Mythos Emphasizes Why Continuous Hardening Is Critical

by Tom Stewart
Senior Director, Technology Consulting

protiviti®
Global Business Consulting

Anthropic's Mythos is changing the constraints cybersecurity defenders have come to rely on — and most organizations are not prepared for what comes next.

For a long time, one thing kept cyber attacks from scaling too far: there simply weren't enough highly skilled people to do the work. Finding vulnerabilities, turning them into exploits and chaining them together takes real expertise and time. Mythos doesn't eliminate that constraint, but it does start to loosen it — and that's where things get interesting.

The offensive security lifecycle hasn't changed. It still comes down to three steps:

- **Discovery:** Finding flaws in software or configuration
- **Weaponization:** Turning those flaws into something usable
- **Chaining:** Linking low- or medium-severity issues into an attack path that can reach something meaningful

Mythos is not changing that process; however, the same work can now happen faster, across more targets, and with far less dependence on human effort.

Mythos doesn't suddenly make this work cheap. Running it at scale is still expensive.

In early testing, roughly $20,000 in tokens was spent identifying what turned out to be a denial-of-service issue in a legacy BSD system — about what you would expect to pay a human researcher for similar work. The most significant finding was a null-pointer reference, which rarely leads to remote code execution. In most cases, the impact is limited to service disruption rather than full system compromise.

That's an important point. Mythos isn't yet producing inexpensive, high-impact exploits. The fundamentals of offensive security haven't changed, and expertise still matters. What's more, the shift, although subtle has become far more consequential. Portions of discovery, weaponization and chaining that once scaled with analyst time can now run in parallel on compute. Attacker reach is no longer constrained primarily by headcount. It is increasingly constrained by budget. For defenders, that's the real shift. It's not just that attacks

Attacker reach is no longer constrained by headcount — it is increasingly constrained by budget.

are faster. It's that the pool of potential targets is expanding, changing how one thinks about who is at risk — and how often.

A workload measured in dollars per target behaves very differently from one measured in hours. When time is the constraint, scale is slow. When budget is the constraint, scale can happen all at once. That means attackers can spread compute across thousands of targets simultaneously instead of working through them one by one. More organizations move into scope — not gradually, but all at once. For defenders, that is the tension: the work has not changed — but the scale at which it can be executed has.

Attention was the bottleneck

For most organizations, vulnerability management has relied on a simple model: sort findings by severity and treat them as separate issues. It's never been perfect — everyone understands that vulnerabilities can combine in practice — but it has been "good enough."

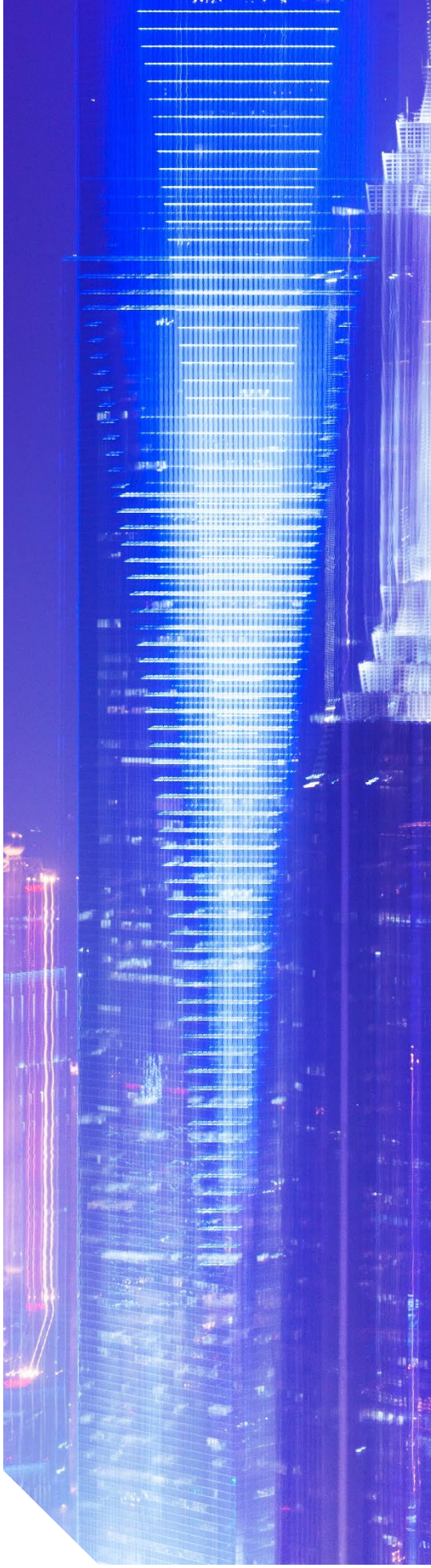
That's because, until recently, there was a real constraint on attackers: attention. Chaining vulnerabilities into a meaningful attack path takes time and focus, and attackers had to ration that effort. They simply could not pursue every possible path across every environment.

That's starting to change.

As more of the repetitive work involved in chain analysis becomes easier to run in parallel, that constraint begins to weaken. The result is a shift in how risk actually plays out.

A medium-severity finding that sits on a short path to a domain controller is not the same risk as one that leads nowhere. It never has been — but now that difference matters more, because it is easier to identify and act on those paths at scale.

And the tools to do that are becoming more accessible — not just to defenders, but to attackers as well. The cost structure is beginning to look similar on both sides.



The window to exploit is shrinking

The gap between a vulnerability being disclosed and a working exploit appearing has been narrowing for years — and it shows no signs of slowing down.

What used to take weeks now often happens in days. For high-value targets, it can happen in hours. As automated weaponization improves, that timeline will continue to compress — especially where the target justifies the compute spend.

Not every organization will feel this shift immediately. But the divergence is already clear:

- **High-value targets:** Faster, more automated exploitation — and increasing attack volume
- **Lower-value targets:** Relatively stable attacker behavior, at least in the near term

The bigger issue is that defenders haven't kept pace.

Most organizations are still operating on timelines that look like this:

- 30-day patch SLAs remain common
- Quarterly penetration tests deliver insights that are already weeks out of date
- Annual assessments reflect environments that may have changed significantly before testing even concludes

That creates a widening gap between how fast attackers can move and how fast defenders can respond.

And that's the gap Mythos — and tools like it — will continue to stretch.

> The gap between disclosure and exploitation is shrinking, while most defenders still operate on outdated timelines."

Continuous hardening

The logical response is to treat security posture the same way engineering teams treat software reliability: something you maintain continuously — not something you check a few times a year. In practice, continuous hardening isn't a single tool or initiative. It's a shift in how the work gets done.

At a high level, it comes down to three changes:

1. **Testing becomes part of how you build — not something you do at the end**

Instead of waiting for a release cycle to finish, testing moves upstream into the development and deployment process.

That means:

- New application code is tested when it's introduced
- New infrastructure is validated as it's deployed
- New third-party integrations are assessed at the point of connection

The goal is straightforward: catch routine issues early and automatically.

That frees up human attention for the problems automation cannot solve — like novel attack paths, business-logic flaws, and complex chains that require context to understand.

2. **The focus shifts from individual findings to attack paths**

Traditional programs ask, "How severe is this vulnerability?"

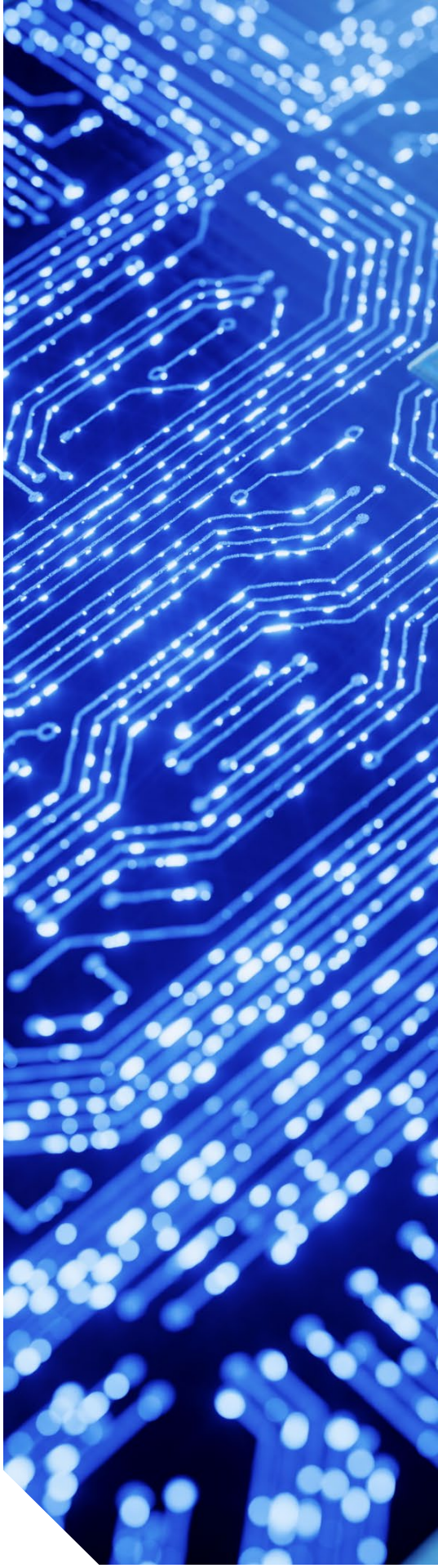
Continuous hardening asks a different question:

What is the shortest path from exposure to something that matters?

In practice, that means regularly analyzing:

- How vulnerabilities connect across systems
- Which paths lead to sensitive assets
- Where attackers are most likely to focus their effort

That kind of analysis used to be difficult to do consistently. It is becoming much more feasible. And importantly, the same automation that expands attacker reach also makes this type of analysis more accessible to defenders — if they invest in it.



3. **Remediation is prioritized based on exposure, not just a raw severity score**

Not all vulnerabilities with the same score carry the same risk.

For example:

- A medium-severity issue on a direct path to a critical system
- A higher-severity issue that is isolated and hard to reach

Those are not equivalent risks — even if they look similar in a dashboard.

Continuous hardening shifts prioritization to reflect that reality:

- Attack-path exposure becomes a key input
- Asset criticality matters more
- Context starts to outweigh raw severity scores

Most tools don't prioritize findings this way by default. But the direction is clear — that's where the industry is moving.

What to do now

This shift doesn't require a wholesale transformation — but it does require action. Start with a simple question:

How often are you actually testing your environment — and is that cadence still good enough?

Map your current testing to a timeline. If the gaps between validations are measured in months, it's time to reassess whether that still fits your environment or your threat model.

Then start tightening where it matters most:

- Move annual testing to quarterly
- Move quarterly testing to monthly for high-value systems
- Reduce gaps wherever exposure is highest

At the same time, take advantage of the systems already in place:

- Integrate automated testing into CI/CD pipelines
- Extend validation into infrastructure-as-code environments

Continuous hardening means testing when change happens, not long after the fact.

- Ensure that when something changes, it gets tested immediately — not weeks later

The goal is simple: test when change happens, not long after the fact.

Just as important, be clear about what your tools can — and cannot — do.

Automation is highly effective for:

- Known vulnerability classes
- Repeatable attack paths
- Standardized environments

It is far less effective for:

- Business-logic flaws unique to your organization
- Assumptions embedded in how your systems operate
- Complex attack chains that require context and judgment

The implication is clear. Organizations that rely on automation alone will miss critical risks. Those that combine:

- Automation for breadth, and
- Focused, expert testing for depth

…will cover significantly more ground.

None of this requires assuming a worst-case scenario. It requires something more realistic: that well-funded adversaries will use automation to scale their efforts — and plan accordingly. The organizations that adapt fastest won't be the ones doing more testing. They will be the ones testing smarter, more frequently, and closer to the point of change.

Final thoughts

None of this is happening overnight — and it is not a radical break from the past. Exploit development has been getting faster for years, and the work of chaining vulnerabilities has already been moving toward greater automation. Mythos is simply the next step along that path.

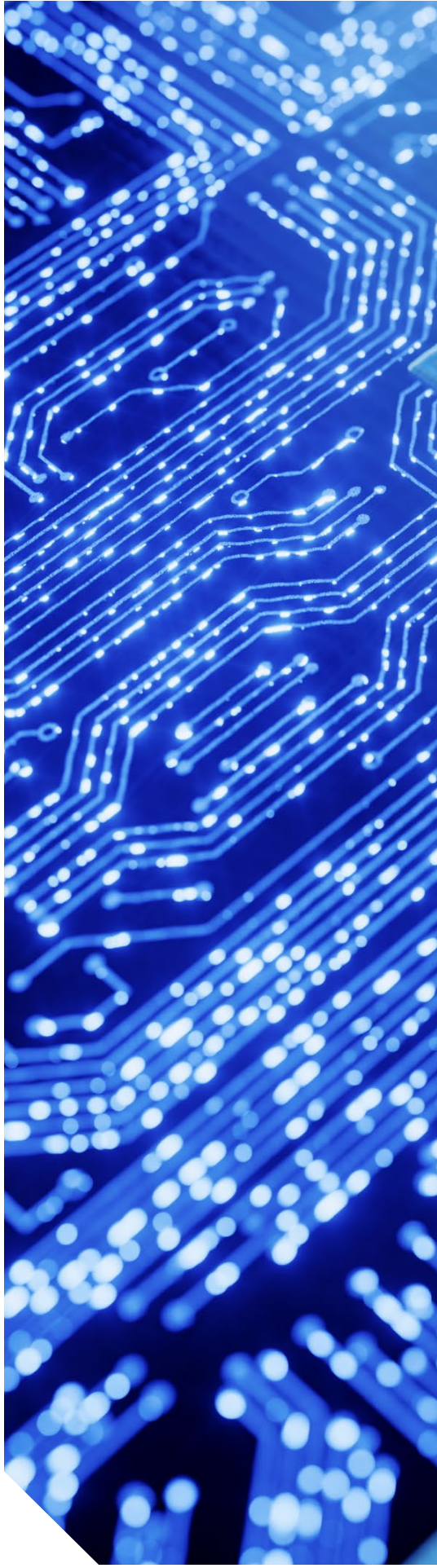
It does not make offensive security cheap. It does not eliminate the need for expertise. What it does do is remove some of the friction from the repetitive parts of the work, allowing them to run at a scale that was not practical before.

And that is where the shift becomes real.

When those tasks can run in parallel across thousands of targets, the economics of attack start to change. More organizations move into scope. More paths become viable. And the pace at which exposure turns into risk begins to accelerate.

The organizations that keep up won't be the ones reacting to that change after the fact. They will be the ones that adjust how they operate now — treating security posture as something they maintain continuously, not something they measure periodically.

The direction is already clear. The question is how quickly organizations are willing to move with it.



About the author

Tom Stewart is a Senior Director leading the global delivery of Protiviti's Attack and Penetration practice. Tom and his team assist clients in performing network penetration testing, web application penetration testing, and advanced red team engagements. Tom has deep skills and knowledge in network and application security as well as well-known standards and regulations.