

# Resúmenes 7 y 8

## *Bases de Datos II*

Luis Diego Delgado Muñoz

Prof. Nereo Campos

Fecha de Entrega: 01/11/2022

---

### Cinco de los principales casos de uso de la tecnología de grafos:

1. **Detección de fraudes:** El almacenamiento de detalles voluminosos de transacciones en una base de datos de gráficos captura las conexiones que ya existen. Luego, los datos están maduros para las técnicas de aprendizaje automático de grafos nativos que revelan el fraude.
2. **Recomendaciones de Tiempo Real:** Al ejecutarse en Neo4j, el sistema de recomendación de Bechtle genera un modelo detallado de todas las acciones que los clientes realizan en el sitio web. Al ejecutarse en Neo4j, el sistema de recomendación de Bechtle genera un modelo detallado de todas las acciones que los clientes realizan en el sitio web.
3. **Lista de Materiales:** El ejército utiliza Neo4j para recopilar y combinar rápidamente esta información masiva de BoM y ahorrar a los analistas una gran cantidad de tiempo. Las respuestas son inmediatas. Con Neo4j, el ejército ahora puede pronosticar la necesidad de piezas de repuesto, calcular con precisión el tiempo medio de falla y responder preguntas vitales de "qué pasaría si" sobre el costo del despliegue de fuerzas. Los beneficios incluyen no solo un costo total de propiedad más predecible, sino también la garantía de que el equipo enviado al campo no se averiará gracias a un mantenimiento predictivo eficaz.
4. **Seguimiento y rastreo:** Los modelos de datos grafos permiten la trazabilidad para una variedad de industrias y casos de uso, incluidos el enrutamiento, la logística, la gestión de la cadena de suministro y el cumplimiento. Seguimiento y rastreo nos permite conocer el estado de un producto (lote, correo/paquete, tren, barco, contenedor, etc.) dondequiera que se encuentre en la cadena de suministro e identificar y verificar su ruta.
5. **Operaciones de red y TI:** Las bases de datos de gráficos son una opción natural para las operaciones de red y TI; después de todo, una red es un grafo. La estructura conectada del grafo permite a los administradores de red realizar análisis de impacto sofisticados. Por ejemplo, pueden descubrir de qué partes de la red (aplicaciones, servicios, máquinas virtuales, máquinas físicas, centros de datos, enrutadores y conmutadores) dependen determinados clientes. Por el contrario, una simple consulta de grafo determina qué aplicaciones, servicios y clientes se verán afectados si falla un elemento de red en particular.

### Las conexiones desbloquean el valor de los datos

El valor de los datos proviene de la capacidad de una organización para comprenderlos en relación con otros datos. Por sí mismos, los datos ofrecen un valor finito. Cuando están conectados, el valor de los datos es infinito. Debido a que las bases de datos

relacionales no conservan la información de la relación en el almacenamiento ni en ninguna otra etapa de sus ejercicios analíticos, encontrar conexiones requiere una enorme cantidad de procesamiento adicional. Y mantener estas conexiones a lo largo de su vida es casi imposible.

Los datos no se conectan simplemente descargándolos en un lago de datos central. Los datos se conectan cuando trata la información de la relación como una entidad de primera clase: la persiste, le asigna propiedades y la usa como un medio para desarrollar contexto para las aplicaciones. Por lo tanto, es mejor pensar en los grandes datos que necesitan estar conectados, no simplemente contenidos. Sin datos conectados, las organizaciones carecen de la información clave necesaria para obtener una vista de 360 grados de un cliente, construir una topología de red completa, brindar recomendaciones relevantes en tiempo real u obtener la visibilidad necesaria para prevenir el fraude. El panorama de datos ha crecido, al igual que la cantidad y los tipos de conexiones de las que dependen estos datos, lo que hace que la necesidad de sacar a la luz datos conectados sea mucho más urgente.

## **El poder de los datos conectados**

El mayor beneficio de los datos conectados es la capacidad de proporcionar una vista conectada de los datos a sus aplicaciones analíticas y operativas, y así obtener y aumentar la inteligencia en sentido descendente. Las conexiones pueden ponerse a disposición o revelarse a aplicaciones o usuarios comerciales para tomar decisiones operativas. Los datos conectados son más poderosos cuando brindan información operativa en tiempo real y no solo análisis posteriores. Los conocimientos en tiempo real permiten a los usuarios y aplicaciones comerciales tomar decisiones comerciales y actuar en tiempo real. Por lo tanto, los motores de recomendación aprovechan los datos de la sesión actual del usuario para ofrecer opciones muy relevantes que el cliente quiere o necesita en ese momento.

Los datos y las aplicaciones de hoy requieren elasticidad, agilidad, velocidad e interconectividad. A pesar del nombre, las bases de datos relacionales no son adecuadas para modelar y almacenar los conjuntos de datos ágiles y altamente conectados de la actualidad. Los RDBMS exigen rediseños de esquemas lentos y costosos que perjudican los procesos ágiles de desarrollo de software y dificultan su capacidad de escalar e innovar rápidamente. La tecnología RDBMS tradicional tiene dificultades para expresar y revelar cómo se relacionan las entidades reales y virtuales. Las columnas y las filas no son cómo existen los datos en el mundo real. Más bien, los datos existen como objetos ricos y las relaciones entre esos diferentes objetos.

La teoría de redes conectadas (teoría de grafos) ha sido una disciplina matemática durante casi tres siglos, pero pocas tecnologías han aprovechado estos modelos teóricos con el fin de almacenar y analizar datos.

## **Recolecciones vs. Conexiones**

Los diseños de almacenamiento centrados en la recopilación implementados por las bases de datos SQL y No solo SQL (NoSQL) están diseñados para dividir y almacenar datos de manera eficiente. En el caso de SQL, la normalización de datos en un esquema tabular tiene como objetivo minimizar el almacenamiento de objetos, tipos y valores de datos duplicados. Estos sistemas nacieron durante la era de la escasez de memoria física y el costoso almacenamiento basado en disco, diseñados para evitar la gestión de objetos

de datos a menudo redundantes, como direcciones de ubicaciones físicas para envíos, facturación, hogares, oficinas, destinos, negocios, etc.

Los sistemas NoSQL, como los almacenes de datos de documentos, columnas anchas y valores clave, llevan esos conceptos hacia adelante (y hacia atrás) al simplificar sus modelos a cambio de niveles más altos de escala y simplicidad. Al evitar las relaciones de datos y proporcionar API programáticas simples, los sistemas NoSQL facilitan que los desarrolladores y administradores trabajen con datos simples de una manera que se pueda escalar fácilmente. La falta de preocupación por las relaciones conduce a garantías de datos más flexibles, API simples y esquemas de escala sencillos. Los datos se distribuyen fácilmente y se recuperan con la misma facilidad, sin necesidad de mantener la integridad de los datos relacionados que se escriben en un almacenamiento distribuido o un grupo de máquinas y sin necesidad de preocuparse por el rendimiento de JOIN distribuidos en esas máquinas.

## **Beneficios de las bases de datos de grafos**

- Modelado de datos simple y natural
- Flexibilidad para estructuras de datos en evolución
- Soporte simultáneo para actualizaciones y consultas en tiempo real
- Consultas y análisis mejores, más rápidos y más potentes