

YAMMER POWERS USER NOTIFICATIONS WITH RIAK KV® ENTERPRISE SOFTWARE

A NEW FEATURE PROMPTS A NEW DATA STORE

Yammer is an enterprise social network, delivered as a web application via a software-as-a-service (SaaS) model that allows greater horizontal communication between individuals inside a company. Yammer has similar attributes to Facebook and instant messenger protocols, which means it has a lot of data to manage from its rapidly growing user base. And, because the nature of Yammer is social connections, messages, etc., the data is constantly being updated.

Recently, Yammer looked to build a new Notifications feature, which allows Yammer users to receive alerts when another user in their network follows them, they are invited into a new group, or if they received a reply to a message they published in the network. “This gives the user a sorted set of notifications, with a typical call to action based on the nature of the notification,” explains Coda Hale, Infrastructure Architect at Yammer. “It is a high value data stream, pushing information users care about into an easily actionable to-do list.”

DATA SIZE

2+ Terabytes

CHALLENGE

Duplicate data and stability concerns due to difficulty with replication and database crashes

SOLUTION

Employ a reliable, scalable NoSQL solution based on Riak KV® Enterprise Software, supported by Basho

EVALUATING WITH DYNAMO IN MIND

Most of Yammer’s data is stored in a Postgres data store, but as the Notifications project team was scoping out the project, they discovered that the data powering Notifications was amenable to an eventually consistent key-value data store like the one found in the Riak platform. “Postgres provides consistency of data guarantees at the expense of availability – but we knew we wanted a system that had greater availability by nature,” Hale notes.

So, with availability being an important factor, the project team began looking at projects based on the Amazon Dynamo model. As Hale notes, “The Amazon Dynamo model has probably the best availability story out of all the data stores out there, period.” So, the team looked at various Dynamo-inspired projects and eventually came to Riak KV® Enterprise Software and Voldemort.

Hale states that his team chose Riak KV for a number of reasons, the first being the fact that Riak KV had the most mature footprint. “The operational amenities were really nice,” Hale states. “The architecture and implementation story was well documented, and the tools to do things like open a new node were pretty simple and open.” In addition, Riak KV won out because of its more mature operational tools and packaging, bigger community, and strong commercial support as compared to the Voldemort project.

Of course, Hale notes that Yammer was especially impressed with the stability and very low latency of Riak KV in testing. “The latency profile of Riak was flat as a pancake,” Hale says. “This made issues like capacity planning and SLAs around the project a lot easier.”



“*The system simply does what it is designed to do, and its packaging makes it easy to set up a bulletproof system.*”

- Coda Hale, Infrastructure Architect, Yammer

DEALING WITH REAL WORLD PROBLEMS

A major benefit to using Riak KV, according to Hale, is that the product understands both the needs associated with building distributed systems, as well as the nature of operating distributed systems. “Some traditional database supporters will tell you that non-trivial partitions don’t happen, but they do and having a strong non-relational tool that can provide availability is great,” Hale notes.

But while Hale and his team liked the fact that Riak KV was stable, and was ideal for managing the distributed nature of Yammer’s data model, he notes that not being familiar with Erlang, the programming language of Riak KV, was a bit off-putting. “I’d heard that Erlang could be difficult to deal with,” Hale says, “but I had our production cluster set up in about fifteen minutes. It has been a complete and total non-issue for us.”

THE BENEFIT OF “BORING”

Perhaps the biggest benefit according to Hale is, “The system simply does what it is designed to do, and its packaging makes it easy to set up a bulletproof system.” According to Hale, not having interoperability issues between the components of the systems was a highly refreshing experience. “Rather than the typical two days of staring at strange code, we were up and running quickly and have yet to have an alert or any problem.”

As Hale puts it, “It’s a piece of plumbing; it has never been a root cause of any of our problems.” So, with a bulletproof Riak-backed system in place, Yammer now has a robust Notifications module in its social collaboration tool that did not increase its data footprint on its single point of failure. The Riak-based system has low latency, and the data powering the notifications highly available. And given the amount of data moving around one of the most popular social networking tools on the market, not having to worry about a major component of its infrastructure is a huge benefit.

ABOUT YAMMER

Yammer’s founders David Sacks and Adam Pisoni saw an opportunity to apply the social media revolution pioneered by Facebook and Twitter to the workplace. The company launched to the public in September 2008 at the TechCrunch50 Conference and won the grand prize, despite strong competition from other great startups. Just two years later, Yammer is used by over 100,000 companies and organizations, including over 80 percent of the Fortune 500.

Yammer is headquartered in San Francisco and is well-funded by top tier investors including Charles River Ventures, Founders Fund, Emergence Capital, US Venture Partners, SV Angel’s Ron Conway, and Goldcrest Investments.

ABOUT BASHO

Basho, the creator of the world’s most resilient databases, is dedicated to developing disruptive technology that simplifies enterprises’ most critical distributed systems data management challenges. Basho has attracted one of the most talented groups of engineers and technical experts ever assembled devoted exclusively to solving some of the most complex distributed systems challenges presented by Big Data and IoT.

Basho’s database, Riak® KV, the industry leading distributed NoSQL database, is used by fast growing Web businesses and by one-third of the Fortune 500 to power their critical Web, mobile and social applications. Built on the same foundation, Basho introduced Riak TS, which is the first enterprise-ready NoSQL database specifically optimized to store, query and analyze time series data. Basho also provides Riak integrations for a variety of Big Data technologies like Apache Spark, Redis, Mesos, and Apache Solr.

For more information visit Basho.com which is full of interesting use cases, customer case studies and product detail, or docs.basho.com for technical documentation.

