DATASTAX

coursera

Use Case: Product Catalogs & Playlists

A product catalog is an organized collection of products, services, inventory or other assets that is constantly modified, searched and analyzed by both internal and external user communities. Playlists are user-defined queues of items like songs, movies, games, gift registries and lessons.

Industry

Online Education

Challenges

- Unexpected application downtime due to RDBMS single points of failure
- Limitations in introducing new features due to scale concerns
- Unstable performance with MySQL

Solution

- Masterless architecture delivers 100% application uptime
- Ease of scale to enable storage of growing user data
- Reliable read and write performance that exceeds expectations

Results

- On-demand courses with 24x7 availability to users
- Expanding to multiple data centers to support demand
- Reduced time-to-market on new product features

PRODUCT CATALOGS + PLAYLISTS

COURSERA DELIVERS ON-DEMAND EDUCATION ON ANY DEVICE WITH DATASTAX

Coursera is an education platform that partners with top universities and organizations worldwide, to offer courses online for anyone to take, for free. Coursera has a mission to provide users around the globe access to a world-class education wherever they are, whenever they want it. Coursera currently offers some 700 classes on their platform to more than 9.9 million students around the world.

The Challenge

Coursera's online education platform provides access to collegiate education to anyone, anywhere, for free, in partnership with some of the world's top universities and organizations.

Coursera had been using MySQL to drive class interactions in its scheduled remote learning platform, which suffered from regular blips in availability due to MySQL's master-slave architecture. Failing over from the master to slave, either for frequent maintenance or due to hardware issues, resulted in periods of unavailability.

Furthermore, there was a desire to build features that did not fit into the existing MySQL sharding scheme, which was designed at a very early stage in the company's development.

The Solution

One of the new features they wanted to offer was an on-demand model that would enable users to learn courses at their own pace, take in-video quizzes, and get feedback from peer assessment. With their prior MySQL experience, the Coursera team knew they needed to spend time choosing the right database to make this new offering successful.

"More than half of our customers are located internationally in countries such as China, India, United Kingdom, and Brazil," said Daniel Chia, software engineer at Coursera. "What this means to us is that there's no good time to take downtime, since our global nature means our platform is always being used. High availability is extremely important to us and our users, and that was the first thing that caught our eye with Apache Cassandra and DataStax."

Coursera evaluated other emerging database technologies including HBase and MongoDB. "We found it very challenging to manage MongoDB operationally," Chia explained, "and HBase had some latency issues along with difficulty in deployment."

Cassandra outperformed the others with its masterless architecture that ensures 100% application uptime, and flexible linear scalability, which will let them easily grow to any size.

"High availability with reliable performance is a big win for us. With DataStax Enterprise, our customers around the world are able to take any course anytime through our on-demand model," Chia said.

> *"High availability with reliable performance is a big win for us. With DataStax Enterprise, our customers around the world are able to take any course, anytime through our on-demand model."*
>
> *Daniel Chia*
> *Software Engineer*
> *Coursera*

The Results

Coursera currently has 3 nodes on AWS in the U.S. east region and plans to expand to multiple data centers across different regions in the future to allocate location-based traffic for optimal performance and to take advantage of multi-data center replication.

"Apache Cassandra and DataStax Enterprise has opened the door to greater possibilities for Coursera," Chia noted. "Migrating to DataStax Enterprise helps us innovate and add new features to Coursera's platform that was previously not possible due to our existing MySQL sharding model."

Since the launch of its on-demand class model, Coursera continues to shift more new courses to its on-demand application, with 9 courses currently running at this point in time and growing, allowing Coursera's internationally-based customers to access more courses at any time, anywhere with no restrictions or downtime.
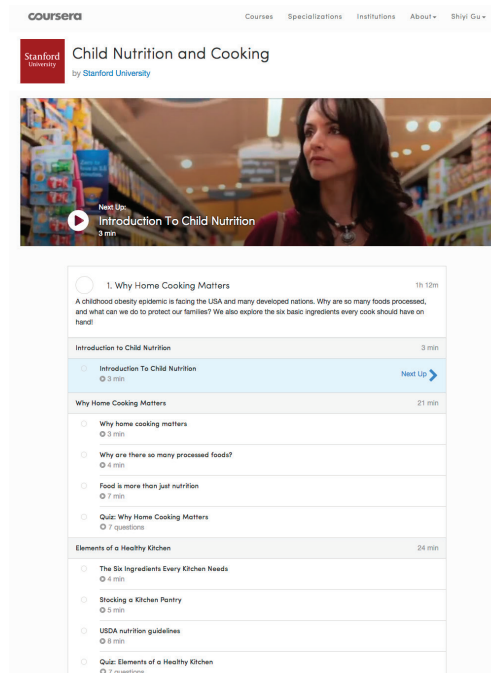
Coursera's on-demand class platform powered by DataStax

DataStax Headquarters

San Francisco Bay Area
3975 Freedom Circle
Santa Clara, CA 95054
650.389.6000 | www.datastax.com

About DataStax

DataStax is the fastest, most scalable distributed database technology, delivering Apache Cassandra to the world's most innovative enterprises. Datastax is built to be agile, always-on, and predictably scalable to any size.

With more than 500 customers in 45 countries, DataStax is the database technology and transactional backbone of choice for the worlds most innovative companies such as Netflix, Adobe, Intuit, and eBay. Based in Santa Clara, Calif., DataStax is backed by industry-leading investors including Lightspeed Venture Partners, Meritech Capital, and Crosslink Capital. For more information, visit DataStax.com or follow us @DataStax.
© 2014 Datastax, All Rights Reserved.