Ensuring availability and performance of World’s largest repository of Java open source software components

Sonatype is a private company founded in 2008 by Jason van Zyl, the creator of Apache Maven, to deliver innovative solutions that increase developer productivity, improve project management, and reduce business risk. Sonatype is based on two transformative open source innovations: the Apache Maven build system; and the Central repository. With a global user community of more than 4 million users, Maven has changed the way Java developers build software.

In the open source world, the Central Repository has become the primary exchange for open-source Java components or artifacts. The Central Repository houses more than 300,000 open source software (OSS) components and receives more than 4 Billion requests per year, making it one of the most widely accessed services on the web. It is used by 42,000 development organizations per month, including more than half of the Global 2,000 and virtually all of the world’s largest financial institutions and software vendors.

“We were stunned by the almost immediate improvements. It is not often that you can install a tool and see benefit instantly. With New Relic, that’s exactly what we found — and that’s a huge value to Sonatype and to the open source community.”

—Charles Gold, CMO

An integral part of maximizing the usability of the Central Repository is providing both publishing and search mechanisms. Thousands of open source projects use Sonatype’s infrastructure (Sonatype OSS Repository Hosting, or OSSRH) for publishing their software to Central, where Java developers from around the world can then download it. They also provide a search mechanism for searching Central at search.maven.org.

The Central Repository, the Sonatype OSSRH, and search.maven.org together form a Java open source ecosystem for open source projects and Java developers around the globe.
Environment

This ecosystem runs on a combination of Java, MySQL, Apache Lucene, and Apache Solr. Hosting services are provided by Contegix and DNS global load balancing is provided by Dyn Inc.

Challenges

As the use of open source continues to grow, some 42,000 software development organizations access the Central repository each month to download Java artifacts for use in their applications. Central has become so integral to the software development and build process, that any lapse in performance or availability could mean significant delays for thousands of development organizations around the world.

Sonatype faced the difficult challenge of monitoring and managing every aspect of the ecosystem. The team was using a combination of tools including Hyperic, but they were still not able to see what was going on inside the parts of the ecosystem. They could scrape through logs after something had gone wrong but the time required to do that was significant, and more importantly, they had no way to discover problems in real time. This limited their ability to be proactive.

To continue facilitating successful use of OSS components, they needed a tool that would provide deep visibility into the Central Repository, the publishing mechanism, and the search mechanism. The tool would allow them to successfully meet the scalability challenges of maintaining a system that supports:

- 300,000 OSS components in the Central Repository — and growing
- 4 Billion access requests to the Repository each year
- 40,000 + companies access the Repository each month
- 320 new components a day are published to the repository

Solution

Sonatype tried New Relic and was impressed with its deep integration into Java, Lucence, Solr, and MySQL. With no comparable solutions available, Sonatype chose New Relic to monitor all aspects of the Java OSS component ecosystem — the Central Repository, Sonatype OSS RH, and search.maven.org. As Brian Fox, Sonatype VP of Product Management, says, “The deep integration that New Relic provides for Java based applications, in addition to the deep analytics of MySQL, Apache Lucene and Apache Solr, were the critical deciding factors. This support lets us really see what’s going on inside our applications and critical third party components, and was unique across all the tools we reviewed.”
In November of 2011, when New Relic added Server Level Monitoring, Sonatype was thrilled with the additional insight into their overall end-to-end environment.

Results

Using New Relic, Sonatype has made several important advances in monitoring their Java OSS ecosystem — in preventing problems and outages and also in diagnosing problems when they do occur.

As Brian Fox says, “We have visibility into what’s going on inside and across the entire ecosystem and that’s allowed us to make a range of performance improvements. We can monitor our memory profiles so we know when and where we need to add memory to avoid excessive swapping. Our Central Repository is one of the largest repositories on the planet and it just keeps getting bigger. Using New Relic, we have improved control, insight, and management ability to help ensure that the more than 40,000 companies that access the Repository every month can do so quickly and easily with no problems.”

In Sonatype’s old world, they were left scraping through log files after a problem arose. Once New Relic was installed, they find they can prevent many problems before they cause any type of an outage. This has been a significant productivity boost for the company.

With more and more companies choosing to look for open source components rather than building their own, performance is even more important. As the Central Repository continues to grow, using New Relic has quickly become essential to the team. Delivering high performance is an ongoing challenge — but an easier one now that they are using New Relic to monitor and manage the entire ecosystem.

Charles Gold, Sonatype CMO, sums up their enthusiasm and delight with the New Relic APM saying, “We were stunned by the almost immediate improvements. It is not often that you can install a tool and see benefit instantly. With New Relic, that’s exactly what we found — and that’s a huge value to Sonatype and to the open source community.”

About New Relic

New Relic, Inc. is the all-in-one web application performance management provider for the cloud and the data center. Its SaaS solution, which combines real user monitoring, application monitoring, and availability monitoring in a single solution built from the ground up, changes the way developers and operations teams manage web application performance in real-time. More than 14,000 organizations use New Relic to optimize over 6 billion transactions in production each day. newrelic.com