#225135

April 2025

Commissioned by Checkmarx

Checkmarx One Platform

SAST and SCA Application Security Efficacy vs. Competitor

EXECUTIVE SUMMARY

The vast number of applications being coded and updated daily opens a vast attack surface for hackers. Exploiting software applications can be a very effective way for hackers to infiltrate businesses. It is essential that businesses be aware of potential security vulnerabilities in their applications so that they can prioritize the appropriate remediation to protect their business assets.

Checkmarx commissioned Tolly to work with them on reviewing and documenting a comparison between Checkmarx and a competitor. The test encompassed scanning two applications using Static Application Security Testing (SAST) & four applications using Software Composition Analysis (SCA) testing. Results were analyzed to compare true positives, false positives, and false negatives.

Checkmarx demonstrated significantly better results - higher true positives (TP), lower false positives (FP), and lower false negatives (FN) - than the competing solution in tests of both SAST and SCA. See Figure 1.

THE BOTTOM LINE

1 Checkmarx SAST demonstrated higher accuracy with significantly higher TP, lower FP, and lower FN rates.

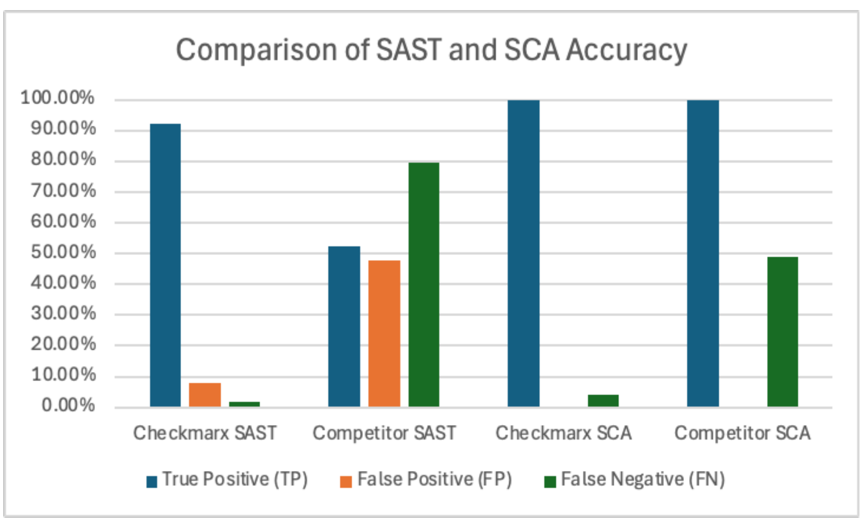
2 Checkmarx SCA identified more TP in the tested packages, vs. the competitor which showed higher FP and FN.

3 Exploitable vulnerabilities provides an example of integration between SAST and SCA. Checkmarx identified more of the exploitable vulnerabilities vs. the competition.

Application Security Efficacy - Checkmarx vs. Competitor

Static Application Security Testing (SAST) & Software Composition Analysis (SCA) Testing

SAST SCA



Note: Based on analysis of several hundred results for each solution. SAST based on examining two source code projects, SCA based on examining four source code projects. Checkmarx had zero false positives in the SCA test.

Source: Tolly, March 2025

Figure 1

Application Security Efficacy - SAST & SCA Testing Tabular Results

Solution	True Positive (TP) What it found (and was right)		False Positive (FP) What it found (but was wrong)		False Negative (FN) What it missed		Total Results	Total without FPs
Checkmarx SAST	*93*	*92.08%*	*8*	*7.92%*	*2*	*1.94%*	*103*	*95*
Competitor SAST	*12*	*52.17%*	*11*	*47.83%*	*89*	*79.46%*	*112*	*101*
Checkmarx SCA	*204*	*100.00%*	*0*	*0.00%*	*9*	*4.23%*	*213*	*213*
Competitor SCA	*101*	*100.00%*	*0*	*0.00%*	*96*	*48.73%*	*197*	*197*

Source: Tolly, March 2025 Table1

SAST & SCA Overview

Enterprise applications will generally consist of both proprietary code and open-source code. Because open-source code is subjected to scrutiny by many developers and security analysts, vulnerabilities are identified over time, catalogued, and available to app security vendors and others in third-party databases such as https://nvd.nist.gov. This is not the case with proprietary, custom source code.

SAST testing is focused on identifying vulnerabilities in proprietary code. Each vendor's SAST scanning implementation uses rules and other methods to identify potential security vulnerabilities.

SCA focuses on scanning an application's open-source components to identify security vulnerabilities, aging components, and potential license conflicts. These components are compared against established databases of vulnerabilities.

Checkmarx analyzed all of the identified potential vulnerabilities to determine true and false positives, and Tolly spot-checked and validated the results.

Test Results

Summary

As shown in Figure 1, Checkmarx demonstrated better overall results than the competitor. Table 1 also contains those results with the addition of a column that provides the number of individual results in each category. Each test type will be discussed separately. See the Test Setup & Methodology section for details of the software codebase used in the evaluation and the testing & analysis process.

SAST Results

The results of the SAST scan done with Checkmarx identified 103 potential vulnerabilities across the two applications tested, where the competitor identified 112.

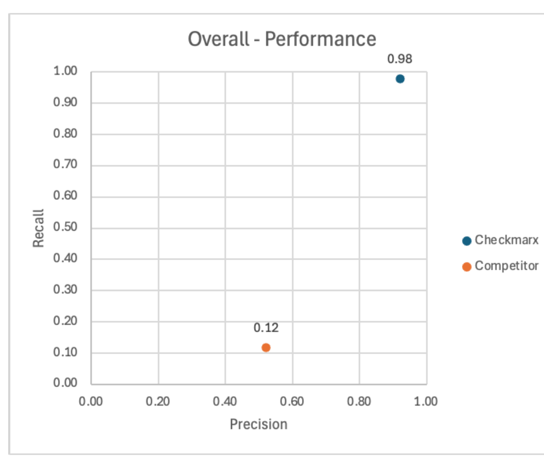
Checkmarx analysts deemed 92.08% of those items to be accurately found as true

SAST Testing: F-Score Calculation & Results

F-score Calculation

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$



Source: Tolly, March 2025 Figure 2

positives (TP) and 7.92% to be false positives (FP).

For the competitor, analysts deemed 52.17% of those items to be accurately found as TPs and 47.83% to be FPs.

Precision Score

Many people often focus on FP as the indicator of accuracy.

FP provides an incomplete picture. It only focuses on the % of positives that were wrong. It completely misses the FN, or vulnerabilities that were never detected in the first place.

That's why it's better to think about accuracy as a combination of TP, FP, and FN.

This testing highlights this well, where at first glance, the competitor shows a smaller number of FP. However, because the competitor identified a lower number of total positives (with a corresponding high FN rate), the FP % is actually much higher than the Checkmarx solution.

Another way to look at accuracy 'at a glance' is by comparing precision and recall. These calculations take into account predicted and actual results and are useful in comparing solutions where different numbers of results were generated. A perfect score for each is 1.0.

Figure 2 contains a visual representation of how precision and recall are calculated and, more importantly, the results of this SAST test.

Checkmarx has a recall of .98 which is over 8x that of the competitor's .12, and a precision of .92 which is 77% higher than the competitor's .52. This indicates that the overall accuracy of Checkmarx is dramatically higher than that of the competing solution.

SCA Results

The results of the SCA scan done with Checkmarx identified 204 vulnerabilities across the two tested applications, where the competitor identified 101. See Figure 3.

Checkmarx

Checkmarx One Platform

Application Security

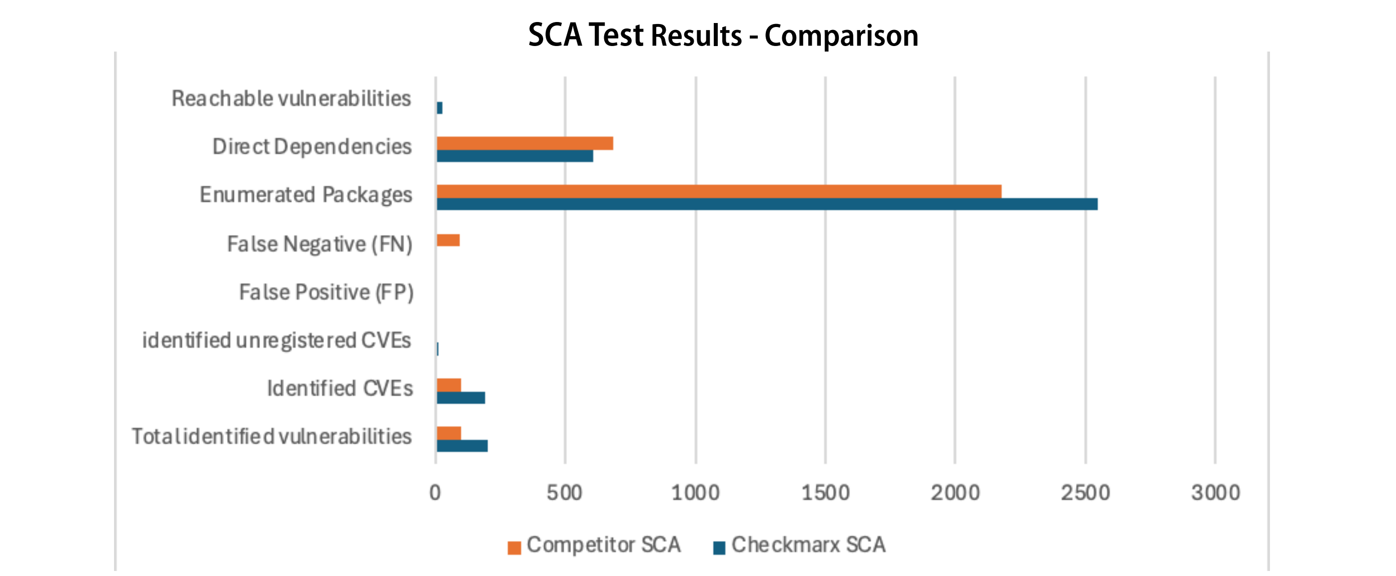
Tolly Certified

Tested March 2025

For Checkmarx, analysts deemed 100% of those items to be accurate as true positives (TP) and 0% to be false positives (FP).

For the competitor, analysts also deemed 100% of those items to be accurate as TPs and 0% to be FPs.

While every SCA solution starts with access to the same vulnerability databases, different solutions may take additional actions for more in-depth analysis. As a result, one can see that Checkmarx identified over 2x more known vulnerabilities than the competitor. Looking into the details of the results



Source: Tolly, March 2025

Figure 3

provides additional insights into the Checkmarx results.

Some SCA products can also identify vulnerabilities that aren't in public CVE databases, such as through their own threat research teams. In this case, Checkmarx identified three additional vulnerabilities in the two applications where the competitor identified one.

As a possible explanation for the results seen, Checkmarx identified 366 more enumerated packages than the competitor when scanning the open-source programs. This means that Checkmarx "looks deeper" into the source code to find additional called modules/packages that the competitor did not identify.

This additional level of inspection likely translates to the higher number of identified vulnerabilities across the various categories shown in Figure 3.

Ultimately, Checkmarx identified more unique Common Vulnerabilities and Exposures (CVEs), packages, and directed dependencies.

Exploitable Path

Because not every vulnerability in an open source library is in a function or method actually called by the application code, not every vulnerability reported by an SCA tool is actually exploitable in the application.

Some SCA solutions include a feature that Checkmarx calls exploitable path. This feature works by looking at the functions/methods in open source libraries that are actually called by the application code

The vulnerabilities in those functions/methods are the ones that are "exploitable" or "reachable."

Identifying exploitable or reachable vulnerabilities helps customers prioritize remediation on the ones that are actually exploitable.

As shown in Figure 3, Checkmarx identified significantly more exploitable paths inside the examined code. I.e., identifies which lines in the project code actually reach the vulnerable method in the vulnerable package. This means that the competitor's feature is less effective at helping customers prioritize what vulnerabilities to remediate. In fact, one will miss vulnerabilities that are actually exploitable.

Test Setup & Methodology

Codebase

In order to allow users to reproduce this test using Checkmarx or another application security tool, open-source projects were chosen for both the SCA and SAST testing - even though the real-world use case for SAST is for proprietary/custom code. See Table 2 for source file names and GitHub links for each project.

Application Source Programs Examined

Program Name	Version Examined	Language	Exam. Type	Source URL (Github)	Checkmarx Processing Options
evershop	*1.0.0*	*Javascript*	*SCA*	*https://github.com/evershopcommerce/evershop/releases/tag/1.0.0-rc.9*	*SCA exploitable true.*
istio	*1.20.3*	*Go*	*SCA*	*https://github.com/istio/istio/wiki/Istio-Release-1.20*	
Mojoportal	*2.9.0.1*	*C#*	*SCA*	*https://github.com/i7MEDIA/mojoportal*	*SAST preset ASA Premium C#. Disabled: SCA, IaC, and API*
OpenMRS-Core	*2.6.2*	*Java*	*SAST & SCA*	*https://github.com/openmrs*	*SCA exploitable true. SAST preset ASA Premium Java. Disabled: IaC and API*
SimplCommerce	*1.0.0rc*	*C#*	*SAST*	*https://github.com/simplcommerce/SimplCommerce/releases*	

Source: Tolly, March 2025 Table 2

Testing Process

Checkmarx staff ran the test using their solution and specified their examination options. See Table 2 for the options.

A third-party application security company familiar with the competitor's product ran the test of the competing solution.

Analysis & Tolly Review Process

Hundreds of vulnerabilities were logged by each of the two solutions. Checkmarx analysts reviewed each finding from the two vendors and classified every item.

A finding was confirmed to be a true positive (TP) if analysts deemed the finding to identify a vulnerability. A positive vulnerability finding was deemed a false positive (FP) if analysts did not agree that the positive identified an actual vulnerability.

To identify false negatives (FN) analysts relied on TPs that one solution identified and the other missed. (Results were categorized as "common true positives" detected by both solutions and "unique true positives" that were identified only by one of the two solutions.

Tolly spot-checked these results with an in-house application programming/security subject matter expert. With that expert, Tolly chose, at random, multiple TP, FP, and FN results for each vendor and navigated through the source code to review the specific line of source code cited. Tolly and its expert concurred with all of the findings of the analysts.

About Tolly

The Tolly Group companies have been delivering world-class IT services for more than 35 years. Tolly is a leading global provider of third-party validation services for vendors of IT products, components and services.

You can reach the company by E-mail at info@tolly.com, or by telephone at +1 561.391.5610.

Visit Tolly on the Internet at:
http://www.tolly.com

Terms of Usage

This document is provided, free-of-charge, to help you understand whether a given product, technology or service merits additional investigation for your particular needs. Any decision to purchase a product must be based on your own assessment of suitability based on your needs. The document should never be used as a substitute for advice from a qualified IT or business professional. This evaluation was focused on illustrating specific features and/or performance of the product(s) and was conducted under controlled, laboratory conditions. Certain tests may have been tailored to reflect performance under ideal conditions; performance may vary under real-world conditions. Users should run tests based on their own real-world scenarios to validate performance for their own networks.

Reasonable efforts were made to ensure the accuracy of the data contained herein but errors and/or oversights can occur. The test/audit documented herein may also rely on various test tools the accuracy of which is beyond our control. Furthermore, the document relies on certain representations by the sponsor that are beyond our control to verify. Among these is that the software/hardware tested is production or production track and is, or will be, available in equivalent or better form to commercial customers. Accordingly, this document is provided "as is," and Tolly Enterprises, LLC (Tolly) gives no warranty, representation or undertaking, whether express or implied, and accepts no legal responsibility, whether direct or indirect, for the accuracy, completeness, usefulness or suitability of any information contained herein. By reviewing this document, you agree that your use of any information contained herein is at your own risk, and you accept all risks and responsibility for losses, damages, costs and other consequences resulting directly or indirectly from any information or material available on it. Tolly is not responsible for, and you agree to hold Tolly and its related affiliates harmless from any loss, harm, injury or damage resulting from or arising out of your use of or reliance on any of the information provided herein.

Tolly makes no claim as to whether any product or company described herein is suitable for investment. You should obtain your own independent professional advice, whether legal, accounting or otherwise, before proceeding with any investment or project related to any information, products or companies described herein. When foreign translations exist, the English document is considered authoritative. To assure accuracy, only use documents downloaded directly from Tolly.com. No part of any document may be reproduced, in whole or in part, without the specific written permission of Tolly. All trademarks used in the document are owned by their respective owners. You agree not to use any trademark in or as the whole or part of your own trademarks in connection with any activities, products or services which are not ours, or in a manner which may be confusing, misleading or deceptive or in a manner that disparages us or our information, projects or developments.

Dm-2-wt-2025-04-24 VerC