neotechnology

ZEPHYR TRANSFORMS DATA

Case Study

Cloud Analytics Platform Using Graph Database

ZEPHYR HEALTH

INDUSTRY

Health Care

CHALLENGE

- Enable end users to "be their own data scientist", discovering new connections between their data
- Develop an easy-to-use Cloud-based engine to mine data from disparate sources

SOLUTION

New architecture leveraging Neo4j and MongoDB for BI reporting

IMPLEMENTATION

- Rapid implementation
- 12-week POC
- Model validated & tuned via load testing

RESULT

High-speed, scalable complex analytics platform

Supports traversals with 16 intersections (hops) in under 200ms

The new system can rapidly adapt to a changing business

Zephyr Health Inc. is headquartered in San Francisco. Its cloud-based *Zephyr analytics platform* applies robust analytics to a diverse variety of data to find unique connections that can advance its customers' R&D, Clinical Trials, and Marketing. For instance, Zephyr's engine helps pharmaceutical companies find the right doctors for a clinical trial by linking private and public data–such as specialty, geography, and clinical trial history.

Brian Roy leads Zephyr's analytics platform engineering group, and evaluated database options for the initial prototype. Neo4j was chosen for its flexibility and ease of use. The result was a 12-week POC, covering modeling to integration, analytics application overlay, and performance testing to fine tune the graph model. The result was an engine that exceeded load and scale requirements.

Zephyr's analytics platform gives customers complex, customized queries against unified data, in the cloud, helping them stay on the leading edge of product development.

"Traditional BI looks at a segment of the business, giving a limited perspective on data. Zephyr's analytics platform first unifies the data to see connections that lead to more insightful questions. The result is 360-degree intelligence to drive better decisions." ***–Brian Roy, Director, Platform Engineering & Architect***

Challenge

The business challenge for Zephyr Health was to build a platform that unifies data in order to provide insight into the dynamics of a given market. The challenge was to get

a view not merely of market entities— for example research physicians conducting clinical trials— but to see the shared attributes around them. The query engine reveals connections that provide deeper business insights, allowing Zephyr's customers to better serve their markets.

Why Neo4j?

Brian Roy's design team knew they had to organize their data model according to an ontology specific domain which would expand and adapt over time. They wanted to develop a "cocktail" app that could query co-mingled, multi-sourced data—a combination of customer and public data sources.

But combining data posed to two complex challenges. The first was the disparity of data sources, and the joins required to bring them together. The second was the data diversity. New attributes were coming in regularly—from doctor surveys, for example. The old sytem required a large and non-performant key, character large object (CLOB) table—creating operational and performance issues. To allow external, non-technical customer users to create queries, Zephyr created its own domain query language—Zquery-- that sits over its two databases: Neo4j and MongoDB. With Neo4j, the implementation went to production in just under four months, with excellent results.

"Neo4j is the best open solution for our need, offering flexibility, scalability and a viable community committed to creating the building blocks that accelerate our business."

Benefits

Neo4j's property graph data model, combined with its Cypher Query query language, allowed Zephyr to develop its platform in record time. Says Roy, "You can dive right in. In a week or two of development we were up and running with data on millions of physicians and hospitals." The platform engine was faster than expected—traversals with 14-16 intersections (hops) completed in fewer than 200ms. "There is a revolution happening in analysis and analytics, creating demand for business users to be their own data scientist. Clinicians want to know how to take useful information out to the market. We can help them solve real problems without taking time to go deep into the IT stack." Zephyr is meeting and exceeding market demands, and its customer base and data volumes are growing exponentially.

About Neo Technology

Graphs are everywhere. From websites adding social capabilities to telecommunications companies providing personalized customer services to innovative bioinformatics research, organizations are adopting graph databases as the best way to model and query their connected data. Neo Technology researchers were the early pioneers of graph databases in 2000 and have been instrumental in bringing the power of graphs to numerous organizations worldwide, including a rapidly-growing number of Global 2000 customers, such as Cisco, HP, Accenture, Deutsche Telekom, and Telenor. Serving customers in production for over a decade, Neo4j is the world's leading graph database with the largest ecosystem of partners and tens of thousands of successful deployments.