Pentaho Improves Software Delivery Speed, Consistency and Reliability

CHALLENGE
Shorten release cycles by minimizing manual steps in complex builds across multiple product releases in the enterprise

SOLUTION
Use the CloudBees Jenkins Platform to automate and accelerate builds, improve reliability and adopt CI practices

RESULTS
» Nine person-day build process shortened to four hours

» New build environments set up in two hours, not two weeks

» Security, reliability and consistency improved

PRODUCTS
» CloudBees® Jenkins Platform™

» Role Based Access Control Plugin™

» Folders Plugin™

» Folders Plus Plugin™

Pentaho, a Hitachi Group Company and a leader in data integration and business analytics, provides its customers with an enterprise-class platform to extract value from data. This open source-based platform simplifies the process of capturing, blending and analyzing a broader array of data – including big data and IoT – from multiple data sources to derive new business insights.

For the past two years, Pentaho has been improving the speed, consistency and reliability of software delivery processes for its core platform by implementing continuous integration (CI) using the CloudBees Jenkins Platform. "At Pentaho, our goal is to help customers gain more value from their data, and the CloudBees Jenkins Platform is central to helping us do that in a fast and reliable way," says Larry Grill, engineering operations manager at Pentaho. "Now, with CloudBees, my team is producing three different code lines of builds continuously on a nightly basis with a high level of success," adds Lee Cheng, director of engineering operations at Pentaho.

CHALLENGE: MINIMIZE MANUAL STEPS IN COMPLEX BUILDS FOR MULTIPLE RELEASE WHILE MAINTAINING SECURITY
In its early days, Pentaho's build processes were based on the original processes used to build the various grassroots open source projects that formed the pillars of its platform. The Pentaho engineering team initially tried to automate some of these processes by implementing CI with Bamboo and CruiseControl, but ultimately found the tools had difficulties handling their complex builds. "At the time we had about 80 different source code repositories that we were pulling together into one product," recalls Cheng. "We wanted to minimize manual steps in the process. Managing the dependencies and orchestrating a build flow from start to finish for 80 projects was difficult with those other tools."

"The CloudBees Jenkins Platform enabled us to execute faster, remove manual headaches from our build process and establish an operational environment for builds that is more controllable, repeatable and reliable."

Lee Cheng
Director of Engineering Operations, Pentaho

As Pentaho sought a solution to address the challenges of automating complex legacy process, security was identified as a priority. "Security requirements from corporate IT stated that all of our systems needed to comply with the Pentaho enterprise security model and work with our corporate Microsoft Active Directory authentication environment," says Cheng. In addition, the unique aspects of simultaneously developing both open source and proprietary code resulted in additional security requirements.

Pentaho also needed a way to effectively manage the multiple releases simultaneously, including the future release of its platform and at least two maintenance releases. "We were constantly duplicating our build environment for these different code lines, and that required numerous tedious tasks that we performed by hand," Cheng recalls. "We needed a better way to manage and automate our various code line builds," Grill adds.

SOLUTION: IMPROVE BUILD AUTOMATION AND RELIABILITY BY IMPLEMENTING CI PRACTICES

Today, Pentaho is managing and automating its build processes with CI and the CloudBees Jenkins Platform.

After first getting started with open source Jenkins, Pentaho moved to the CloudBees Jenkins Platform to address challenges faced in automating complex builds across multiple projects and release streams.

Pentaho engineering used the Role-based Access Control (RBAC) plugin to securely link Jenkins to user data in the corporate Microsoft Active Directory service. "With the RBAC plugin, we don't have to do any user management within Jenkins. Everything is controlled by group policy access and we assign security based on the Active Directory groups that our users belong to. Our developers get one set of access permissions, our QA group gets a different set and our contract workers get a third set," says Cheng. "RBAC makes it easy to control access; we don't have to manually add or remove users from jobs."

The group also set up the Folders plugin and Folders Plus plugin to better manage builds for multiple code lines. "Folders Plus provided a way to segment and containerize our different code line builds while improving automation," says Grill. "It also enabled us to quickly spin off exact copies of build environments without having to tweak settings manually. It's a big time-savings to just copy 120 jobs from one place to another instead of doing it all by hand on the file system."

To manage the complex build process required for large systems, Pentaho is currently using the open source Build Flow plugin, but is preparing to migrate to the Pipeline plugin to better orchestrate long-running jobs. "We got Build Flow to work, but it has some idiosyncrasies, so we are looking forward to moving to the Pipeline plugin," says Cheng.

The CloudBees Jenkins Platform is enabling the group to increase the reliability and consistency of build processes that have steadily grown more sophisticated as new projects and features are added to the Pentaho platform. "We now have development teams in four locations and 120 different projects. Our builds produce more than 3,500 Java components and our distributables are about 6 GB compressed," says Grill. "Every single build produces dozens of artifacts, and we do multiple builds nightly for the various code lines. The releasable builds are deployed automatically to test servers via Puppet, so when QA arrives in the morning they have last night's builds ready to go." Final releases are handed off to the Pentaho support organization, which make them available to customers on a scheduled release cycle.

Cheng notes that the entire process runs continuously with almost no downtime. "It just runs nonstop, and the stability of CloudBees Jenkins Platform is a big part of that. I can't remember the last time we had a Jenkins crash. I don't think it has ever happened. It's a solid product and its reliability is never even a point of discussion."

Pentaho currently runs two masters on CloudBees Jenkins Platform, with plans to add more. "As we continue to expand and look for new ways to help our customers gain more value from big data and the Internet of Things, we are considering additional opportunities to partner with CloudBees to help with the centralized management of our environments," says Cheng.

Pentaho's success to date with CI and CloudBees has the development organization well-positioned to adopt continuous delivery practices and achieve deeper integration with other development teams. "We're striving towards the best practices of doing continuous delivery, and one of our first steps was moving away from highly customized, unique builds to more standardized builds with Jenkins," says Grill. Cheng adds, "At the same time, there are other teams working on IoT initiatives and building products that include the Pentaho platform. With CloudBees we are in a good position to scale our offering to these other groups."

RESULTS

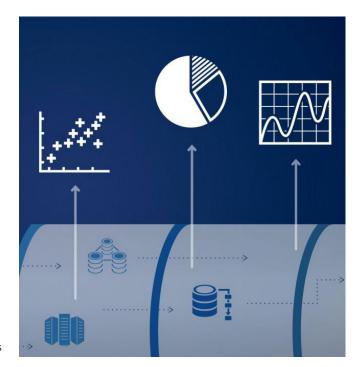### Nine person-day build process shortened to four hours.

"Five years ago it took three of us three days to get through one release build," says Cheng. "We've been able to operationalize much of that process with Jenkins, so it's now automated and done in four hours, even though we have a lot more happening in our builds than we did five years ago."

New build environments set up in two hours, not two weeks – or more.

"Before we started using the CloudBees Folders Plus plugin, setting up a new build environment would take at a minimum two weeks by the time we ironed out all the issues," says Grill. "Now that we've eliminated manual steps, we can do it within two hours."

Security, reliability and consistency improved.

"CloudBees enabled us to improve greatly in our ability to produce builds in a manner that is exactly the same every time without fail, without any question," says Cheng. "We now have more control and more confidence in the reliability of our process because we know that every single build has followed the same process, built the same components and put the release together in the same way."

Learn More About Pentaho

www.pentaho.com