

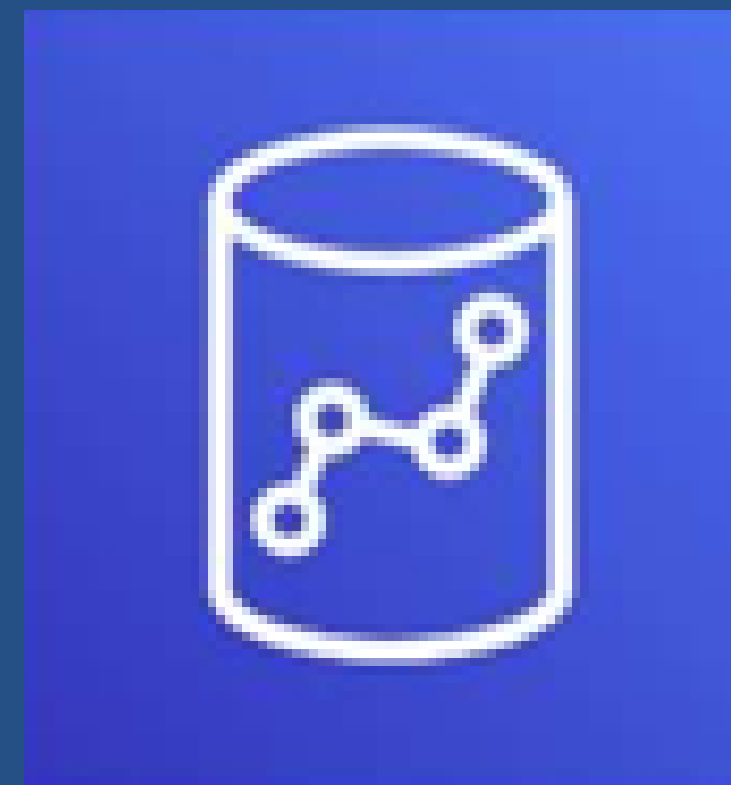
LECCIÓN 3:

AMAZON

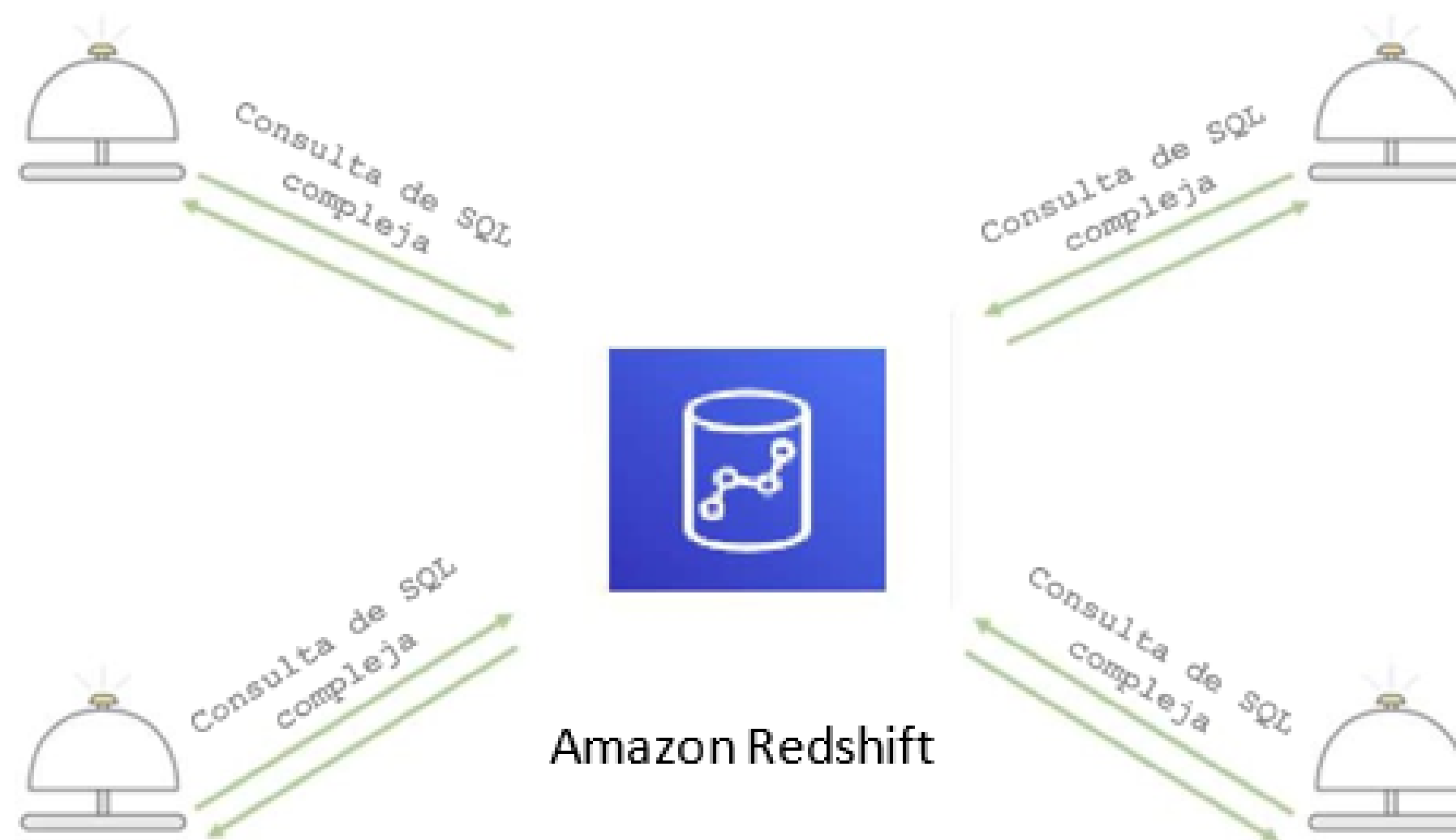
REDSHIFT



Amazon Redshift es un almacén de datos rápido y completamente administrado que simplifica y rentabiliza el análisis de todos sus datos mediante SQL estándar y sus herramientas de inteligencia empresarial (BI) existentes. A continuación, presentamos Amazon Redshift y cómo se puede utilizar para aplicaciones analíticas.



INTRODUCCIÓN A AMAZON REDSHIFT

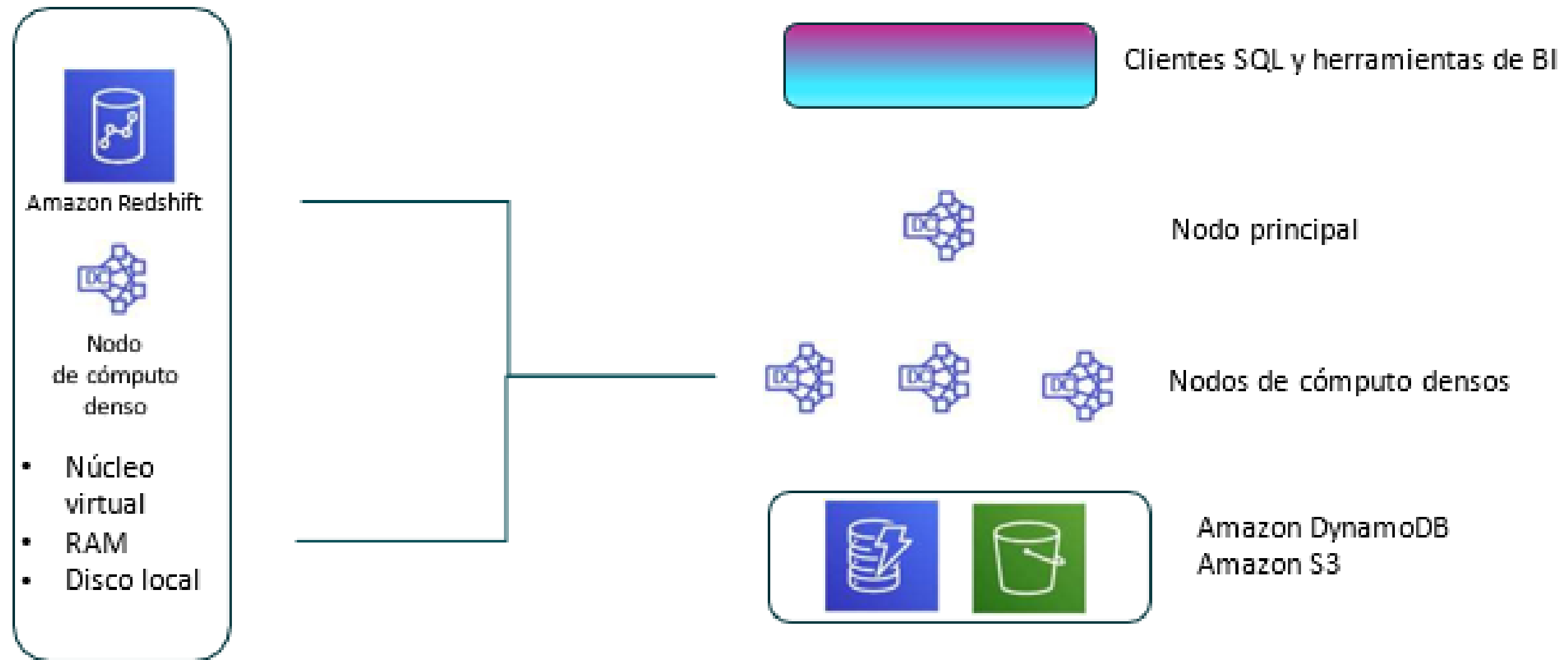


En la actualidad, el análisis es importante para las empresas; sin embargo, crear un almacén de datos es complejo y costoso. La creación de almacenes de datos puede llevar meses e importantes recursos financieros.

Amazon Redshift es un almacén de datos rápido, potente y completamente administrado que se puede configurar, utilizar y escalar de manera simple y rentable. Permite ejecutar consultas analíticas complejas a partir de petabytes de datos estructurados mediante una sofisticada optimización de las consultas, el almacenamiento en columnas en discos locales de alto rendimiento y el procesamiento masivo de datos en paralelo. La mayoría de los resultados se obtienen en segundos.

A continuación, revisaremos un poco más en detalle las funciones principales de Amazon Redshift y algunos casos prácticos comunes.

ARQUITECTURA DE PROCESAMIENTO PARALELO





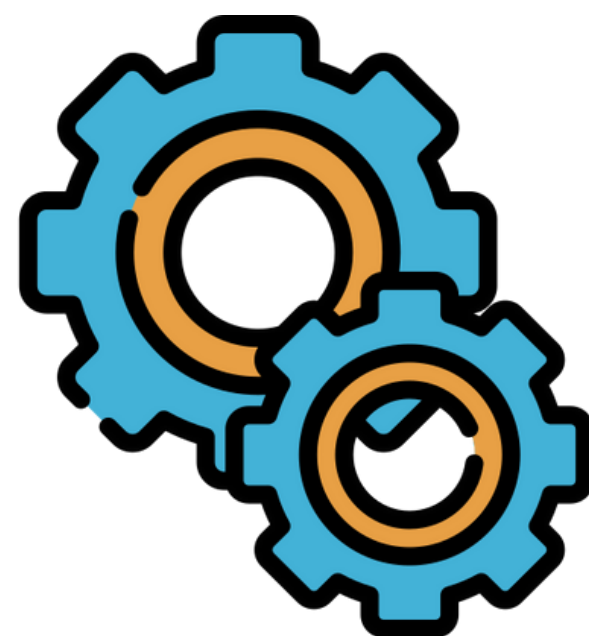
El nodo principal administra las comunicaciones con los programas cliente y todas las comunicaciones con los nodos de cómputo. Analiza y desarrolla los planes de consulta para realizar las operaciones de bases de datos, en particular, la serie de pasos necesarios para obtener resultados de consultas complejas. El nodo principal compila código para los elementos individuales del plan y lo asigna a los nodos de cómputo individuales. Los nodos de cómputo ejecutan el código compilado y devuelven resultados parciales al nodo principal para su agregación final.

Al igual que otros de los servicios de AWS, solo paga por lo que usa. Puede comenzar por tan solo 25 centavos por hora y escalar. Amazon Redshift ofrece almacenamiento y procesamiento por casi 1000 USD por terabyte al año (con un precio de instancia reservada parcial por adelantado de 3 años).

La característica Amazon Redshift Spectrum le permite ejecutar consultas en exabytes de datos de manera directa en Amazon S3.



AUTOMATIZACIÓN Y ESCALADO



ADMINISTRAR

SUPERVISAR

ESCALAR



Es sencillo automatizar la mayoría de las tareas administrativas comunes para administrar, supervisar y escalar el clúster de Amazon Redshift, lo que permite centrarse en los datos y el negocio.

La escalabilidad es inherente a Amazon Redshift. El clúster puede escalarse verticalmente y se puede reducir su escala vertical a medida que cambian las necesidades, con apenas unos clics en la consola.

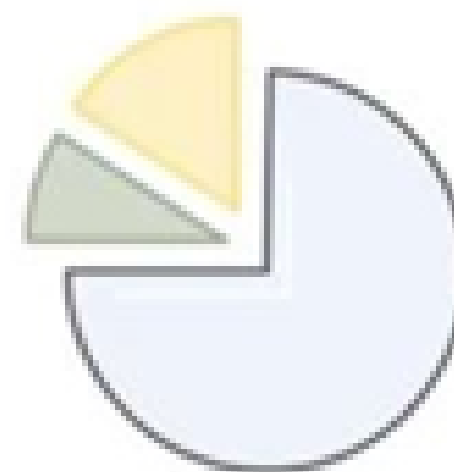
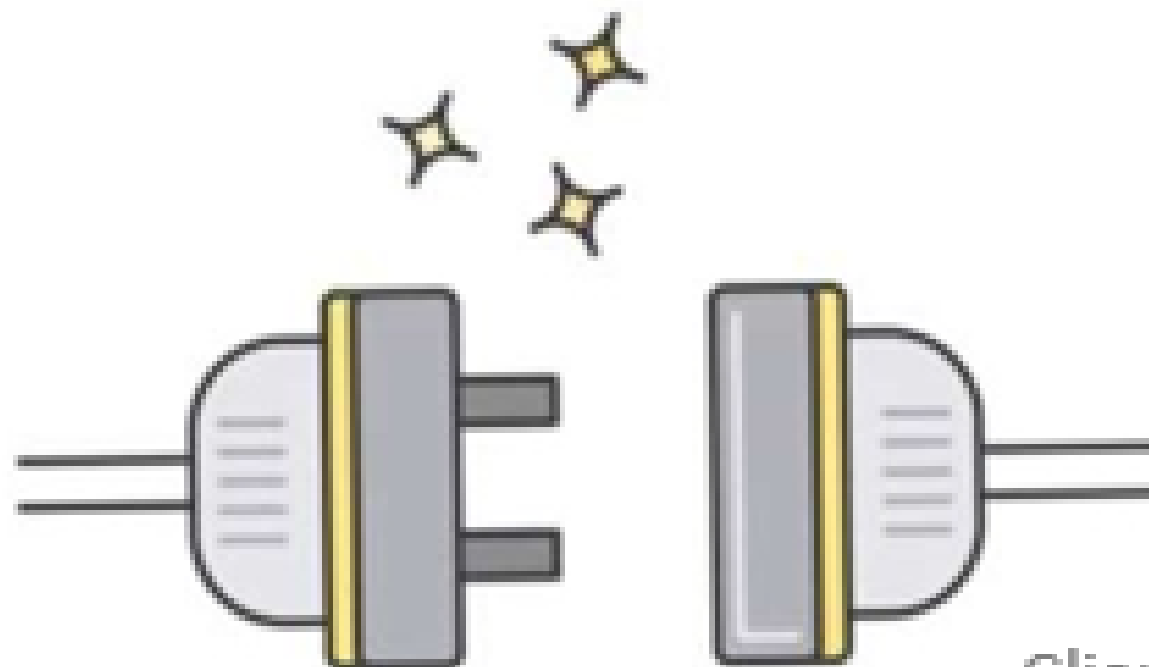
La seguridad es la máxima prioridad de AWS. En Amazon Redshift, la seguridad está integrada, y está diseñado para brindar un cifrado sólido de los datos en reposo y en tránsito.



COMPATIBILIDAD

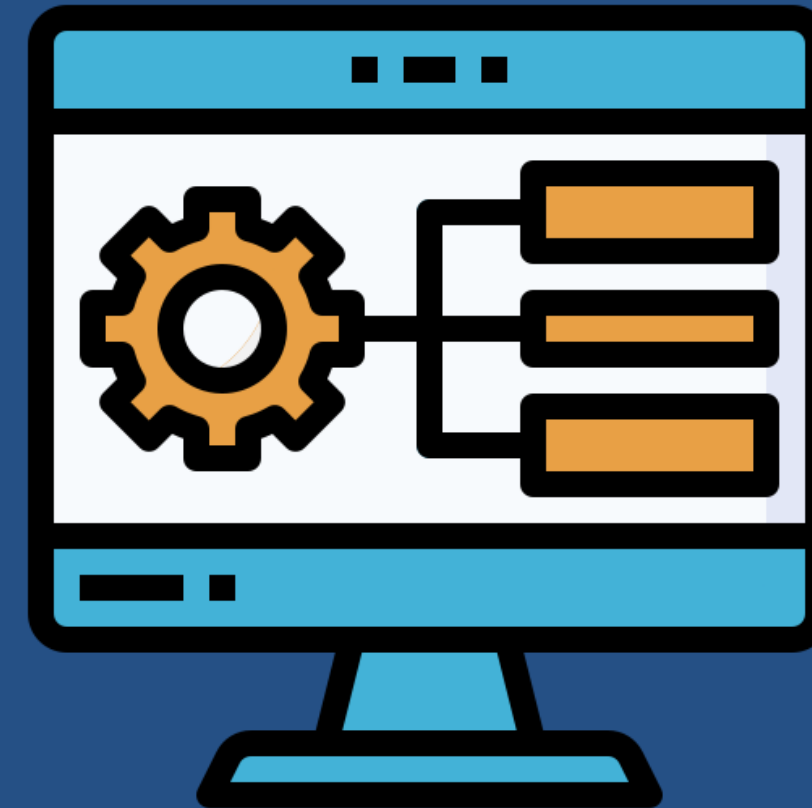


Amazon Redshift



Cientes de QSL y herramientas
de inteligencia empresarial (BI)

Por último, Amazon Redshift es compatible con las herramientas que ya conoce y utiliza. Amazon Redshift es compatible con SQL estándar. También proporciona conectores de alto rendimiento Java Database Connectivity (JDBC) y Open Database Connectivity (ODBC), que permiten utilizar los clientes SQL y las herramientas de BI de su elección.



A continuación, se presenta un repaso de algunos casos prácticos comunes de Amazon Redshift.



ALMACÉN DE DATOS EMPRESARIALES (EDW)

- 
- Migración a un ritmo agradable para los clientes
 - Experimentación sin un gran costo o compromiso inicial
 - Respuesta más rápida a las necesidades del negocio



BIG DATA

- Precios bajos para clientes pequeños
 - Servicio administrado para una implementación y mantenimiento sencillos
 - Mayor enfoque en los datos, menor enfoque en la administración de la base de datos
- 

Muchos clientes migran los almacenes de datos tradicionales de su empresa a Amazon Redshift con el objetivo principal de lograr una mayor agilidad. Los clientes pueden empezar a la escala que deseen y experimentar con sus datos sin necesidad de depender de procesos complicados con sus departamentos de TI para adquirir y preparar su software.

Los clientes de big data tienen algo en común: cantidades ingentes de datos que ponen a prueba sus sistemas actuales. Es posible que los clientes más pequeños no cuenten con los recursos para adquirir el hardware y los conocimientos necesarios para ejecutar estos sistemas. Con Amazon Redshift, los clientes más pequeños pueden configurar y utilizar rápidamente un almacén de datos a un precio relativamente bajo.

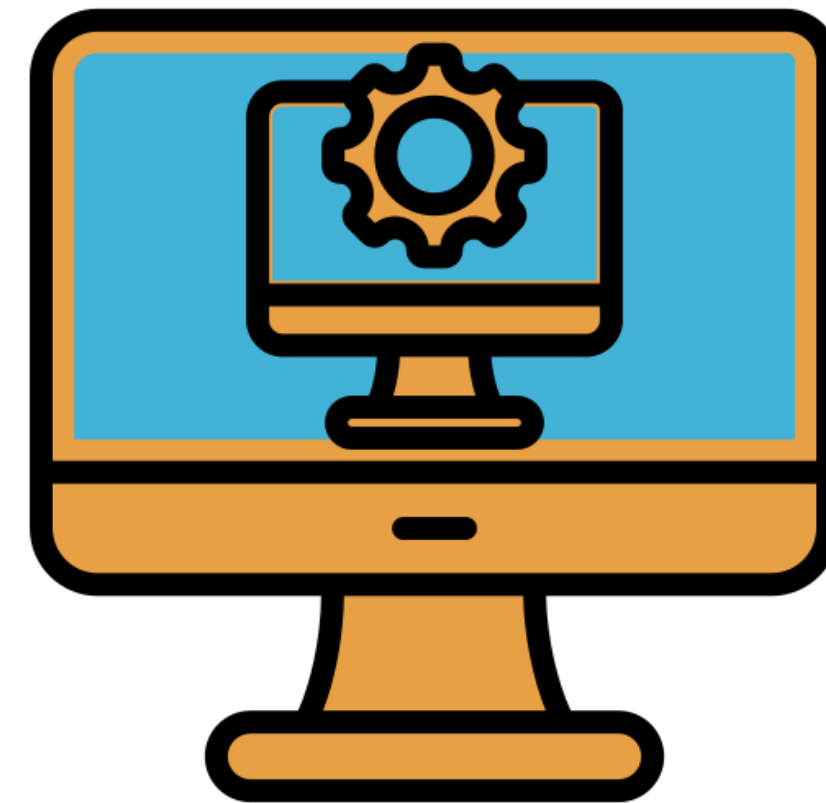
Como servicio administrado, Amazon Redshift se encarga de muchas de las tareas de implementación y mantenimiento continuo que suelen requerir un administrador de base de datos. Esto permite a los clientes centrarse en consultar y analizar sus datos.



CASOS PRÁCTICOS DE AMAZON REDSHIFT (2 DE 2)

SOFTWARE COMO SERVICIO (SAAS)

- Escalar la capacidad del almacén de datos a medida que crece la demanda
- Agregar funcionalidad de análisis a las aplicaciones
- Reducir los costos de hardware y software



Los clientes de software como servicio (SaaS) aprovechan las funciones escalables y fáciles de administrar que proporciona Amazon Redshift. Los clientes utilizan Amazon Redshift para dotar a sus aplicaciones de habilidades analíticas. Otros usuarios implementan un clúster por cliente y utilizan el etiquetado para simplificar y administrar la facturación y los acuerdos de nivel de servicio (SLA). Amazon Redshift permite reducir los costos de hardware y software.



CONCLUSIONES LECCIÓN 3

En resumen, Amazon Redshift es un servicio de almacén de datos rápido y completamente administrado. A medida que una empresa crece, puede escalar fácilmente sin tiempo de inactividad añadiendo más nodos. Amazon Redshift añade los nodos automáticamente a su clúster y redistribuye los datos para obtener el máximo rendimiento.

Amazon Redshift está diseñado para ofrecer un alto rendimiento de forma constante.

Amazon Redshift utiliza almacenamiento en columnas y una arquitectura de procesamiento paralelo masivo. Estas funciones paralelizan y distribuyen los datos y las consultas entre varios nodos. Amazon Redshift también supervisa automáticamente el clúster y realiza respaldos de los datos para que pueda restaurarlos con facilidad en caso necesario. El cifrado está integrado, solo tiene que activarlo.

Para obtener más información sobre Amazon Redshift, consulte <https://aws.amazon.com/efs/>