

Extracting data from an AWS RDS MySQL database, storing it locally as a CSV and uploading it to S3.

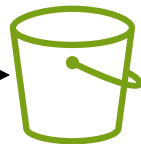
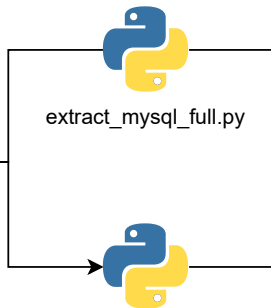
This can be done doing a full load ("extract_mysql_full.py") or an incremental refresh ("extract_mysql_incremental_binlog.py") using binlog. In my case, binlog needed to be set up within AWS RDS MySQL by turning on automated backups, creating a parameter group with the relevant binlog parameters (binlog_format, binlog_row_image and binlog_row_metadata among others) and attaching the parameter group to the database.



Amazon Web Services (AWS)



AWS RDS MySQL Database



S3 Bucket

extract_mysql_incremental_binlog.py