# Lab 4 - Loading Real-Time Data Into Your Amazon Redshift Cluster

Objs -

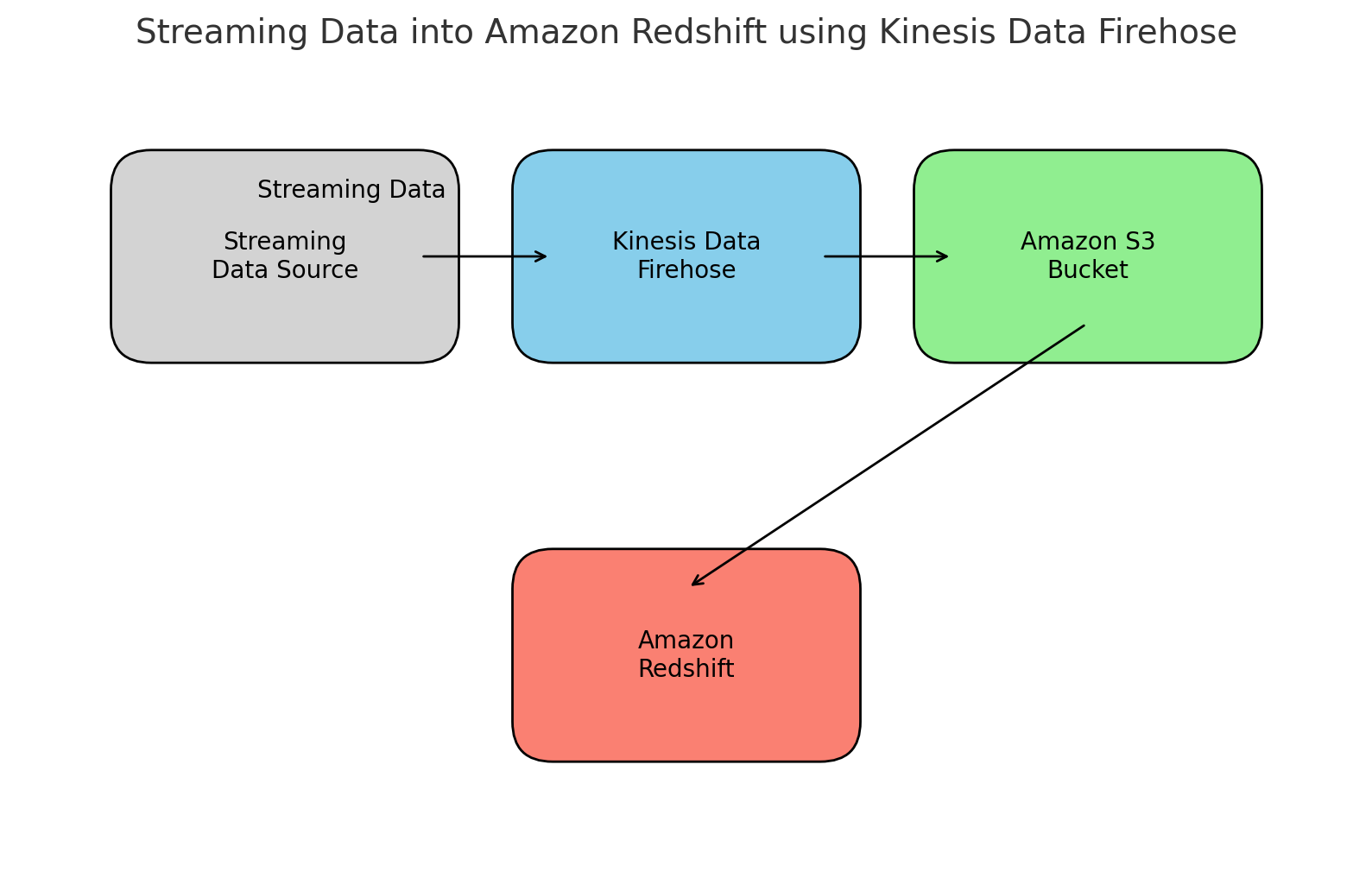
1. Create a table in Amazon Redshift
2. Create an Amazon Kinesis Data Firehose delivery stream and use it to load data into Amazon Redshift
3. Confirm receipt of Streaming Records

The lab environment includes an Amazon Elastic Compute Cloud (Amazon EC2) instance that automatically sends data to your Firehose delivery stream once it has been created. This simulates the transmission of game scores coming into a stream.

KDS part -

1. Set up a delivery stream to receive streaming data.
2. Configure the stream to use an Amazon S3 bucket as a temporary storage area.
3. Load the data from the S3 bucket into the Amazon Redshift table.

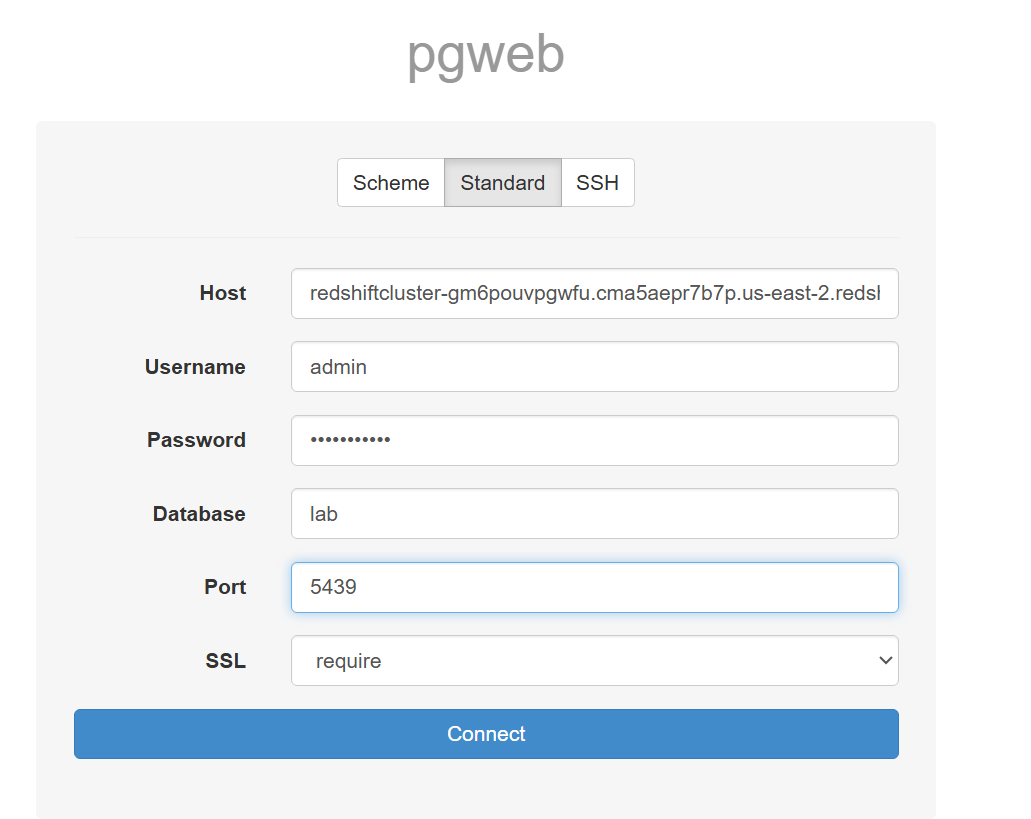
GPT generated Img-



An Amazon S3 bucket is used to stage the data before loading it into Amazon Redshift.

**Task 1: Creating a table in Amazon Redshift**

1.1 open pgweb,. Login using details, connect to cluster and create a new table using:



1.2

*CREATE TABLE game\_score (*

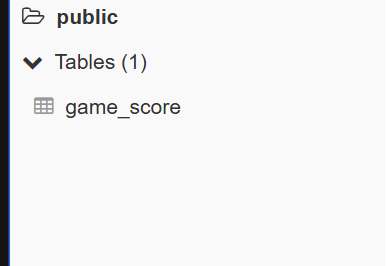
*record\_time TIMESTAMP,*

*user\_id INT,*

*game\_id INT,*

*score INT*

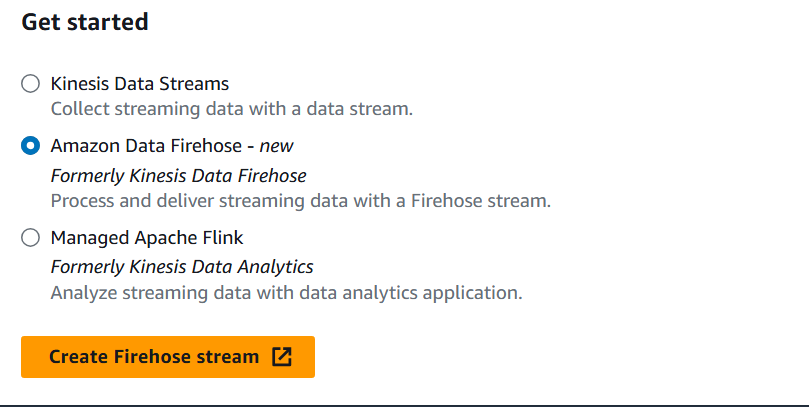
*)*

table created

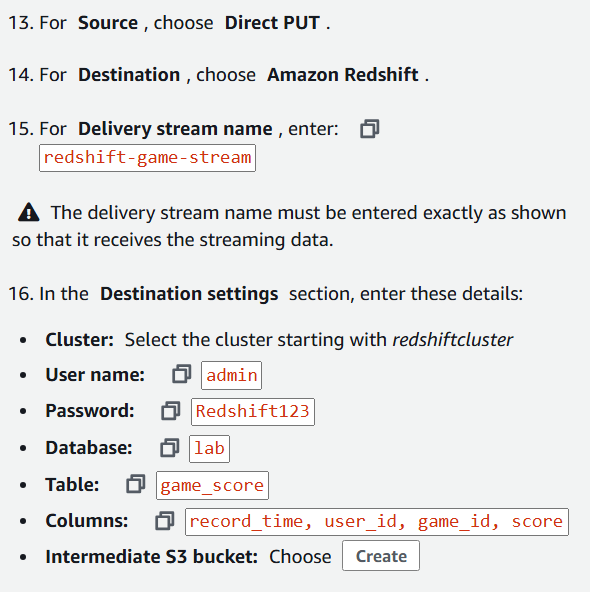
**Task 2: Creating an Amazon Kinesis Data Firehose delivery stream**

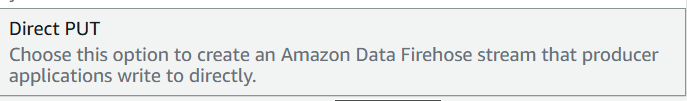
2.1 open kinesis in console

2.2 In the Get started section, select the radio button next to Kinesis Data Firehouse.

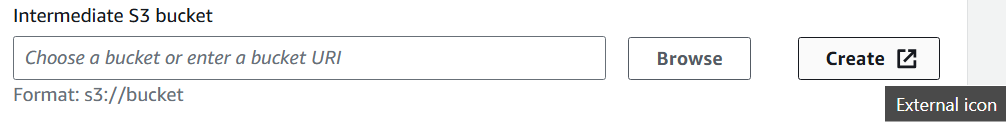


2.3 use these to configure the stream

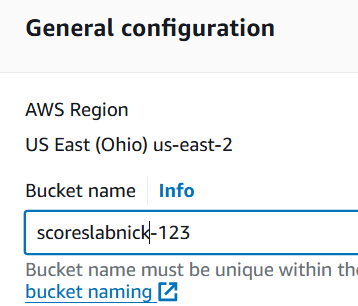




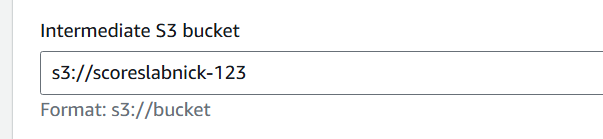
2.4 for s3 bucket, click choose



2.5 name the bucket and go ahead



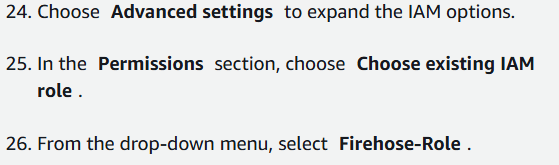
2.6 in kinesis, choose the bucket created



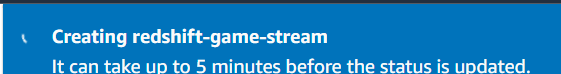
2.7 choose buffer interval as 60s

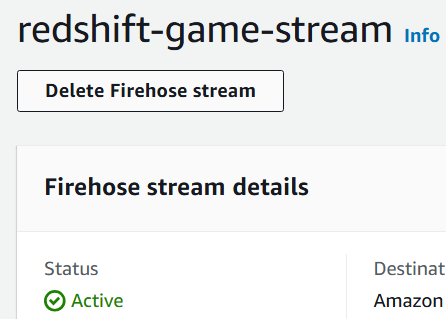
This sends data to Amazon S3 every 60 seconds for loading into Amazon Redshift.

2.8 for IAM



2.9 create it



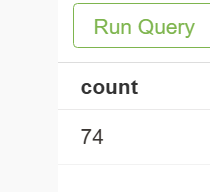


**Task 3: Confirming receipt of Streaming Records**

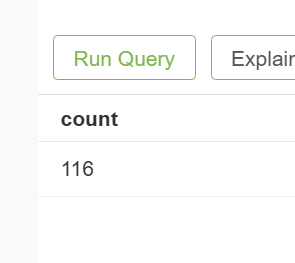
EC2 -> KFH ->S3 -> RS

3.1 use this to verify that data has started to get into Redshift

*SELECT COUNT(\*) FROM game\_score*



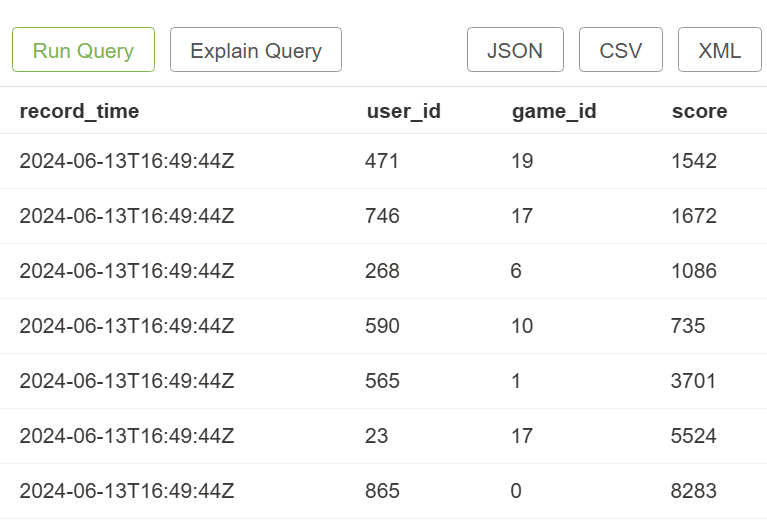
3.2 run after 60 secs, more data comes in



3.3 to view the data

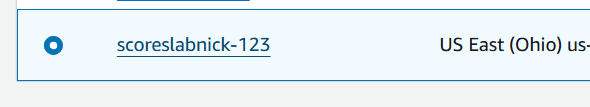
*SELECT \* FROM game\_score*

*LIMIT 50*



**Task 4 - optional - view files in S3**

4.1 open S3, and open the bucket



4.2 go to this path



To view the data files

