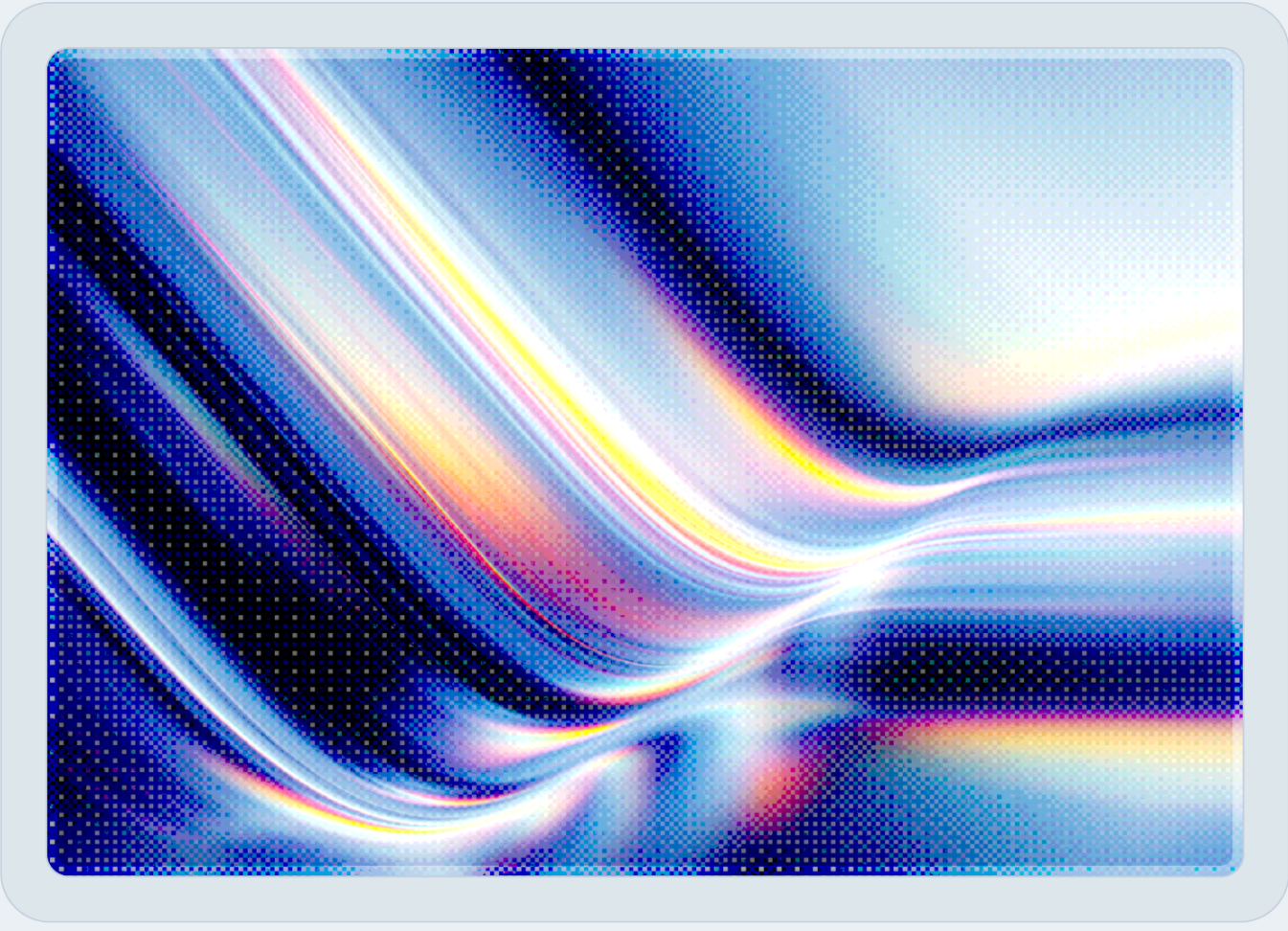
CSAI Foundation | Cloud Security Alliance

The Collapsing Patch Window: Project Zero's Full X and AI-Speed Remediation

What Fixed-Deadline Disclosure Means for Enterprise Patch Deployment

2026-11-17

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

© 2026 Cloud Security Alliance. Some rights reserved.

You may download, store, display, view, print, redistribute, and link to this document in its original, unmodified form, provided that attribution to the Cloud Security Alliance is maintained and all trademark and copyright notices remain intact.

This document may not be modified or altered. You may quote portions of the document as permitted by the Fair Use provisions of the United States Copyright Act, provided that attribution is given to the Cloud Security Alliance.

This document may be shared on professional and social media platforms in its original form with attribution.

This document was generated with AI assistance and has not undergone official CSA review and approval processes.

Key Takeaways

Google Project Zero retired its "90 + 30" disclosure policy on November 13, 2026, and replaced it with a fixed-deadline model it calls "Full X." Vulnerability details are now published a set number of days after the vendor report, whether or not a fix has shipped. That deadline is 90 days today, falls to 60 days on February 2, 2027, and falls to 30 days on January 3, 2028 [1][2]. Project Zero's stated reason is that AI-assisted discovery and exploitation have made the old timeline "clearly insufficient," and that "patch soak time is no longer meaningful" because attackers can now reverse-engineer a fix almost as soon as it ships [1].

For enterprises, the change matters less for what Project Zero publishes than for what it acknowledges. The 30-day soak period was never fully protected, since skilled attackers have reverse-engineered patches for years. What has changed is cost and speed: AI-assisted patch diffing has compressed work that once took specialists days or weeks into hours. Anthropic's Frontier Red Team reported a working Firefox exploit within about an hour of a patch becoming public, and eight Windows kernel privilege-escalation exploits produced at roughly $2,000 each [4]. Meanwhile, the median enterprise took 43 days to fully remediate known exploited vulnerabilities in the 2026 Verizon DBIR reporting period, up from 32 days the year before [6][8].

Four conclusions follow from this evidence. First, enterprises should treat a vendor's patch release, not a researcher's write-up, as the moment of public disclosure, and should measure deployment speed from that moment. Second, under Full X a vendor that ships a fix near the deadline leaves customers days, not weeks, before technical details are public, and a vendor that misses the deadline leaves them with no patch at all, so that compensating controls become the only recourse. Third, vendor severity and exploitability ratings are a weak guide to deployment urgency when AI models can produce exploits for bugs rated "Exploitation Unlikely" [4]. Finally, the workable response is a continuous remediation pipeline that combines compressed deployment rings, mitigations prepared before a patch exists, and disclosure-date tracking, informed by the risk-tiered model in CISA BOD 26-04 [8].

Background

Project Zero's disclosure policy has long set the de facto pace of coordinated vulnerability disclosure (CVD) for the largest software vendors. Since 2021, the team gave vendors up to 90 days to fix a reported vulnerability. If the fix landed within that window, it added 30 days of "soak time" before publishing technical details, so users had time to install the update [1]. In July 2025 the team added a "Reporting Transparency" trial. About a week after reporting a bug, it publicly lists the affected vendor, the product, the report date, and the deadline, without technical details [3].

The 2026 policy post, authored by Natalie Silvanovich and published November 13, 2026, describes the new model and why it was adopted [1]. Under Full X, disclosure happens X days after the vendor report. Most special cases are gone, including the grace period and the variant policy. The remaining exceptions cover bugs actively exploited in the wild, which fall under Google's broader application security disclosure policy; bugs whose details others have already published; early publication at the vendor's request; and, occasionally, vulnerabilities such as hardware flaws that require longer coordination, for which Project Zero has historically granted exceptions (as with Spectre and Meltdown) [1]. Project Zero will still track whether bugs are fixed and will say when they are not, but fix status "will not be a decision point in our policy" [1]. Table 1 summarizes the transition.

Table 1: Project Zero disclosure timeline transition

Policy phase	Effective date	Disclosure trigger	Soak time after fix
90 + 30 (legacy)	2021 to Nov 12, 2026	90 days after report, or 30 days after an in-deadline fix	Up to 30 days
Full 90	November 13, 2026	90 days after report, regardless of fix status	None
Full 60	February 2, 2027	60 days after report, regardless of fix status	None
Full 30	January 3, 2028	30 days after report, regardless of fix status	None

Sources: Project Zero [1]; BleepingComputer [2].

Project Zero gives three reasons for the change, and each bears on enterprise remediation. The first is volume: "We are finding several vulnerabilities a week in targets where we once found a few vulnerabilities a year" [1]. The second is duplicate discovery: with AI tools widely available, other parties, including malicious ones, are increasingly likely to find the same bug independently. The third is that patches now reveal their bugs almost immediately, which turns any period between patch release and installation into an exploitation window [1]. Project Zero also reported that, when it solicited vendor feedback, vendors warned that an abrupt move to 30 days could leave unfixed bugs exposed while they improved their processes, and could push them to prioritize Project Zero reports over more severe flaws reported by others; the phased schedule is Project Zero's response to those concerns [1][2].

CSA published a governance-focused analysis of Full X on November 15, 2026. It covers board briefings, supplier risk, and SLA alignment [9]. This note focuses on a narrower operational question: how enterprise patch deployment pipelines need to change when both the disclosure clock and the exploitation clock run faster than current remediation practice.

Security Analysis

The Patch Is the Disclosure

Removing soak time acknowledges a long-standing reality, that a published patch is a blueprint for exploitation, which AI has now made far cheaper to act on. Human-driven binary patch diffing and "1-day" exploitation have been documented for well over a decade; what recent research adds is evidence about speed and cost. Anthropic's Frontier Red Team, publishing in June 2026, measured how quickly frontier models could turn patches into working proofs of concept (PoCs) and exploits [4]. On 18 SpiderMonkey security patches from Firefox 148 and 149, Claude Mythos Preview produced its first working PoC in about 12 minutes and its first working exploit in just under an hour. It completed eight exploits in about 12 hours. Because the median gap between a patch landing in the public source tree and the corresponding Firefox release was 19 days, working exploits could exist more than two weeks before the fixed browser shipped [4].

The closed-source results matter more for most enterprises, whose fleets are dominated by Windows endpoints. On 21 Windows kernel local privilege-escalation vulnerabilities from January and February 2026, the same model produced crash PoCs for 18 within six hours, at roughly $2,200 in API credits. It built eight full privilege-escalation exploits for about $15,700 in total [4]. The researchers compared these timings with Windows Autopatch deployment schedules, under which roughly 90% of devices receive a patch about seven days after release and forced reboots occur around day 11. They concluded

that the exploits would likely have been finished before most devices were patched [4]. In the authors' words, "A lone operator can now turn a month's worth of patches into working exploits in a single afternoon" [4].

These results deserve appropriate weight. They come from a single study by an AI developer evaluating its own preview model on 39 selected vulnerabilities in two codebases. They indicate what frontier capability can achieve rather than what typical attackers do today, and they have not yet been independently replicated.

Incident data points in the same direction, though it measures a different phenomenon. Mandiant's M-Trends 2026 estimated the mean time-to-exploit for 2025 at -7 days, which Mandiant interprets as exploitation routinely occurring before a patch is released. Exploits remained the most common initial infection vector for the sixth consecutive year, at 32% of intrusions [5]. VulnCheck found that 23.43% of the 495 known exploited vulnerabilities it identified in the first half of 2026 had evidence of exploitation on or before the day their CVE was published, a measure taken against CVE publication rather than patch release [7]. Taken together, these findings suggest two compounding pressures: a growing share of exploited vulnerabilities are attacked before or at disclosure [5][7], and, for those that are not, laboratory evidence indicates that capable AI-assisted adversaries could weaponize a patch within hours of release [4].

How Full X Reshapes the Enterprise Window

Full X changes the enterprise exposure profile in distinct ways, depending on when the vendor ships. When a vendor fixes a bug well before the deadline, the patch itself has already exposed the vulnerability to capable adversaries, so Project Zero's publication mainly widens access to less sophisticated attackers, which still matters for organizations with long deployment tails. When a vendor ships close to the deadline, the legacy policy would have held technical details for another 30 days. Under Full X, a fix released on day 85 of a 90-day window leaves customers five days before full details are public, and that margin shrinks as X moves toward 30. When a vendor misses the deadline, enterprises face a published vulnerability with no patch, and only compensating controls stand between disclosure and exploitation.

Table 2 compares these timelines with current enterprise performance. The comparison is illustrative: the figures come from different populations and measurement methods, and their clocks start at different points, so they should not be read as directly comparable. The gap between attacker and defender timelines, however, holds across every source.

Table 2: Selected attacker and defender timelines (start points vary)

Measure	Clock starts at	Reported value	Source
AI-generated working exploit (Firefox, first exploit)	Patch release	~1 hour	Anthropic [4]
AI-generated crash PoCs for 18 Windows kernel bugs	Patch release	~6 hours	Anthropic [4]
Windows Autopatch: ~90% of devices patched	Patch release	~7 days	Anthropic, citing Autopatch timelines [4]
CISA BOD 26-04 highest-risk tier (federal)	Directive trigger	3 days	CyberScoop [8]
Median enterprise time to fully remediate KEVs	KEV listing	43 days	Verizon DBIR via Tenable and CyberScoop [6][8]
Share of KEVs fully remediated in reporting period	Reporting period	26%	Verizon DBIR via Tenable and CyberScoop [6][8]
Full X deadline (from January 2028)	Vendor report	30 days	Project Zero [1]

Severity Ratings Are a Weak Deployment Signal

Many enterprises still decide deployment urgency from vendor severity and exploitability ratings. Recent evidence suggests this practice now underestimates risk. Of the 21 Windows kernel vulnerabilities in Anthropic's study, Microsoft had rated 14 as "Exploitation Less Likely" or "Exploitation Unlikely." The model produced PoCs for 13 of those 14, including a full privilege escalation for one rated "Exploitation Unlikely" [4]. Project Zero makes a related argument: severity describes user impact, not the likelihood that an attacker will find the bug, so it is a poor basis for disclosure timing [1]. Neither source shows that vendor ratings are wrong in general. Both indicate that exploitability assessments made before AI-assisted exploitation should be treated as lower bounds on risk, not reliable predictors.

Volume Pressure on Remediation Capacity

Volume adds a further pressure. When a single research team reports finding several vulnerabilities a week in targets that once yielded a few a year [1], and other teams with similar tooling may be seeing comparable increases, vendors respond with more frequent and larger security releases. Anthropic notes that Mozilla has moved Firefox dot releases from a monthly to a roughly weekly cadence [4]. CSA's August 2026 research note documented record patch volumes from major vendors linked to AI-assisted discovery [11], and CSA's analysis of the resulting patch debt examines how that backlog accumulates across enterprise environments [14]. Enterprise change management built around monthly maintenance windows was designed for a lower patch arrival rate. The 2026 DBIR's lengthening time to remediate KEVs (32 to 43 days), a fully remediated share that fell from 38% to 26%, and a reported increase of nearly 50% in the number of KEVs organizations had to patch are consistent with remediation capacity already under strain [6][8]. More patches arriving under shorter disclosure deadlines will likely widen this gap unless deployment becomes more automated.

Regulatory Convergence

The policy landscape is moving in the same direction. CISA's Binding Operational Directive 26-04, issued June 10, 2026, introduced a risk-based remediation model that formally supersedes BOD 22-01 and its fixed KEV deadlines [8][12]. Vulnerabilities on publicly exposed assets that allow fully automated exploitation and system takeover, and that show evidence of active exploitation, must be remediated within three days, with forensic triage to check for compromise [8]. CISA cited AI's compression of the time from discovery to weaponization as part of the rationale [8]. Federal agencies are the only parties bound by the directive, but taken together with Full X it indicates that both researchers and regulators are converging on remediation timelines measured in days.

Recommendations

Immediate Actions

Security teams should first change how they define the start of the remediation clock. For internet-facing systems and widely deployed endpoints, the clock should start at vendor patch release, not at public advisory, researcher publication, or KEV listing. The evidence from AI-assisted patch diffing indicates that capable adversaries can begin work at that moment [4]. Teams should also measure their current

time from patch release to 90% deployment coverage for each major platform. Many organizations track mean time to remediate rather than deployment coverage curves, yet the coverage curve is what determines exposure during the first week.

Building on that baseline, teams should start tracking disclosure dates as an input to scheduling. Project Zero's Reporting Transparency list publishes the vendor, product, and deadline for each report [3]. Matching those entries against the asset inventory and software bills of materials (SBOMs) lets teams anticipate which products may face disclosure without a patch, and prepare mitigations before the deadline. Finally, teams should pre-approve an emergency change path for the highest-risk tier so that a three-day remediation target is achievable without convening a change advisory board for each patch.

Short-Term Mitigations

Before the February 2, 2027 move to Full 60, about 11 weeks from publication, organizations should begin compressing their staged deployment rings, with the goal of completing the change before Full 30 takes effect in January 2028. Patch rings were designed to catch regressions through soak time, and that soak time now has a security cost. A practical pattern is a canary ring measured in hours, backed by automated functional and regression testing, followed by broad deployment within days for internet-facing and high-exposure assets. Hotpatching, where the platform supports it, reduces the reboot dependency that delays coverage in Autopatch-style schedules, where forced reboots arrive around day 11 [4].

Ring compression addresses the case in which a patch exists; organizations also need a plan for the case in which disclosure arrives first. Mitigation-first playbooks should identify, for each critical product family, the controls available without a vendor fix: feature flags or component disablement, network segmentation and access restriction, WAF or IPS virtual patching, egress controls, and targeted detection rules. Exploit-class mitigations such as Control Flow Guard and hardware-enforced shadow stacks raise the cost of exploitation across whole bug classes and are worth enabling broadly [4].

Prioritization logic should be updated in parallel so that vendor exploitability ratings are no longer a sole input. Prioritization should combine exposure, exploitability (including whether exploitation can be automated), evidence of active exploitation, technical impact, asset criticality, and, newly, the disclosure deadline. The first four correspond to the criteria in BOD 26-04 [8]; asset criticality and disclosure deadline are enterprise additions. This structure is also consistent with the exploitability-first model in CSA's remediation guidance [10].

Strategic Considerations

Over the longer term, the trends described here favor a continuous remediation architecture over periodic patch cycles. CSA's *The AI-Driven Patch Wave* describes such an architecture: exploitability-first prioritization, continuous validation, tiered AI-assisted patch generation, and formal compensating controls for systems that cannot be patched quickly [10]. Defenders can apply the same AI capabilities attackers use. Running AI-assisted analysis against incoming vendor patches can identify which fixes are most readily weaponizable and should be deployed first. This is an inference from the offensive research, not an established practice, and organizations should validate it before relying on it.

Procurement and supplier governance should treat vendor remediation speed as a measurable attribute. Relevant evidence includes whether vendors ship fixes inside disclosure deadlines, whether they offer interim mitigations for unpatched issues, and how quickly their update channels deliver fixes. Organizations should also reduce the patch burden itself through software minimization, retirement of unsupported components, and preference for memory-safe implementations, which Anthropic identifies as a way to reduce the supply of exploitable bugs [4]. Finally, organizations should assume that Project Zero will not remain the only researcher operating on fixed AI-era deadlines. Other research teams and AI-driven discovery platforms may adopt similar models, though none had publicly done so as of this writing.

CSA Resource Alignment

The findings in this note build on a sequence of CSA analyses of AI-accelerated vulnerability discovery and remediation. The most directly related is CSA's research note *Project Zero's Full X Policy: Governing for a 30-Day Disclosure World* (November 2026) [9], which addresses the governance side of the same policy change: board communication, supplier risk questions, and SLA realignment. The present note complements it with an operational focus on deployment pipelines, ring compression, and pre-patch mitigation. The two should be read together.

CSA's *The AI-Driven Patch Wave* (May 2026) [10] supplies the target-state architecture that the Strategic Considerations above reference. Its argument that periodic remediation cannot keep pace with AI-accelerated discovery is consistent with Project Zero's later conclusion that soak time no longer protects users, and its exploitability-first prioritization and compensating-control requirements correspond to the short-term mitigations recommended here. CSA's *AI Vulnerability Discovery Is Outpacing Patch Capacity* (August 2026) [11] documents the vendor patch-volume increases that make the volume pressure described in this note concrete, while *AI-Accelerated Vulnerability Discovery and the Patch Debt Crisis* (June

2026) [14] examines the remediation backlog that results. CSA's research note on CISA BOD 26-04 [12] analyzes the risk-tiered federal remediation model that this note recommends as a reference structure for enterprise prioritization.

At the control-framework level, the CSA AI Controls Matrix (AICM) v1.1 [13] provides the auditable baseline. Its Threat and Vulnerability Management (TVM) domain covers remediation timelines, prioritization, and compensating controls. Its Application and Interface Security (AIS) domain addresses secure development practices that reduce the inflow of vulnerabilities. Organizations updating remediation SLAs in response to Full X should document the changes against the relevant AICM TVM controls so the new timelines carry through to assurance and audit processes.

References

[1] Natalie Silvanovich, Google Project Zero. "Policy and Disclosure: 2026 Edition." Project Zero, November 13, 2026.

[2] Bill Toulas. "Google Project Zero cuts bug disclosures to 30 days amid AI threats." BleepingComputer, November 16, 2026.

[3] Tim Willis, Google Project Zero. "Policy and Disclosure: 2025 Edition." Project Zero, July 29, 2025.

[4] Anthropic Frontier Red Team (W. Xiao, T. Abbott, N. Carlini, et al.). "Measuring LLMs' impact on N-day exploits." Anthropic, June 8, 2026.

[5] Jurgen Kutscher, Mandiant. "M-Trends 2026: Data, Insights, and Strategies From the Frontlines." Google Cloud Blog, March 23, 2026.

[6] Scott Caveza, Tenable. "Key findings from the Verizon DBIR 2026: Slower vulnerability remediation meets faster exploitation." Tenable Blog, May 19, 2026.

[7] Patrick Garrity, VulnCheck. "VulnCheck State of Exploitation 1H-2026." VulnCheck, July 28, 2026.

[8] Tim Starks. "CISA directive orders agencies to prioritize vulnerability patching in a new way." CyberScoop, June 10, 2026.

[9] Cloud Security Alliance AI Safety Initiative. "Project Zero's Full X Policy: Governing for a 30-Day Disclosure World." Cloud Security Alliance, November 15, 2026.

[10] Cloud Security Alliance AI Safety Initiative. "The AI-Driven Patch Wave." Cloud Security Alliance, May 10, 2026.

[11] Cloud Security Alliance AI Safety Initiative. "AI Vulnerability Discovery Is Outpacing Patch Capacity." Cloud Security Alliance, August 2, 2026.

[12] Cloud Security Alliance AI Safety Initiative. "CISA's BOD 26-04: A Risk-Based Reset of Patch Rules." Cloud Security Alliance, August 20, 2026.

[13] Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." Cloud Security Alliance, June 22, 2026.

[14] Cloud Security Alliance AI Safety Initiative. "AI-Accelerated Vulnerability Discovery and the Patch Debt Crisis." Cloud Security Alliance, June 17, 2026.

© 2026 Cloud Security Alliance. AI-assisted Rapid Research.