Goldman Sachs doubles code coverage for a legacy application overnight

Saves over a year of manual effort with Diffblue Cover and AI for Code

Challenge

Unit test suites are a key part of effective software development, but creating them for legacy code is resource-intensive. Goldman Sachs aimed to efficiently boost legacy code coverage and allow engineering teams to refocus their efforts on the development of innovative, business-critical new features.

Solution

Goldman Sachs has a long history as a technology leader among global banks. The latest advance is the company's use of Diffblue Cover, an AI-powered tool that enables the engineering teams to improve code quality and efficiency through the automatic generation of unit tests.

Results

Since starting the legacy modernization program using Diffblue Cover, the engineering teams have increased test coverage for the first batch of applications from 36% to 72% in less than 10% of the time it would take to do manually.[1]

For some applications, this result was achievable in hours or days rather than man-years.[1] Higher code coverage has also provided greater confidence in application stability when adding new code, improving the speed at which the engineering teams can deliver business value.

Key Benefits

Code Coverage
With Diffblue Cover, new code coverage levels exceed industry standards.

Agility
Increase in agility and efficiency of new development.

Confidence
Engineering teams can modify code with confidence.

Savings
Create test suites in less than 10% of the manual time,[1] saving years of work and sharply reducing cost.

Modernization
Legacy applications can now migrate to cloud or be included in other critical transformation programs.

Coverage was boosted by 100% in less than 24 hours for one repo

Achieved a time savings of 1 man-year[1] for a core application

[1]Manual effort assumes industry averages of 30 minutes per manual test and 6 hours productive time per day.

diffblue
AI for Code

The Legacy Code Challenge

Unit tests—tests that confirm the functionality of individual units of code—are a small but critical part of the software development lifecycle. These fast, lightweight tests make it possible to track the connections between units of code so developers can write and refactor with confidence. When an organization has an automated suite of unit tests that covers all existing code, any new code a developer adds is immediately checked against the entire codebase, and the developer is alerted if their modifications cause any issues or breaking changes in the code's behavior.

Achieving high code coverage—the percent of a codebase covered by unit tests—has been a longstanding challenge for Goldman Sachs and other banks with significant legacy software, most of which incorporate code that was written before unit testing became an established practice. Editing or adding to poorly document-ed legacy code without unit tests can result in unexpected bugs and headaches, but writing the required quantity of new unit tests is time- and labor-intensive. As a result, meeting unit testing goals and new feature development goals simultaneously can be an uphill battle.

In the 2018 book *Accelerate: Building and Scaling High Performing Technology Organizations*, au-thors Nicole Forsgren, Jez Humble and Gene Kim found a relationship with high software delivery performance and the use of automated test-ing, and Goldman Sachs has been automating the execution of their unit tests for years. Until recently, however, the technology did not exist for automating the writing of unit tests them-selves, leaving organizations with only the option of using the manual efforts of internal or external development teams.

Goldman Sachs' QAE Team

Goldman Sachs has a long history as a technology lead-er among global banks, and the Goldman Sachs Quality Assurance Engineering (QAE) team is responsible for empowering the company's engineers to proactively deliver quality software and services. The processes put in place by the QAE team enable the early identi-fication of quality gaps with low-touch controls across Goldman Sachs technology.

The QAE team has been working towards reaching industry best coverage levels. However, given the company's large legacy estate and the volume of unit tests required to increase the average level of code coverage, the team had also been looking for ways to efficiently bolster productivity. Artificial intelligence (AI) was a natural avenue to explore.

Proposed Solution: Diffblue Cover

While working towards the goal of bringing every application in Goldman Sachs to higher levels of code coverage, the QAE team landed on Diffblue Cover, a tool that automatically and intelligently writes unit tests for Java applications using AI for code. One of Diffblue Cover's primary benefits is its unique ability to rapidly generate a test suite for legacy codebases.

"We decided to use Diffblue Cover because of the potential it offered for helping us meet our most ambitious code coverage targets, while also freeing up developers' time for the work only they can do," says Matt Davey, Managing Director, Technology QAE & SDLC. "Diffblue Cover is enabling us to im-prove quality and build new software faster."

"Diffblue Cover is enabling us to improve quality and build new software, faster."

—Matt Davey, Managing Director, Technology QAE & SDLC at Goldman Sachs

---

[1]Manual effort assumes industry averages of 30 minutes per manual test and 6 hours productive time per day.

Time Spent Writing Unit Tests

	Manual Effort[1]	Diffblue Cover
Number of tests	3,211	3,211
Average time to write each	30 minutes	10 seconds
Days spent writing tests per application	268 workdays	1/3 day (run overnight)

Results: Doubled Code Coverage in a Fraction of the Time

Diffblue Cover has been implemented on various applications within Goldman Sachs; for each software product, a suite of high-quality tests has been generated in less than one day. For one module within an important backend system, existing unit test coverage was boosted from 36% to 72% in less than 24 hours. Creating the same number of unit tests manually would have taken more than eight days of developer time[1], compared to three-quarters of a workday with Diffblue—a time savings of more than 90%. Diffblue Cover also picked up on edge cases in other applications that could have led to customer-impacting incidents.

Another back-end application has fifteen thousand lines of code. Diffblue Cover created over three thousand tests overnight. Compared to the time it would have taken to write these 3,211 unit tests manually, Diffblue Cover was more than 180 times faster.[1]

Diffblue Cover not only increased the quantity of tests, but also passed the quality bar for application owners. The tests were immediately ready to be integrated into the test suite, and the review of these generated tests took one day.

"We are thrilled with these results," adds Jonathan Goodfellow, Managing Director, QAE. "They have definitely exceeded our expectations and we're excited about how much time and work this has saved our engineers so they can refocus on increasing Goldman Sachs' feature velocity, code quality, and software security. It's great to have higher confidence in the integrity of our existing codebase."

Next Steps for Code Quality

To further streamline the development of quality code at Goldman Sachs in the future, the QAE team will be introducing Diffblue Cover across the company to help improve code coverage. With the confidence and reduced operational risk conferred by high coverage, the company expects to continue to see the transformation of legacy code into accessible and highly functional modern software.

"We expect this to be a key technology for our transformation and a game-changer for Goldman Sachs," Matt Davey concluded.

"We're excited about how much time and work this has saved our engineers so they can refocus on increasing Goldman Sachs' feature velocity, code quality, and software security. It's great to have higher confidence in the integrity of our existing codebase."

—Jonathan Goodfellow, Managing Director, QAE

About Diffblue
Founded by leading computer scientists from the University of Oxford, Diffblue is changing the way code is developed. The company's flagship developer tool, Diffblue Cover, intelligently writes unit tests that help software development teams and organizations efficiently improve their code coverage and quality.

Diffblue: AI for Code.
Learn more at **Diffblue.com** or contact us at **info@diffblue.com**

[1]Manual effort assumes industry averages of 30 minutes per manual test and 6 hours productive time per day.

diffblue
AI for Code