

BLACK DUCK®

READINESS REPORT FOR 2027

Benchmarks for Remediation Readiness in the AI Era

TABLE OF CONTENTS

EXECUTIVE SUMMARY 1

THREE KEY TAKEAWAYS 2

THE MYTHOS READINESS FRAMEWORK 4

- Visibility 5
- Coverage 6
- Prioritization 7
- Remediation Velocity 9

BUSINESS IMPACTS 11

- Data Theft and Breach Exposure 11
- Financial and Operational Drag 11
- Regulatory and Compliance Exposure 11
- Customer Trust and Organizational Confidence 11

RECOMMENDATIONS 12

- Build Complete, Current Visibility 12
- Close the Coverage Gaps 12
- Prioritize by Exploitability, Not Severity Alone 12
- Measure and Accelerate Remediation 12
- Pair Deterministic Testing with AI-Assisted Triage and Agentic Remediation 12

WHY MACHINE-SPEED DEFENSE MATTERS 13

HOW BLACK DUCK CAN HELP 14

- A Unified View of Application Risk with Black Duck Polaris 14

CONCLUSION 15

METHODOLOGY 16

APPENDIX A: SUMMARY STATISTICS 17

APPENDIX B: MONTHLY FINDING TRENDS BY SEVERITY 18

APPENDIX C: TIME-TO-CLOSE PERCENTILES BY METHODOLOGY AND SEVERITY 19

APPENDIX D: COMPOSITION BY COMPANY SIZE 20

APPENDIX E: COMPOSITION BY REVENUE BAND 21

APPENDIX F: SCAN FREQUENCY VS. TIME-TO-CLOSE BY METHODOLOGY 22



EXECUTIVE SUMMARY

For most of the last decade, security teams have had time on their side. In 2018, the mean time between a vulnerability's disclosure and its first exploit was 2.6 years, but today, that span has shrunk to under one day, according to the Zero Day Clock. Frontier AI models are accelerating the trend, pushing the exploit window from hours to minutes. In Anthropic's lab tests, Claude Mythos Preview produced 181 working Firefox exploits under the same conditions in which its predecessor produced 2. In February 2026, Sysdig documented an AI-driven attack that reached admin-level access in eight minutes. Vulnerability management programs weren't built for this speed.

> **The exploit window has shrunk so drastically that traditional, human-speed patching processes break under the pressure.**

To document the widening gap between discovery and action, Black Duck analyzed 65.5 million application security testing findings across roughly 13,100 applications at 533 organizations between June 2025 and May 2026.

Here's what we found.

- **Remediation runs on a monthly clock. Exploitation runs on an hourly one.** Our data shows that static application security testing (SAST) findings were closed in a median of 28 days and software composition analysis (SCA) findings were closed in a median of 39. Together, these two methods account for 99.7% of all closures in our study. In stark contrast, closing dynamic application security testing (DAST) and manual testing findings took months: 316 and 232 days, respectively.

- **Most findings don't get fixed at all.** Only about 1 in 10 findings in the dataset had been closed by the end of the study.

- **The gap is widening.** By June 2025, organizations were discovering 11 new findings for each 1 they closed. By May 2026, that ratio was 18 to 1, and it peaked at 22 to 1 in December 2025. In April 2026 alone, organizations added 10.4 million unresolved findings to their backlogs.

- **Critical- and high-severity findings take a bigger share.** Critical- and high-severity findings made up 16.8% of new findings in May 2026, the highest share of the year, up from 13.1% in June 2025.

Some organizations are responding successfully to the challenge—they've built systems to act on what they find, including continuous scanning, security embedded in developer workflows, and repeatable processes that turn findings into fixes. But for every organization, the central question is what to fix first. Our study shows that the best way to answer it is to combine deterministic AST with AI-assisted triage, and then to use agentic remediation to close the loop at machine speed. The advantage belongs to organizations that can determine what to fix first and can fix it before an attacker gets there.

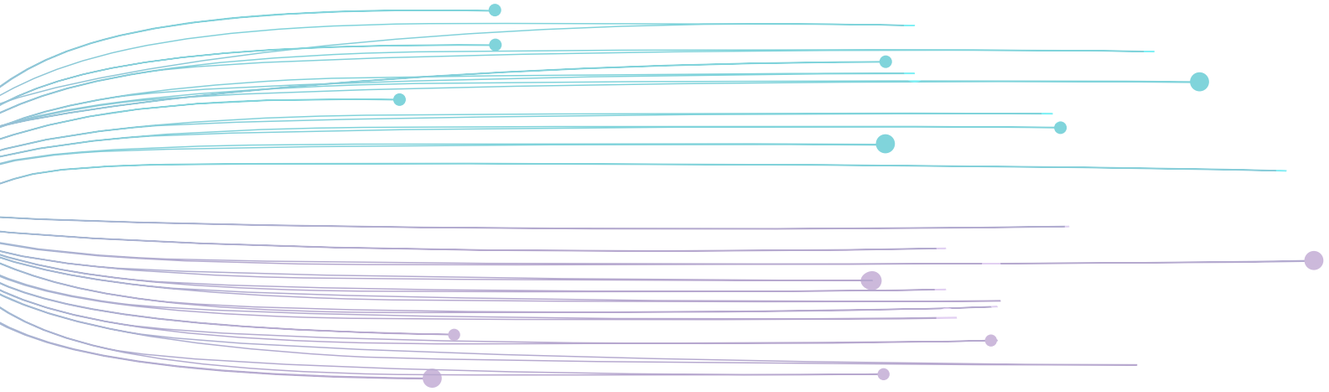
THREE KEY TAKEAWAYS

1

Frontier AI models are compressing exploit windows to hours, and manual vulnerability management can't keep up.

The traditional patch cycle assumed months between disclosure and weaponization, but that assumption no longer holds. Time-to-exploit has fallen below a day, and AI models now generate working exploits with far less effort than before.

Defensive processes haven't changed at the same rate as discovery, with the discovery-to-closure ratio moving from 11:1 to 18:1 over the study period. Even the fastest-remediated findings (those found by SAST and SCA) stay open for about a month at the median. The bottom line is that remediation workflows that move at human speed can't close a gap that grows by millions of findings a month.



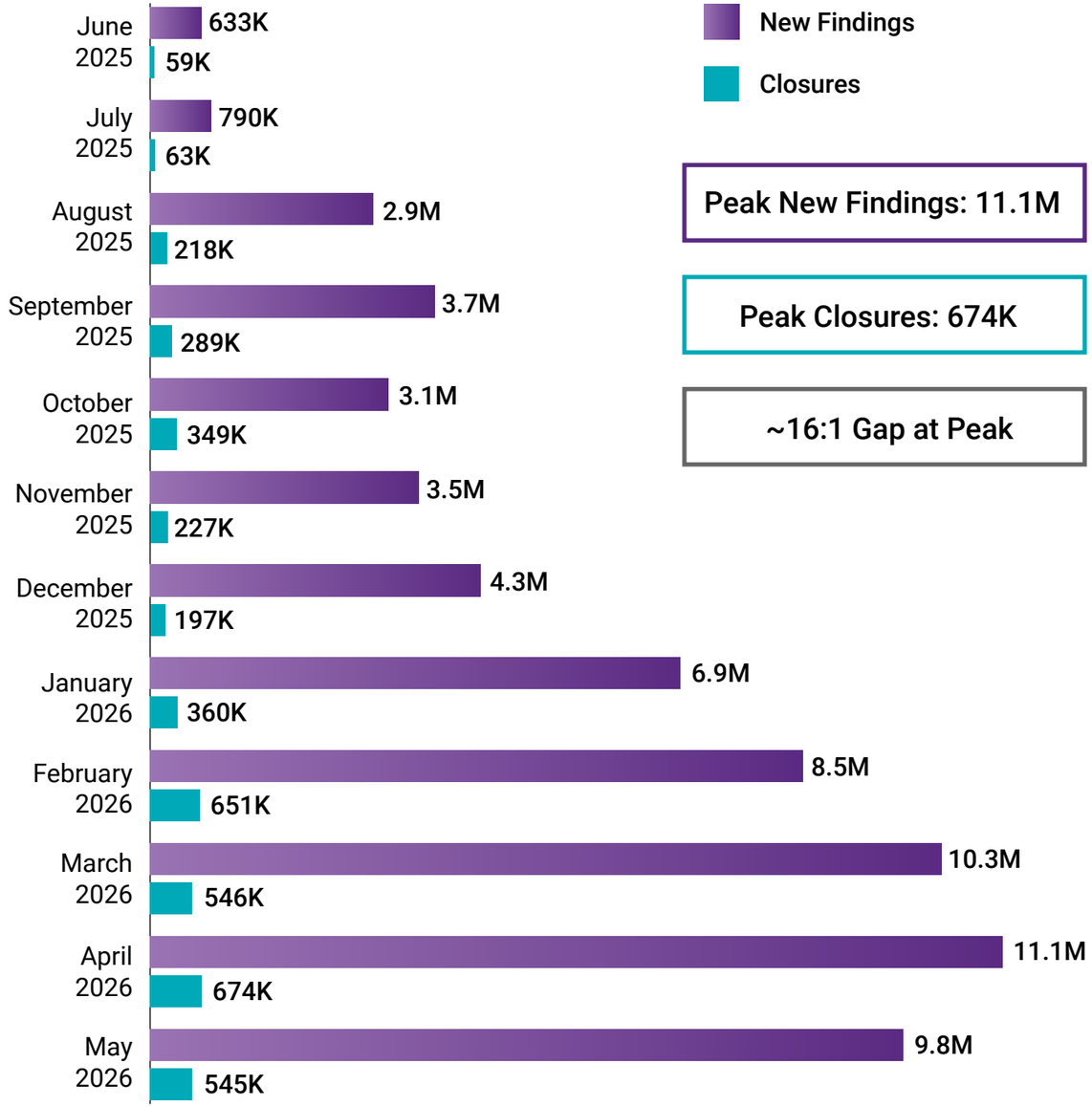


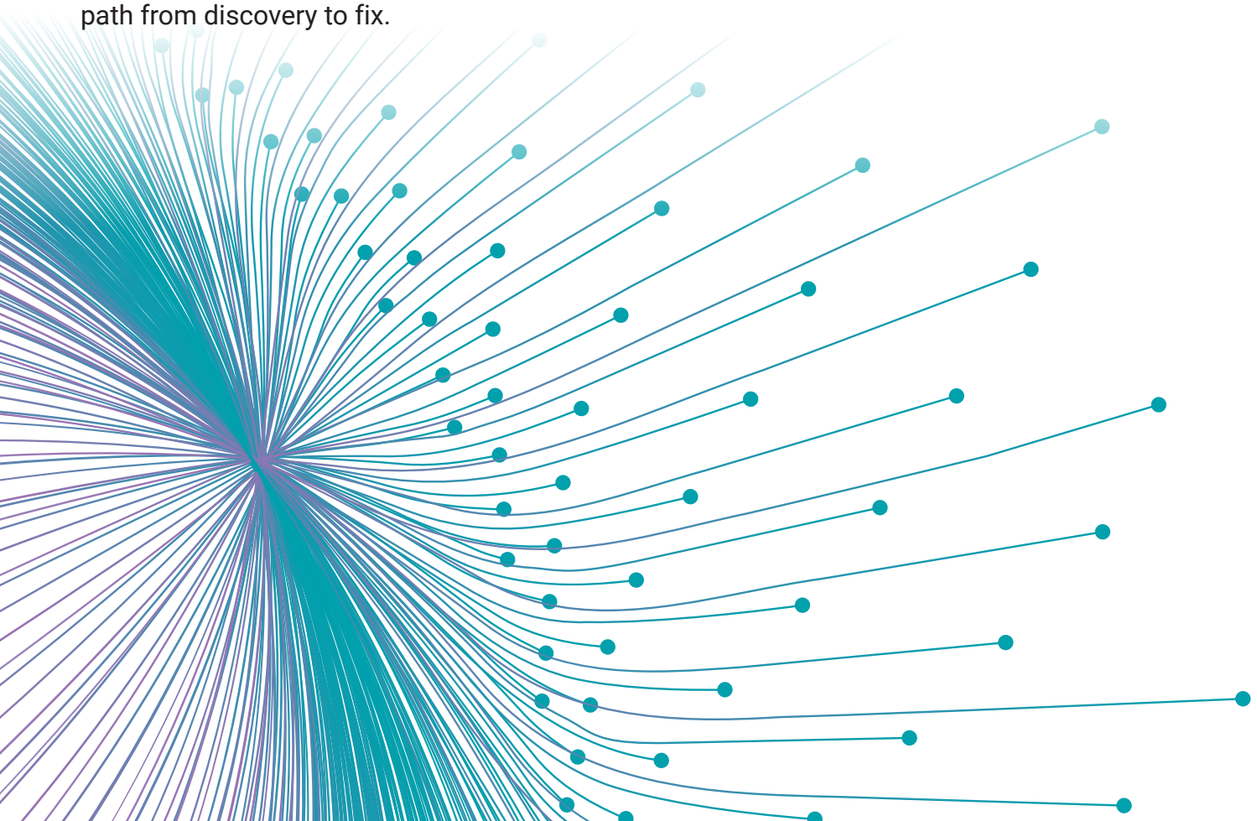
Figure 1. New findings vs. closures per month (June 2025–May 2026).

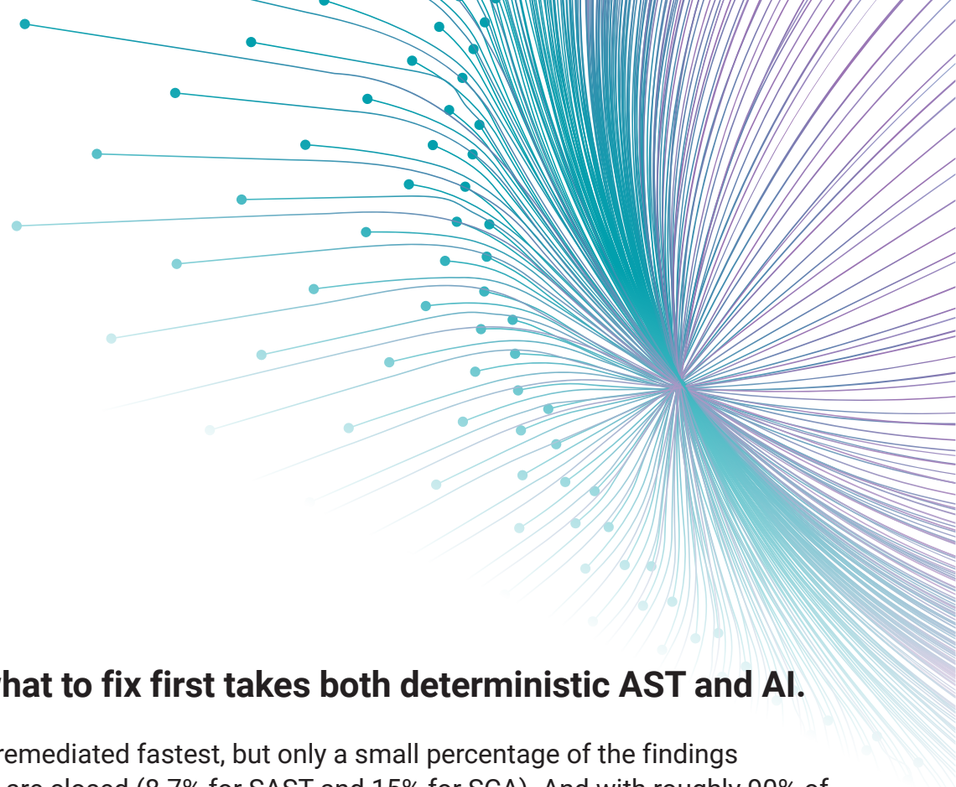
2 The organizations adapting the best have built the systems to act on what they find.

Remediation speed tracks to operational practices more closely than it tracks to vulnerability volume, a pattern that shows up in several places.

- **Scan frequency.** Organizations scanning open source dependencies 10 to 29 times a month closed SCA findings 54% faster (in 30 days versus 66 days) than those scanning less than monthly.
- **Type of testing.** Findings from automated scans that feed directly into developer workflows close quickly; e.g., 53% of critical SCA findings closed within 30 days. Findings from manual testing lag far behind, with only 18% of critical findings closing within 30 days.

What separates leaders from laggards is whether an organization has engineered a repeatable path from discovery to fix.





3 Knowing what to fix first takes both deterministic AST and AI.

SAST and SCA findings are remediated fastest, but only a small percentage of the findings identified by those methods are closed (8.7% for SAST and 15% for SCA). And with roughly 90% of findings still open, prioritization isn't optional.

Severity scores alone can't settle which vulnerability to fix first. A lower-severity vulnerability with a known exploit is more urgent than a theoretical critical one. And because AI models can chain lower-severity weaknesses into serious attacks, "lower priority" no longer means "never."

Determining what to fix first at scale requires three levels of insight.

- **Deterministic AST.** SAST and SCA produce consistent, reproducible, auditable results that hold up to compliance review.
- **AI-assisted triage.** AI helps weigh the exploitability, reachability, and business context when deciding what to fix first.
- **Agentic remediation.** Fix recommendations, pull requests, and validation must run at the same pace as the threat.

THE MYTHOS READINESS FRAMEWORK

Traditional application security programs were built for a world where time was on the defender's side. Teams could find a vulnerability, assess it, and schedule a fix well before attackers weaponized it. They could focus on visibility, because seeing a flaw still left time to deal with it, but that is no longer the case. Visibility alone doesn't reduce risk, and neither does ensuring all your applications are tested. Risk falls only when an organization can identify, prioritize, and remediate threats before they're exploited.

Black Duck built the Mythos Readiness Framework to address what Anthropic's Claude Mythos Preview now can do: find vulnerabilities at scale, chain lower-severity weaknesses into multistep exploits, and produce working exploits with far less effort than attackers needed before. Within the next 6 to 12 months, Black Duck expects comparable capabilities to be widely available among AI models.

Mythos readiness measures how quickly an organization can act across these four pillars.

- **Visibility**
- **Coverage**
- **Prioritization**
- **Remediation velocity**

4 PILLARS

The gap between what organizations can see and what they can fix is the defining risk of the AI era.

The four pillars work together, not as a menu to select from. Fast remediation doesn't help if you can't see what's in your applications, and broad visibility doesn't help if you can't determine what to fix. The four pillars also map directly to this report's three takeaways.

- **Visibility and coverage provide the deterministic foundation.** These two pillars identify what's in your software and whether you're testing all of it, with results that are consistent and auditable.
- **Prioritization answers the "what to fix first" question.** AI-assisted triage adds the value here, weighing exploitability and context alongside severity.
- **Remediation velocity is where you can address the compressed exploit window.** The organizations closing findings fastest share a common trait: They've built systems to act on what they find.

Pillar	The question it answers	What this report measures
Visibility	Do you know what's in your applications, both components and vulnerabilities?	• Volume and severity of identified findings • Percentage of critical- and high-severity findings from open source
Coverage	Are you testing all your applications with the methods they need or just a subset?	• Volume and severity of findings by testing methodology
Prioritization	Are you fixing what matters most first?	• Time-to-close (TTC) by severity and methodology • TTC by severity and company size • Exploit availability by severity
Remediation velocity	Are you closing vulnerabilities faster than they accumulate—and faster than attackers can exploit them?	• Median TTC by methodology • Discovery-to-closure ratio • Percentage of findings closed • Net backlog growth

Figure 2. The four pillars of the Mythos Readiness Framework.

PILLAR 1

VISIBILITY

Visibility means knowing two things: which components are in your applications, and which vulnerabilities they contain. Everything else in the framework depends on this visibility, because you can't prioritize or fix what you can't see. And in a world of AI-assisted attackers, you need to see everything.

Organizations Are Finding More Vulnerabilities than Ever

Organizations in our dataset identified 65.5 million vulnerabilities across roughly 13,100 applications in 12 months, including 9.2 million critical- and high-severity findings.

Open source is a large part of that picture. SCA found 18.9 million vulnerabilities in open source and third-party components, accounting for 52% of all critical- and high-severity findings in the dataset. Much of the serious risk in these applications sits in code the organizations didn't write.

A Software Bill of Materials (SBOM) is how organizations formally track what's in their software. This inventory matters most when a new vulnerability is disclosed, because the first question is always "Are we affected?" An SBOM that is complete and current answers this question quickly. When organizations miss a vulnerable component, it's almost always because their inventory didn't show that the component was there.

> **Organizations are finding vulnerabilities at unprecedented scale, but seeing vulnerabilities isn't the same as knowing what's in your software.**

Two changes have raised the stakes.

- **Exploit windows have collapsed.** With exploits now arriving less than a day after discovery, you need to know within hours if you are affected.
- **Public vulnerability intelligence is thinning out.** In April 2026, NIST announced that it would prioritize enrichment for only a narrow subset of CVEs. The rest are still listed in the National Vulnerability Database, but they aren't scheduled for immediate analysis. This leaves most new disclosures without the analysis many programs rely on for prioritization, just as AI-driven discovery drives disclosure volumes up.

Keeping up with the attacker's clock takes a modern application security program that includes three levels of insight.

- **Deterministic testing is the foundation.** SCA inventories open source components and their known vulnerabilities, and SAST finds flaws in proprietary code, producing consistent, auditable results.
- **AI-powered code analysis finds what the rules miss.** AI can identify novel threats, risky application behavior, business-logic flaws, and AI-generated code patterns that traditional analysis may not detect on its own.
- **Agentic remediation turns inventory into answers.** It determines which newly disclosed vulnerabilities affect which applications and what to do about them.

Visibility starts the find/analyze/fix loop. The pillars that follow cover whether organizations can finish it.

PILLAR 2 COVERAGE

AI-assisted attackers don't limit themselves to one layer of an application. Frontier models can combine memory corruption, logic flaws (such as authentication bypass and privilege escalation), and cryptographic weaknesses into a single multistep exploit. Testing only one layer leaves the others open, and each untested layer is a link an attacker can use in the chain.

Each testing methodology examines a different part of the application and identifies a different kind of risk.

- **SAST** analyzes proprietary source code without running it. It finds flaws such as injection vulnerabilities, insecure cryptography, and hard-coded credentials.
- **SCA** identifies open source and third-party components and the known vulnerabilities in them.
- **DAST** simulates an attack on a running application from the outside, the way an attacker would.
- **Manual application security testing (manual AST)** combines tooling with expert-led assessment. It uncovers business logic flaws, authorization weaknesses, and multistep attack chains that automated tools miss.

Figure 3 shows that the methods differ sharply in how much risk they find, and how serious that risk is.

Methodology	What it examines	Applications tested	Findings	Percentage of critical- or high-severity findings
SAST	Proprietary source code	6,521	46.6 million	9.5%
SCA	Open source and third-party components	4,223	18.9 million	25.3%
DAST	Running applications	6,383	35,410	5.1%
Manual AST	Business logic and authorization	475	3,024	53.1%

Figure 3. Testing methodologies.

> **More than half of what manual testing finds is critical or high severity. The methods that find the most issues aren't necessarily the ones that find the most serious risk.**

Manual AST is the least-used methodology by far in our dataset, but more than half of what it finds is critical or high severity. These are exactly the logic and authorization flaws that AI models can now chain into working attacks.

Different Methods Identify Different Risk

Two-thirds of applications in the dataset (66%) were tested with a single methodology, and roughly one-third were tested with two. Only 0.4% were tested with three, and none with all four. Most enterprises run several application security tools, often from different vendors, so this data doesn't measure overall coverage, but what it does show is that each method highlights a different slice of risk.

The findings in Figure 3 show why no single method is enough. SAST and SCA produce more than 99.9% of all findings, yet for manual AST, the smallest source of findings, more than half of what's found is critical or high severity. A program that uses only high-volume methods will identify the most findings, but it might miss some of the most serious ones—the logic and authorization flaws that AI-assisted attackers are best placed to exploit.

Coverage Is Part of the Deterministic Foundation

Broad coverage doesn't happen through good intentions or team-by-team choices, it takes automation—security policies applied consistently across every repository, pull request, and build, without anyone having to remember to run a scan. That's the deterministic foundation highlighted in this report's third takeaway: consistent, repeatable testing that produces results that AI-assisted triage can build on.

Coverage tells you what risk exists across the attack surface. The question that follows is what to do about the findings.

PILLAR 3

PRIORITIZATION

Visibility and coverage tell you what your risk is—prioritization determines what you do about it. With roughly 9 in 10 findings remaining open, every security program is already choosing what gets fixed and what waits. The question is whether it's choosing deliberately.

Where Volume Is Highest, Severity Stops Driving the Order

If severity was determining the order, critical findings would be closed first. But in the results from the highest-volume testing methods, they don't.

Methodology	Critical	High	Medium	Low
SCA	58 days	22 days	47 days	25 days
SAST	n/a*	56 days	35 days	37 days
DAST	201 days	420 days	552 days	347 days
Manual AST	133 days	237 days	n/a**	351 days

Figure 4. Median time-to-close by severity and methodology.

**SAST critical findings were too few, and closed too uniformly, for comparison.*

***Manual AST had too few medium-severity findings to report.*

- **High-volume methods: Severity doesn't determine the priority.** Critical open source vulnerabilities take more than twice as long to close as high-severity ones. In SAST, high-severity findings close more slowly than medium- or low-severity results.
- **Low-volume methods: Severity largely drives the order.** DAST and manual AST are slower overall, but critical findings close first.

Some of the slower close times reflect harder fixes; e.g., a critical vulnerability in an open source component may require a major version upgrade, whereas a low-severity fix is often a simple patch. But the pattern is consistent: Teams work through the flood of findings by fixing what's easiest. The problem isn't that severity order is being ignored—severity alone has oversights of its own—it's that nothing deliberate seems to be setting the order at all.

The same split shows up in the backlog of critical- and high-severity findings.

- **SAST** added about 13 new critical- and high-severity findings for each one closed.
- **SCA** added about seven for each one closed.
- **DAST and manual AST** stayed much closer to even.

The slowest methods are keeping pace, but the fastest ones are being overwhelmed.

Measuring Severity Alone Isn't the Answer

Fixing critical vulnerabilities first would be an improvement, but it isn't enough on its own. Exploitability doesn't track to severity scores neatly. In the vulnerability advisories Black Duck's researchers published in 2025, known exploits existed for 45% of high-severity vulnerabilities, 39% of medium-severity ones, and 31% of low-severity ones.

AI raises the stakes further. Frontier models can chain lower-severity weaknesses into serious attacks, and a medium-severity flaw with a working exploit in a reachable part of a production application can be more urgent than a theoretical critical weakness. "Lower priority" can no longer mean "never fix."

What to fix first is determined by

- **Exploitability:** Is there a known or likely exploit?
- **Reachability**: Can an attacker actually reach the vulnerable code?
- **Business context:** What does the affected application do, and what data does it hold?

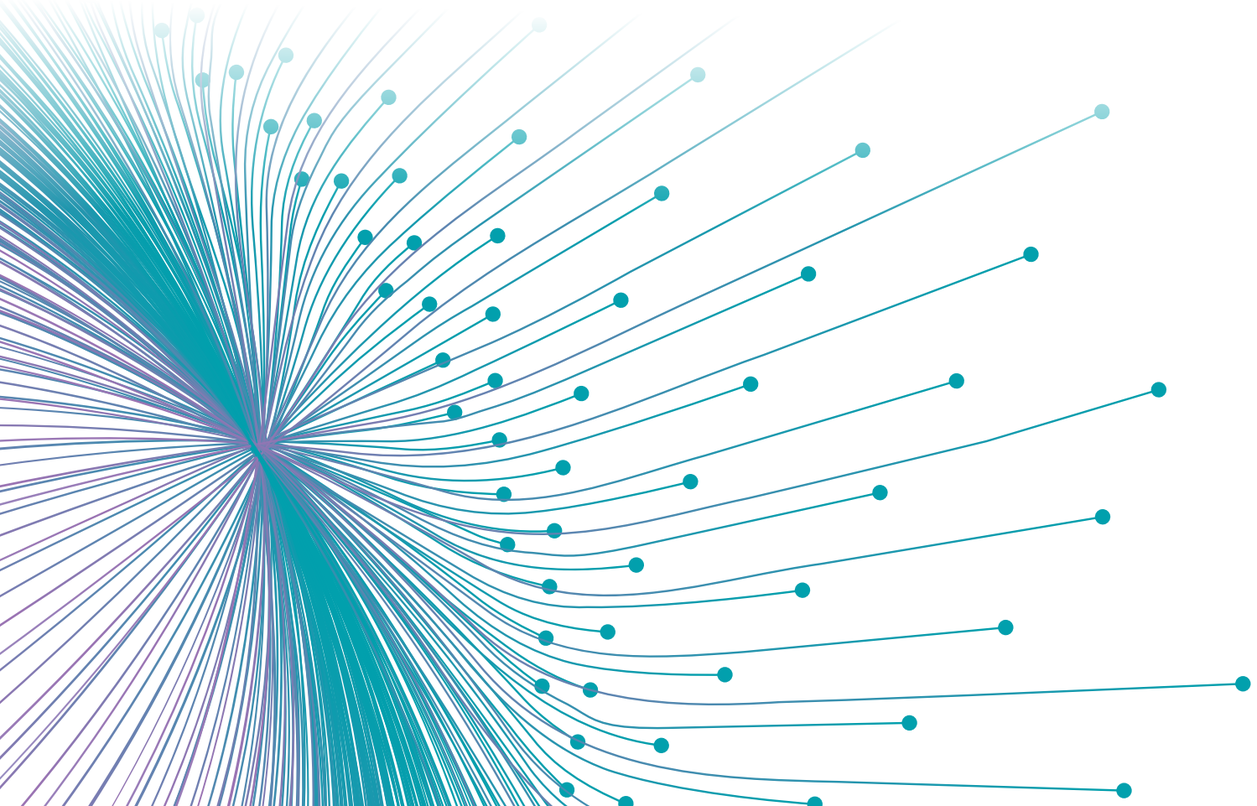
Some Organizations Already Prioritize Deliberately

Severity is one input, but it's not the complete answer. The largest organizations in the dataset show what deliberate prioritization looks like.

- **Organizations with more than 10,000 employees** close critical- and high-severity findings in a median of 12 days, compared to 19 days for all findings.
- **Organizations with 5,001 to 10,000 employees** show the same pattern, closing critical- and high-severity findings in a median of 17 days versus 22 days for all findings.

Smaller organizations show the reverse.

- **Organizations with 2,501 to 5,000 employees** take longer on their most serious findings (28 days) than on everything else (21 days).
- **Organizations with 501 to 1,000 employees** also take longer (28 days versus 24 days).



What Prioritization Takes in the AI Era

Knowing what to prioritize at the volume and speed of the AI era requires both deterministic testing (SAST, SCA, DAST) and AI-powered code analysis. Deterministic testing supplies consistent, trustworthy findings; AI code analysis finds the novel threats and business-logic flaws that rule-based tools can miss.

Prioritization determines what to fix first. The last pillar asks whether organizations can fix it fast enough.

> **The critical vulnerabilities identified by the highest-volume testing methods take longer to fix than lower-severity ones. When findings arrive by the millions, severity stops setting the order, and what gets fixed first is what's easiest.**

PILLAR 4 REMEDIATION VELOCITY

Visibility, coverage, and prioritization only reduce risk if they end in a fix—and the fix has to arrive before an exploit does. When AI is writing the exploits, that means remediating at machine speed, not human speed.

Two measures show where remediation stands: How quickly do findings close once teams act on them? And how many get closed at all?

Methodology	Median time-to-close	Percentage of findings closed
SAST	28 days	8.7%
SCA	39 days	15.0%
DAST	316 days	49.2%
Manual AST	232 days	43.8%

Figure 5. Remediation speed and closure rate by methodology.

As Figure 5 shows, the two measures are pulling in opposite directions: SAST and SCA findings close in about a month once teams act, but fewer than one in six of these findings are closed at all. DAST and manual AST findings take the better part of a year to close, but they are closed at several times the rate of SAST and SCA findings.

At the aggregate level, discovery is outpacing remediation. Organizations closed more findings over the year, but the discovery-to-closure ratio increased from 11:1 to 18:1.

How quickly a finding closes depends heavily on what fixing it involves. Open source vulnerabilities often close the fastest, with 53% of critical SCA findings closed within 30 days. Many need only a dependency update, which fits directly into a developer's workflow.

On the other hand, manual testing findings are slowest, with only 18% of critical manual AST findings closed within 30 days. Business logic and authorization flaws often require design changes rather than a simple patch.

It appears that the closer a fix sits to the developer's everyday workflow, the faster it happens. And that's also where automation can be the most help.

The clearest sign of this comes from fixing open source vulnerabilities. Organizations that run SCA scans 10 to 29 times a month closed findings in an average of 30 days, versus 66 days for those scanning less than monthly. Speed improves steadily as scan frequency rises, and then slows again at the highest frequency (41 days, at 30 or more scans a month).

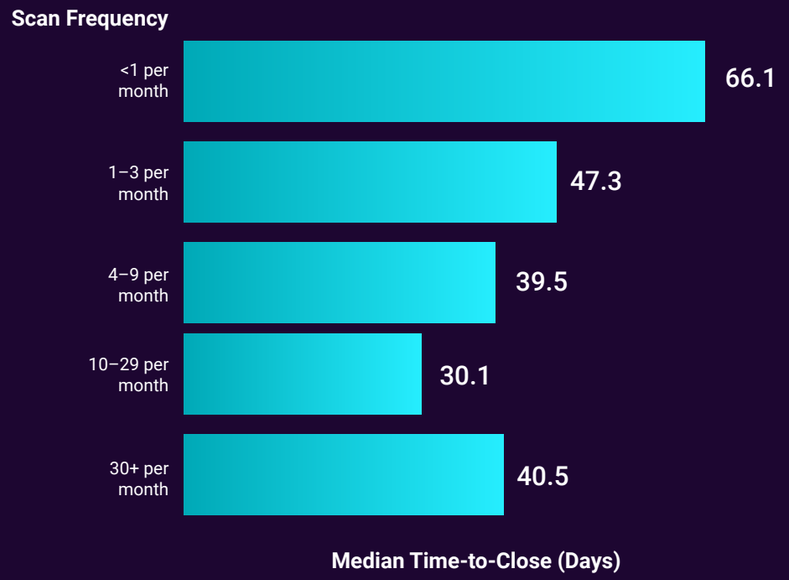


Figure 6. SCA time-to-close by scan frequency.

The relationship is correlational, and for SAST, it doesn't appear at all. Frequency alone isn't the lever. It's better read as a marker of maturity—frequent scanning usually indicates testing that is wired into pipelines and developer tools, with findings routed straight to the people who can fix them.

Organization size shows a similar picture (see Figure 7). The largest organizations (more than 10,000 employees) close findings in a median of 19 days and the smallest (100 or fewer) close in 17.

Number of employees	All findings	Critical- and high-severity findings
100 or fewer	17	18
101–500	34	22
501–1,000	24	28
1,001–1,500	34	15
1,501–2,500	56	50
2,501–5,000	21	28
5,001–10,000	22	17
More than 10,000	19	12

Figure 7. Median time-to-close by company size.

Why Agentic Remediation Matters

Human-paced remediation can't close a gap that grows by millions of findings a month. Keeping up takes agentic remediation built on the foundation described in the earlier pillars.

- **Fix recommendations and pull requests** generated from trusted, deterministic findings
- **Validation in the pipeline,** so fixes are confirmed without waiting for the next manual review
- **Human judgment saved for the calls that need it,** e.g., design changes, business logic flaws, and risk decisions

The find/analyze/fix loop is now complete. The organizations adapting the best to the AI era have the systems to act on what they find, at the speed attackers now move.

> **Fewer than one in six SAST and SCA findings get fixed. The ones that do take about a month, while attackers need less than a day to exploit.**

BUSINESS IMPACTS

The data we studied covers tens of millions of vulnerabilities across thousands of production applications that support real business operations, customer interactions, and sensitive data. Not every finding is exploitable, but the scale of the backlog, the age of what's in that backlog, and the direction of the trend all point to real exposure.

Now AI is changing what exposure means—when exploits arrive in hours rather than months, a vulnerability left open for weeks is no longer a routine backlog item, it's an open window. The greatest business risk is no longer failing to find vulnerabilities, it's failing to fix the right ones fast enough.

Data Theft and Breach Exposure

Discovery outpaced remediation throughout the study period, by 18 to 1 in May 2026 and 22 to 1 at the December peak. In April 2026 alone, organizations added 10.4 million unresolved findings to their backlogs.

At these ratios, even doubling remediation capacity would leave the backlog growing. What builds up is a larger and larger pool of known, unfixed vulnerabilities competing for limited engineering time. And frontier AI models are cutting the time and cost needed to find and exploit exactly those flaws.

Financial and Operational Drag

Unresolved vulnerabilities don't only create security risk, they create work. Every finding that stays open gets triaged, reassessed, reprioritized, and sometimes formally accepted as an exception, often more than once. Some stay open for an inconceivably long time: 10% of DAST findings remain open for 3.8 years or more; 10% of manual AST findings remain open for 3.0 years or more. Even with faster methods such as SAST and SCA, 10% takes roughly 10 months or more to close.

Regulatory and Compliance Exposure

Regulators increasingly expect timely remediation and fast disclosure.

- **PCI DSS 4.0.1** requires critical security patches within one month of release. In this dataset, nearly half of the critical open source vulnerabilities that were fixed took longer than 30 days. Most were never fixed during the study period at all.
- **The EU Cyber Resilience Act (EU CRA)** requires manufacturers to report actively exploited vulnerabilities in their products with digital elements, including an early warning due within 24 hours. Those obligations took effect September 2026.
- **The Digital Operational Resilience Act (DORA)** requires EU financial entities to demonstrate effective information and communications technology risk management, including how they identify and address vulnerabilities.
- **U.S. Securities and Exchange Commission (SEC) rules** require public companies to disclose material cybersecurity incidents within four business days of determining they're material.

Customer Trust and Organizational Confidence

A breach traced to a vulnerability that sat open for months won't be seen as a purely technical failure; it raises questions about governance, prioritization, and operational discipline. In other words, how long was this known, and why wasn't it fixed?

Organizations that can show they know what's in their software, decide what to fix first, and act on it quickly will be better positioned to keep customer trust and meet regulatory expectations.

RECOMMENDATIONS

The data points to three conclusions that we've examined repeatedly through this report.

- **Exploit windows have collapsed.** Frontier AI models are compressing exploits to hours, and manual vulnerability management can't keep pace.
- **Systems differentiate the leaders.** The organizations adapting the best have built systems to act on what they find.
- **What to fix first is the central question.** Answering this question, and acting on the answer, requires deterministic testing paired with AI-assisted triage and agentic remediation.

Build Complete, Current Visibility

When a new vulnerability is disclosed, you should be able to know if you are affected in hours, not weeks.

- **Maintain a current SBOM for every application.** Update it continuously, not at release time only.
- **Don't rely on the National Vulnerability Database alone.** With NIST now prioritizing enrichment for only a narrow subset of CVEs, you need vulnerability intelligence that doesn't depend on it.

Close the Coverage Gaps

Test each application from the inside and the outside. This means that for

- **Code and components,** run SAST and SCA.
- **Running behavior,** run DAST.
- **Business logic,** extend manual testing to your highest-risk applications. Logic and authorization flaws are the weaknesses that AI-assisted attackers can chain.

Enforce testing policies automatically across every repository, pull request, and build, so coverage doesn't depend on individual teams remembering to scan.

Prioritize by Exploitability, Not Severity Alone

Weigh exploitability, reachability, and business context alongside severity scores.

Start by checking your own data: Are your critical findings actually closing before your low-severity ones? If so, set timeframes for lower-severity findings, too. When attackers can chain minor weaknesses into major ones, "lower priority" can't mean "never."

Measure and Accelerate Remediation

Track three numbers together.

- Median time-to-close
- Percentage of findings closed
- Discovery-to-closure ratio

A fast median means little if only a small fraction of findings ever gets fixed.

Route fixes directly into developer workflows. Automate the fixes that can be automated, such as dependency upgrades and AI-generated code fixes. Each fix should be automatically rescanned and tested, and then submitted as a pull request for a developer to approve.

Pair Deterministic Testing with AI-Assisted Triage and Agentic Remediation

Knowing what to fix first and then acting on that answer takes all three levels working together.

- **Deterministic AST provides the foundation.** It delivers consistent, reproducible, auditable findings you can trust and defend.
- **AI-assisted triage determines what to fix first.** It weighs exploitability and context across millions of findings.
- **Agentic remediation acts on that decision.** It works at a pace that matches the attacker's clock.

Keep people in the loop for the judgment calls, i.e., validating risk, resolving complex vulnerabilities, and making business tradeoffs. Start where volume is greatest, with SAST and SCA findings.

WHY MACHINE-SPEED DEFENSE MATTERS

The numbers in this report should feel familiar to anyone running an application security program. Your backlog keeps growing, your team works harder every quarter, yet the gap between what you discover and what you can remediate keeps widening. The economics behind vulnerability management have shifted, and the operating model most programs were built on can no longer keep pace.

Across the data, we're seeing a structural mismatch play out in real time. The number of vulnerabilities being found is increasing thanks to expanded test coverage, growing software complexity, AI-assisted development, and evolving attacker capabilities. But remediation capacity has not kept up, and security teams are now processing more risk than their workflows were ever designed to handle.

This isn't a staffing problem—no amount of headcount can keep pace with vulnerability growth measured in millions of findings per month. Rather, it's an operational scale problem that demands a fundamentally different approach to how organizations identify, prioritize, and remediate risk.

The organizations closing this gap share a common characteristic: They've introduced automation into the remediation loop. The top performers in this dataset scan more frequently, use multiple testing methodologies, and act on findings without waiting for manual review. Their advantage comes from a different operating model, supported by better tools.

For Mythos readiness, the implication is straightforward: Machine-speed attacks require machine-speed defenses. Programs that stay rooted in manual, linear remediation workflows will watch the gap between discovery and closure grow wider every month, with exposure windows growing alongside it.

But automation doesn't mean you should remove humans from the loop, it means you should redirect human expertise toward the decisions that require judgement, like validating risk, resolving complex vulnerabilities, and making business tradeoffs. Automation absorbs the volume so people can handle the judgment calls.

Mythos readiness is about building your operational capacity so you can act on risk at the pace your software development and threat environment demand. The strongest performers in our dataset have already started this transition. They don't face less risk than their peers, they're just able to act on it faster.

HOW BLACK DUCK CAN HELP

The recommendations in this report describe the capabilities organizations need to act on vulnerabilities at the speed the AI era demands.

- A deterministic testing foundation paired with AI code analysis to find novel threats
- AI-assisted triage
- Agentic remediation

Black Duck provides the platform and intelligence to put those capabilities into practice.

A Unified View of Application Risk with Black Duck Polaris

Black Duck Polaris™ Platform brings SAST, SCA, and DAST together with a common risk and policy framework, giving you a unified view of application risk across proprietary code, open source dependencies, and runtime behavior. By correlating findings across testing methodologies, Polaris helps reduce noise, improve prioritization, and focus remediation on the vulnerabilities most likely to impact the business. It provides the consistent, auditable testing foundation that AI-assisted triage and remediation build on.

Polaris also integrates security testing directly into software development workflows and provides portfolio-level reporting on remediation performance, including closure rates, backlog growth, and exposure trends. This enables organizations to identify risk earlier, measure remediation effectiveness, and determine whether remediation efforts are keeping up with discovery.

Finding and Filtering Risk with Black Duck Signal

Signal AI Code Analysis extends deterministic testing with AI-powered code analysis that finds what rule-based tools may miss, including novel threats, risky application behavior, business-logic flaws, and AI-generated code patterns.

Large-scale application security programs often spend significant effort validating and triaging discoveries. AI SAST Triage automatically detect false positives, reducing the volume of findings your teams need to review before prioritization can begin. By filtering out false positives early, It helps teams spend less time eliminating noise and more time investigating and remediating vulnerabilities.

Agentic AppSec at Scale

Signal also extends application security workflows with agentic AI capabilities that help you analyze findings, reduce noise, accelerate remediation, and improve developer productivity. With automated routine security analysis and actionable guidance, your teams can spend less time managing findings and more time reducing risk. This is the operational scale that Mythos-era vulnerability volumes demand.

A Foundation for Continuous Risk Management

Together, Polaris and Signal pair deterministic testing with AI-powered scanning and AI-assisted triage and remediation. This combination provides the answer to what to fix first, and it lets you measure remediation performance as rigorously as discovery performance, tracking closure rates, backlog growth, and exposure windows over time. The result is a security program designed not just to find vulnerabilities but close them faster than they accumulate.

CONCLUSION

Between June 2025 and May 2026, 533 organizations identified 65.5 million vulnerabilities, including 9.2 million rated critical or high severity. Discovery outpaced remediation throughout the year, and the gap widened, from 11 to 18 new findings for every 1 closed. Only about 1 in 10 findings were closed by the end of the study—those that were closed through SAST and SCA took about a month at the median.

A month used to be fast. It isn't anymore. The time between disclosure and exploitation has fallen from years to less than a day, and frontier AI models are pushing it to minutes. Vulnerability management built around human-paced triage and quarterly patch cycles can't close that gap no matter how hard the people running it work.

Some organizations are pulling ahead, though. They've built systems to act on what they find, and their testing is wired into pipelines, with frequent scanning and findings routed straight to developers. They're focused on prioritization that puts the most serious risk first. The largest organizations close critical- and high-severity findings in 12 days, compared to 19 days for all their findings.

For every organization, the central question is the same: what to fix first. Answering it takes deterministic AST, which delivers findings you can trust, paired with AI-assisted triage that weighs exploitability and context. Then agentic remediation acts on the decision at machine speed.

Mythos readiness comes down to one capability: turning what you find into what you fix, faster than attackers can turn it into an exploit.

METHODOLOGY

The findings data reflects aggregated, anonymized telemetry from Black Duck customers. It covers 533 organizations, 13,127 applications, and 65,521,627 identified vulnerabilities collected between June 2025 and May 2026. Findings come from four testing methodologies.

- Static application security testing (SAST)
- Software composition analysis (SCA)
- Dynamic application security testing (DAST)
- Manual application security testing (manual AST), including business logic assessments

Exploit data on vulnerabilities by severity comes from Black Duck Security Advisories authored by Black Duck researchers in 2025.

The report uses four core measures.

- **Time-to-close** is the number of calendar days between a finding's initial identification and its confirmed closure. The report uses the 10th percentile (the fastest closures), the median, and the 90th percentile (long-lived findings) to show the full range of outcomes.
- **Percentage of findings closed** is the proportion of identified findings that had been closed at the time of analysis.
- **Discovery-to-closure ratio** is the number of new findings identified in a month divided by the number of findings closed that month.
- **Scan frequency** is the average number of scans per month per customer, grouped into tiers from less than monthly to 30 or more.

Time-to-close is reported separately for each methodology. The four methodologies produce very different volumes of findings, so this report doesn't combine their time-to-close statistics into a single overall figure. Severity-level time-to-close figures exclude findings that were deleted rather than remediated. Scan frequency results are averages of each customer's median time-to-close.

Severity. Findings are categorized in a five-level model.

- Critical (CVSS 9.0–10.0)
- High (7.0–8.9)
- Medium (4.0–6.9)
- Low (0.1–3.9)
- Informational (0.0)

Critical- and high-severity findings are combined in some analyses.

Segments. Company size and revenue are based on customer firmographic data.

Interpretation. Each methodology identifies different classes of vulnerabilities. Comparisons across SAST, SCA, DAST, and manual AST therefore shouldn't be read as direct measures of methodology effectiveness. Differences in volume, severity, and remediation speed often reflect the nature of the vulnerabilities found rather than the capabilities of the method. Findings from DAST and manual AST frequently require more complex remediation than those from SAST or SCA.

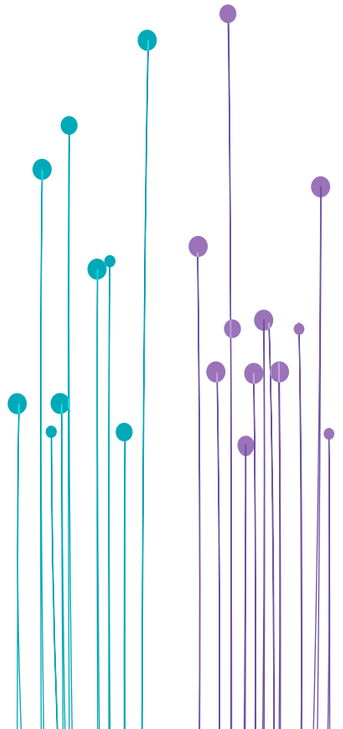
Limitations. The dataset comprises organizations that have invested in commercial application security tooling, and they may have higher baseline maturity than the broader market.

- Organizations that remediate effectively may be more likely to remain in the dataset.
- Organizations may use multiple testing tools for some methodologies. The same software may also be recorded as separate applications in different testing products.
- Monthly volumes reflect new customer onboarding as well as changes in discovery, so the report relies on ratios and percentages rather than raw volume growth where possible.
- Relationships between scan frequency, organization size, and remediation speed are correlational. They may share common drivers such as development maturity and automation investment.

External sources. Exploit timing and AI capability data come from the Zero Day Clock, Sysdig, and Anthropic. Regulatory references are drawn from PCI DSS 4.0.1, the EU Cyber Resilience Act, the Digital Operational Resilience Act, and U.S. SEC cybersecurity disclosure rules. The NVD enrichment change is from NIST.

APPENDIX A: SUMMARY STATISTICS

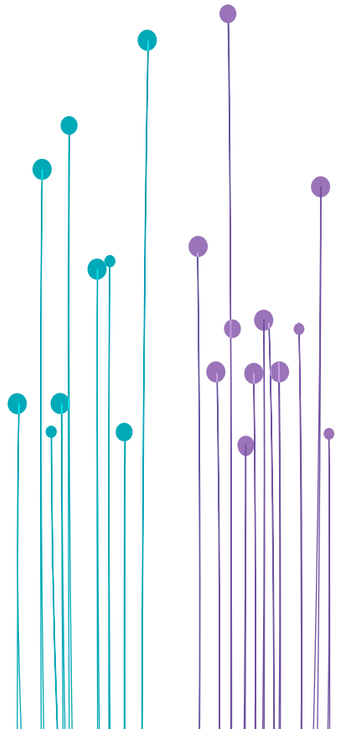
The following table summarizes key metrics for each application security testing methodology. The total counts each customer and application once, even when tested by more than one methodology. Time-to-close statistics are reported per methodology only.



Metric	Total	SAST	SCA	DAST	Manual AST
Total findings	65,521,627	46,601,777	18,881,416	35,410	3,024
Critical- and high-severity findings	9,199,730	4,425,073	4,771,242	1,809	1,606
Customers covered	533	293	226	271	68
Applications analyzed	13,127	6,521	4,223	6,383	475
Applications with findings	10,459	5,178	3,209	4,340	454
Percentage of findings closed	10.5%	8.7%	15.0%	49.2%	43.8%
10th percentile TTC (days)	—	0	0	1	34
Median TTC (days)	—	28	39	316	232
Mean TTC (days)	—	95.1	109.2	547	424.8
90th percentile TTC (days)	—	299	328	1,394	1,113
95th percentile TTC (days)	—	420	526	2,182	1,649
Maximum TTC (days)	—	1,181	1,431	5,433	2,871

APPENDIX B: MONTHLY FINDING TRENDS BY SEVERITY

The following table shows new findings identified each month by severity, alongside closures, the discovery-to-closure ratio, and the percentage of new findings rated critical or high. Monthly volumes partly reflect customer onboarding, so ratios and percentages give the clearest view of the trend.

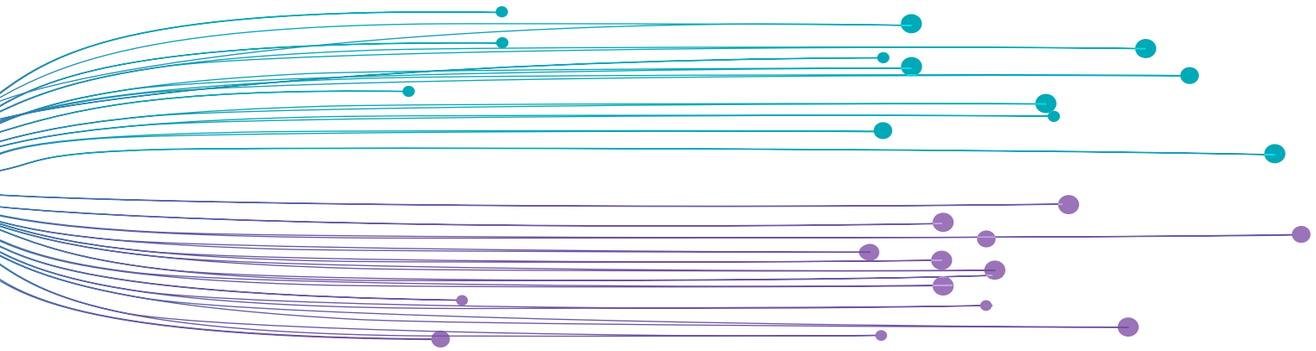


Month	Critical	High	Medium	Low	Informational	Total new	Closed	Ratio	Percentage of critical- and high-severity findings
Jun 2025	2,840	84,130	272,962	294,173	9,373	663,478	58,569	11.3:1	13.1%
Jul 2025	3,603	102,202	388,192	273,238	23,091	790,326	63,455	12.5:1	13.4%
Aug 2025	20,431	392,679	1,297,204	1,186,302	31,836	2,928,452	218,456	13.4:1	14.1%
Sep 2025	28,854	423,093	1,835,311	1,353,287	24,258	3,664,803	288,551	12.7:1	12.3%
Oct 2025	22,600	399,105	1,338,690	1,286,726	15,950	3,063,071	349,486	8.8:1	13.8%
Nov 2025	18,208	374,527	1,675,930	1,341,550	40,760	3,450,975	227,113	15.2:1	11.4%
Dec 2025	44,258	501,473	2,043,759	1,648,448	99,399	4,337,337	197,042	22.0:1	12.6%
Jan 2026	33,711	819,399	3,544,334	2,365,258	179,459	6,942,161	359,762	19.3:1	12.3%
Feb 2026	60,219	1,225,668	3,867,825	3,045,994	251,371	8,451,077	651,320	13.0:1	15.2%
Mar 2026	67,077	1,355,541	3,869,600	4,469,554	581,958	10,343,750	545,669	19.0:1	13.8%
Apr 2026	120,273	1,486,849	5,572,056	3,203,607	709,940	11,092,725	674,437	16.4:1	14.5%
May 2026	42,870	1,597,213	4,541,298	2,939,210	641,197	9,761,788	545,183	17.9:1	16.8%

Note: March 2026 includes 20 findings with no severity assigned.

APPENDIX C: TIME-TO-CLOSE PERCENTILES BY METHODOLOGY AND SEVERITY

The following table shows the 10th percentile, median, and 90th percentile time-to-close for each methodology by severity. Values are in calendar days and exclude deleted findings.

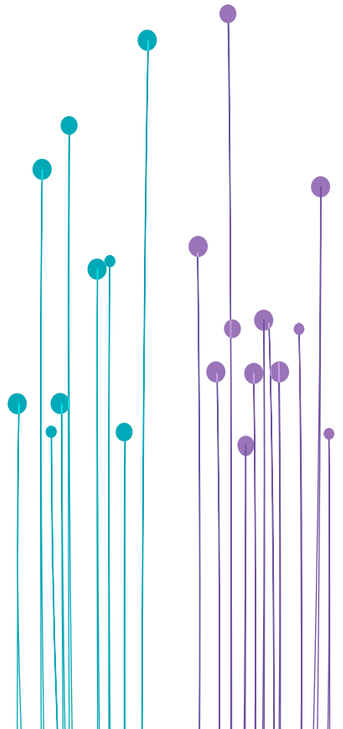


Methodology	Severity	10th	Median	90th percentile
SAST	Critical	0	0	0
SAST	High	1	56	298
SAST	Medium	1	35	336
SAST	Low	0	37	315
SAST	Informational	2	76	252
SCA	Critical	2	58	538
SCA	High	0	22	407
SCA	Medium	2	47	308
SCA	Low	0	25	230
SCA	Informational	20	20	42
DAST	Critical	7	201	1,736
DAST	High	2	420	2,469
DAST	Medium	40	552	846
DAST	Low	2	347	1,411
DAST	Informational	3	297	1,286
Manual AST	Critical	27	133	544
Manual AST	High	34	237	1,022
Manual AST	Medium	56	395	824
Manual AST	Low	63	351	917
Manual AST	Informational	126	586	1,816

Note: SAST critical findings were few, and nearly all closed within a day. This suggests automated resolution rather than remediation, deduplication, suppression, or other factors. Manual AST medium-severity values are based on eight closed findings.

APPENDIX D: COMPOSITION BY COMPANY SIZE

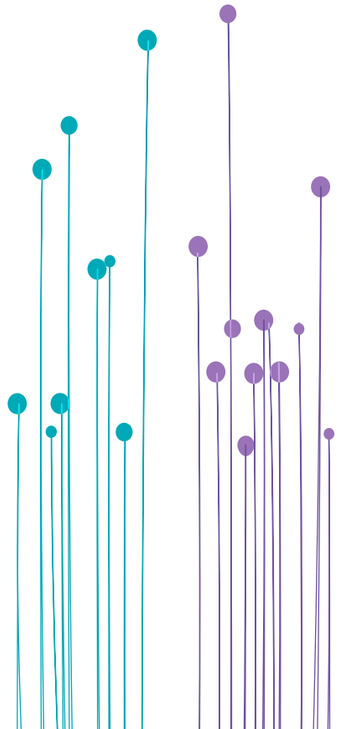
The following table shows findings and time-to-close by company size. The slowest remediation appears in midsize organizations, particularly those with 1,501 to 2,500 employees.



Company size (employees)	Customers	Applications	Total findings	Critical- and high-severity findings	Median TTC (all)	Median TTC (critical- and high-severity findings)
100 or fewer	99	789	1,575,420	266,987	17	18
101–500	109	828	3,987,713	534,699	34	22
501–1,000	46	494	4,578,300	681,372	24	28
1,001–1,500	18	235	6,929,997	533,705	34	15
1,501–2,500	44	2,395	2,603,320	239,944	56	50
2,501–5,000	45	479	13,328,031	1,324,952	21	28
5,001–10,000	31	1,420	8,260,113	1,912,656	22	17
More than 10,000	110	6,316	22,387,095	3,383,608	19	12
Unknown/ unclassified	31	171	1,840,010	312,405	27	13

APPENDIX E: COMPOSITION BY REVENUE BAND

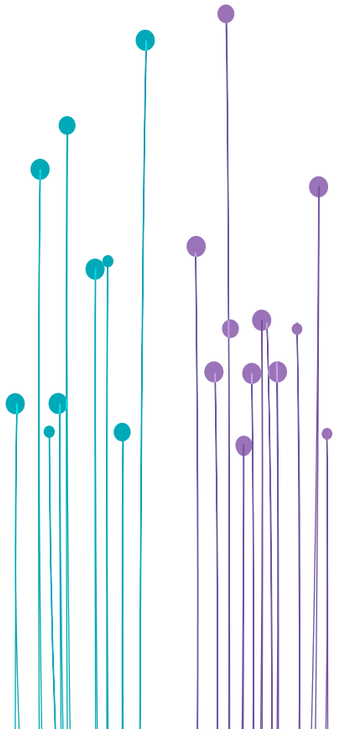
The following table shows findings and time-to-close by annual revenue band. Some bands have small samples and very low medians, so individual values should be read with caution.



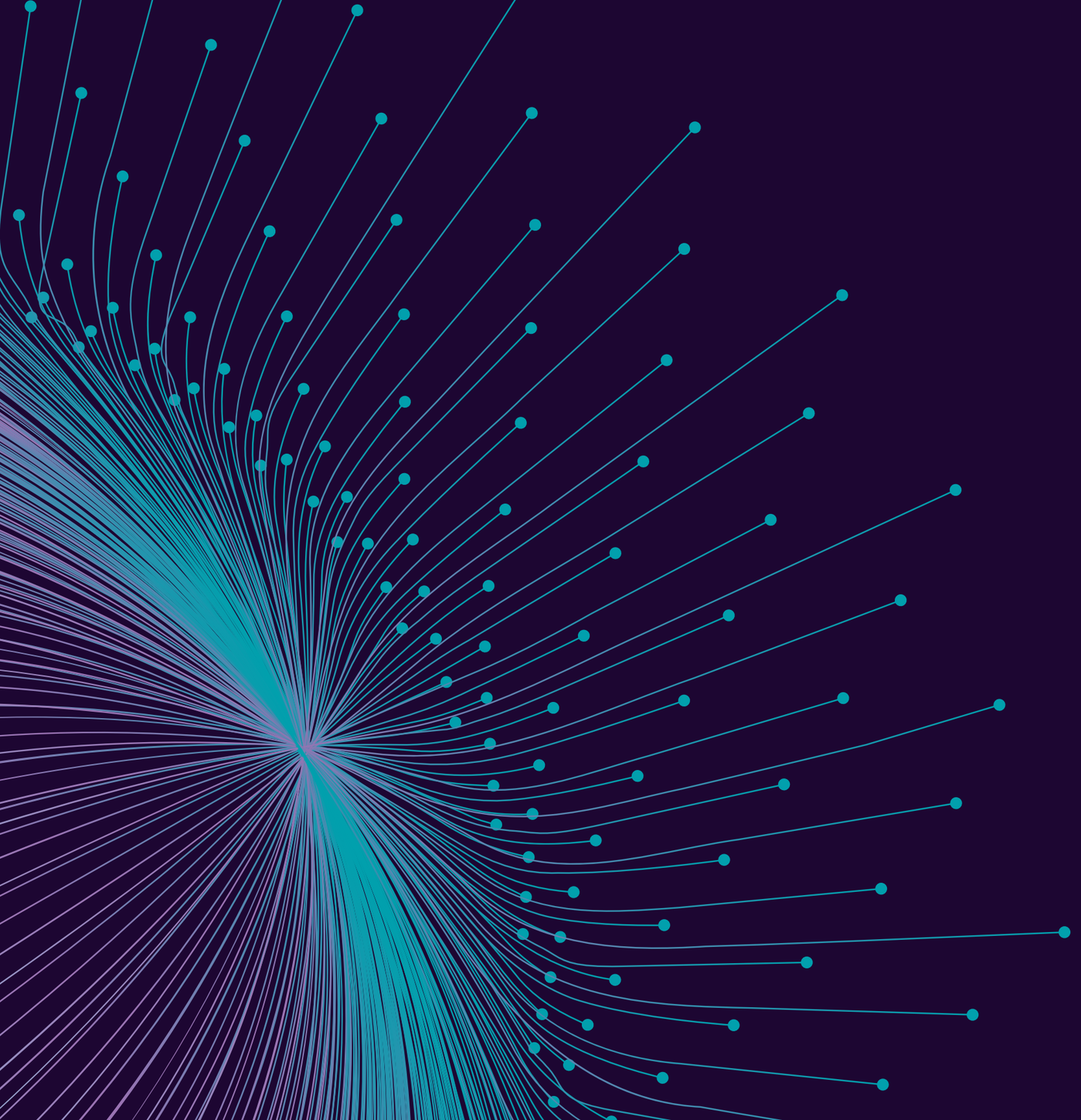
Revenue band	Customers	Applications	Total findings	Critical- and high-severity findings	Median TTC (all)	Median TTC (Critical- and high-severity findings)
$1M–$5M	30	136	1,154,102	45,372	3	2
$5M–$10M	25	149	733,332	91,128	21	41
$10M–$25M	47	519	853,050	191,545	24	26
$25M–$50M	37	176	277,725	45,880	3	0
$50M–$100M	32	470	1,375,014	298,291	45	21
$100M–$250M	59	2,424	6,130,950	888,028	25	29
$250M–$500M	40	432	7,497,403	621,760	45	42
$500M–$1B	54	725	11,947,417	1,206,704	24	23
$1B–$5B	87	2,445	15,854,654	3,020,499	16	14
Over $5B	84	5,433	17,872,562	2,475,021	22	14
Unknown/ unclassified	38	218	1,793,790	306,100	19	6

APPENDIX F: SCAN FREQUENCY VS. TIME-TO-CLOSE BY METHODOLOGY

The following table shows the relationship between scan frequency and remediation speed. Values are averages of each customer's median time-to-close, so they aren't directly comparable with the medians elsewhere in this report. Tiers with fewer than five customers are omitted. The relationship is correlational: Scan frequency and remediation speed may share common drivers, such as development maturity and automation investment.



Methodology	Scans per month	Average of customer median TTC (days)	Customers
SAST	Less than 1	49.6	11
SAST	1–3	35.0	24
SAST	4–9	52.9	26
SAST	10–29	45.3	36
SAST	30+	42.3	149
SCA	Less than 1	66.1	20
SCA	1–3	47.3	27
SCA	4–9	39.5	24
SCA	10–29	30.1	27
SCA	30+	40.5	110
DAST	30+	65.4	113
Manual AST	Less than 1	96.9	28
Manual AST	1–3	127.7	9



ABOUT BLACK DUCK

Black Duck® is application security for the AI era. Our approach is hybrid by design. Deterministic analysis delivers verifiable results and serves as the system of record for governance, while AI-powered reasoning prioritizes what's exploitable and agentic remediation fixes it at machine speed — so your team ships as fast as AI writes code, without inheriting the risk that comes with it. Trusted by thousands of organizations in many of the world's most demanding industries, including aerospace, defense, automotive, semiconductor, and healthcare, and backed by more than 20 years of human-verified vulnerability intelligence. Learn more at blackduck.com.

©2026 Black Duck Software, Inc. All rights reserved. Black Duck is a trademark of Black Duck Software, Inc. in the United States and other countries. All other names mentioned herein are trademarks or registered trademarks of their respective owners. October 2026


BLACK DUCK®