

Resumen 5

Edgar André Araya Vargas 2020142856

Bases de datos de grafos para principiantes.

El capítulo 1 del texto habla sobre las bases de datos gráficas y por qué son importantes para el futuro de la tecnología. Estas bases de datos se distinguen de las bases de datos tradicionales en que priorizan las relaciones entre los puntos de datos en lugar de los propios datos. Esto hace que sean más rápidas y flexibles para manejar grandes conjuntos de datos, ya que pueden adaptarse a medida que cambian los requisitos empresariales. Las bases de datos gráficas están compuestas por nodos (entidades) y relaciones (cómo están conectados los nodos). Por ejemplo, en Twitter, cada usuario es un nodo y las relaciones son quién sigue a quién.

Las bases de datos gráficas son útiles en una amplia variedad de campos, como la optimización de rutas de logística, la detección de fraudes, la monitorización de redes sociales y la sugerencia de productos al por menor. Las bases de datos gráficas también son más intuitivas de entender que las bases de datos relacionales y NoSQL, lo que las hace más fáciles de utilizar en situaciones de desarrollo ágil. Además, las bases de datos gráficas nativas, como Neo4j, son más rápidas y eficientes que las no nativas. Las bases de datos gráficas son una tecnología emergente que está cambiando la forma en que se manejan grandes conjuntos de datos y se comprenden las relaciones complejas entre los datos. Son más rápidas, flexibles e intuitivas que las bases de datos tradicionales y son útiles en una variedad de campos. Las empresas que adopten las bases de datos gráficas pueden obtener una ventaja competitiva en el mercado.

El capítulo 2 habla sobre la importancia de las relaciones de datos y cómo las bases de datos relacionales, a pesar de su eficiencia en la codificación de formas y estructuras tabulares, tienen limitaciones en el manejo de relaciones de datos, especialmente cuando éstas son añadidas o ajustadas de manera ad hoc. La mayor debilidad de las bases de datos relacionales es que su esquema es demasiado inflexible para adaptarse eficientemente a las cambiantes necesidades empresariales, lo que obliga a los desarrolladores a escribir más código para manejar las excepciones en los datos. Además, a medida que los datos se multiplican en complejidad y diversidad, las bases de datos relacionales se ven obstaculizadas por grandes tablas JOIN que afectan el rendimiento y dificultan el desarrollo adicional.

A diferencia de las bases de datos relacionales y las NoSQL, las bases de datos de grafos almacenan las relaciones de datos como relaciones, lo que significa que hay menos desconexiones entre el esquema en evolución y la base de datos real. La flexibilidad de un modelo de grafo permite añadir nuevos nodos y relaciones sin comprometer la red existente o migrar los datos de forma costosa. Además, las bases de datos de grafos son extremadamente eficientes en cuanto a la velocidad de las consultas, incluso para consultas profundas y complejas. Las relaciones de datos son vitales y las bases de datos de grafos son la mejor opción para manejar conexiones de datos complejas y cambiar esquemas de manera flexible sin comprometer los datos existentes.

El capítulo 3 de este artículo se centra en los conceptos básicos de modelado de datos y cómo difieren los enfoques para modelar datos en bases de datos relacionales y en bases de datos de grafos. El modelado de datos es un proceso de abstracción. Se comienza con las necesidades de negocio y de los usuarios y se mapean esas necesidades en una estructura para almacenar y organizar los datos. Sin embargo, para las bases de

datos relacionales, el modelado puede ser complejo. Después de haber elaborado las ideas iniciales en una pizarra, se requiere crear un modelo lógico y luego forzar esa estructura en un modelo físico tabular. Para el momento en que se tiene una base de datos de trabajo, ésta no se parece en nada al boceto original (lo que hace difícil determinar si se están cumpliendo las necesidades de los usuarios). Por otro lado, el modelado de datos para una base de datos de grafos es mucho más simple. Es fácil imaginar que una estructura en una pizarra sería una colección de círculos y cajas conectados por flechas y líneas. Lo sorprendente es que ese modelo ya es un grafo. Crear una base de datos de grafos a partir de ahí es sólo cuestión de ejecutar algunas líneas de código. Se enriquece el modelo de la pizarra según las necesidades de los usuarios y de negocio, en lugar de alterar el boceto inicial para ajustarlo a tablas y relaciones.

El capítulo 4 de un artículo habla sobre los problemas que pueden surgir al modelar datos en bases de datos de grafos. A pesar de que estas bases de datos son altamente expresivas para problemas complejos, no garantizan que el modelo de datos sea correcto desde el primer intento. Incluso los expertos en bases de datos de grafos cometen errores, y los principiantes tienen más probabilidades de cometerlos. El artículo presenta un ejemplo de modelo de datos que se utiliza para detectar fraudes en las comunicaciones de correo electrónico. Este modelo no es suficientemente sólido para detectar comportamientos indebidos, por lo que se mejoró agregando relaciones para realizar un seguimiento de las respuestas y reenvíos de correos electrónicos. Sin embargo, esta mejora resultó insuficiente. Para crear un modelo de datos más efectivo, se debe considerar los fundamentos del dominio específico.

Por ejemplo, en el caso de las respuestas a los correos electrónicos, una respuesta es un nuevo correo electrónico y una respuesta al original. Los dos roles se pueden representar mediante dos etiquetas: "correo electrónico" y "respuesta" unidas a los nodos apropiados. Finalmente, se presenta una tarea para los estudiantes de modelado de datos para documentar cómo se modelarían los datos de correo electrónico reenviados, rastreando las relaciones con los remitentes, los destinatarios directos, los destinatarios en CC, los destinatarios en CCO y el correo electrónico original. En resumen, aunque la modelación de datos se ha vuelto más fácil con el uso de bases de datos de grafos, es importante evitar errores comunes para extraer información valiosa de los datos.