column is added at the beginning of each row. The file below shows updates (U) to the table along with the values after the update. Inserts (I) show data after the insert and Deletes (D) show data before the delete.

Database Migration Services Lab

	A	B	C	D	E	F	G	H	I	J
	Op	id	sporting_event_id	sport_location_id	seat_level	seat_section	seat_row	seat	ticketholder_id	ticket_price
U		145192591	3931	4	2	10 A		2	2898028	98
U		145192601	3931	4	2	10 A		1	2898028	98
U		145192581	3931	4	2	10 A		3	2898028	98
U		145192501	3931	4	2	10 B		1	2898028	98
U		145187751	3931	4	2	13 B		2	2898028	49
U		145187741	3931	4	2	13 B		3	2898028	49
U		145187721	3931	4	2	13 C		2	2898028	49
U		145187711	3931	4	2	13 C		3	2898028	49
U		145187731	3931	4	2	13 C		1	2898028	49
U		145187701	3931	4	2	14 A		1	2898028	49
U		145187681	3931	4	2	14 A		3	2898028	49
U		145187691	3931	4	2	14 A		2	2898028	49
U		145187471	3931	4	2	14 B		3	2898028	49
U		145187671	3931	4	2	14 B		1	2898028	49
U		145187481	3931	4	2	14 B		2	2898028	49
U		145187451	3931	4	2	14 C		2	2898028	49
U		145187461	3931	4	2	14 C		1	2898028	49
U		145190341	3931	4	2	14 C		3	2898028	49
U		145183201	3931	4	2	15 A		4	2898028	49
U		145179691	3931	4	2	15 A		1	2898028	49
U		145179661	3931	4	2	15 A		4	2898028	49
U		145179671	3931	4	2	15 A		3	2898028	49
U		145179681	3931	4	2	15 A		2	2898028	49
U		145190321	3931	4	2	15 A		2	2898028	49

Explore the objects in the S3 directory further.