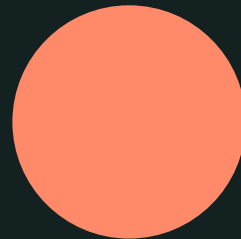




Sauce Labs Continuous Testing Benchmark Report



Introduction

In today's digital-first world, business success hinges on delivering flawless digital experiences. Yet, broken software is all too common.

We've all experienced banking app malfunctions preventing us from accessing our accounts or e-commerce site crashes stopping shopping sprees. Even minor glitches, like a mobile app freezing or crashing, often result in over half of users uninstalling the app.

These issues highlight the potential for significant financial losses, with just an hour of software downtime costing millions in revenue. Really, the foundation of any successful business is robust software quality.

Achieving this level of quality requires continuous testing, a core principle of DevOps. Unlike traditional testing that occurs at the end of the development cycle, continuous testing provides frequent feedback to all teams involved in the delivery process. To execute this, organizations need a centralized testing platform that seamlessly integrates into current tools, workflows, and development cycles.

“Even minor glitches, like a mobile app freezing or crashing, often result in over half of users uninstalling the app.”

We wanted to understand how organizations are approaching continuous testing and how it impacts their software delivery and business goals. Leveraging our expertise as a leading provider of test infrastructure, we analyzed anonymized data from 100,000 customers. With over 15 years of experience and more than seven billion tests conducted, we have deep industry knowledge in software testing.

This report serves as a benchmark for assessing and improving test performance. It's designed to aid in strategizing and decision-making, ultimately leading to improvements in efficiency, velocity, risk management, and superior digital experiences.

Continuous Testing Benchmark

The Sauce Labs Continuous Testing Benchmark (CTB) identifies crucial best practices for successful continuous testing and assesses how organizations stack up against these standards using actual test data. This benchmark encompasses both desktop and mobile testing, deriving insights from over seven billion tests on the Sauce Labs test platform.

For this report, we analyzed anonymized Sauce Labs test data over a 30-day period. While test volumes vary significantly across organizations, each one's influence on the overall metric performance is weighted equally, ensuring fairness in both desktop and mobile test evaluations.

Benchmark Components



Test Quality

Benchmark for Excellence: Pass at least 90% of tests run

Desktop

24.37%



Mobile

25.45%



Test Run Time

Benchmark for Excellence: Test run times averaging 2 minutes or less

Desktop

45.68%



Mobile

43.60%



Test Platform Coverage

Benchmark for Excellence: Test against at least 5 platforms on average

Desktop

74.69%



Mobile

63.48%



Only 7.79% of organizations meet the benchmark for excellence across categories, demonstrating the challenging yet critical nature of achieving high standards in continuous testing.

Benchmark Metric #1: Test Quality

To ensure your QA practice sends out reliable signals, you must start with high-quality data. And to achieve that, you need high-quality tests. Test quality is fundamental in continuous testing — it serves as the primary indicator of its effectiveness and the foundation upon which other testing practices are built.

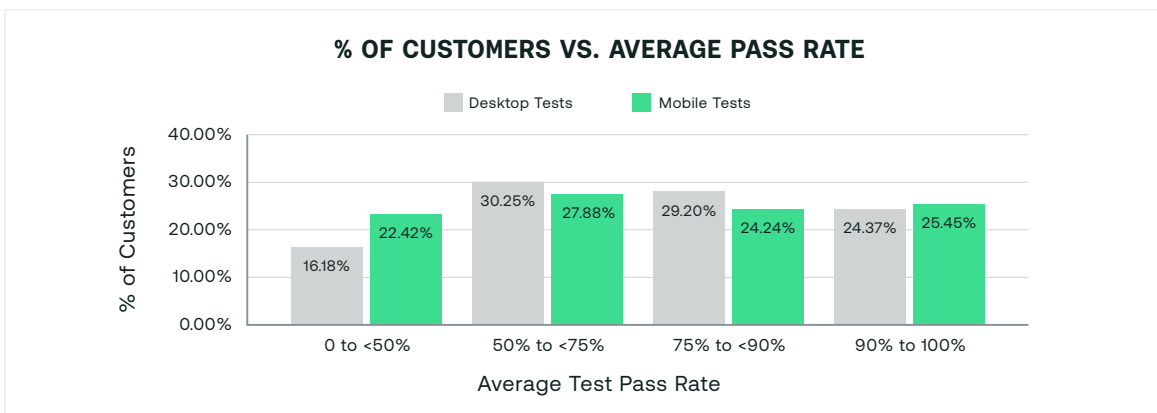
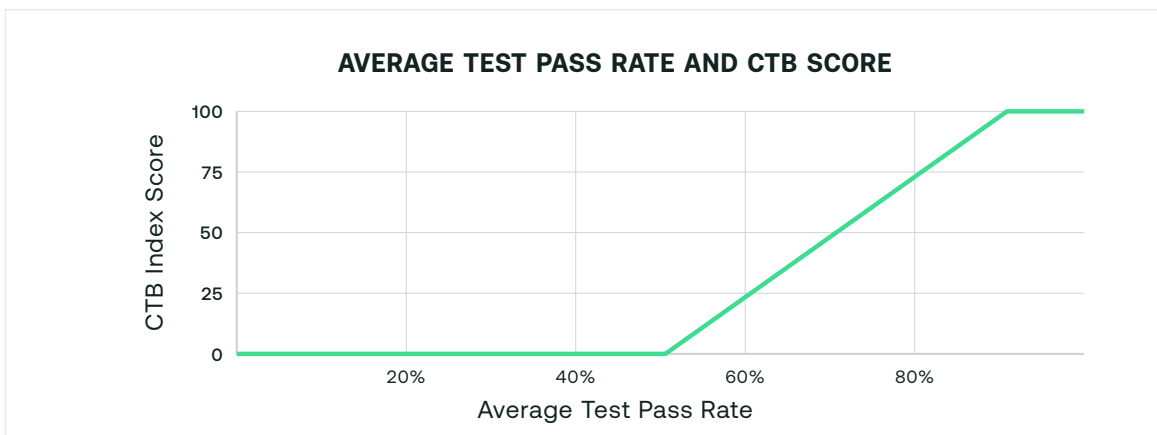
Why it matters: Automated testing aims to speed up development and reduce release cycles, requiring a high test pass rate. When tests fail, it's vital for developers to swiftly identify and fix the problem. High failure rates, often caused by test “flakiness” (tests that produce both passing and failing results despite no changes to the code or test), can disrupt this process. This necessitates manual intervention, slowing down the delivery of quality software. A low failure rate ensures that each failed test indicates a real issue rather than just inconsistencies in the test suite.

The Benchmark for Excellence: The CTB evaluates organizations based on the typical pass rate of their tests, highlighting the importance of not just the pass rate but also an organization's capacity to handle manual follow-ups on failures. **To achieve a test quality score of 100, symbolizing excellence, the average pass rate should be at least 90% across both desktop and mobile tests.**

Conversely, pass rates of 50% or lower receive a score of 0, indicating that failures are as frequent as successes. Organizations need to ensure their test suites can keep up with rapid application changes to avoid compromising the pace of delivery and the quality of the final product.

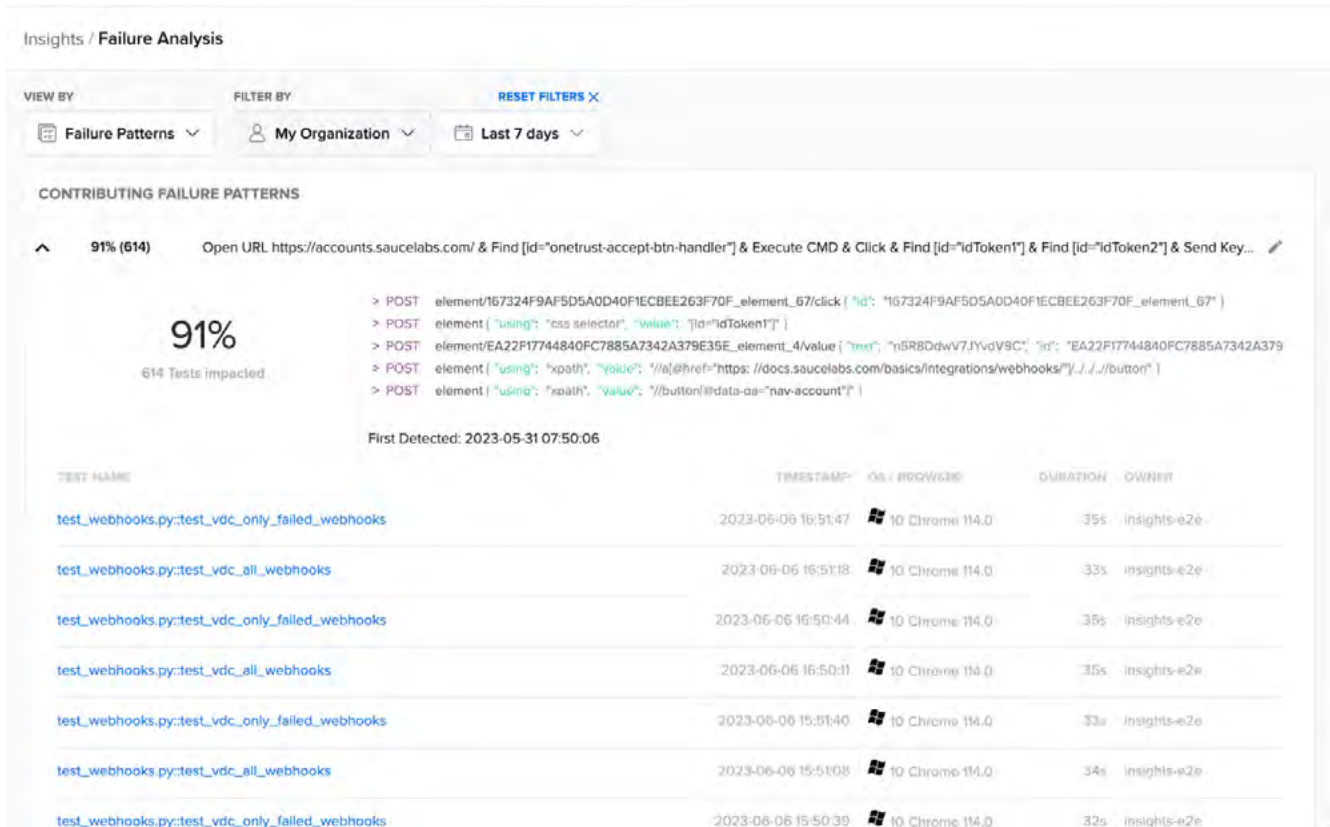
How Organizations Performed:

- **Desktop: 24.37% of organizations pass the test quality benchmark**
- **Mobile: 25.45% of organizations pass the test quality benchmark**



The Takeaway: There's an urgent need for improvement in test quality across the board — especially as organizations increase the speed of development and scale their test automation. Failure analysis stands as an innovative approach to combat poor test quality, leveraging ML to analyze pass/fail data and enabling developers to quickly address issues and improve test quality.

In a recent study, Sauce Labs analyzed 757,000 failed tests to assess the impact of failure analysis on their test environments. The analysis revealed that the most frequent issue accounted for 27% of all failures, with the top three issues responsible for almost half of all failures. These results show that by identifying and correcting a few key issues, organizations can achieve marked improvements in test quality — potentially seeing a 27% enhancement from a single fix.



Benchmark Metric #2: Test Run Time

If test quality is the area most in need of improvement, test run time is right behind it. This is important given that one of the main reasons organizations invest in automated testing is to accelerate velocity of innovation.

The longer tests take to run, the more time developers spend waiting around for feedback. And the longer it takes to push code through the delivery cycle and deliver new software to users.

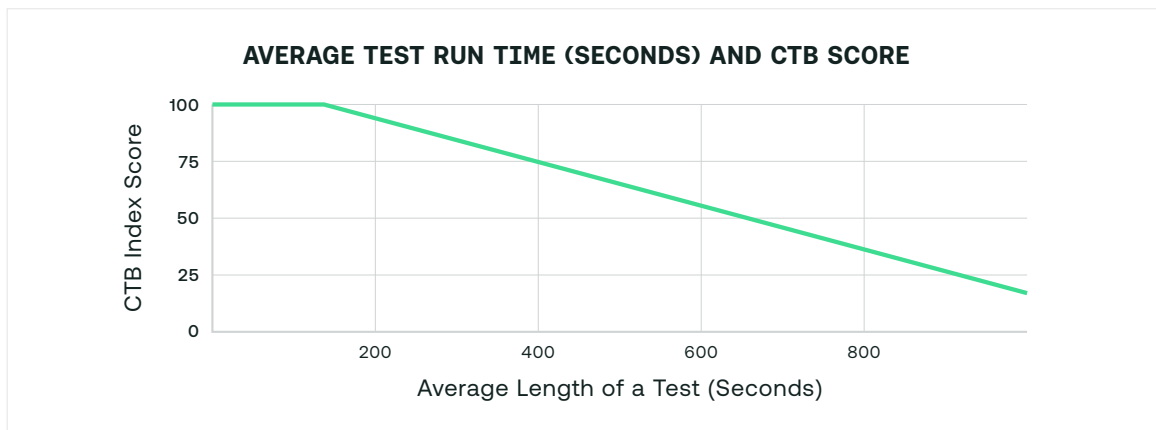
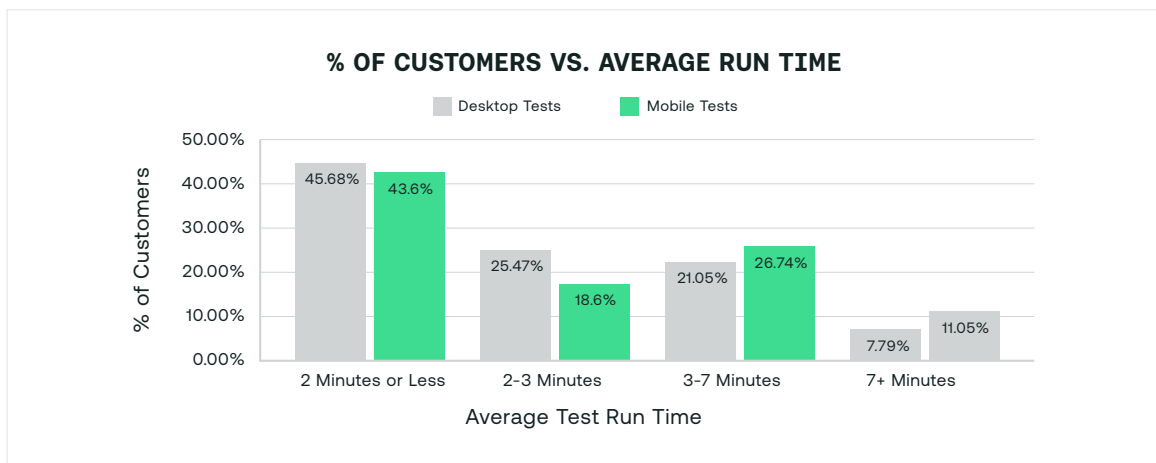
Why it Matters: Long, slow tests make it tough to facilitate development. The benefits of keeping test run times short extend beyond just speed alone. The data confirms that longer tests lead to poor test quality, as tests that complete in two minutes or less are nearly twice as likely to pass.

Simply put, the longer a test takes to run, the more likely it is to fail. This is often because longer tests cover more parts of the application, increasing the chances of encountering issues. Additionally, longer tests are more challenging to troubleshoot. They often attempt to assess multiple functionalities in one script, making it harder to pinpoint the source of failure when they don't pass.

The Benchmark for Excellence: The CTB Test Run Time metric looks at the average run time of the tests run on Sauce Labs. To achieve a score of 100 signifying excellence, the average run time of an organization's tests should be two minutes or less.

How Organizations Performed:

- **Desktop:** 45.68% of organizations complete in an average of two minutes or less
- **Mobile:** 43.60% of organizations complete in an average of two minutes or less



The Takeaway: Over half of organizations still don't meet the benchmark for test run times, across both desktop and mobile. However, the noticeable improvement in desktop test speeds — a nearly 10% increase in organizations achieving the two-minute standard — is a positive trend, indicating progress towards faster testing run times.

Benchmark Metric #3: Test Platform Coverage

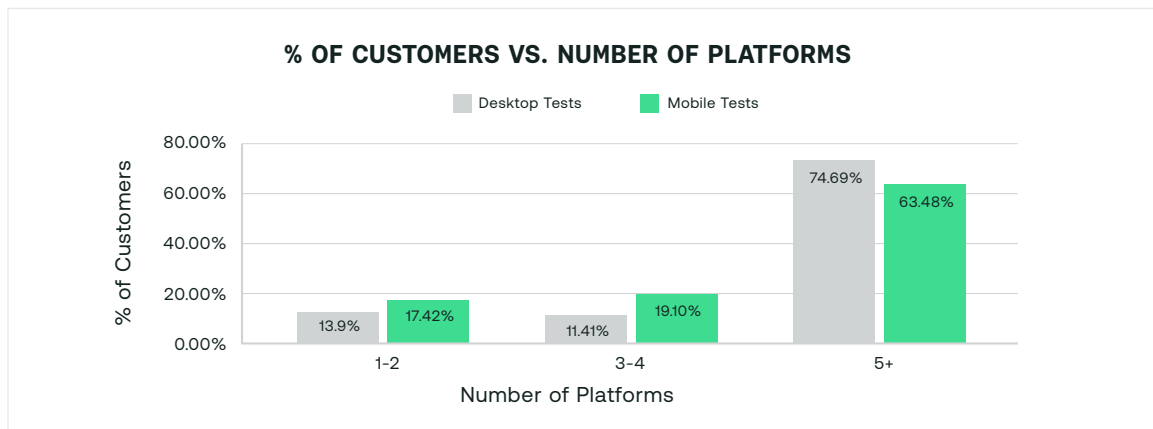
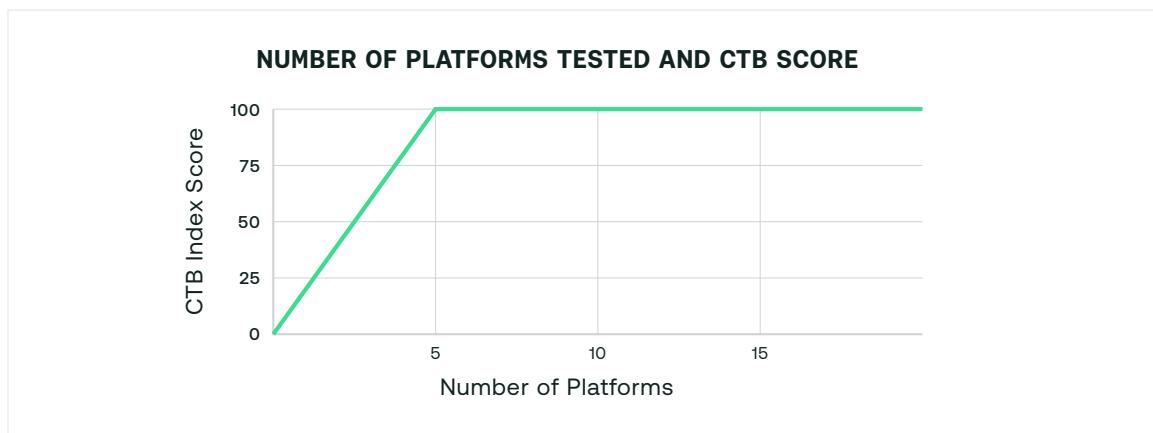
As the universe of mobile and laptop computing devices, operating systems and browsers available to consumers continues to expand on a seemingly daily basis, the ability to quickly determine if an application looks, functions, and performs exactly as intended across as many of those platforms as possible is a vital component of digital confidence.

Why it matters: Every business is now a digital business, and web and mobile applications are the new storefront. The more devices, operating systems, and browsers on which a business can “open” their store, the more opportunities they have to connect with customers. You can even take your test strategy one step further by collecting data about which devices, versions and operating systems your end-users engage with most often, and build a coverage plan, rooted in data.

The Benchmark for Excellence: The Test Platform Coverage metric looks at the number of platforms against which an organization tests, and gives full marks if at least five platforms are included in their typical testing activity. For desktop tests, a platform is defined as any combination of an operating system and a browser. For mobile tests, a platform is defined as any device type.

How Organizations Performed:

- **Desktop:** 74.69% of organizations test across at least five devices
- **Mobile:** 63.48% of organizations test across at least five devices



The Takeaway: Enterprises have clearly gotten the memo: Surviving in the all-digital era means delivering apps that work as intended whenever, wherever, and however customers wish to access them. The percentage of organizations running desktop tests across five or more platforms increased, and just 13.9% of organizations still fail to test across at least three desktop platforms.

The numbers are slightly lower across the board with respect to mobile testing. This is not unexpected, however, given that just a small handful of makers (namely Apple and Samsung) account for the lion’s share of the mobile device market. In either case, the benchmark data once again confirms that the vast majority of organizations now fully understand the importance of testing against multiple platforms.

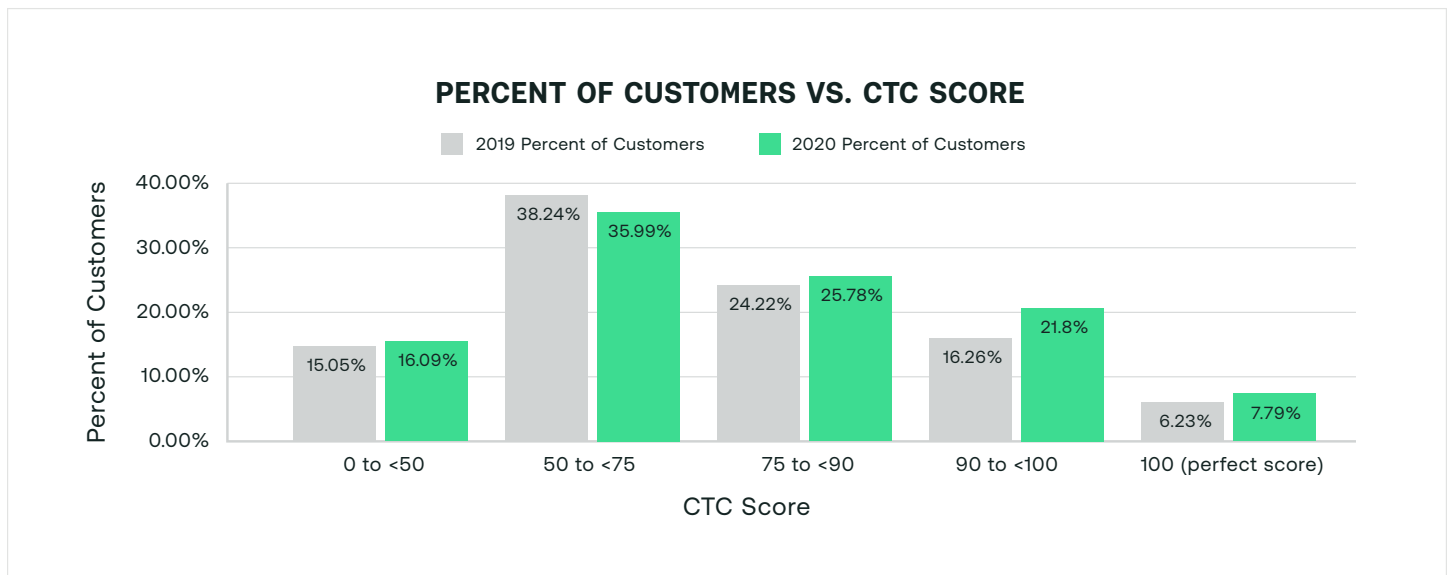
Additionally, businesses have started to leverage data from Google Analytics about which devices and systems their end-users engage with to further refine their Coverage strategy. If you know that your end users primarily engage on iPhone 13, 14, and 15, and Google Pixel 8 Pro, you’ll want to make sure that your app is thoroughly tested on those devices.

Conclusion

Organizations continue to prioritize implementing continuous testing throughout the software development lifecycle, and they continue to make considerable progress and improvements in their efforts to do so.

From the previous report, the number of organizations achieving the standard for excellence grew by at least 5% across each of the three benchmark metrics. The fact that mobile testing performance metrics closely mirrored desktop testing metrics is also encouraging. Despite mobile being a more nascent area of testing, organizations are doing well to carry over established best practices.

Staying on the positive side of the ledger, at least 75% of organizations are now meeting the benchmarks for excellence with respect to test platform coverage. As a collective group, we're testing across a wide variety of platforms and running tests in parallel as we do. In addition, although the overall numbers for test run time are not yet where they need to be, the nearly 10% improvement in the number of organizations meeting the benchmark for excellence is highly encouraging.

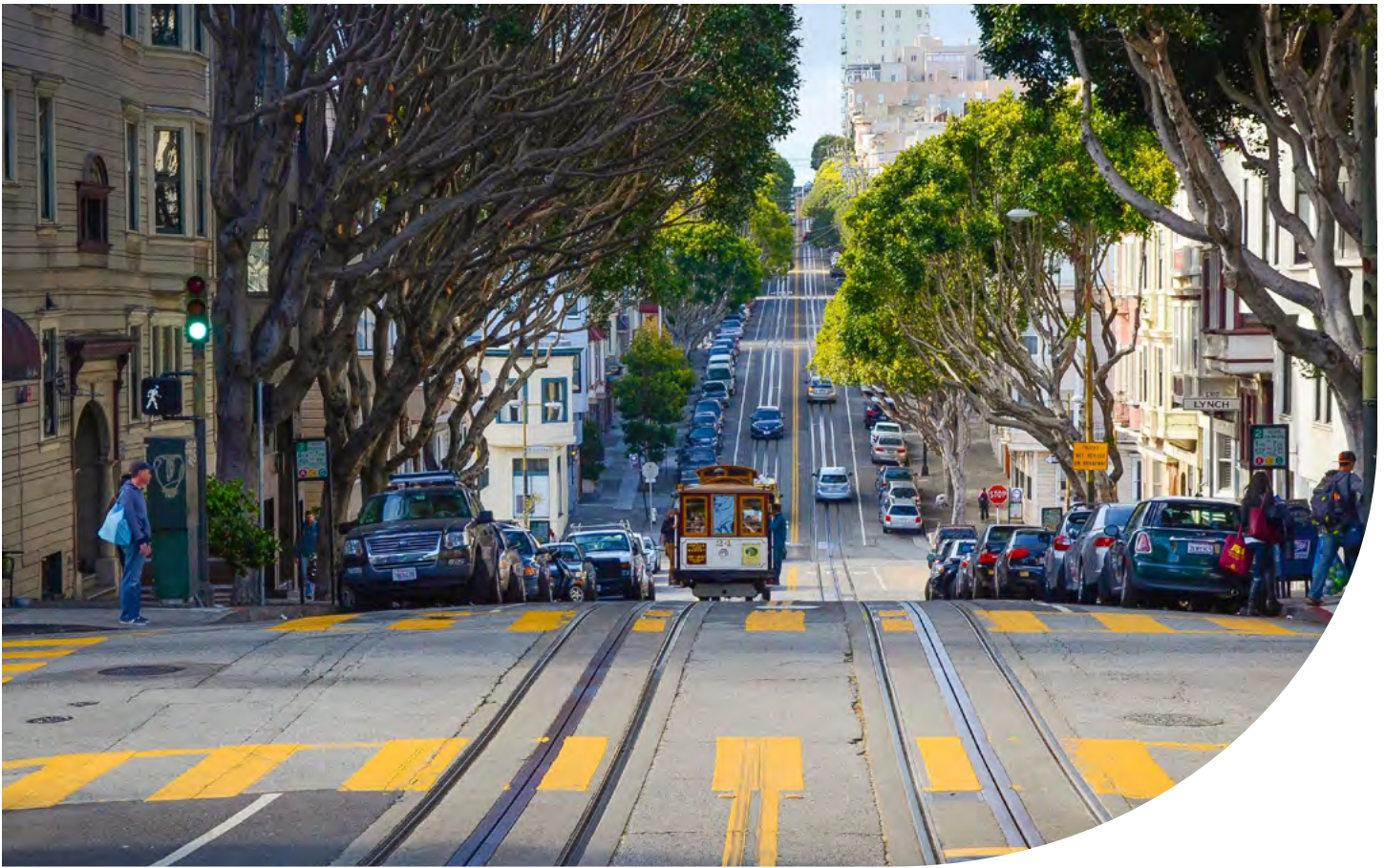


Continuous Testing Recommendations

- Run atomic tests to improve test quality, reduce test run times, and achieve parallelization.
- Leverage failure analysis to surface your most common test errors, prioritize how frequently they are occurring, and gain insight into where you can focus efforts to improve your pass rates.
- Be sure to send the status of your tests as pass/fail data, in order to capture insights over time.
- Adopt a mobile-first strategy in which your investment in mobile testing and development is proportional to current and anticipated customer usage.
- Seek the help of experts. Testing is hard. Experienced testing experts can help you navigate the challenges.

How is my organization doing?

Want to know how your test performance compares to these benchmarks? [Schedule a demo to see how you stack up.](#)



About Sauce Labs

Sauce Labs is the leading cloud-hosted platform for automated testing of web and mobile applications, enabling fast delivery of high-quality software across the development lifecycle. Founded by the creators of Selenium, Sauce Labs has been the testing leader for over 15 years and now runs over 1 billion tests annually. Trusted by Fortune 500 companies like Toyota, Walmart, Verizon, Gannett, and Fidelity Investments, its scalable, secure platform supports testing across thousands of operating systems, browsers, and devices while meeting the highest compliance standards.

For more information, please visit
→ saucelabs.com

SAUCE LABS INC. - HQ 450 Sansome Street, 9th Floor, San Francisco, CA USA 94111

© 2020-2024 Sauce Labs. All Rights Reserved.

