What's New in Neo4j 5

October 2022

This document summarizes the new capabilities included in the Neo4j 5 release.

What Is Neo4j 5?

In a world where the amount of data created, replicated, and stored continues to grow exponentially, Neo4j helps the world make sense of data, especially when relationships between data are paramount. Neo4j allows you to ask complex questions from multiple layers of data from different data sources, unearthing hidden insights that allow you to support new use cases and business models.

Neo4j 5 is the next-generation graph data platform from Neo4j Inc. that makes it effortless to take your superior graph performance, scale it in the cloud of your choice, and meet the growing demands of your intelligent applications.

Areas of Investment

We continue to build on making Neo4j meet the scale expectations of a world powered by the insights found in connected data, with particular emphasis on these three areas:

- **Performance and scalability.** As a pioneer and leader in the graph database space, Neo4j has a 1000x query performance advantage over relational databases, especially when relationships are paramount.

- **Developer and data scientist experience.** Neo4j makes it easy for developers and data scientists to build applications and explore connected data on our graph data platform using their preferred languages and tools. We also make it easy to interoperate the database with other data platforms and applications.

- **Operational trust.** Neo4j meets the security standards and reliability expectations of the world's top companies. Neo4j removes the friction of running our database by bringing the capabilities of the world's number 1 graph data platform to the cloud, and by infusing the discipline of cloud operations into the self-managed product.

What's New?

In version 5, we focus on making it operationally effortless to scale Neo4j and achieve even faster queries in analyzing data, even when it involves very large datasets or when it is deployed in multiple databases or clusters in the cloud.

Performance & Scalability:
Widening the performance lead, with unbounded scalability

Neo4j 5 widens its performance lead over relational and other databases by accelerating deep queries and making it easy to scale that performance to unlimited heights.

- **Autonomous Clustering.** Neo4j 5 provides a cloud-ready architecture for global clusters that provide scale out for your data and the number of databases without scaling up your cluster. It automatically handles the placement of primary and secondary copies to the most optimal server in the cluster, reducing manual effort and optimizing your infrastructure capacity utilization.

- **Multi-cluster Fabric.** Neo4j 5 allows you to use Cypher to quickly create Fabric, a composite graph of other graph databases or shards. Use Fabric in an Autonomous Cluster to execute queries across databases in a cluster, and even those in remote clusters without the need for separate Fabric proxies.

- **Incremental import.** Neo4j 5 allows you to merge bulk data incrementally into an existing database, drastically reducing data loading time and providing flexibility for loading large datasets.

- **Faster K-Hop queries.** K-Hop is a type of deep query that involves a large and variable number of hops – where K is the number of hops – in order to find all unique nodes, typically in combination with aggregation functions to count properties. In Neo4j 5, this type of query has been optimized to go breadth first, resulting in blazingly fast response times; for example, the response time for 8-hop queries has been improved by 1000x.

- **Faster runtime for Community Edition.** The runtime component has been upgraded from Interpreted to Slotted, which allows Community Edition to provide ~30% faster reads.

Developer Experience: Easier, more powerful queries

Neo4j 5 includes improvements and optimizations that make writing queries in Cypher easier and query execution more powerful.

- **Graph pattern matching improvements and query plan optimizations.** MATCH clauses are easier to write and read, since users can now express relationship filters inline in the path and take advantage of more sophisticated label expressions. Query plans have been optimized for faster plan execution for some types of Cypher queries.

- **Enhanced indexes.** Indexes are essential for quickly identifying the most efficient starting point, node(s) or relationship(s) for a query. Neo4j 5 has extended the matching capabilities of indexes:
    - FULLTEXT now indexes lists and arrays of strings to improve the quality of text search results
    - RANGE allows you to specify or compare values, e.g., find reviews rated 3-5 by users in postal codes 94000-95000
    - POINT, often used in routing and supply chain analysis, allows you to find and compare spatial data like longitude and latitude.

Operational Trust: Agility for operating in the cloud

It has always been possible to deploy Neo4j in the cloud, and Neo4j is available as a service in the form of AuraDB and AuraDS. Neo4j 5 brings the discipline and expectations of cloud operations to self-managed Neo4j and enables seamlessness between self-managed Neo4j and Aura.

- **Neo4j Ops Manager.** Neo4j Ops Manager provides an intuitive operations console for monitoring and administering Neo4j deployments (database, instance, or cluster).

- **Cloud-first, continuous release support model.** The latest and greatest innovations are now available on a frequent release schedule to self-managed users, just as they have been for Aura users.

- **Any-to-any rolling upgrade.** Upgrade single servers or perform upgrades without downtime on Neo4j 5 clusters from any Neo4j 5 version to any later minor version without going through intermediary versions.

- **Backup and restore.** A more efficient backup engine provides more control and a simpler and faster backup experience:
    - Differential backup and a single compressed file archive
    - Point-in-time restore
    - APIs for inspecting and managing backup files
    - Option to turn on consistency checking