ETL WITH AWS — GLUE & REDSHIFT ...

Data Explosion- 21st century

2.5M TB
DATA CREATED
EVERY DAY

40 TIMES

AS MANY DIGITAL BYTES AS STARS IN UNIVERSE

200B+ TB

IN CLOUD
STORAGE DEMAND
GLOBALLY

What is ETL?

Data integration process that combines data from multiple sources into a single, consistent data store.









Data Warehousing





ETL USE CASES





Data Integration



What is GLUE?

What is REDSHIFT?





AWS Glue is a serverless data integration service that makes it easy for analytics users to discover, prepare, move, and integrate data from multiple sources.

It is a fully managed, petabyte-scale data warehouse service in the cloud. Just a few hundred gigabytes of data and scale to a petabyte or more.

AWS Architecture Analysis of Customer Reviews





