

Hochperformante Abfragen im Graphen

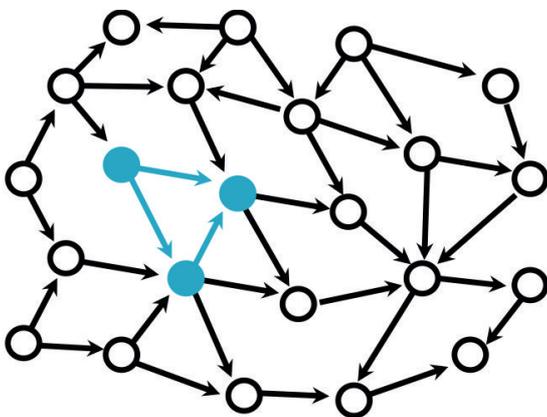
Graphendatenbanken bieten eine moderne und effiziente Lösung, um Datenstrukturen natürlich abzulegen. Schemafrei werden strukturierte und verknüpfte Daten in Knoten und Kanten abgelegt. Mit bekannten Algorithmen der Graphentheorie gelingt eine problemlose und hochperformante Auswertung der Daten.

Im Zeitalter von Big Data sammeln viele Unternehmen Daten über das Verhalten ihrer Produkte und Kunden. Unter dem Schlagwort IOT (Internet of Things) werden auch Messdaten von Sensoren beispielsweise aus Produktionsstrassen, Smarthomes, Mobiltelefonen oder Wearables erfasst. Diese Daten sind möglichst im Rohformat abzuspeichern und in Zukunft mit Fragestellungen zu analysieren, die derzeit noch nicht erschlossen sind. Doch wie ist diese Flut an Daten zu speichern und anschliessend performant auszuwerten? Relationale Datenbanken kommen in vielen dieser Anwendungsfälle an ihre Leistungsgrenzen.

Unter dem Begriff NoSQL vereinen sich neue Arten von Datenbanken. Sie orientieren sich nicht an relationalen Datenbanken, sondern beschreiten einen neuen Weg. Diese Datenbanken sind hauptsächlich Variationen des Key-Value-Modells und speichern unstrukturierte, schemalose Daten. Darunter existiert jedoch ein Sonderfall, der bei der Modellierung der Daten eine stärkere Ausdrucksfähigkeit besitzt und sich auf zusammenhängende Daten spezialisiert hat.

Graphendatenbanken

Graphendatenbanken speichern Informationen und deren Beziehungen. Eine Technologie, die aktuell die Datenspeicherung und -verarbeitung erfolgreich verändert. Dies bestätigten auch Analysten von Forrester im vergangenen Jahr, als sie prognostizierten, dass 25 Prozent aller Unternehmen bis 2017 Graphendatenbanken nutzen werden. Graphendatenbanken kommen in Onlinesystemen mit



Informationen werden im Graphen über «Muster» gefunden.

höchsten Anforderungen an Performance und Verfügbarkeit zum Einsatz. Mit einem flexiblen, auf den Anwendungsfall ausgelegten Datenmodell erzielt man hochperformante Leistungen. Informationen, repräsentiert durch Knoten, sind durch Beziehungen (Kanten) miteinander verbunden. Kanten besitzen die Fähigkeit, gerichtet und benannt zu sein und ebenfalls Eigenschaften zu enthalten. Die Datenbank speichert diese Daten in einer Netzstruktur. Eine Abfrage erfolgt somit ohne Berechnung der Beziehungen und entspricht einer Navigation von Knoten über Kanten zu Knoten. Diese Traversierung, beschrieben mit gängigen Algorithmen aus der Graphentheorie, wird direkt in der Datenbank ausgeführt.

Einsatz in Unternehmen

Graphen eignen sich optimal, um logische Probleme zu lösen. Für gewöhnlich sind dies beispielsweise soziale Netzwerke, Empfehlungserstellung, Betrugserkennung, Netzwerkmanagement, Routenplanung, Logistik, Wissensdatenbanken, Time-Series-Daten oder ACLs. Um ein Beispiel zu nennen: Sie planen eine Route mit dem Auto im Strassennetz. Ihr GPS versucht, die Strecke zwischen A und B zu berechnen, und ist imstande, die schnellste, ökologischste oder kürzeste Strecke zu finden. Dies ist ein praktischer Anwendungsfall, den Reiseroutenplaner, Logistikfirmen oder auch IT-Infrastrukturanbieter beanspruchen.

NoSQL-Datenbanken, insbesondere Graphendatenbanken, haben nicht den Anspruch, relationale Datenbanken abzulösen. Vielmehr verstehen sie sich als Lösung für spezialisierte Fragestellungen, bei denen klassische Datenbanken an ihre Grenzen stossen. Die Datenmodellierung im Graphen ist für jeden verständlich. Ein Projektteam entwirft und diskutiert die Konzepte mit dem Domänenexperten, zeichnet diese als Elemente und Pfeile auf ein Whiteboard und übernimmt dieses Modell – meist ohne Anpassungen – in die Datenbank. Fragestellungen und Probleme lassen sich direkt am Entwurf analysieren und diskutieren.

Graphendatenbanken eignen sich für die unterschiedlichsten Anwendungsgebiete. Das zugrundeliegende Datenmodell bietet viel Flexibilität, Dateneinsicht und Ausdruckskraft. Graphendatenbanken basieren auf einer reifen und verlässlichen Technologie. Es ist lohnenswert, diese Technologie in Betracht zu ziehen, bevor eine Datenspeicherung mit klassischen Systemen festgelegt wird.

