REDSHIFT

Petabyte scale data warehouse (1000 TB)

Suitable for OLAP systems

ADVANTAGES:

Hardware failures = no data loss due to network backups

Speed to market is fast

Savings in cost of servers and equipment

Pay for what you use

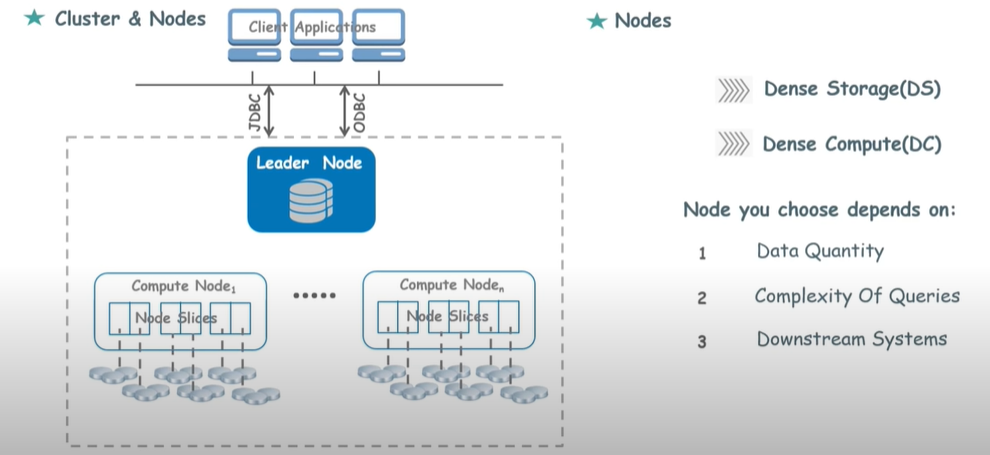
Frequent updates

Virtual private cloud, encryption help keep data secure

Allows massively parallel processing

Allows to query from data lake through Redshift spectrum

ARCHITECTURE:



COLUMN COMPRESSION:

Data in each dish block is same type (e.g. int, char, Boolean); block data uses compression scheme.

Compression scheme includes RAW, BYTE, Mostly, Runlength, LZO

EFFICIENT VACUUMING:

Enhances query performance by vacating space for re-use and sort table

COLUMNAR STORAGE:

Data stored in disk blocks in columnar fashion; saves space and reduces query execution time

Size of each disk block = 1 MB

