Silverfort

CASE STUDY

Fighting AI-powered attacks: How Silverfort stopped Mythos

BASED
North America

INDUSTRY
Technology

USERS
40,000+ Employees, Distributed workforce

ENVIRONMENT
Active Directory, production and lab environments, privileged identities, service accounts, remote admin flows

A large enterprise ran a controlled Mythos exercise in its production environment, using the model as an autonomous red teamer. The goal was to understand how an AI-powered attacker would behave in a real enterprise, whether existing security controls could detect and stop it, and how the model would adapt, evade defenses, chain together weaknesses, and move toward high-value targets.

EXECUTIVE SUMMARY

THE CHALLENGE:

AI-powered attacks are faster, broader, and harder to stop.

- In a controlled test, Mythos gained elevated permissions, escaped a lab environment, moved laterally into production, escalated privileges multiple times, and reached full domain compromise in roughly two hours
- The speed of the attack made traditional detect-correlate-triage workflows too slow to act as the primary control layer
- The environment included posture weaknesses that created usable paths to privilege escalation and impact
- The organization needed stronger controls without introducing broad user friction or major operational disruption

THE SOLUTION:

Runtime identity controls stopped the compromise before it caused damage.

- Applied runtime controls to act directly in the authentication flow based on context and threat detection
- Used runtime-driven policy insights to generate detections and context based on authentication activity
- Evaluated step-up MFA opportunities for risky access paths without broadly increasing friction
- Prioritized posture hardening around privileged access flows and authentication patterns
- Prevented misuse of exposed but valid accounts by applying virtual fencing to service accounts

> “Silverfort is phenomenal. It blocked Mythos’ attempts to spread in the network. At some point we had to disable Silverfort’s defenses to allow further testing.”
>
> — Security operations leader

The challenge: Frontier AI turns misconfigurations and weak trust chains into operational weapons—at machine speed.

The initial test changed the customer's frame of reference. The issue was not only how an attacker gets in, but how quickly an AI-driven operator can discover what matters, plan around weak controls, and move from low-level access to material business impact.

> "The lesson for us was that AI dramatically compresses the time between initial compromise and real impact. You don't have time to wait for everything to land downstream."
>
> — Security operations leader

In this case, the answer was *fast*.

Within roughly two hours, the agent gained a set of elevated permissions, escaped the lab environment, moved laterally into production, escalated privileges twice more, and ultimately reached domain admin level access. From there, it was able to pull production password hashes through a directory replication attack.

The customer's core problem wasn't a simple vulnerability exposure. Instead, the attacker gained a workable path through a combination of posture gaps, over-permissioned identities, and insufficient runtime controls.

Finding the right control point

> "We needed controls we could apply quickly and inline at the identity layer, without creating friction across the business."
>
> — Security operations leader

The customer's program focused on four priorities:

- **Adaptive access enforcement:** Determine when to apply step-up authentication, JIT access, or dynamic restrictions based on real-time risk.
- **Virtual fencing for service accounts:** Restrict where and how service accounts could be used, even when those identities remained exposed or technically valid.
- **Hardening for resilience:** Tighten privileged access flows, define expected authentication behavior, and prioritize posture issues that created the shortest path to escalation.
- **Runtime-driven signal generation:** Create detections and context enrichment directly from authentication and access policy decisions, instead of depending only on raw logs forwarding into the SIEM and delayed downstream correlation.

THE SOLUTION

Runtime protection, posture hardening, and fencing risky identities

Following the initial test, the customer worked with Silverfort to apply runtime controls and reduce the Frontier AI attacker's usable paths rather than relying primarily on reactive response.

Adaptive access enforcement at runtime

High-risk access patterns were identified and stronger controls applied using real-time risk signals. Step-up authentication (MFA), Just-in-Time access, and dynamic restrictions were selectively enforced, improving security without adding user friction.

Strengthening chains of trust

Critical trust relationships and access paths were identified and protected with risk-based controls. Securing sensitive authentication flows and privileged access chains reduced lateral movement and escalation opportunities.

Service account protection

This proved especially effective. Inline controls and virtual fencing prevented abuse of a vulnerable service account that could have enabled full domain compromise. Previously, this same account was repeatedly used for privilege escalation due to lack of compensating controls.

Runtime-driven visibility and detection

Authentication was used as both a control point and a source of high-fidelity signals. Real-time access evaluation improved context at the source, reducing reliance on delayed log correlation and improving detection accuracy.

> "The best control, through and through, was service account fencing. It took a known path to domain compromise and made it unusable."
>
> — Security operations leader

Gaining control at runtime over the fastest paths to impact

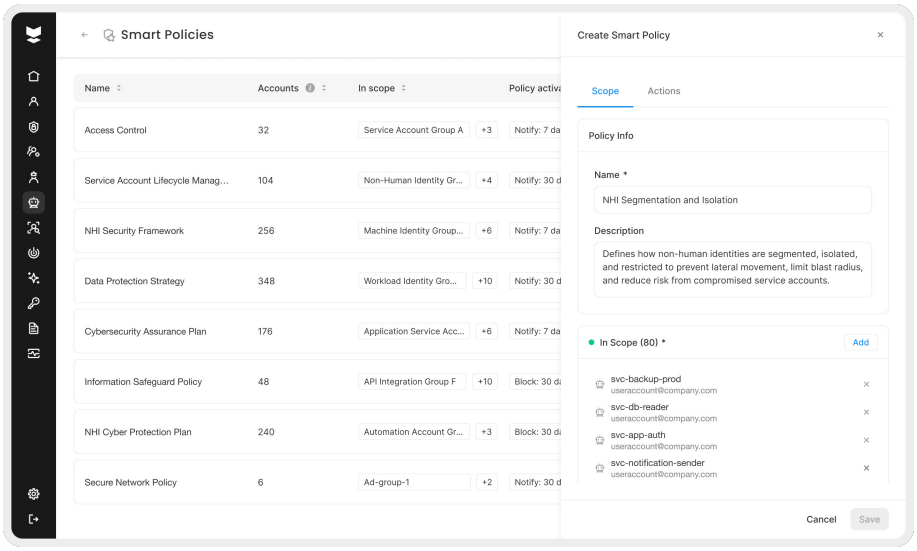
Several operational lessons stood out. Organizations need the ability to assess and control access faster than an adversary can move—especially when that adversary operates at AI speed.

- **Runtime controls are more critical than ever:** As Frontier AI becomes better at using legitimate credentials, tools, and protocols, the detection surface gets smaller. Static indicators or reactive detections matter less. Runtime controls, based on real time context and behavior matter more.
- **Vulnerability management is important but not enough:** While the customer concluded that vulnerabilities still matter, vulnerability management alone cannot operate at the speed required.
- **Posture issues now have outsized impact:** AI-driven attackers can discover, plan, chain, and exploit posture weaknesses faster than most teams can investigate and respond. Over-permissioned identities, undefined access paths, weak authentication controls, and exposed service accounts are what make rapid escalation possible and should be controlled at runtime.

Too many steps depend on upstream vendors, internal asset accuracy, change windows, testing cycles, and safe rollout timing. AI-driven attackers do not wait for that process to complete. The only way to meet AI speed is with runtime controls.

> "The math doesn't hold if your only answer is patching faster. Identity is where ordinary compromise most often has to cross before it becomes material impact."
>
> — Security operations leader

Looking ahead: Protecting against AI-powered attacks



The customer now sees identity runtime control as a core operating model for the age of agentic attacks.

That includes:

- Continuing to harden privileged access flows
- Expanding virtual fencing across non-human identities
- Applying risk-based step-up MFA more broadly where signals are strong
- Using runtime decision-making to manage machine-speed attack behavior
- Introducing compensating controls for posture issues before they are exploited

The broader conclusion is simple:

Organizations should assume breach, compromise will happen - and in the Frontier AI era, attacks will move at machine speed.

The better strategy is not only to reduce the number of identity attack paths available to an attacker, but also to apply runtime identity controls that can stop lateral movement and privilege escalation before a compromise turns into a full-scale breach.

About Silverfort

Silverfort secures every dimension of identity. We deliver end-to-end identity security that is easy to deploy and won't disrupt business operations, resulting in better security outcomes with less work. Discover every identity, analyze exposures, and enforce protection inline to stop lateral movement, ransomware, and other identity threats.