

Rubicon Project Deploys MapR for Mission Critical Real-time Advertising Platform

“The problem with Hadoop was instability. We had to solve that problem. That’s where MapR comes in. ”

Jan Gelin
*Vice President, Engineering and Chief System Architect
 Rubicon Project*

“MapR is at the core of all things Big Data at Rubicon Project.”

Byron Dover
*Big Data Engineer
 Rubicon Project*

The Business

Rubicon Project offers real-time trading technology that automates the selling and buying of online advertisements. Rubicon Project’s automated advertising platform has surpassed Google in U.S. audience reach and is used by more than 500 of the world’s premium publishers to transact with over 100,000 ad brands globally. The company’s customers include eBay, TIME, ABC News, the Wall Street Journal, Tribune Company, Virgin Media, People, Universal and many other Fortune 500 companies.

Rubicon Project performs 125 billion real-time auctions on their global transaction platform per day. That translates to about 4.0 petabytes of data that needs to be managed and analyzed.

The Challenge

Rubicon Project has relied on Hadoop for several years to store and analyze its data. However, the company’s business was growing rapidly and they needed to move to a fault-tolerant, mission-critical Hadoop production system. They needed high availability of services and jobs as well as data protection and disaster recovery capabilities.

MapR Solution

“We made a decision that Hadoop was going to be a critical piece of our day-to-day operations,” explains Jan Gelin, Rubicon Project’s VP of Engineering. “The problem with Hadoop was instability. We had to solve that problem. That’s where MapR comes in.”

Rubicon Project chose the MapR Distribution including Apache™ Hadoop® because of its enterprise features including high availability, data protection and recoverability, disaster recovery and advanced monitoring features. “Redundancy and the availability of support are critical to our business,” says Gelin.

continued on next page



Rubicon Project is Gaining Multiple Benefits from their MapR Solution

Rubicon Project's MapR solution is providing many benefits to their business including a centralized data source, accelerated application development, automated disaster recovery, increased efficiency and productivity and new opportunities for their business.

"We use our data to drive our business, to increase revenue and do a better job in the marketplace. Our company cannot run without Hadoop and MapR. It's critical to our business."

Jan Gelin
*Vice President, Engineering and Chief System Architect
 Rubicon Project*

Benefits

Automated disaster recovery

MapR provides automated failover of critical services from multiple failures along with automated recovery of the job tracker preventing service disruption. This allows Rubicon Project to run Hadoop along with the rest of their enterprise infrastructure in a lights out data center. With such automation, Rubicon Project manages hundreds of Hadoop nodes with just one dedicated administrator and responsive support from MapR. "I don't want to have to hire a staff of Hadoop experts. We have to be focused on our business," says Gelin.

A centralized data source

One of Rubicon Project's key goals was to create a centralized data source that the entire development organization could access. "We have about 150 engineers, all in different divisions with different initiatives, agendas and products. If we didn't have a centralized resource, our engineers would start to create silos with different technologies," explains Gelin. "We want to make it easy and cost effective to access data every time any developer wants to develop a new feature or product."

Industry standard enterprise tools

MapR also supports industry standard APIs such as NFS and ODBC allowing users to directly access Hadoop and integrate it with existing enterprise software. This allows Rubicon Project developers to access production data easily and develop applications at a much faster pace. Their developers and administrators have easy access to Hadoop through standard enterprise tools and save time in building applications and administering Hadoop.

Increased efficiency

Rubicon Project is very focused on increasing productivity and efficiency. They wanted to make it more efficient for its administrators and developers to access and use Hadoop. "Everything our company does is driven by metrics. We need to make sure transactions are as efficient and profitable as possible. We know how much revenue each server is making," he explains.

Competitive advantage

Rubicon Project sees how smart access to data provides competitive advantage. "Some companies are not clear about what data they have and how to use it. Some are only using traditional data warehouse technology," says Gelin. "Hadoop is established as the technology people want to use to get to the next level in Big Data. Rubicon Project is on the forefront of this because of the size of our Hadoop cluster and the centralized way we use our data."

New capabilities and opportunities

Rubicon Project is now able to leverage Hadoop directly for both Buyer and Seller Cloud transaction metrics. "We are continually innovating and powering new data analytics tools, including Algorithmic Revenue Optimization, on our MapR infrastructure," says Byron Dover, Big Data Engineer at Rubicon Project. "Our latest research and development efforts include Apache Spark trials, via MapR 4.0 on Amazon Elastic MapReduce, as we seek to close the gap on real-time global transaction reporting. MapR is at the core of all things Big Data at Rubicon Project."

MapR delivers on the promise of Hadoop with a proven, enterprise-grade platform that supports a broad set of mission-critical and real-time production uses. MapR brings unprecedented dependability, ease-of-use and world-record speed to Hadoop, NoSQL, database and streaming applications in one unified distribution for Hadoop. MapR is used by more than 500 customers across financial services, government, healthcare, manufacturing, media, retail and telecommunications as well as by leading Global 2000 and Web 2.0 companies. Amazon, Cisco, Google and HP are part of the broad MapR partner ecosystem. Investors include Google Capital, Lightspeed Venture Partners, Mayfield Fund, NEA, Qualcomm Ventures and Redpoint Ventures. MapR is based in San Jose, CA. Connect with MapR on Facebook, LinkedIn, and Twitter.

