

Data Warehousing on AWS

Duración: 3 días

TEMARIO

Data Warehousing on AWS presenta conceptos, estrategias y prácticas recomendadas para diseñar una solución de almacenamiento de datos basada en la nube con Amazon Redshift, el almacenamiento de datos que permite escalar a petabyte en AWS. Este curso muestra de qué manera recopilar, almacenar y preparar datos para el almacenamiento de datos con otros servicios de AWS, como Amazon DynamoDB, Amazon EMR, Amazon Kinesis Firehose y Amazon S3. Además, muestra cómo utilizar herramientas de inteligencia empresarial para analizar sus datos.

DIRIGIDO A

Este curso está dirigido a:

- Arquitectos de bases de datos
- Administradores de bases de datos
- Desarrolladores de bases de datos
- Científicos y analistas de datos

OBJETIVOS DEL CURSO

Este curso le enseñará a:

- Analizar los principales conceptos del almacenamiento de datos.
- Evaluar la relación entre Amazon Redshift y otros sistemas de big data.
- Evaluar casos de uso de cargas de trabajo de almacenamiento de datos y revisar estudios de caso que muestran la implementación de servicios de análisis y datos de AWS como parte de una solución de almacenamiento de datos.
- Elegir un tamaño y tipo de nodo de Amazon Redshift adecuado para sus necesidades relacionadas con datos.
- Analizar características de seguridad relacionadas con Amazon Redshift, como cifrado, permisos IAM y permisos de bases de datos.
- Lanzar un clúster de Amazon Redshift y utilizar los componentes, las características y la funcionalidad para implementar un almacén de datos en la nube.
- Utilizar otros servicios de análisis y datos de AWS, como Amazon DynamoDB, Amazon EMR, Amazon Kinesis Firehose y Amazon S3 para complementar la solución de almacenamiento de datos.
- Evaluar estrategias y metodologías para diseñar almacenes de datos.
- Identificar orígenes de datos y evaluar requisitos que influyan en el diseño de un almacén de datos.

- Diseñar el almacén de datos para realizar un uso efectivo de métodos de compresión, distribución de datos y clasificación.
- Cargar y descargar datos y realizar tareas de mantenimiento de datos.
- Escribir consultas y evaluar planes de consulta para optimizar su rendimiento.
- Configurar la base de datos para asignar recursos, como memoria, a colas de consultas y definir criterios para derivar determinados tipos de consultas a sus colas de consultas configuradas para un mejor procesamiento.
- Utilizar características y servicios, como registro de auditoría de base de datos de Amazon Redshift, Amazon CloudTrail, Amazon CloudWatch y Amazon Simple Notification Service (Amazon SNS) para auditar, monitorear y recibir notificaciones de eventos acerca de actividades del almacén de datos.
- Prepararse para tareas operativas, como ajustar el tamaño de clústeres de Amazon Redshift y utilizar instantáneas para realizar copias de respaldo y restablecer clústeres.
- Utilice una aplicación de inteligencia empresarial (BI) para realizar tareas de visualización y análisis de datos de sus datos.

CONTENIDO

Día 1

- Introducción al curso
- Introducción al almacenamiento de datos
- Introducción a Amazon Redshift
- Comprensión del funcionamiento de los recursos y componentes de Amazon Redshift
- Lanzamiento de un clúster de Amazon Redshift

Día 2

- Revisión de las estrategias de almacenamiento de datos
- Identificación de requisitos y orígenes de datos
- Diseño de almacén de datos
- Carga de datos en el almacén de datos

Día 3

- Escritura de consultas y ajuste de rendimiento
- Mantenimiento del almacén de datos
- Análisis y visualización de datos
- Resumen del curso

PRE-REQUISITOS

Aconsejamos que los asistentes a este curso cumplan con los siguientes requisitos previos:

- Cursos realizados: Fundamentos técnicos de AWS (o experiencia equivalente en AWS)

- Estar familiarizado con conceptos de diseño de bases de datos y bases de datos relacionales