

Redshift

Data Warehousing

- Used to pull in very large and complex data sets. Usually used by management to do queries on data.
 - eg: Cognos, Jaspersoft, SQL Server Reporting Services....

OLTP vs OLAP

OLTP(Online Transaction Processing) vs OLAP (Online Analytics Processing)

- OLTP: eg- I need particular value in so and so row and column
- OLAP: eg- I need sum of all values in table

Redshift

Amazon Redshift is a fast and powerful, fully managed, petabyte-scale data warehouse service in the cloud.

Redshift Configuration

- Single Node (160Gb)
- Multi-Node
 - Leader Node (manages client connections and receives queries)
 - Compute Node (store data and perform queries and computations). Up to 128 Compute Nodes

Massive Parallel Processing (MPP)

- Amazon Redshift automatically distributes data and query load across all nodes.
- Redshift makes it easy to add nodes to your data warehouse and enables you to maintain fast query performance as your data warehouse grows.