



# Data Science Academy

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

## Design e Implementação de Data Warehouses

### Dense Compute x Storage Compute



Um cluster do Amazon Redshift consiste em nós. Cada cluster tem um nó de liderança e um ou mais nós de computação. O nó de liderança recebe consultas de aplicativos cliente, analisa as consultas e desenvolve planos de execução de consulta. Em seguida, o nó principal coordena a execução paralela desses planos com os nós de computação e agrega os resultados intermediários desses nós. Então, ele retorna os resultados de volta para os aplicativos cliente.

Nós de computação executam planos de execução de consultas e transmitem dados entre si para atender essas consultas. Os resultados intermediários são enviados ao nó de liderança para agregação antes de serem enviados novamente para os aplicativos clientes.

Quando você executa um cluster, uma opção que você especifica é o tipo de nó. O tipo de nó determina a CPU, RAM, capacidade de armazenamento e o tipo de unidade de armazenamento de cada nó. Os nós do tipo armazenamento denso (DS) são otimizados para armazenamento. Os nós do tipo computação densa (DC) são otimizados para computação.

Você pode criar um cluster usando tipos de nós de armazenamento denso (DS) ou computação densa (DC). Os tipos de nós de armazenamento denso permitem criar data warehouses muito grandes usando discos rígidos (HDDs) por um preço bastante reduzido. Os tipos de nós de computação densa permitem criar data warehouses de altíssimo desempenho usando CPUs rápidas, grandes quantidades de RAM e discos de estado sólido (SSDs).

Os nós do tipo DS2 são otimizados para grandes cargas de trabalho de dados e usam armazenamento com base em unidade de disco rígido (HDD).

Os nós DC1 e DC2 são otimizados para cargas de trabalho de desempenho intensivo. Como eles usam armazenamento de disco de estado sólido (SSD), os tipos de nó DC1 e DC2 fornecem uma E/S muito mais rápida em comparação aos nós do tipo DS, mas fornecem menos espaço de armazenamento.

Você executa clusters que usam os tipos de nó DC2 em uma nuvem privada virtual (VPC). Você não pode executar clusters DC2 no modo EC2 clássico.

O tipo de nó escolhido depende intensamente de três informações:



- A quantidade de dados que você importa para o Amazon Redshift
- A complexidade das consultas e operações que você executa no banco de dados
- As necessidades de sistemas downstream que dependem dos resultados dessas consultas e operações

Os tipos de nó estão disponíveis em diferentes tamanhos. Os nós DS2 estão disponíveis nos tamanhos xlarge e 8xlarge. Os nós DC2 estão disponíveis nos tamanhos large e 8xlarge. O tamanho do nó e o número de nós determinam o armazenamento total de um cluster.

Alguns tipos de nó permitem um nó (single-node) ou dois ou mais nós (multi-node). A quantidade mínima para clusters 8xlarge são dois nós. Em um cluster de single-node, o nó é compartilhado para a funcionalidade principal e de computação. Em um cluster de multi-node, o nó de liderança é separado dos nós de computação.

O Amazon Redshift aplica cotas aos recursos para cada conta da AWS em cada região. Uma cota restringe o número de recursos que sua conta pode criar para determinados tipos de recursos, como nós ou Snapshots, em uma região.