**[Diapositiva 1: Título]**

**Introducción a Big Data y CassandraDB**

**Hola a todos**, hoy les hablaré sobre **Big Data** y cómo **CassandraDB** juega un papel crucial en el manejo de grandes volúmenes de datos. En particular, veremos cómo **Cassandra** se utiliza en plataformas que todos conocemos y usamos a diario, como **Netflix**, **WhatsApp** y **Spotify**.

**[Diapositiva 2: ¿Qué es Big Data?]**

**Big Data**

* **Definición**: Big Data hace referencia a conjuntos de datos que son demasiado grandes o complejos para ser procesados por sistemas tradicionales de gestión de bases de datos.
* **Características clave**:
  + **Volumen**: Grandes cantidades de datos.
  + **Variedad**: Diferentes tipos de datos (estructurados, no estructurados, semi-estructurados).
  + **Velocidad**: La rapidez con la que los datos se generan y deben ser procesados.

**[Diapositiva 3: Introducción a CassandraDB]**

**¿Qué es Cassandra?**

* **Cassandra** es una base de datos distribuida diseñada para gestionar grandes cantidades de datos en tiempo real sin perder rendimiento.
* Se caracteriza por:
  + **Escalabilidad horizontal**: A medida que aumenta el volumen de datos, Cassandra puede añadir nodos para distribuir la carga.
  + **Alta disponibilidad**: Si un nodo falla, el sistema sigue funcionando sin interrupciones.
  + **Desempeño optimizado para escrituras rápidas**.

**[Diapositiva 4: Características clave de Cassandra**

**Características Principales**

* **Distribuida y descentralizada**: No existe un solo punto de fallo.
* **Escalabilidad**: Se puede escalar fácilmente añadiendo más nodos.
* **Tolerante a fallos**: Asegura que los datos siempre estén disponibles, incluso si algunos nodos fallan.
* **Desempeño optimizado para grandes volúmenes de datos**.

**[Diapositiva 5: Aplicaciones famosas que usan Cassandra**

**¿Dónde se usa Cassandra?**

Cassandra es utilizada por algunas de las plataformas más grandes y conocidas, manejando grandes volúmenes de datos distribuidos de manera eficiente y a gran escala. A continuación, les hablaré de algunas aplicaciones famosas que dependen de Cassandra.

**[Diapositiva 6: Netflix y Cassandra**

**1. Netflix**

* **Uso de Cassandra**: Netflix usa Cassandra para gestionar las interacciones de los usuarios y ofrecer recomendaciones personalizadas.
  + **Almacenamiento de preferencias**: Los datos sobre lo que vemos, cuándo lo vemos y nuestras calificaciones se almacenan en Cassandra.
  + **Recomendaciones personalizadas**: Cassandra ayuda a generar recomendaciones en tiempo real basadas en el comportamiento de los usuarios.
  + **Alta disponibilidad**: Asegura que los datos de los usuarios estén siempre disponibles, incluso si una parte de la infraestructura falla.

**[Diapositiva 7: WhatsApp y Cassandra**

**2. WhatsApp**

* **Uso de Cassandra**: WhatsApp maneja más de 100 mil millones de mensajes diarios, y Cassandra es clave en esta infraestructura.
  + **Mensajes en tiempo real**: Los mensajes, chats y metadatos se almacenan y distribuyen en Cassandra para garantizar la entrega rápida de mensajes.
  + **Escalabilidad**: Gracias a Cassandra, WhatsApp puede gestionar el crecimiento exponencial de usuarios sin afectar el rendimiento.
  + **Alta disponibilidad**: Asegura que los mensajes estén disponibles y se entreguen correctamente, incluso si algunos nodos fallan.

**[Diapositiva 8: Instagram y Cassandra**

**3. Instagram**

* **Uso de Cassandra**: Instagram maneja miles de millones de interacciones (likes, comentarios) cada día, y Cassandra ayuda a gestionarlas eficientemente.
  + **Interacciones del usuario**: Datos como **likes** y **comentarios** se almacenan en Cassandra y se actualizan en tiempo real.
  + **Escalabilidad**: Con millones de interacciones diarias, Cassandra asegura que la plataforma pueda crecer y seguir funcionando sin problemas.
  + **Alta disponibilidad**: Los usuarios siempre tienen acceso a sus datos y contenido, incluso si un servidor falla.

**[Diapositiva 9: Spotify y Cassandra**

**4. Spotify**

* **Uso de Cassandra**: Spotify gestiona grandes volúmenes de datos relacionados con las preferencias musicales de los usuarios.
  + **Datos de usuario y música**: Los registros de escucha, las listas de reproducción y las preferencias musicales se almacenan en Cassandra.
  + **Recomendaciones musicales**: Usan Cassandra para generar recomendaciones basadas en el historial de escucha de los usuarios.
  + **Escalabilidad global**: Cassandra distribuye los datos en varios centros de datos, garantizando que todos los usuarios puedan acceder a sus canciones sin importar su ubicación.

**\*\*[Diapositiva 10: Ejemplo práctico con Docker y Java]**

**Ejemplo práctico con Cassandra, Docker y Java**

Ahora que entendemos cómo Cassandra está presente en plataformas globales, vamos a ver un ejemplo simple usando **Docker** y **Java** para interactuar con Cassandra.

1. **Levantar Cassandra en Docker**:

bash

Copiar código

docker run --name cassandra -d cassandra:latest

Este comando crea un contenedor con Cassandra en tu máquina local.

1. **Conectar Cassandra con Java**: Utilizaremos la **Java Driver API** de Cassandra para conectar nuestra aplicación Java con la base de datos:

java

Copiar código

Cluster cluster = Cluster.builder().addContactPoint("127.0.0.1").build();

Session session = cluster.connect();

session.execute("CREATE KEYSPACE IF NOT EXISTS test WITH REPLICATION = {'class': 'SimpleStrategy', 'replication\_factor': 3}");

session.execute("CREATE TABLE IF NOT EXISTS test.users (id UUID PRIMARY KEY, name text)");

session.execute("INSERT INTO test.users (id, name) VALUES (uuid(), 'John Doe')");

1. **Ejecutar consultas**:
   * Podemos ejecutar consultas en Cassandra como insertar y recuperar datos, todo gestionado a través de Java.

**[Diapositiva 11: Conclusión]**

**Conclusión**

* **Cassandra** es una herramienta poderosa para manejar grandes volúmenes de datos, y su capacidad para escalar y ofrecer alta disponibilidad la convierte en una opción ideal para empresas con necesidades de Big Data.
* Empresas como **Netflix**, **WhatsApp**, **Instagram** y **Spotify** confían en Cassandra para mantener sus plataformas en funcionamiento y mejorar la experiencia del usuario.

Gracias por su atención. ¿Alguien tiene alguna pregunta?