

Documentation

1. Log in to AWS (Amazon Web Services) account and create Database Instance in RDS (Relational Database Service).
2. Install DBeaver and connect the Database instance to it.
3. Create clusters in Redshift and connect them in DBeaver.
4. Add VPC (Virtual Private Cloud) for RDS & Redshift and update your current IP.
5. Load data into MySQL RDS using Dbeaver.
6. Create S3 Bucket to store the imported data (initialload, incrementalload, table1.csv).
7. Create IAM (Identity & Access Management) roles for each tool.
8. Configure the IAM roles for Data Pipelining.
9. Open AWS Glue and create a new job.
10. Generate tables in Redshift using the copy command.
11. Setup AWS Pipeline for incremental data load.
12. Edit script for the Glue job using Python Shell.
13. AWS Lambda function is triggered using the Glue job for visualizing the logs and metrics using Cloudwatch.
14. Lambda function is deployed and the dashboard is published.
15. AWS Crawler is set up in Glue jobs.
16. Athena is configured for data and table scanning using query data in S3.
17. Quicksight account is created with AWS Glue and S3 configured.
18. Dashboards are connected with Redshift and analyses are created.
19. Quicksight Dashboard is published and shared publicly.
20. Cloudwatch and Quicksight dashboards are visible on the website created for the application.