Zscaler Is Proud to be Part of Project Glasswing: AI Can’t Breach What It Can’t Find

April 21, 2026

Overview

Anthropic has been at the forefront of AI innovations. Dario Amodei, Anthropic CEO, has always been mindful of the dangers of very powerful AI models and has advocated for their responsible use. Recognizing the power of their Mythos model to uncover long-hidden software vulnerabilities, Anthropic took a responsible approach. Through Project Glasswing, they made the model available only to a select group of organizations that either operate or protect our country's critical infrastructure. Zscaler is proud to collaborate with Anthropic on Project Glasswing, which has provided us with access to Claude Mythos Preview.

The premise is simple, frontier AI models have reached a point where they can find software vulnerabilities faster than humans can. Mythos Preview understands code the way a skilled human researcher does, reading logic, chaining multiple weaknesses together, and producing working exploits in hours, at machine speed, instead of weeks. It has already uncovered thousands of high-severity flaws across major operating systems and browsers. The ability of AI to rapidly uncover vulnerabilities and produce working exploits is going to accelerate, and when it does, defenders need to be ahead.

Reactive patching is no longer a viable defense strategy. You cannot outpace AI-driven vulnerability discovery, and you cannot out-hire the efficiency of an automated adversary. The only durable answer is founded on architecture. This means simply adding another tool on top of your security stack won’t cut it. You cannot patch, detect, or respond your way out of a problem created by exposing applications to the internet in the first place; you have to stop exposing them.

The Old Game Is Lost

For thirty years the industry has played the same game. Put a firewall at the edge. Put a VPN in front of your applications. Scan for known vulnerabilities. Patch what you find. Hope you find them before the adversary does.

That game assumed a human-speed attacker. Mythos Preview ends that assumption. If your application is exposed to the internet behind a firewall or a VPN, a frontier model can already see it. It can scan every internet-facing surface parallel, test for weaknesses no human team has the bandwidth to check and do it continuously. Once that capability is in the hands of a nation-state or a ransomware group, your patch cycle is irrelevant.

Legacy security was built on the hope that we could outrun the attacker. In an era of AI-driven exploits, that race is over. We now have to assume the attacker is already inside.

A Fundamentally Different Architecture

Zscaler was built for exactly this moment, and we have been saying it for more than 18 years. If you are reachable, you are breachable.

Zero Trust is not a feature. It is not a firewall with a new label. It is a fundamentally different architecture, built on a different principle. Users never connect to the network and applications are never exposed to the internet. Endpoint context is understood, and devices are verified before they connect. Data is protected the moment it is accessed. Every connection, whether human or AI agent, is brokered one to one with a verified identity in real time, with no lateral path to anything else.

When an application is hidden behind the Zscaler Zero Trust Exchange, it has no public IP, open port, or discoverable surface. An attacker scanning the internet cannot find what is not there. The vulnerability may exist in the code. It may even be cataloged in a CVE (Common Vulnerabilities and Exposures) database. But the adversary has no way to reach it.

This is the difference between detecting attacks and taking your applications off the public internet entirely, so there is nothing for attackers to target. Both matter. Only one scales against machine-speed offense.

What Zscaler Brings to Project Glasswing

Zscaler is the platform that 40% of the Global 2000 trust to run their businesses. Our contribution is grounded in how the Zero Trust Exchange platform already operates at the core of the enterprise.

- **The largest security cloud in the world:** Zscaler processes over 500 billion transactions every day and hundreds of trillions of signals. That scale is what lets our AI distinguish a benign request from a reconnaissance probe. We do this inline, before a connection is ever established.
- **Attack surface elimination:** The Zscaler Zero Trust Exchange makes internal applications invisible to the internet. Whether those applications are running in your data center, or in the public cloud, Zscaler hides them from attack. No firewalls or VPNs to exploit, and nothing for a frontier model to find.
- **Data protection at the point of use:** The new risk is not someone breaking in. It is your own AI tools quietly taking sensitive data out. Zscaler’s AI guardrails see every request as it happens, across SaaS, private apps, email, and encrypted traffic, and stops the data before it leaves.
- **Zero trust for AI agents:** Agents are now acting autonomously on behalf of users. They are authorized to access data, they take action and connect to other systems. They must be governed with the same architecture we apply to human users. Every agent gets a verified identity, access to one specific application, and a full record of what it did.

How Zscaler Will Use Mythos Preview

We are integrating Mythos Preview into our secure software development lifecycle. It will enable us to rapidly find vulnerabilities in our software stack and Zero Trust Exchange, further hardening our environment and reducing risk for our customers. As a proud member of the Project Glasswing coalition, we will share our findings back to the community, helping everyone improve security outcomes for the world. Additionally, we will integrate Anthropic's Opus 4.7 model into our AI Red Teaming and Agentic SecOps offerings, to help fight AI threats with advanced AI security capabilities.

A Familiar Pattern

When the cloud arrived, the industry said the old perimeter would hold. It did not. When mobile and SaaS arrived, the industry said VPNs would adapt. They did not. Every twenty to thirty years the architecture has to change, and the companies that adapt win the next decade.

AI is that inflection, and it is moving faster than any shift before it. The adversary already has the model. So do we. The question is whether the enterprise will keep defending a perimeter that no longer exists, or take its applications off the public internet entirely.

There is no such thing as a Zero Trust firewall or an AI-proof VPN. There is only the architecture you choose before the next breach.

Zscaler is that choice. Project Glasswing is how we accelerate it across the industry. The time to act is now.

Where to Learn More

Watch the webinar recording form Wednesday, April 22 or Thursday, April 23, where we discussed how to protect your organization against vulnerabilities found by frontier AI models like Claude Mythos.