



Use Case: Internet of Things

Refers to the revolution of a growing number of internet-connected devices that can network and communicate with each other.

Industry

Multi-Media Machine Learning

Challenges

- Large volumes of high velocity unstructured data (images and anonymized customer account info)
- Difficult and costly to scale with relational databases
- Need to deliver continuous availability for PhotoTime app and OEM customers

Solution

- Linear scale and 100% uptime with DataStax
- Real time analytics with integrated Apache Spark
- Operational simplicity with OpsCenter

Results

- Lightning fast performance with over 1000 calls per second
- 100% uptime even with peak traffic
- Easily handle 3-5 million photos per day, or 300-500GB of data per day

INTERNET OF THINGS

ORBEUS DELIVERS “PHOTOTIME”, A REVOLUTIONARY SMART PHOTO MANAGEMENT APP, BUILT ON DATASTAX

Orbeus makes finding and sharing your memories easier with PhotoTime, the new real-time, smart photo management app that automatically tags, sorts and organizes, all of your photos to make it fast and easy to find the one you want, when you want it. PhotoTime is the first and only app that uses image analysis to makes sense of faces, scenes, and objects in your photo and then labels it for you. (phototime.com)

The Challenge

Orbeus, a pioneer in image recognition technology recently launched PhotoTime, a smart photo management app, and had to hit the ground running by supporting tens of thousands of new users in the days after launch. PhotoTime uses Orbeus’ proprietary face, scene and objects recognition technology to identify people, places and things in customer photos, regardless of where they are stored. That means each customer potentially has tens of thousands of photos and each photo may have many faces, like a group shot.

With photos being the core source of data fueling these applications, Orbeus ran up against scale and performance issues and quickly realized that they needed a database backbone that was built to handle massive volumes of high velocity unstructured data. Orbeus handles on average 3 to 5 million photo and video posts per day - creating a data load of 300-500GBs of data per day that needs to be stored and analyzed without impact to performance.

As PhotoTime continues to grow in popularity, the need to ensure continuous availability has been a growing concern, particularly when dealing with this level of scale. “Our goal is to continually increase users, traffic and user engagement of PhotoTime,” explained Tianqiang Liu, Co-founder & VP of Engineering, Orbeus. “So we needed to ensure our infrastructure is built to support that workload.”

The Solution

Orbeus embarked on their journey for a database technology that would meet their requirements. They researched a number of solutions including MySQL, MongoDB, HBase and Apache Cassandra. Their requirements were scalability, performance and availability.

“We evaluated a number of solutions and chose Cassandra due to its ability to scale and remain available, even during heavy workloads,” said Tianqiang. “MySQL can’t easily scale, MongoDB doesn’t perform, and HBase doesn’t offer the same breadth of tools.”

Orbeus, a member of the DataStax Startup Program, started out with the community version of Cassandra, and realized the value DataStax Enterprise provided with integrated Spark analytics and OpsCenter, its visual management console that makes operating and maintaining DataStax Enterprise a breeze.

“OpsCenter was a big reason why we migrated to DataStax Enterprise,” said Tianqiang. “OpsCenter

“When speed, scalability, and ease of use are all critical, DataStax is better than any other database offering on the market.”

*- Tianqiang Liu
Co-founder & VP of Engineering
Orbeus*

allows us to easily visualize our database environment and provides advanced tools to help us scale and maintain performance and availability.”

The Results

Today, Orbeus relies on DataStax for a number of services that feed into their applications. One of their core services is the ability to help users find their photos through tags, stored in DataStax Enterprise, that are tied to meta data. DataStax Enterprise is also the primary data store for the applications.

“Our API needs a database to save all the photos for our engine to recognize faces and objects in a photo,” explained Tianqiang. “DataStax is the ideal solution because of its ability to store huge amounts of small files.”

Storing the photos and tags is only part of the equation though. Orbeus leverages DataStax Enterprise’s integration with Spark for fast, real-time analytics - allowing them to analyze visual information of the images and then upload the information to the database. And due to the sheer speed of its analytics, Orbeus’ users get the information they are looking for in real time. In fact, Orbeus cites that DataStax is able to act on over 1000 calls in less than 1 second, far faster than MySQL, MongoDB and others.

But the primary benefit that Orbeus is experiencing is performance and availability. “You can imagine how many photos you have on your smart phone, on social media and in the cloud” explained Richard Michaels, COO at Orbeus. “DataStax allows us to quickly scale out as fast and reliable as possible without having an impact on performance and availability. If users don’t have a positive experience, they may not come back, a major lost opportunity.”

So where does Orbeus go from here? Their business model is based on technology that leverages machine learning and learning techniques. In order for their models to continue to improve, you need a huge amount of data to analyze and a data store that can handle that volume without compromising on performance.

This type of technology can be used in many different ways from helping people organize their massive library of photos to allowing companies leverage their facial recognition technology to deliver a more secure and personalized experience. Their OEM customers have come up with creative ways to implement our technology, which they look forward to announcing in the future.

“We deal with millions of photos every day,” said Richard. “In order for us to continue to innovate and move the needle, speed and scalability of the database will be as important as the accuracy and performance of the application. DataStax ensures we can have both.”

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.

© 2014 Datastax, All Rights Reserved.