REPORT REPRINT

DataStax delivers on graph, focuses on cloud applications

JAMES CURTIS

28 APRIL 2016

The company, which is known for providing enterprise support around the open source NoSQL Cassandra database, has added graph to its lineup, driving further into multi-model territory. Graph is starting to catch on, so how will this move impact DataStax's bottom line?

451 Research®

 |

When DataStax acquired Aurelius in January 2015, it was a clear signal that the company was taking graph seriously and moving aggressively to offer a multi-model NoSQL database. Fast-forward to today, and the company has delivered on that promise by not only adding graph to DataStax Enterprise (DSE) but also introducing a product called DataStax Studio, which is designed to help organizations that are less familiar with what graph provides leverage graph capabilities. The addition of graph also further positions DataStax in the direction of cloud applications, a refined focus area for the company going forward.

THE 451 TAKE

When DataStax announced the acquisition of Aurelius, our take was that the move made a lot of sense for the company, particularly because of the graph knowledge possessed by the target's team and its track record with the Titan open source database. The result is a solid graph offering, DSE Graph, that nicely complements the company's enterprise offering and plays well with Cassandra's inherent strengths of scalability and reliability. The recent focus on cloud applications is likely a natural progression for DataStax as it pursues enterprise clients, particularly as it finds itself frequently competing with the incumbent relational database vendors. With multi-model and graph on the rise, notably among NoSQL providers, we expect greater competition to ensue.

CONTEXT

DataStax is best known for providing commercial support for the NoSQL Cassandra database. The company was founded in 2010 and has been steadily adding greater functionality to its primary product, DataStax Enterprise, including search, analytics, Hadoop and Spark integration, indexing and security, among others. As such, the purchase of Aurelius has led to graph being integrated with DSE, which furthers DataStax's development efforts to broaden its offerings.

When we last covered the company, it had just released DSE 4.7 with the addition of Live Indexing, meant to speed up search and analysis functionality. At the time, DataStax counted 500 paying customers, and it now reports a bit more than 500, although it declined to provide exact numbers. The company does, however, note that it has altered its sales strategy – it used to target midmarket businesses, but now pursues larger enterprise accounts. DataStax claimed 450 employees in May 2015 and says that number has stayed steady. The company has amassed $190m in funding to date.

CLOUD APPLICATIONS

DataStax believes DSE caters particularly well to cloud applications. What are cloud applications? According to the company, they allow for the convergence of real-time capabilities with what it calls 'epic scale,' the notion of a distributed environment that is continuously available and built on commodity hardware. Some core attributes of cloud applications include the ability to handle significant data inputs, such as from the Internet of Things, manage multiple geographies for distributing data, provide low-latency response time, and scale based on unpredictable load.

What might a DSE-based cloud architecture look like? Similar to other cloud architectures, there is a client that interacts with an HTTP application, which then inserts messages into a message queue. The components and how they are assembled to interact with the message queue is where DataStax claims its differentiation. The streaming analytics component ingests data using Spark Streaming, which pulls the data out of the message queue. The data then gets persisted to the Cassandra data store followed by virtual datacenters. The data is also used to augment the machine-learning models to drive real-time recommendations. Another component for batch analytics facilitates the long-term recommendations. Lastly, there is the real-time element that handles customer-facing queries.

At the center of DataStax's cloud application pitch is that it includes Cassandra, with all of its inherent advantages – scaling and reliability, for instance. However, the company also notes that for enterprises to truly roll a successful cloud application, value-add components are required. This is where streaming analytics, batch analytics, search, indexing and the new graph capability come into play – in essence, these are the conduits to drive value from data.

PRODUCTS

DataStax Enterprise is currently at version 4.8. The recently announced DataStax Enterprise Graph offering will be integrated with DataStax's upcoming product release. This means that all of Cassandra and DSE's capabilities are available, including security, built-in analytics, search, automated workload management, drivers, and so forth. As graph functionality is starting to pick up, DataStax claims that DSE Graph is differentiated by its scaling and real-time capabilities. DSE Graph is based on the Apache TinkerPop graph computing framework, and is also compatible with the Gremlin graph programming language.

Accompanying the DSE Graph is a new tool dubbed DataStax Studio. The company introduced Studio realizing that there could be a skills gap among some organizations in working with graph databases. That being the case, Studio is a Web-based environment that enables developers to visually interact with DSE Graph.

Additionally, DataStax has released OpsCenter 6.0. OpsCenter is an administrative tool for managing and monitoring DSE and Cassandra. DataStax rewrote the tool's provisioning capability, thus enabling it to provision new nodes, easily handle configuration management and carry out rolling upgrades. The company also rewrote the UI and made it easier to manage large clusters. OpsCenter now supports DSE Graph for graph database configuration, cluster provisioning and monitoring.

COMPETITION

DataStax competes with the established relational database providers, specifically Oracle, IBM, Microsoft and SAP. The company also notes that it has displaced some of these players in a number of situations. While the established relational database vendors hold a commanding market share, it should be noted that they are adopting NoSQL characteristics and including them as part of their portfolios. For example, Oracle currently offers Oracle Spatial and Graph as an option for Oracle Database. A companion offering, Oracle Big Data Spatial and Graph, includes graph capabilities to run workloads on Hadoop and NoSQL. Meanwhile, IBM recently announced IBM Graph as a managed service on Bluemix. This is built on the TinkerPop framework, and comes with a visually based graph development tool.

DataStax can also be categorized as a multi-model database supplier. The whole idea of a multi-model database is to address the polyglot persistence challenge – negating the need to maintain multiple databases with difference models. Some players here, including those with graph capabilities, are OrientDB, ArangoDB and MarkLogic (with its Semantics functionality). Other multi-model providers include MongoDB, Basho, Couchbase and InterSystems.

Neo Technology is perhaps the best-known graph database purveyor and a strong proponent of graph technologies. The company has a partnership with IBM for its Neo4j graph database to run on Big Blue's POWER8 architecture. While DataStax has entered the graph arena, it's not as likely to vie directly with Neo because it is more focused on multi-model capabilities, whereas the latter caters to those looking specifically at graph databases.

SWOT ANALYSIS

STRENGTHS
DataStax is the most well-known supplier of Cassandra support and has built a solid portfolio of enterprise value-add products to take advantage of Cassandra's benefits.

WEAKNESSES
The company's success relies largely on the adoption of Cassandra. Customer numbers have stayed relatively constant since our previous coverage, but that could change with graph and other enhancements.

OPPORTUNITIES
Graph offers DataStax some added functionality and complements its existing core functionality of indexing, search and the ability to do batch and streaming analytics, giving it a number of entry points among prospective customers.

THREATS
While DataStax is categorized as a NoSQL database vendor, it likes to target larger enterprise clients and finds itself competing with the incumbent relational database providers. Other NoSQL firms are also entering the multi-model fray, and we expect graph to be on their radar.