

Healthx

Healthx wanted a certified Apache Cassandra™ database for managing real-time data, so it turned to DataStax – and become one of the first companies to use DataStax Enterprise 2.0.

Healthx develops and manages online portals and applications for the healthcare market that help support strategic activities such as enrollment, reporting, claims management, and business intelligence. Formed in 1998, the privately held company has grown rapidly over the past 14 years to become the healthcare industry's leading developer of self-service communication and data integration portals for more than 130 payers and 39,000 groups.

Scott McPheeters is chief security officer and manager of IT for Indianapolis-based Healthx. "We help payers deliver benefits faster, more effectively, and at a lower cost," he explains. "Healthx is a B2B firm that works primarily with mid-tier level healthcare companies that don't have resources to manage these types of services on their own. You don't really 'see' us when interacting with the services we provide because we're working behind the scenes."

All Healthx clients are supported on a cloud-based platform housed at a data center in Indianapolis. The company has a vast array of ever-changing data in its system, including benefits, physician, and prescription information. Querying that data – particularly, searching for specific information about physicians – was difficult for the company's Microsoft SQL servers to handle efficiently. So McPheeters and his team, which includes senior software engineer Mark Hodson, decided it was time to introduce an open source database solution into what had long been a "100 percent Microsoft shop."

McPheeters says Healthx looked at "every single big data platform" that hit the market over a two-year period, but it wasn't until the company evaluated Apache Cassandra in 2011 did everything "click and make sense." Although the primary motivation for finding a NoSQL solution for Healthx was to improve search, McPheeters says he instantly saw the value of using Cassandra from an operational standpoint.

"What I found to be the most intriguing and exciting thing about Cassandra is the idea that my database becomes a farm, just like my web servers are a farm," he says. "Our Microsoft SQL servers do just about everything, but when we have to upgrade or run maintenance on them, it's a big ordeal. But if a web server blows up in the middle of the night in the web farm, I don't worry about it. I just turn off my alarm and go back to sleep. That's what I want from my database."

Improving text search with Cassandra and Solr

Healthx engineers initially are using Cassandra for one specific but significant project: improving the text search process for the company's provider directory system. "The provider lookup allows people to search for a new doctor," McPheeters explains. "For example, say you want to find a podiatrist who is female, speaks German, and has an office close to where you live. You can search the provider directory and find doctors who meet those criteria. Trying to do this type of text search with a Microsoft SQL server was causing us a lot of pain. We weren't getting good indexes that would help us to retrieve the data quickly, and performance was becoming an issue."

Healthx wanted a certified Cassandra database for managing real-time data, so the company turned to DataStax in the summer of 2011 and became one of the first companies to use DataStax Enterprise 2.0, a complete big data platform built on Cassandra and architected to manage real-time, analytic, and enterprise search data in the same database cluster. DataStax Enterprise includes continuously available Hadoop, along with strong enterprise search support via Lucene and Apache Solr.



Healthx

Company
Healthx

Data Size
1 3-node cluster

Business Challenge
Healthx looked to improve the text search process of its provider directory system, and explore other ways to enhance the user experience through more efficient search capabilities.

Technical Challenge
The need for a highly reliable but low maintenance NoSQL distributed database solution that would allow Healthx to reduce the burden on its Microsoft SQL servers

Solution
DataStax Enterprise, a fully integrated big data platform for managing real-time, analytic, and enterprise search data all in the same database, which is powered by Apache Cassandra™ software and features continuously available Apache Hadoop™ and powerful search support with Apache Solr™.

"Our need to improve those text searches is why we looked at DataStax and DataStax Enterprise," says McPheeters. "We really like the integration with Solr. We get the full redundancy that you'd expect out of Cassandra as well as the full text indexing of Solr. The two things together make a win."

McPheeters credits his colleague Mike Rowe, chief architect for the application development group at Healthx, with ensuring a smooth transition to Cassandra. "He changed our applications data access layer so we could route requests to either Microsoft SQL server or DataStax Enterprise," he says. "This allows us to flip back and forth as needed – and we didn't have to make any other changes to the application."

Making plans—and learning new skills

Once the provider directory project with Cassandra is complete, McPheeters says his team will explore other ways to use the NoSQL solution. "We're already talking about moving some of our logging over to Cassandra to take better advantage of Hadoop reporting in DataStax Enterprise," he says. "This will help us spread the load on reporting – and get it off of our production database server."

McPheeters says one of the "dreams" he and his team would like to realize by using Cassandra is to develop a "one search box." He explains, "Like a Google-type search, you could type in free-form text. For example, a doctor could type, "Tell me every single patient who has the flu," and the system would quickly retrieve those results. Try to do that today and you'd have to pick about five different fields just to build up the search before you could perform the search."

According to McPheeters, transitioning to Cassandra has been easier than expected – and whenever the Healthx team has needed guidance, DataStax techs have been there to offer advice and help solve problems. "Cassandra was our very first foray into open source," he says. "We're a Microsoft shop, so all of our developers use C# language and have limited or no background in open source. Learning how everything works together – it's been challenging, but fun at the same time."

“We really like the integration with Solr. We get the full redundancy that you'd expect out of Cassandra as well as the full text indexing of Solr. The two things together make a win.”

Scott McPheeters
Chief Security Officer and
Manager of IT
Healthx

About DataStax

DataStax powers the apps that transform businesses. DataStax powers over 250 Big Data apps for startups and 20 of the Fortune 100 with its flexible and massively scalable big data platform built on Cassandra, through multi-data centers.

DataStax Enterprise delivers enterpriseready Cassandra, then goes one step further by integrating the best of breed Big Data technologies – Apache Hadoop for analytics, and Apache Solr for search across multiple datacenters and the cloud.

Top companies such as Adobe, HealthCare Anytime, eBay, and Netflix rely on DataStax to transform their businesses. Based in San Mateo, Calif., DataStax is backed by industry-leading investors: Lightspeed Venture Partners, Crosslink Capital and Meritech Capital Partners. For more information, visit <http://www.datastax.com/> and follow@DataStax.