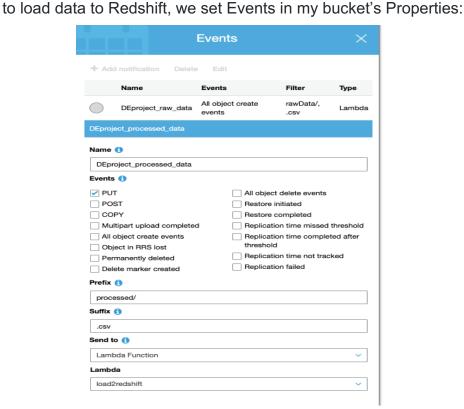
1. AWS lambda psycopg2 package:

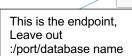
https://github.com/qiaoqiaohejianjian/ETL--Project/tree/master/lambda redshift

AWS redshift and lambda configuration:
 S3: after generating processed data, to trigger lambda function "load2redshift"



Lambda: a) wrap up code and package, and then upload in Function code -> Actions -> Upload a zip file

- b) Set Environment Variables for your redshift database information, such as: database name, host(end point:), password etc.
- (New console) In the navigation pane, choose CLUSTERS. Then choose the cluster name
 from the list to open its details. On the Properties tab, in the Database configurations
 section, record the Database name and Port. View the Connection details section and
 record the Endpoint which is in the following form:



- endpoint:)port/databasename
 - c) set up lambda function basic info
 - d) choose same VPC as your Redshift cluster, and same subnet

e) In Permissions: attach s3 Read-only and AWS Lambda VPC Access Execution Role policies

Redshift: a) create your cluster and store your database login information:
user and password in a safe place
In your cluster -> properties, you can find more detail information

b) cluster permissions: attach a s3 read-only policy to the cluster

- c) for the Network and security: edit Security Group and Network ACLs, make sure related IP address can access you cluster with your database information.
- d) create an endpoint for S3 service.

3. Connect your Redshift cluster:

https://towardsdatascience.com/redshift-from-the-command-line-5d6b3233f649 https://docs.aws.amazon.com/redshift/latest/mgmt/connecting-from-psql.html https://forums.aws.amazon.com/thread.jspa?threadID=310081

