



Providing Relevant News to Readers Using Graph Technology

In a world where content is key, one of the main goals for startup company [New.sc](#), was a simple concept with a complex execution. Their objective: to provide highly tailored, yet fully customizable news results to its users with as little input as possible. What's more, these results should not simply match topics, but provide relevant results, more likely to be of interest and importance to its readers. In order to achieve this, although having previously stored data in Redis, the most suitable data structure was a graph.

Forming relevant links with graphs

When initially developing New.sc, all data was stored in Redis*. However, their developers quickly realised that in order to determine content relevance with as little user input as possible, a Key-Value Store was not the most suitable technology. Their algorithms needed a quick and convenient method to search for links between elements, detecting topics and determining dependencies between objects and subjects as well as process natural language.

The complexity of such algorithms were highly taxing on their current database and so the decision to use a graph database was made. Following their initial research of multiple vendors, having been widely utilised and discussed online, their choices were narrowed down to two: OrientDB or Neo4j*.

Among the many considerations taken into account when making their final decision, first and foremost was speed. They knew that large amounts of data had to be processed quickly, and based on their research, the winner seemed to be OrientDB. Of course, other criteria had to be met. *"The main selection criteria were: language on which DBMS was developed, License, Graph model, API, query languages, Consistency, scalability."* explained Vadim Savchuk, CTO at New.sc. *"OrientDB met all our criteria."*

Industry

News/media

Challenge

Rapidly analyze growing news data to find relevant news content with little manual input.

Approach

Change DBMS from Key-Value Store Redis* to OrientDB's graph/document Multi-model database.

Solution

Utilised OrientDB's SQL dialect and graph structure to uncover links between elements.

Result

2.5 increase in database performance speeds able to easily form relationships between relevant content and provide up-to-date, appealing content to users.

About New.sc

New.sc is a startup providing highly tailored news content to users based on intuitive algorithms. With Natural language processing, news topic detection system and object-subject relationship discovery, New.sc finds recent, relevant and interesting news from around the globe.

<https://new.sc>

Developing startups with cutting edge technology

For startup companies not seeking outside investment, finding the most efficient yet cost effective technology is not always easy. Licensing models, such as GPL can be limiting as well as costly. When it comes to graph databases, several open source solutions available offer limited features, with tools necessary to run production environments only available in enterprise editions. Furthermore, utilising NoSQL Graph solutions usually require learning new query languages, such as Neo4j's Cypher, or delaying projects while finding those already well versed in these.

“Once we began using OrientDB and working with graphs, our algorithms were 2.5 times faster!”

- Andrey Kozak, CEO, New.sc

Multi-model NoSQL solutions such as OrientDB address these challenges. With an Apache 2 license, there are no restrictions when using their open source community edition. Being widely used by startups and deployed in production environments around the world, OrientDB Community edition Does not only have a commercial friendly license but includes sharding, scaling, replication and security capabilities which most competitors only offer in their Enterprise editions. Their Multi-model engine also enables graph, document, object-oriented capabilities and OrientDB's schema flexibility to be harnessed in order to reduce the need for multiple systems and to optimise performance. This was of particular interest to New.sc. With a graph database comprised of *Vertex* classes and inherited *Edge* connections, News.sc was also able to store keys within vertices and utilise transactions when creating vertex-edge links. Being familiar with OrientDB's extended SQL and using the built-in *Shortest Path* function, the most efficient path between two requested vertices was optimised.

Optimising user experience

By harnessing the power of graphs to quickly form relevant links between rapidly increasing data, New.sc was able to optimise their user experience. The result was a highly personalised and customisable news feed that allows readers to seamlessly pass from one relevant news article to the next with virtually no manual input. As News.sc CEO Andrey Kozak Explained, “Once we began using OrientDB and working with graphs, our algorithms were 2.5 times faster! Thanks to OrientDB, we were able to proceed to the next step in the development of our service.”

*All trademarks are the property of their respective owners.

About OrientDB

The native multi-model database combines the connectedness of graphs, the agility of documents and a familiar SQL dialect. Fortune 500 companies, government entities and startups all use the technology to build large-scale innovative applications. Some of their clients include Accenture, Comcast, Ericsson, the United Nations, Verisign, Pitney Bowes, Sky, Diaku, CenturyLink and Sonatype.

www.orientdb.com

Contact Info

enquiries@orientdb.com