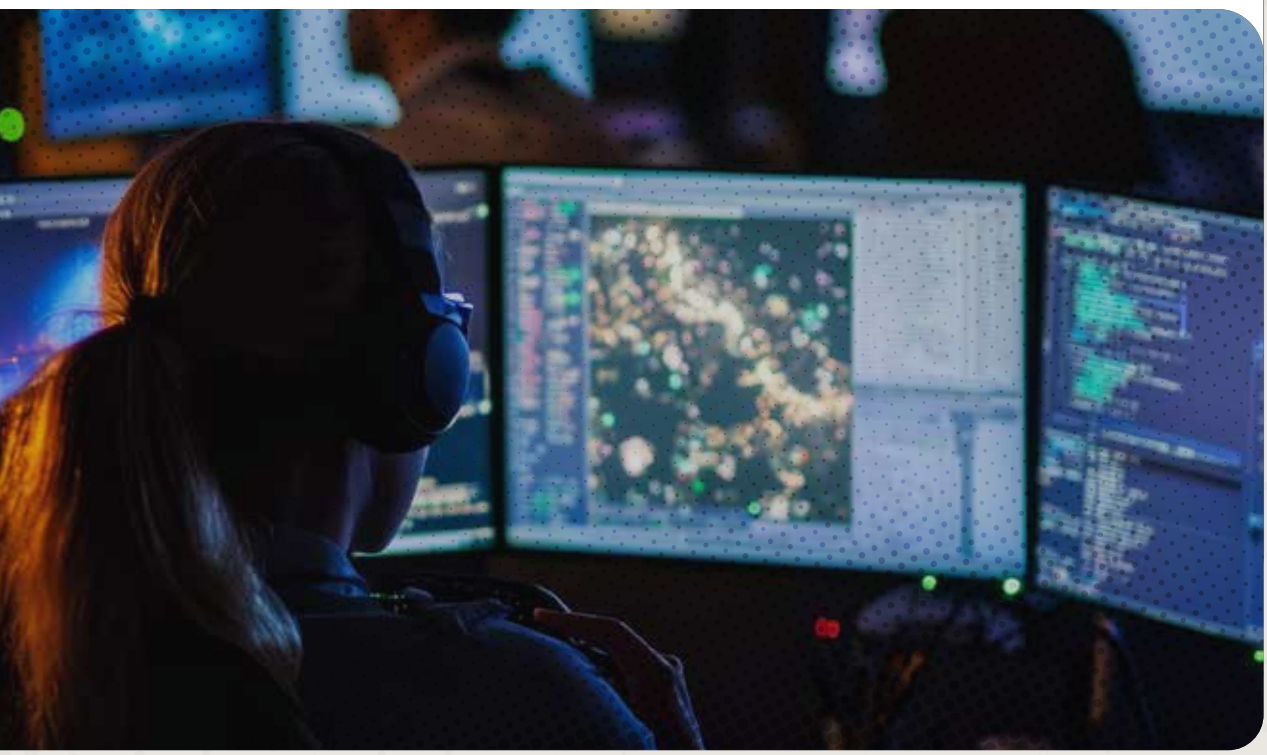
White Paper

halcyon

From Discovery to Eviction

A SecOps Practitioner's View of Project Glasswing and the Halcyon Imperative

BY JORDAN JEFFER | HALCYON FIELD CISO



Introduction

When I read the Glasswing announcement, my first reaction was not awe. It was operational dread. Project Glasswing and the Claude Mythos Preview model represent a legitimate, non-incremental shift in how vulnerabilities are discovered, and by extension, how quickly they can be exploited. Per Anthropic's disclosure, pre-release testing surfaced thousands of previously unknown vulnerabilities across major operating systems and browsers, including flaws more than a decade old that had survived extensive prior testing, some chainable to achieve full system compromise from a low-privilege starting point. That is not a research result. It is a preview of what the threat environment will look like when equivalent capabilities reach adversarial hands.

This paper is written from the ground level, from the perspective of someone running a Security Operations Center (SOC) and an Attack Surface Management program day to day. It is not a strategic overview for the boardroom. It is a practitioner's assessment of what Glasswing changes functionally, what it exposes about the structural limits of traditional vulnerability management, and why the window between vulnerability identification and active exploitation is the most dangerous gap in modern enterprise security.

The central argument is this: **Glasswing and Halcyon's anti-ransomware platform are not competing approaches to the same problem, they are complementary responses to adjacent layers of the same threat lifecycle.** Glasswing works upstream, identifying vulnerabilities at machine speed. Halcyon works downstream, defending the environment when those vulnerabilities are exploited before they are closed. The gap between those two layers is where the risk lives, and it is a gap that no single tool can address alone.

The Vulnerability Management Lifecycle Glasswing Exposes

Every practitioner who has run a vulnerability management program understands its core logic: find the vulnerability, prioritize it, apply the manufacturer's patch in a controlled and sequenced manner, and then investigate the affected asset to determine whether it was compromised during the window of exposure. That four-step process–identify, prioritize, remediate, investigate–has been the procedural spine of vulnerability management for years. What Glasswing changes is not the process. It is the volume, velocity, and severity of what that process must absorb.

In a traditional environment, a skilled penetration tester or internal red team might identify a handful of meaningful vulnerabilities in a targeted assessment. The cycle time from discovery to patch deployment, even in a well-run program, can run from days to weeks. Per Anthropic's Glasswing disclosure, the model surfaced thousands of previously unknown vulnerabilities in a single pass, some more than a decade old, others chainable to full system compromise. It did not just find needles in haystacks. It found that some of the haystacks were on fire.

For a practitioner running an attack surface management program, the immediate implication is stark: the prioritization step becomes the crisis. When the vulnerability queue is measured in the thousands, triage is no longer a process, it is a war of attrition. Traditional CVSS scoring and asset criticality frameworks were not designed for this volume. They were designed for an environment where vulnerability discovery was bottlenecked by human analysis time. Remove that bottleneck with an AI model, and every assumption underneath your prioritization model needs to be reassessed.

The exposure window closes. The investigation opens. Confirming whether a vulnerable asset was accessed or compromised is where response burden compounds fastest, and where the argument for Halcyon stops being an argument and starts being self-evident. Consider the math: if a Glasswing-scale disclosure surfaces vulnerabilities across ten thousand assets simultaneously, and your team has a 48-hour window before active exploitation begins, you have roughly 17 seconds per asset for forensic triage. That is not an investigation. That is a guess. The forensic obligation does not scale with human headcount. It requires a platform that monitors asset behavior continuously and surfaces active compromise in real time so that investigation is triggered by evidence, not conducted as a blind sweep across thousands of endpoints.

If a Glasswing-scale disclosure surfaces vulnerabilities across ten thousand assets simultaneously, and your team has a 48-hour window before active exploitation begins, you have roughly 17 seconds per asset for forensic triage. That is not an investigation. That is a guess.

The Patch Gap Is Not a Failure of Process. It Is a Structural Reality.

Per Anthropic's Glasswing disclosure, fewer than one percent of the vulnerabilities identified had been patched at the time of the initiative's announcement. That number is not an indictment of the organizations involved. It is a reflection of the structural reality every practitioner already knows: the patch ecosystem was not built for the volume of vulnerabilities that AI-driven discovery will generate.

Legacy systems are a major contributor. In most enterprise environments, there are assets that cannot be patched–not because the patch does not exist, but because the asset is too old, too embedded in day-to-day workflows, or running software for which the vendor no longer provides updates. Unmanaged devices, operational technology environments, and shadow IT assets compound this further. Attack surface management programs exist precisely because organizations often lack complete visibility into what they are running, much less a reliable mechanism to patch all of it on demand.

An optimistic counterargument holds that Glasswing doesn't just surface vulnerabilities; it creates an incentive structure and coalition mechanism that accelerates patching. If the twelve-member coalition includes the infrastructure and development partners needed to deploy patches at scale, the window between discovery and remediation might narrow relative to today's baseline. That is true, and it is the right aspiration. But current industry data on mean-time-to-patch, which consistently runs weeks to months even for critical vulnerabilities, suggests that even an optimistic patching scenario leaves a meaningful window. It narrows. It does not close.

Prevention remains essential. Patching remains essential. But the practitioner who builds their entire posture around the assumption that patches will be applied before exploitation occurs is building on a foundation that Glasswing has just demonstrated is structurally compromised.

Mean-time-to-patch consistently runs weeks to months even for critical vulnerabilities. The window narrows. It does not close.

halcyon

Where Halcyon Operates: Inside the Gap

Halcyon's anti-ransomware platform is built on a premise that every experienced SOC operator recognizes as correct: attackers will get in. The question is what happens when they do. Halcyon operates in the space after initial access–the phase of a ransomware attack where the adversary is moving laterally, escalating privileges, staging data for exfiltration, and preparing to deploy encryption. It is precisely in this phase that most conventional endpoint detection tools fail to interrupt the attack before meaningful damage is done.

A technically literate practitioner will immediately ask the right question: why Halcyon specifically, and not CrowdStrike Falcon, SentinelOne, or Microsoft Defender for Endpoint, all of which have behavioral detection, lateral movement alerting, and some form of ransomware-specific response? The answer lies in architectural specificity. Halcyon is not a general-purpose EDR with ransomware detection added as a module. It is a platform built from the ground up around a single threat category. Its eviction workflows are purpose-designed for attacker removal, not generalized remediation playbooks. Its DXP module is a dedicated data exfiltration defense capability, not a DLP bolt-on, with behavioral models trained specifically on ransomware actor tradecraft around data staging and theft. The difference between a generalist platform with ransomware features and a ransomware-specific platform with broader integrations is precisely the difference that matters when Glasswing-scale vulnerability disclosures are putting thousands of assets at simultaneous risk.

The DXP module addresses the second critical layer of modern ransomware operations. Industry reporting from Coveware and Palo Alto Networks Unit 42 consistently documents double extortion as the dominant ransomware model: adversaries steal data before encrypting it and threaten to publish unless the ransom is paid. This means that even organizations with robust backup and recovery capabilities face the extortion threat independently of their ability to restore from backup. DXP detects the behavioral signatures of data staging and exfiltration–large-volume data movement to anomalous destinations, compression and packaging consistent with preparation for theft–and interrupts that process before data leaves the environment.

Halcyon is not a general-purpose EDR with ransomware detection added as a module. It is a platform built from the ground up around a single threat category.

One tension worth acknowledging directly: active attacker eviction, while functionally appealing, carries a forensic risk that mature practitioners will recognize immediately. Evicting a threat actor before forensic capture is complete can eliminate the artifacts needed to understand full breach scope, identify patient zero, or satisfy regulatory reporting requirements. Halcyon's eviction model is most effective when sequenced alongside forensic capture, not as a replacement for it. The practical resolution for a SOC is to treat eviction and forensic documentation as parallel workstreams rather than sequential ones, using Halcyon's continuous monitoring data as the forensic record while the eviction workflow runs. This is a tension the SOC must manage actively, not one Halcyon eliminates entirely.

Evicting a threat actor before forensic capture is complete can eliminate critical artifacts. Eviction and forensic documentation must run in parallel.

The Complementary Architecture

Glasswing Upstream, Halcyon Downstream

The clearest way to describe the practical relationship between Glasswing and Halcyon is through the vulnerability lifecycle: find, prioritize, patch, investigate. Glasswing transforms the first step at unprecedented scale. But it does not change the structural constraints that govern how quickly an organization can move through prioritization, remediation, and investigation across thousands of assets simultaneously.

Halcyon addresses those constraints by operating in parallel, not in sequence. Rather than waiting for the patch cycle to complete before declaring an asset safe, it assumes that some percentage of assets in every environment will be compromised through vulnerabilities that were identified but not yet remediated. It monitors those assets continuously, looking for the behavioral indicators of active ransomware tradecraft, and triggers eviction and containment when those indicators are detected.

The following scenario illustrates what this looks like in practice:

Scenario: Glasswing Disclosure — Hour Zero

A Glasswing-equivalent scan surfaces a critical vulnerability in a widely deployed software library. The vulnerability management team begins triage, assessing asset criticality, patch availability, and deployment sequencing across thousands of affected endpoints. While that process is underway, an adversary who has independently identified the same vulnerability establishes initial access on an unpatched asset and begins moving laterally toward sensitive data.

Halcyon detects the lateral movement, triggers an eviction workflow, and simultaneously engages DXP to block any data staging or exfiltration attempts. The forensic record is captured in parallel. The incident response team receives an alert with active context–asset, behavior, timeline–rather than a breach notification after the fact. The asset is contained. The data does not leave.

This is the gap Halcyon fills: not the vulnerability, but the exposure window between its discovery and its closure.

Operational Implications and Decision Framework

For practitioners running SecOps programs or attack surface management functions, the practical implications of Glasswing-class tooling and Halcyon's response model play out across four dimensions. The table below maps each dimension to the capabilities that address it and the sequencing priority a program should follow.

Dimension	Priority	Glasswing Addresses	Halcyon Addresses
Asset Inventory & Discovery	Immediate	Surfaces unknown assets with vulnerabilities	Identifies assets where ransomware behavior is active
Vulnerability Prioritization	Immediate	Generates threat-informed prioritization at scale	Flags assets showing active exploitation behavior, jumps the patch queue
IR Capacity & Eviction	Immediate	N/A–not an incident response tool	Automates containment and eviction; frees responders for forensics
Data Exfiltration Defense & Discovery	Immediate	N/A–vulnerability discovery only	DXP detects and interrupts data staging before exfiltration occurs
Governance & Regulatory Alignment	Immediate	Provides disclosure and sourcing for regulatory reporting on vulnerability exposure	Provides incident documentation, eviction logs, and DXP activity for breach reporting

Asset Inventory and Continuous Discovery

You cannot defend what you cannot see, and you cannot patch what you do not know exists. Glasswing's ability to identify vulnerabilities in obscure or long-running software components makes complete asset inventory more critical than ever. Attack surface management programs relying on periodic scanning or agent-based coverage will have visibility gaps that adversaries armed with AI-driven vulnerability discovery will find before you do. Continuous, comprehensive asset discovery is no longer aspirational. It is a prerequisite.

Prioritization at Scale

CVSS scores alone are insufficient when vulnerability volume is measured in the thousands. A high-CVSS finding in an internet-isolated legacy system is categorically different from a medium-CVSS finding in an internet-facing asset handling sensitive data. Prioritization frameworks must incorporate asset context, exposure surface, data sensitivity, and real-time threat intelligence about active exploitation. Halcyon's continuous behavioral monitoring provides a critical input: assets showing indicators consistent with active exploitation should immediately jump the patch queue regardless of baseline CVSS score.

Incident Response Capacity

The most immediate staffing implication of a Glasswing-scale vulnerability wave is surge demand on incident response capacity. If thousands of vulnerabilities are disclosed in a compressed timeframe, and a meaningful percentage are actively exploited before patches can be applied, the volume of incidents requiring investigation will exceed the capacity of most internal SOC teams. Halcyon's eviction and DXP capabilities extend effective IR capacity by automating the most time-intensive components of active containment, freeing responders for the analytical work that requires human judgment.

Governance and Regulatory Alignment

Regulated industries face additional pressure in a Glasswing environment. Regulatory expectations around incident notification timelines, third-party risk management, and patch cycle documentation were calibrated to a pre-AI vulnerability discovery environment. Practitioners in regulated industries should begin aligning their incident response playbooks, disclosure documentation, and third-party contractual requirements to anticipated regulatory evolution now, using Glasswing disclosure materials and Halcyon incident logs as the evidentiary foundation for that alignment.

The gap between vulnerability discovery and remediation is where the risk lives, and it is a gap no single tool can address alone.

The Natural Partnership and the Coalition Gap

Glasswing's twelve-member coalition currently covers infrastructure, development, and financial services: AWS, Apple, Google, Microsoft, Nvidia, JPMorgan Chase, and the Linux Foundation among others. What it conspicuously does not include is a purpose-built ransomware response capability. That absence is not an oversight; the coalition's initial mandate is vulnerability discovery and responsible disclosure, not downstream incident response. But as Glasswing's findings create a sustained surge of disclosure obligations and patch gap exposure, the logical next phase of the initiative is resilience and recovery, not just discovery.

That is the specific institutional gap Halcyon fills in the Glasswing ecosystem. While CrowdStrike and Microsoft already provide detectionat scale within the coalition's existing membership,Halcyon provides ransomware-specific eviction and data exfiltration defense, purpose-built for exactly the threat scenario that Glasswing's vulnerability disclosures create: a known, unpatched exposure window during which ransomware actors move toward data before encryption. The coalition's infrastructure partners can accelerate patching. They cannot evict an attacker who is already inside, or intercept data staging that is already underway. Halcyon can.

The structural complementarity is real regardless of whether a formal business relationship exists between Anthropic and Halcyon today. For practitioners and organizations deploying both, the architecture is clear: Glasswing upstream to find and disclose at scale, Halcyon downstream to defend the environment during the remediation window. For Glasswing's coalition as it evolves toward full-lifecycle resilience, the gap in its current membership is equally clear. It is the gap between discovering that the haystacks are on fire and having the capability to put them out before the data burns.

Glasswing finds the exposure. Halcyon defends the environment while the patch window is open. The coalition covers discovery. It does not yet cover what happens when an attacker is already inside.

Conclusion

Project Glasswing represents a threshold event in cybersecurity, not because it introduces a new category of threat, but because it accelerates an existing category to a velocity that existing frameworks cannot absorb without significant adaptation. For practitioners running SecOps programs and attack surface management functions, the implications are immediate: more vulnerabilities, faster exploitation timelines, greater forensic obligations, and a structural gap between discovery and remediation that adversaries will increasingly exploit with their own AI-driven tooling.

The vulnerability management lifecycle–find, prioritize, patch, investigate–remains the correct framework. What changes is the scale and speed at which each step must execute, and the recognition that investigation cannot wait for remediation to complete. Some percentage of vulnerable assets in every environment will be compromised before they are patched. Halcyon's platform, combining continuous behavioral monitoring, ransomware-specific eviction, forensic-aware containment sequencing, and dedicated data exfiltration defense, is purpose-built for that reality in ways that general-purpose EDR platforms are not.

What should you do differently starting Monday? Audit your patch gap against your active behavioral monitoring coverage. For every asset that Glasswing-class tooling would flag as vulnerable and unpatched, ask whether you have a platform watching that asset's behavior in real time and capable of triggering eviction before data leaves. If the answer is no, that is the gap. It is not a product feature gap; it is a strategic exposure. And in a world where AI has permanently altered the velocity of vulnerability discovery, closing it is not optional.

About the Author

Jordan Jeffer is a cybersecurity executive and operator who has built, led, and transformed security programs at scale across financial services and critical national infrastructure. Currently serving as Field CISO at Halcyon, he partners with CISOs and security leadership teams to architect defensible programs against today's most consequential threats, including ransomware and advanced persistent adversaries.

Sources: Anthropic Project Glasswing announcement materials (April 2026); Coveware Quarterly Ransomware Report; Palo Alto Unit 42 Ransomware and Extortion Report; Halcyon.ai platform documentation.

halcyon

Halcyon Operates Where Prevention Ends

Project Glasswing changed the math. When AI can surface thousands of previously unknown vulnerabilities in a single pass (some more than a decade old, some chainable to full system compromise) the patch gap stops being a process problem and starts being a structural reality. Attackers will find unpatched assets before remediation completes. The question is what happens when they get in.

Halcyon is the first anti-ransomware platform purpose-built for that response. A platform designed from the ground up to stop ransomware at the execution layer, capture encryption key material in real time, and recover operations in hours, not weeks, not after a ransom is paid, not after backups that no longer exist are restored. Data Exfiltration Protection intercepts bulk data movement before stolen files become extortion leverage. The Ransomware Operations Center (ROC) watches your environment around the clock at no additional cost. And the Halcyon Ransomware Warranty puts recovery SLAs in writing–120-minute response, 12-hour recovery initiation, guaranteed.

In a world where AI has permanently accelerated vulnerability discovery, the exposure window between detection and remediation is where ransomware actors operate. Halcyon closes that window. Glasswing finds the vulnerability. Halcyon defends the environment while the clock is still running.

halcyon.ai/demo | **halcyon.ai/resources**