DATASTAX

Use Case: Personalization
Systems that understand each person's unique habits and preferences and bring to light products and items that a user may be unaware of and not looking for.

Industry
Streaming Media

Challenges
- Affordable capacity to store and process immense amounts of data (more than 2.1 billion reads and 4.3 billion writes per day)
- Single point of failure with Oracle's legacy relational architecture
- Achieving business agility for international expansion

Solution
- DataStax Enterprise delivers a persistent datastore, 100% uptime and cost effective scale across multiple data centers
- DataStax expert support

Results
- DataStax delivers a throughput of more than 10 million transactions per second
- Effortless creation/management of new data clusters across various regions
- Capture of every detail of customer viewing and log data

PERSONALIZATION

NETFLIX GIVES USERS EXACTLY WHAT THEY WANT – EVERY TIME

Netflix is the world's leading Internet television network with more than 48 million streaming members in more than 40 countries. Based in Los Gatos, California, Netflix has successfully shifted its business model from DVDs by mail to online media and today leads the streaming media industry with more than $1.5 billion digital revenue. Netflix dominates the all peak-time Internet usage and its shares continue to skyrocket with soaring numbers of subscribers.

The Challenge

Netflix revolutionized the customer streaming experience with its "Watch Now" service launched in 2007. From the moment a customer signs up to be a member, Netflix collects and stores every detail about the subscriber including the titles they play, what titles are played before and after, why they abandon after five minutes viewing, and where they pause. Netflix collects all these data points to tailor the overall experience to each customer.

In 2010, Netflix began moving its data to Amazon Web Services (AWS) to offer subscribers more flexibility across devices. At the time, Netflix was using Oracle as the back-end database and was approaching limits on traffic and capacity with the ballooning workloads managed in the cloud. The company knew that a single data center meant a single point of failure, and outages or poor-quality streaming could drive away customers. The problem they found with a central SQL database was that everything was in one place—which is only convenient until it fails. Because the database is expensive, users tend to put everything in there—compounding the impact of a potential failure.

Another problem was that schema changes required system downtime. Netflix found that every two weeks they'd have at least 10 minutes of downtime to put in the new schema. The limitations of a SQL database impacted their availability and scalability, not to mention the reliability and flexibility they needed to create and manage data clusters quickly as the company expended internationally.

The Solution

Netflix learned that traditional relational database technologies were not built to accommodate large volumes of data, neither from a cost-effective perspective nor regarding the needs for continuous availability and flexibility. The Netflix team performed an extensive evaluation of their database options and Apache Cassandra™ was the clear winner. Besides affordable scalability, Cassandra's schemaless architecture ensures no single point of failure and eliminates any downtime incurred by schema changes. Eventually, Netflix decided to migrate from Cassandra to DataStax Enterprise for its enterprise-level security and production. The entire migration of more than 80 clusters and 2500+ nodes was completed with only two engineers.

Netflix tailors content delivery based on viewing data captured and analyzed in Cassandra.

The Results

Today, DataStax Enterprise is the preferred database at Netflix. About 95% of Netflix's non-program data is stored in DataStax Enterprise, which includes customer account information, movie ratings, bookmarks and logs, and viewing history services. Thanks to DataStax Enterprise, Netflix is able to store and perform immense amounts of customer viewing and log data to improve subscribers' searching and streaming experiences. DataStax Enterprise drives throughput to more than 10 million transactions per second. On average, Netflix processes more than 2.1 billion reads and 4.3 billion writes per day.

DataStax Enterprise's multi-data center replication enables Netflix to get more flexibility than ever to create and manage data clusters. Netflix can create a Cassandra cluster in any region of the world in 10 minutes. When the marketing team decides to enter a certain part of the world, the technical underpinning is ready.

Netflix's database engine can deliver more than 76,000 different genre types such as "alien films from the 1970s"—and feeds subscribers' viewing habits back into the data-driven Netflix programmed commissioning machine. When the company decides to sink $100 million into House of Cards without even seeing a pilot, it isn't a blind bet. Netflix knew it would be successful based on the customer data stored in Cassandra.

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 world's 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.