Aim: - To load data from a s3 bucket into a relational database configured on Amazon RDS.

Solution: - Create amazon Aurora database on RDS and upload the data onto s3. Create a lambda trigger which gets triggered when a file is placed in the s3 bucket and eventually that trigger will call a function to load the data into aurora DB instance.

Steps to implement this solution.

Initial Steps

1)Create a s3 bucket and give permissions to read and write to the public.

2)Create an aurora MYSQL instance within the default VPC with same security groups and public access. Ensure we have all traffic and all ip address included by having inbound and outbound rules set to 0.0.000/0

3)Now verify if we can connect to aurora by connecting from mysql workbench. This step is done to ensure mysql instance can be accessed from outside.

Steps to read data from s3

4)Create a new dbcluster parameter group associated with MYSQL instance.

5) Create a new role, and grant it s3 full access.

6)Add role created in step 5 to db cluster parameter group. Select aws\_s3\_default\_load\_role.

7)Create a schema in mysql db.

8)Add csv file to the s3 bucket.

9)Run load command and verify that data is loaded into the db.

Creation of lambda function

10) Go to lambda and create lambda execution role. Grant s3 RDS full access and CloudWatch should be given read/write access.

11) Create a function with function name which should be same as the python/javascript script name and handler name should be same as event handler or function name.

12)Define trigger and use s3 object created template for this.

13)Upload package as a zip. Zip package without parent directory. Zip just the contents.

Function Creation and Testing

14) Use Pymysql to connect to db. This needs to be installed on the instance on which event is running.

15) Go to lambda execution environment and find the AMI used.

16)Create a linux instance of the same AMI and generate a new key pair and download the pem file. Use puttygen to convert it to ppk file

17) Connect to the EC2 and download the relevant packages into a folder and zip the contents.

Upload the function to Lambda function

18) Create the relevant secret key and access key and update the function with the zip.

Testing the function

19) Upload the file onto s3 bucket and monitor the cloudwatch logs for insertion of columns into db.