

Instituto Tecnológico de Costa Rica

Bases de Datos - IC4302

Resumen 7 y 8

Estudiante:

Andrea María Li Hernández - 2021028783

Profesor:

Gerardo Nereo Campos Araya

Fecha de entrega: 5 de noviembre del 2022
Segundo Semestre, 2022

Neo4j and Graph Data Platforms

Introduction to Graph Data Platforms

Graph databases are useful for real-time analysis and they're specifically designed and optimized for highly interconnected datasets to identify patterns and hidden connections. They mirror the way a human brain works by mapping association through neurons (**nodes**) and synapses (**relationships**).

This kind of database efficiently stores and queries connected data in a node-and-relationships format.

When to use them?

- When there is **no a priori knowledge of path length or shape**.
- When we have to solve complex problems with **dynamic and interconnected datasets**.
- When there's **lack of schema flexibility and adaptability**.

Common use cases:

- For **fraud detection and analytics**: Real-time analysis of data relationships is essential to uncovering fraud rings.
- Other common use cases are: AI and Machine learning, real-time recommendation engines, knowledge graphs and Master Data Management (MDM).

Traditional technologies vs Connected data technologies

Traditional technologies	Connected data technologies
a. They have a difficult time expressing and revealing how real and virtual entities are related.	a. They express how data exists in the real world as rich objects and the relationships between those objects.
b. Emphasis on data aggregation & collection.	b. Emphasis on data connections.
c. Designed to minimize storage of duplicative data values since disk space was costly	c. Designed to address interrelated, contextualized data and how it might be traversed to reveal unobvious relationships.

Graph Systems Focus on Data Connections

Connected data is the materialization and harnessing of relationships between data elements, which is modeled as a **property graph**

- **Property graph**: It is a data model designed to express data connectedness as nodes connected via relationships to other nodes, where both nodes and relationships can have properties (name/value pairs) attached to them.
 - Properties can be updated easily since its design eliminates most of the structural overhead of traditional DB schemas.
 - You can visualize and navigate property graphs efficiently by following the relationships on their paths to context.

Data relationships are persisted so they can be navigated along connected paths to gain context. Also, relationships are both typed and directional.

- **Importance of context:** The context provided by the connections is essential to identify friendships, making relevant real-time recommendations and following money trails.

Benefits of Graph Databases

- **Simultaneous support for real-time updates and queries:** This model allows real-time updates on graph data while supporting queries concurrently.
- **Better, faster and more powerful querying and analytics:** They provide high query performance with using native storage and native indexed data structure.

Conclusion about Graph Databases

Nowadays, businesses need to bring together data from many different systems, which makes it necessary to implement new data integration models. Complex data environment often requires real-time data analysis to make split-second decisions to optimize the business, which is why Graph databases are a scalable, secure and flexible platform for enterprises.

The Return on Connected Data

In this section, you'll see information about the relevance of connected data and a brief description of Neo4j.

What is connected data?

- It is the representation, usage and persistence of relationships between data elements. The key is to maintain ongoing knowledge of the relationship.

Connections unlock the value of data

- The value of data comes from an organization's ability to understand it in relation to other data. When connected, data's value is infinite.
- Increasing data's connectedness further increases its value through additional context.
- Business leaders can take advantage of relationships that either live in data or are represented by data as part of how people interact with the processes.

For example, Neo4j **stores connections** with the data, so there's no extra work, the data is already connected.

Data everywhere , nor any drop to drink

Many times, businesses collect all the data they can, but there is no insight since the values are within the relationships or connections between the data. Data becomes connected when you treat relationship information as a first-class entity. Without connected data, organizations lack key information that's necessary to obtain a 360-degree view of a customer.

The power of connected data

- The biggest benefit of connected data is the ability to provide a **connected view of the data to your analytic and operational applications**, thereby gaining and growing intelligence downstream.

- Besides getting better understanding for recommendations, connected data is to understand the flow of money to detect fraud.
- Recommendation engines leverage data from the current user session to deliver highly relevant options that the customer wants or needs at that moment.

Quench your thirst for insights with connected data

A graph database is a highly scalable transactional and analytic database that stores data relationships as first-class entities.

A Graph DB can address the challenge of migrating all the data into one connected place. Also, you can use it to transfer knowledge of what the organization has done across other departments. The graph DB operationalizes the data by making it possible for the app to traverse the data; once you have this ability, the data becomes **actionable**.

Neo4j: Your path to connected data

Neo4j offers an innovative, reliable native graph platform that reveals and maintains the integrity of connected data. It offers a connections-first approach as a new paradigm to reveal and utilize relationships among data. It allows you to easily **query** and **visualize** your connected data for patterns and trends that are essential for a sustainable competitive advantage.

- Native graph store.
- Complete graph platform.
- Flexible schema.
- Performance and scalability.
- High availability.

Conclusion

In order to obtain the deepest and most revealing insights, you must understand the connections between your data and preserve relationships as you go. With Neo4j, enabling connected data is easy so organizations can improve their data quality and prevent fraud.

5 use cases for graph technology

1. **Fraud Detection:** Storing voluminous transaction details in a graph database captures connections that already exist, so the data is ripe for graph-native machine learning techniques that surface fraud.
 - Fintech firm TD01 uses Neo4j to power proprietary fraud detection app.
2. **Real-Time Recommendations:** Recommendations must be contextual and instantaneous.
 - Bechtle AG uses Neo4j to generate a detailed model of all the actions customers take on their website, this is how they can optimize the customers' navigation.
3. **Bill of Materials:** To track every component and its cost, what equipment the components relate to, and the expected product lifespan is important information for an organization like the United States Army.
 - The Army uses Neo4j to rapidly collect Bill Of Materials.

4. **Track & Trace:** Graph data models enable traceability for a variety of industries and use cases, including routing, logistics, supply chain management, and compliance. It lets us find out the status of the product and indentify its path.
 - An example is a pharmaceutical company that uses graph DBs to improve recall management.
5. **Network & IT Ops:** Network is itself a graph, so this enables network managers to conduct sophisticated impact analyses.
 - Commonwealth Bank uses Neo4j to easily map out its business applications and see relationships in a highly visual and intuitive way using graphs.