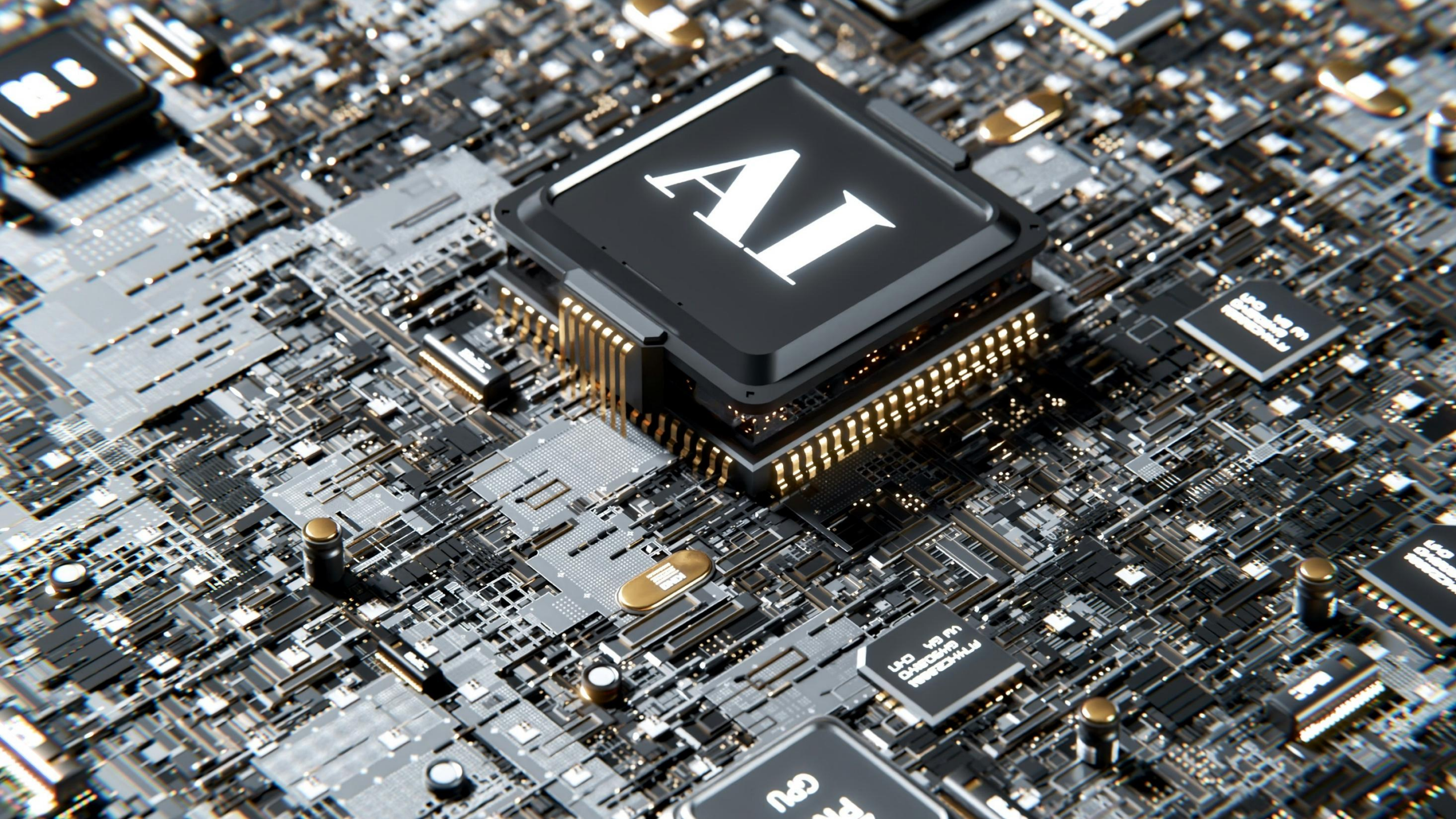
How the world will end


AI

Disrupting the first reported AI-or

https://www.anthropic.com/news/disrupting-AI-espionage

Research Economic Futures Commitments Learn News Try Claude

Policy

Disrupting the first reported AI-orchestrated cyber espionage campaign

Nov 13, 2025

Read the report

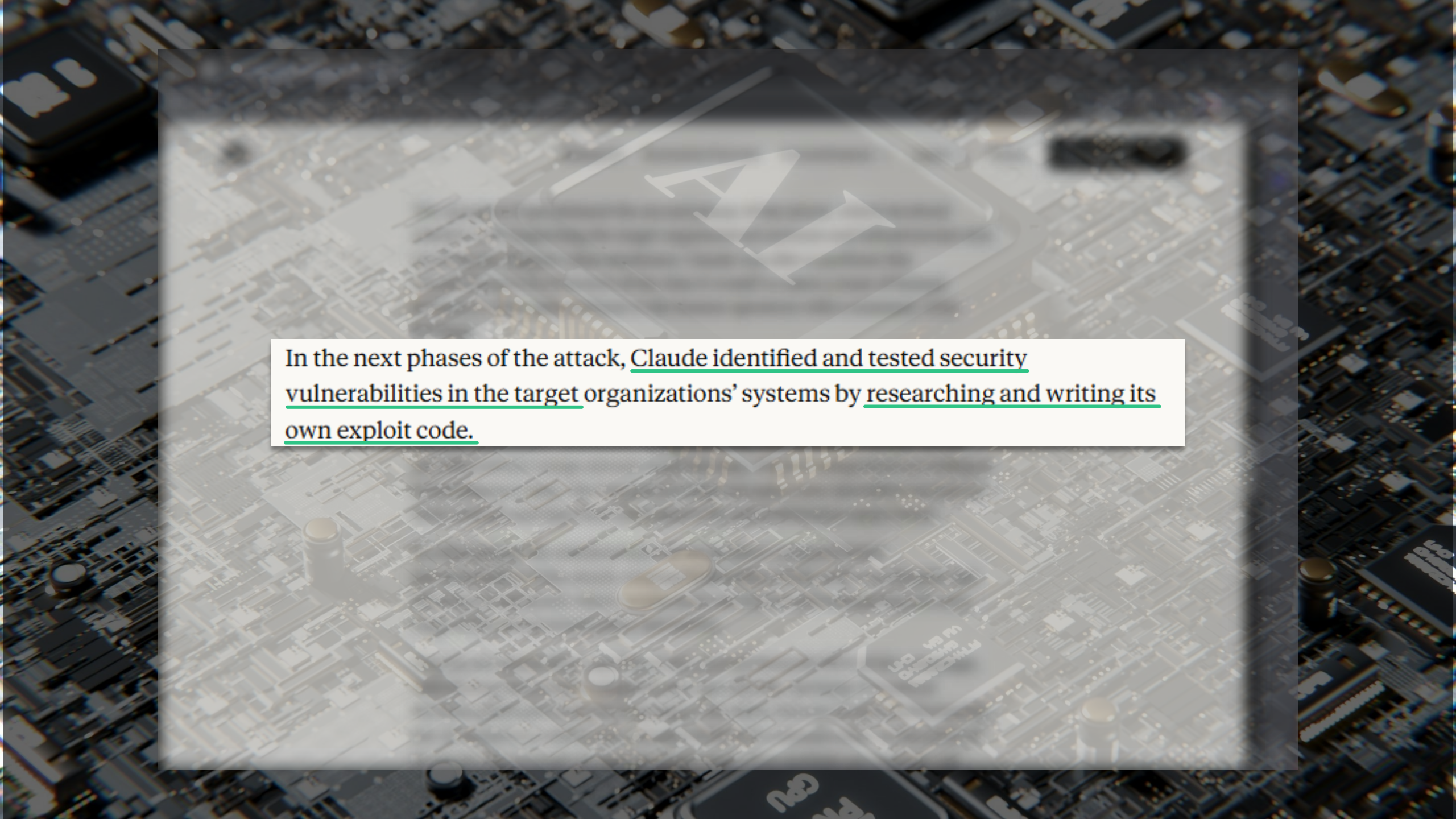
The attackers then initiated the second phase of the attack, which involved Claude Code inspecting the target organization's systems and infrastructure and spotting the highest-value databases. Claude was able to perform this reconnaissance in a fraction of the time it would've taken a team of human hackers. It then reported back to the human operators with a summary of its findings.

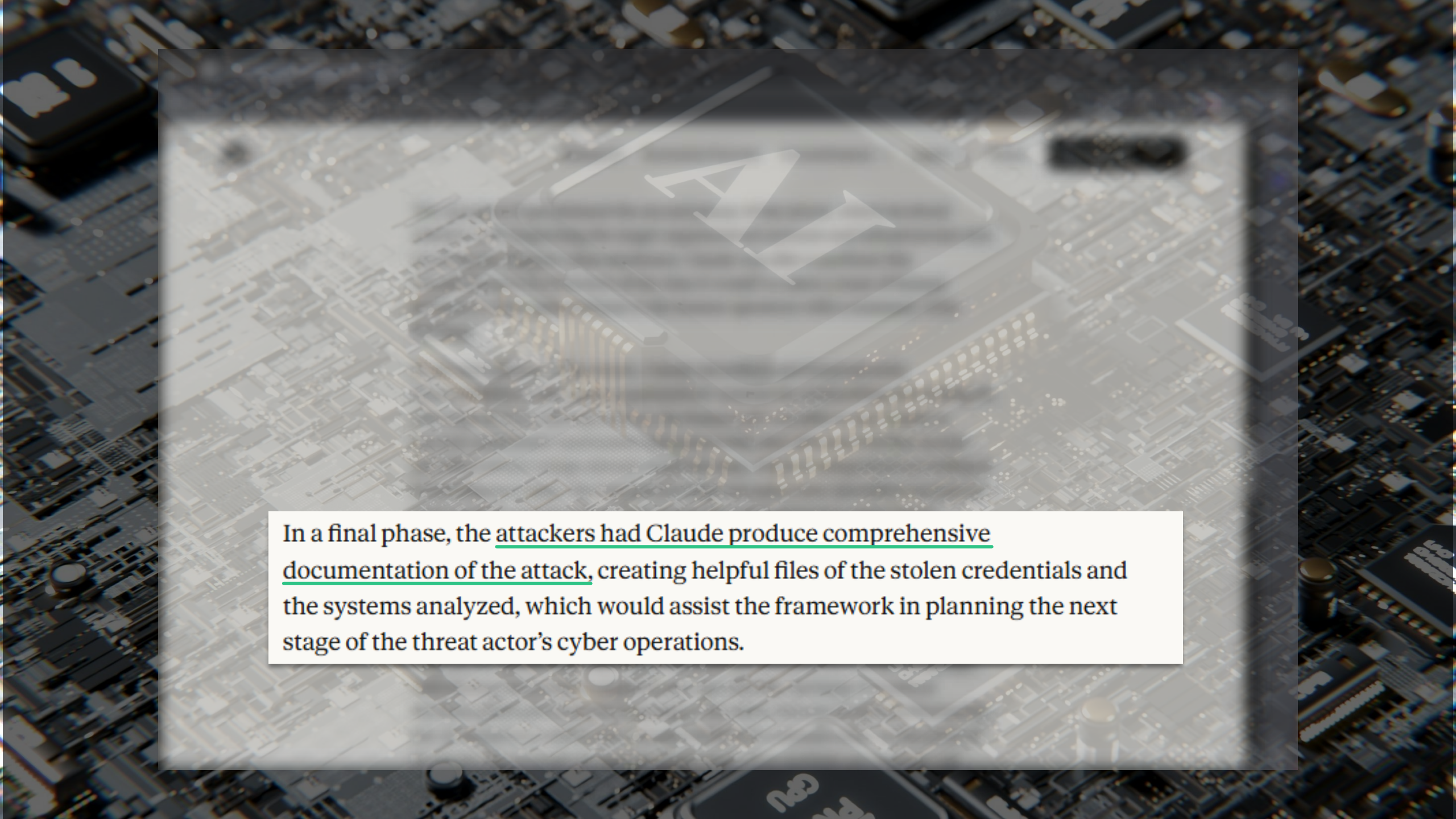
In the next phases of the attack, Claude identified and tested security vulnerabilities in the target organizations' systems by researching and writing its own exploit code. Having done so, the framework was able to use Claude to harvest credentials (usernames and passwords) that allowed it further access and then extract a large amount of private data, which it categorized according to its intelligence value. The highest-privilege accounts were identified, backdoors were created, and data were exfiltrated with minimal human supervision.

In a final phase, the attackers had Claude produce comprehensive documentation of the attack, creating helpful files of the stolen credentials and the systems analyzed, which would assist the framework in planning the next stage of the threat actor's cyber operations.

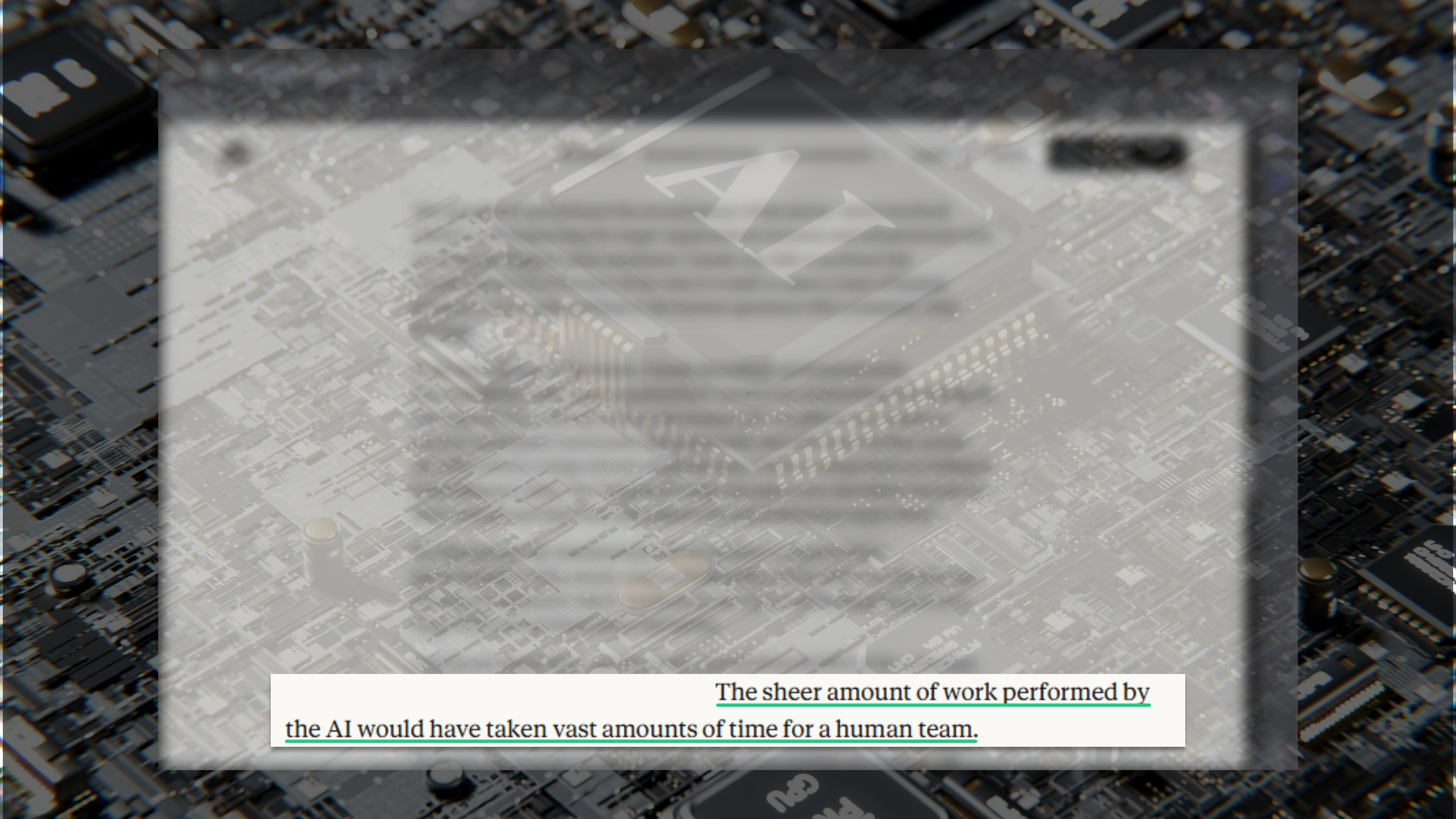
Overall, the threat actor was able to use AI to perform 80-90% of the campaign, with human intervention required only sporadically (perhaps 4-6 critical decision points per hacking campaign). The sheer amount of work performed by the AI would have taken vast amounts of time for a human team. At the peak of its attack, the AI made thousands of requests, often multiple per second—an

The attackers then initiated the second phase of the attack, which involved Claude Code inspecting the target organization's systems and infrastructure and spotting the highest-value databases. Claude was able to perform this reconnaissance in a fraction of the time it would've taken a team of human hackers. It then reported back to the human operators with a summary of its findings.


In the next phases of the attack, Claude identified and tested security vulnerabilities in the target organizations' systems by researching and writing its own exploit code.



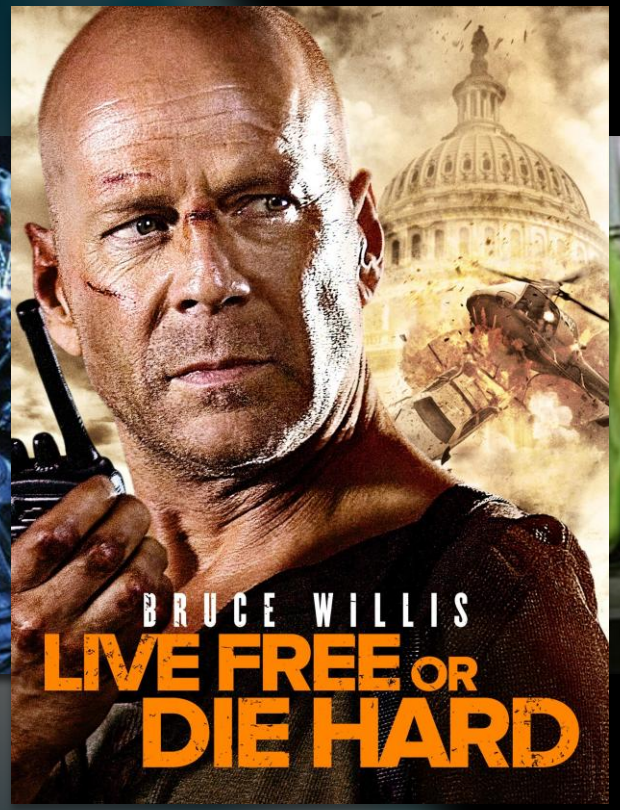
In a final phase, the attackers had Claude produce comprehensive documentation of the attack, creating helpful files of the stolen credentials and the systems analyzed, which would assist the framework in planning the next stage of the threat actor's cyber operations.



The sheer amount of work performed by the AI would have taken vast amounts of time for a human team.

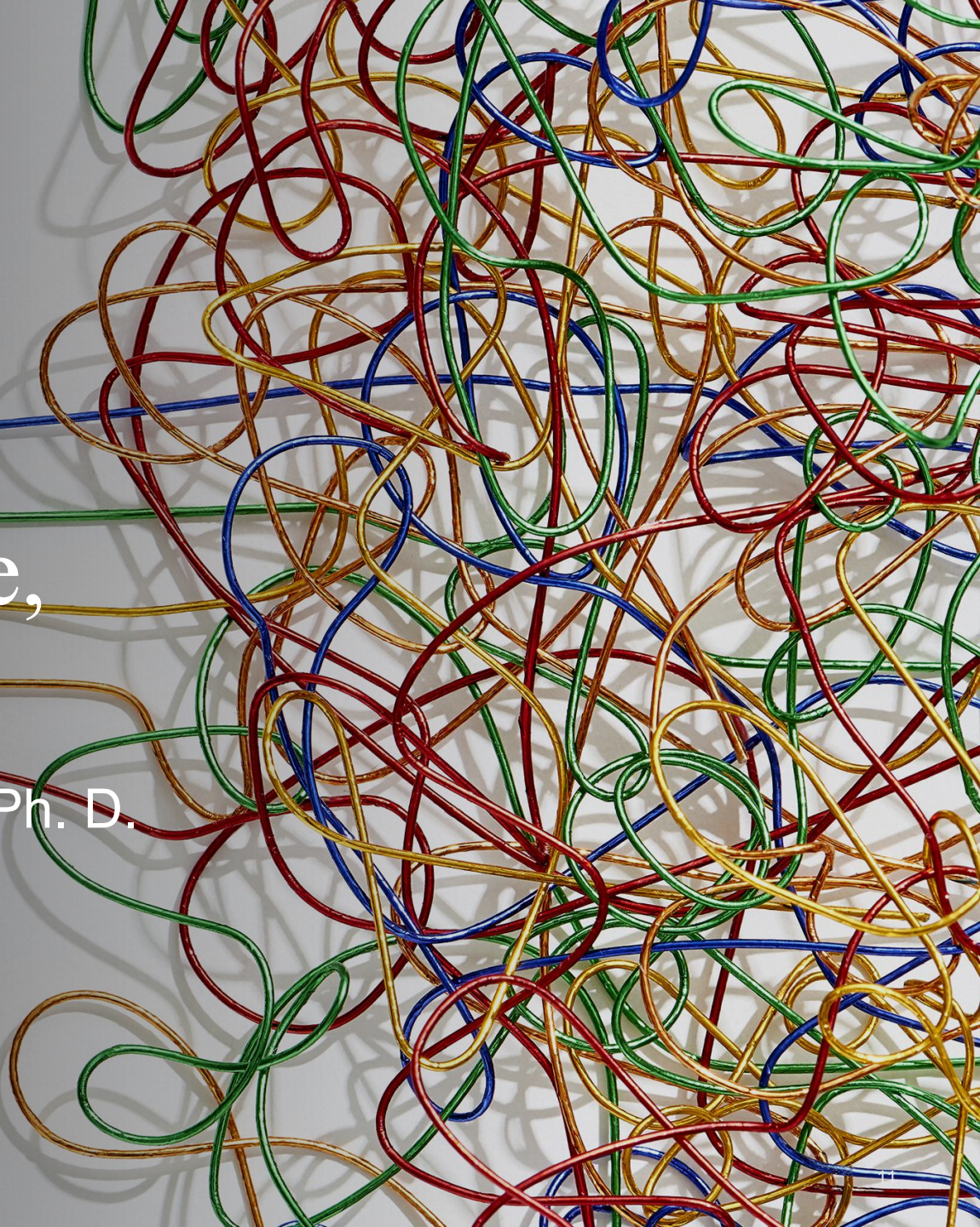
But we had cybersecurity problems *before* AI.

Of course we did…


BRUCE WILLIS
LIVE FREE OR
DIE HARD

Everything, as they say, is connected to everything else, and not neatly.

— Dr. Donella H. Meadows, Ph. D.



The Hacker News

Subscribe – Get Latest News

Home Vulnerabilities Cyber Attacks Expert Insights Awards

Apple Fixes Exploited Zero-Day Affecting iOS, macOS, and Other Devices

Ravie Lakshmanan Feb 12, 2026 Zero-Day / Vulnerability

Apple on Wednesday released iOS, iPadOS, macOS Tahoe, tvOS, watchOS, and visionOS updates to address a zero-day flaw that it said has been exploited in sophisticated cyber attacks.

— Trending News

INDEPENDENT

Podcasts Newsletters ENHANCED BY Google

NEWS SPORTS CULTURE LIFESTYLE MONEY TRAVEL CLIMATE PREMIUM MORE bulletin.

Tech > Security

Criminals hijack thousands of devices to create never-before-seen cyber weapon

Victims of the KadNap botnet are spread throughout the world

7 Comments

TechCrunch

Latest Startups Venture Apple Security AI Apps

SECURITY

Microsoft says hackers are exploiting critical zero-day bugs to target Windows and Office users

CA SECTIONS FORUM SUBSCRIBE

cybernews

Home » Security

Russian gang tied to recent massive cyberattack on Poland's power grid

Published: 29 January 2026 · Last updated: 29 January 2026

Gintaras Radauskas, Senior Journalist

TechRepublic.

Sign In / Sign Up

Topic — Security

Critical Chrome Security Flaws Threaten Billions of Users Worldwide

Published March 13, 2026

Written by Ken Underhill

Google patches two actively exploited Chrome vulnerabilities that could allow attackers to crash browsers or run malicious code. Billions of users urged to update.

ry 11, 2026

BLOCK

Never-before-seen Linux malware is "far more advanced than typical"

VoidLink includes an unusually broad and advanced array of capabilities.

DAN GOODIN – JAN 13, 2026 5:07 PM | 48

Credit: Getty Images

TEXT SETTINGS

Researchers have discovered a never-before-seen framework that infects Linux machines with a wide assortment of modules that are notable for the range of advanced capabilities they provide to attackers.

The framework, referred to as VoidLink by its source code, features more than 30 modules that can be used to customize capabilities to meet attackers' needs for each infected machine. These modules can provide additional stealth and specific

67%

organizations that report their sensitive data was **breached** or compromised ***at least once*** in the previous 12 months

Source: Forrester's Security Survey, 2025
Base: 2643 Security decision makers

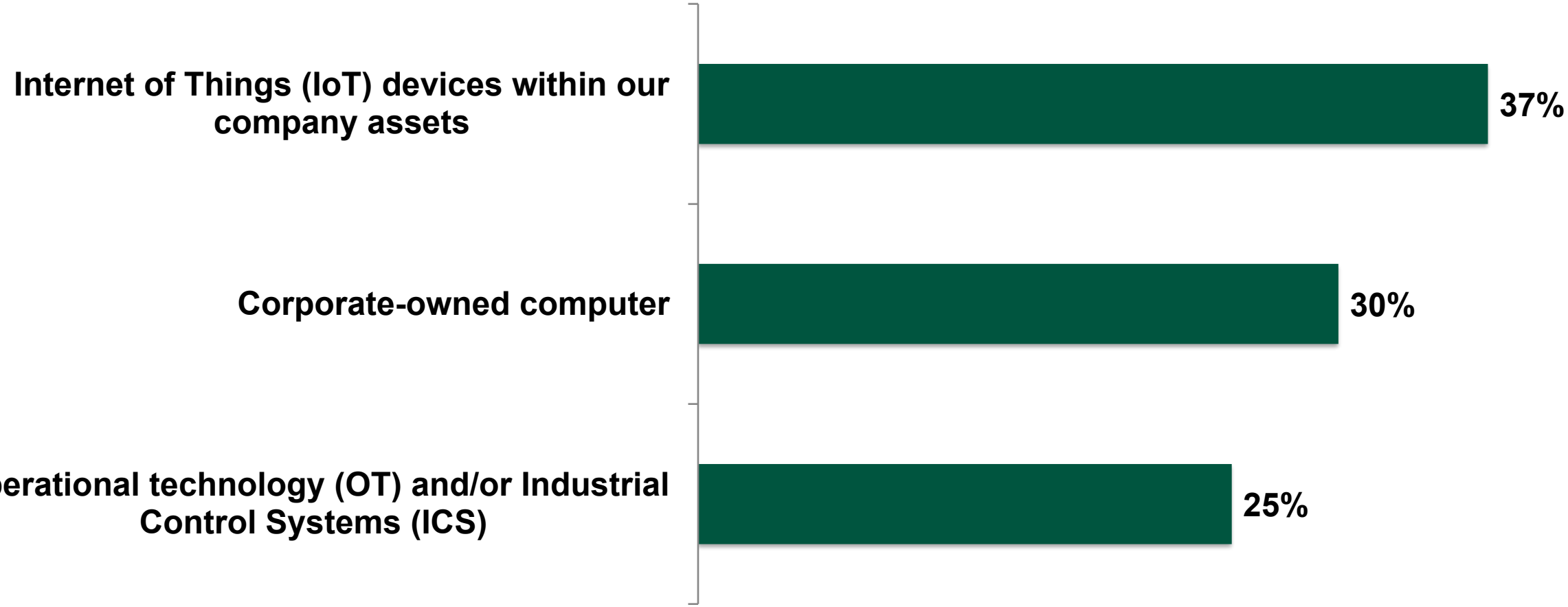
51%

organizations that report their sensitive data was **breached** or compromised ***two or more*** times in the previous 12 months

Source: Forrester's Security Survey, 2025
Base: 2643 Security decision makers

What is being targeted?

(some answers omitted)



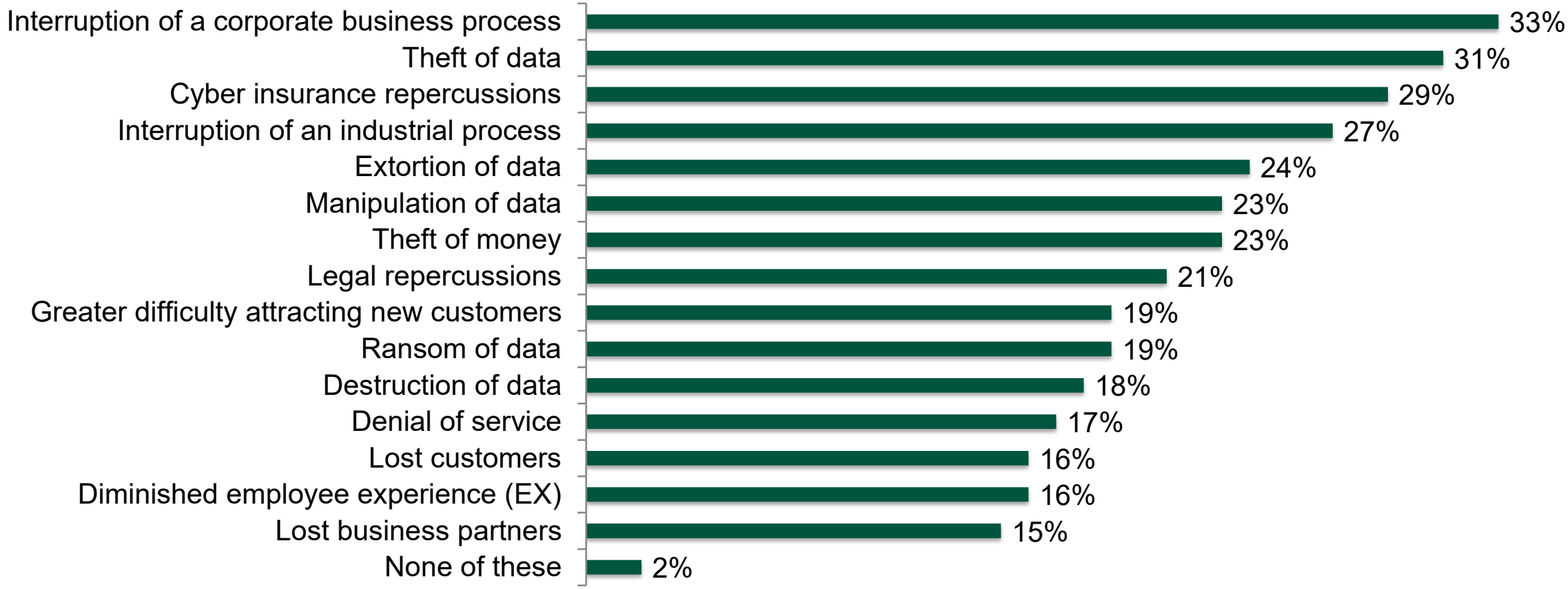
Source: Forrester's Security Survey, 2025
Base: 1138 Security decision-makers who experienced an external attack when their company was breached

$3.4M

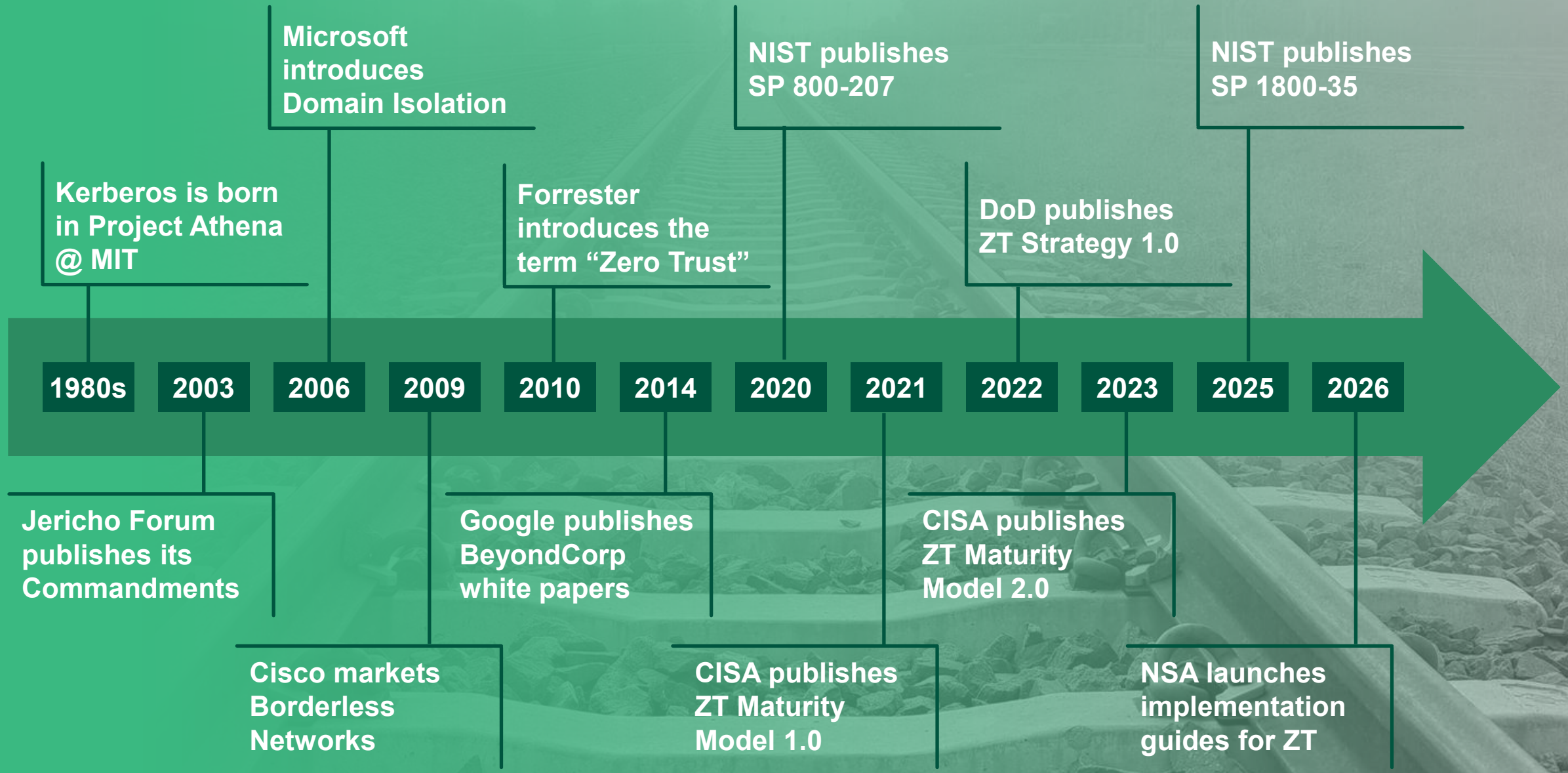
mean cumulative cost of all data breaches experienced by your organization in the last 12 months

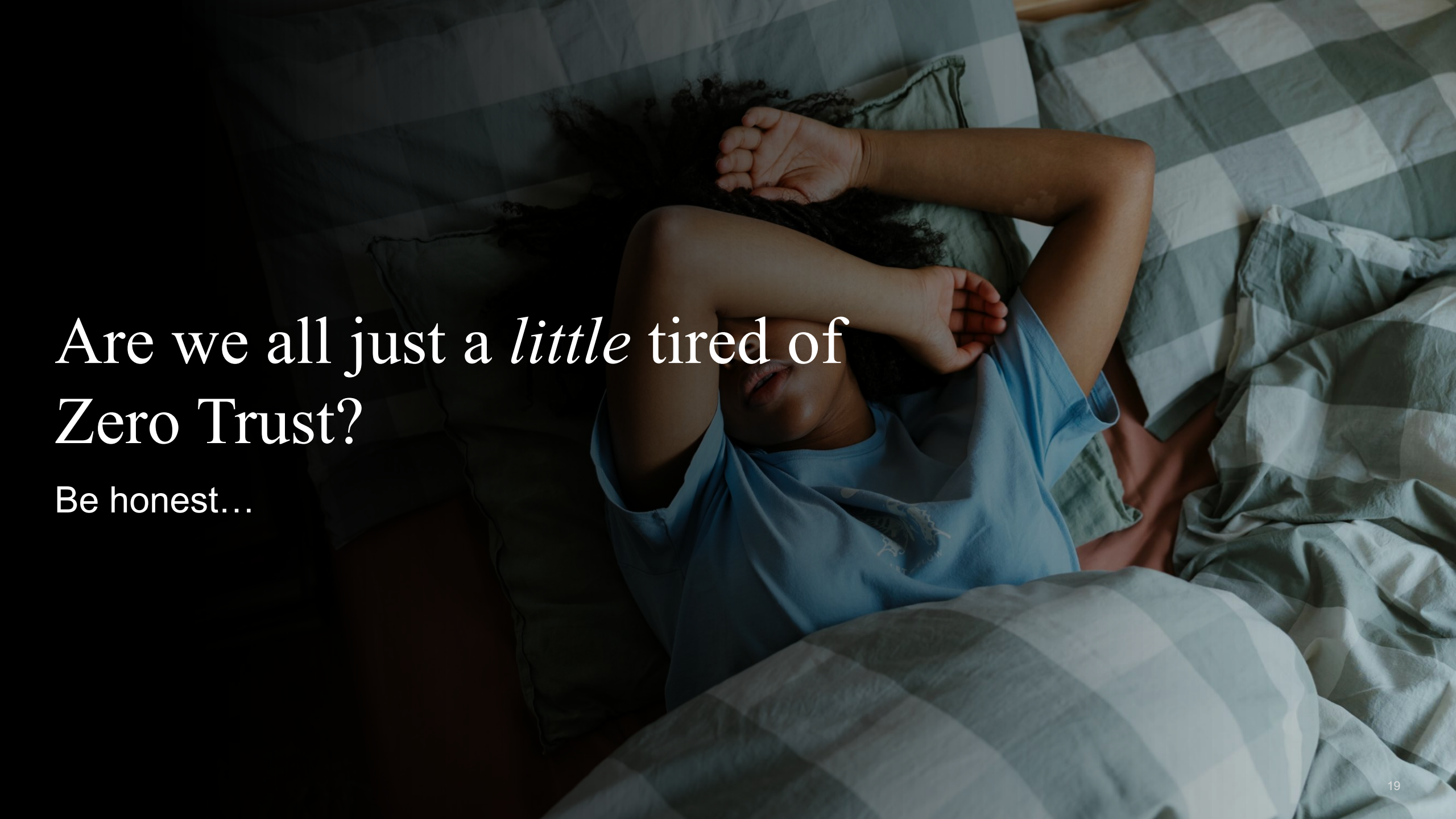
Source: Forrester's Security Survey, 2025
Base: 1,740 Security decision makers who have experienced a breach in the last 12 months

What were the effects of the breach(es)?



Source: Forrester's Security Survey, 2025
Base: 1766 Security decision-makers who have experienced a breach in the last 12 months


Kerberos is born in Project Athena @ MIT
Microsoft introduces Domain Isolation
Forrester introduces the term “Zero Trust”
NIST publishes SP 800-207
DoD publishes ZT Strategy 1.0
NIST publishes SP 1800-35
1980s
2003
2006
2009
2010
2014
2020
2021
2022
2023
2025
2026
Jericho Forum publishes its Commandments
Cisco markets Borderless Networks
Google publishes BeyondCorp white papers
CISA publishes ZT Maturity Model 1.0
CISA publishes ZT Maturity Model 2.0
NSA launches implementation guides for ZT



Are we all just a *little* tired of Zero Trust?

Be honest…

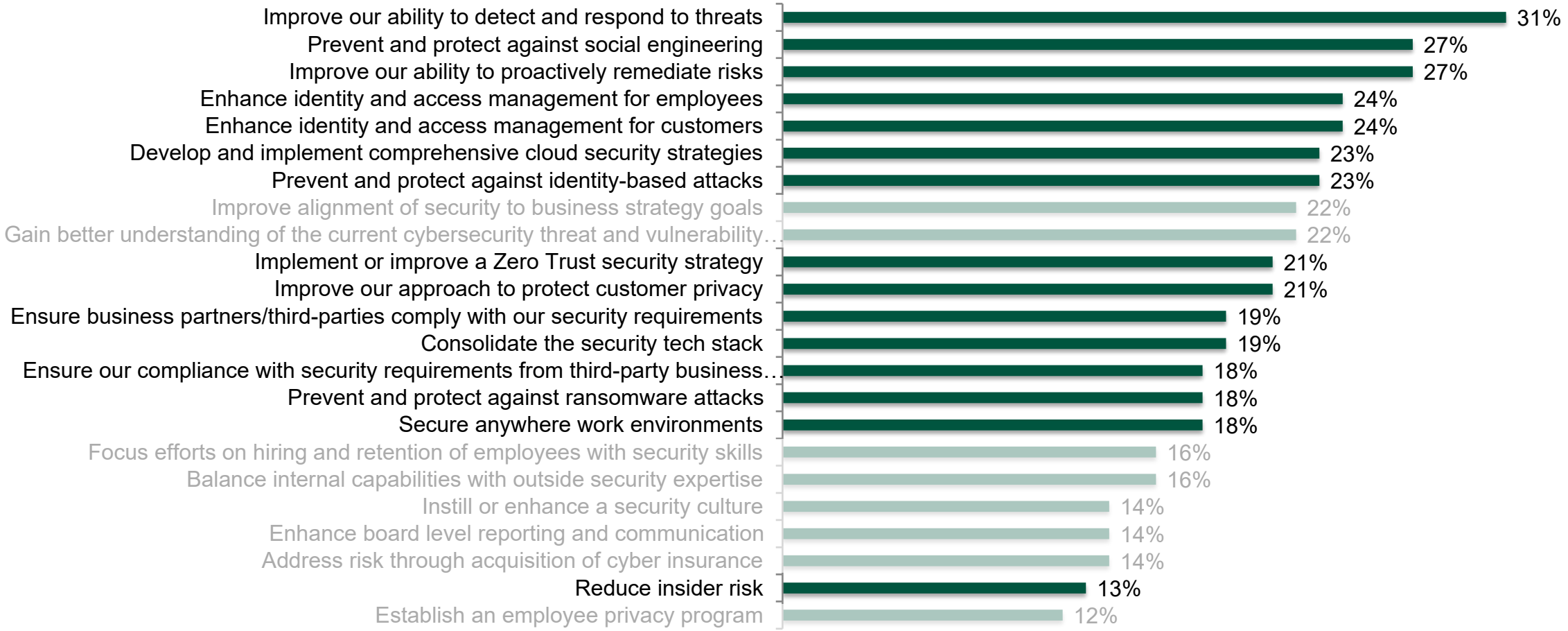
Zero Trust *as* a strategic security initiative

Which of the following are likely to be your top strategic IT security priorities over the next 12 months?

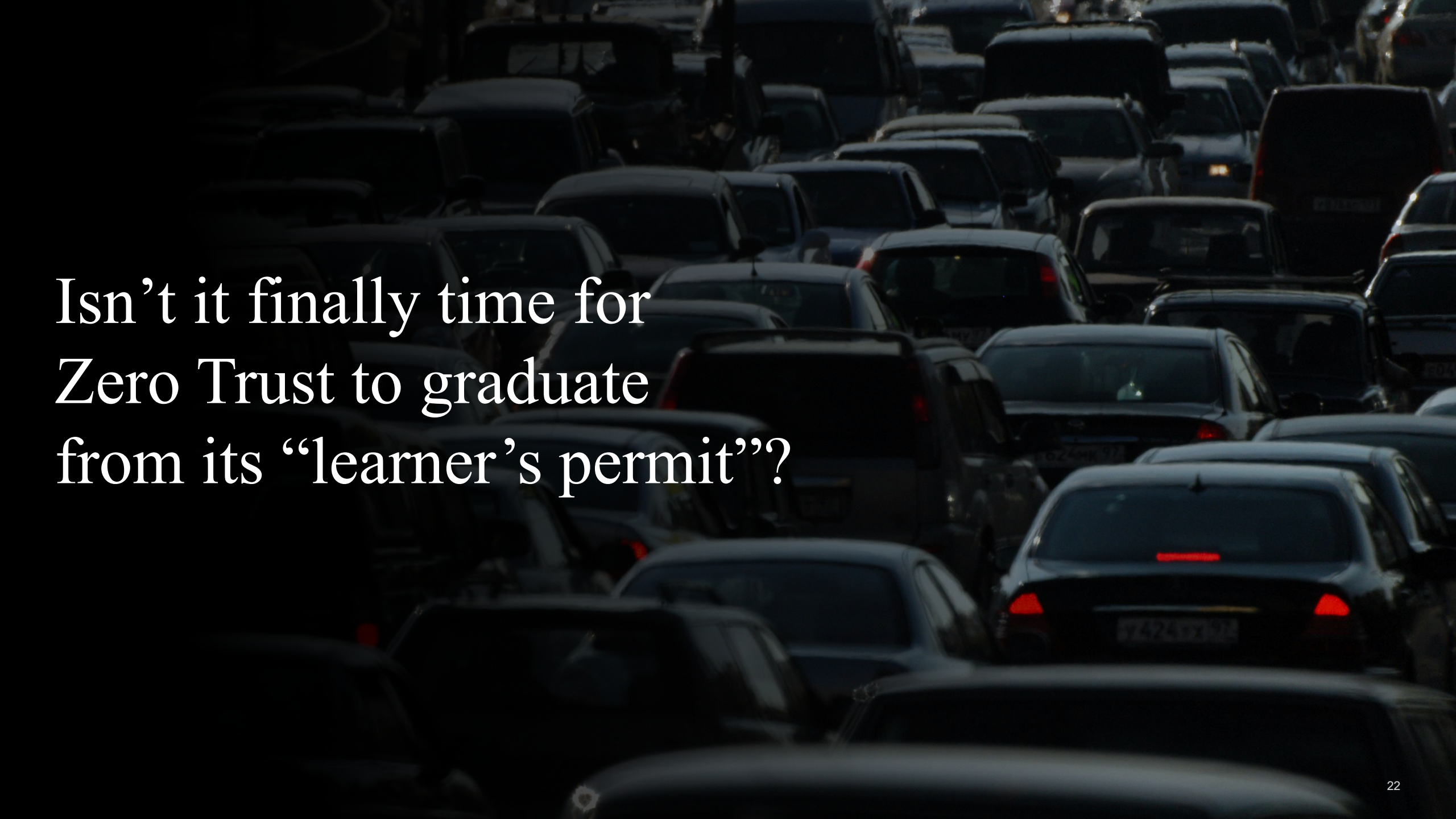
Source: Forrester's Security Survey, 2025
Base: 2665 Security decision-makers

Zero Trust *supports* strategic security initiatives

Which of the following are likely to be your top strategic IT security priorities over the next 12 months?



Source: Forrester's Security Survey, 2025
Base: 2665 Security decision-makers


Isn’t it finally time for
Zero Trust to graduate
from its “learner’s permit”?

“Assume context at your peril.”

— Jericho Forum, Commandment #3

The context we need for Zero Trust

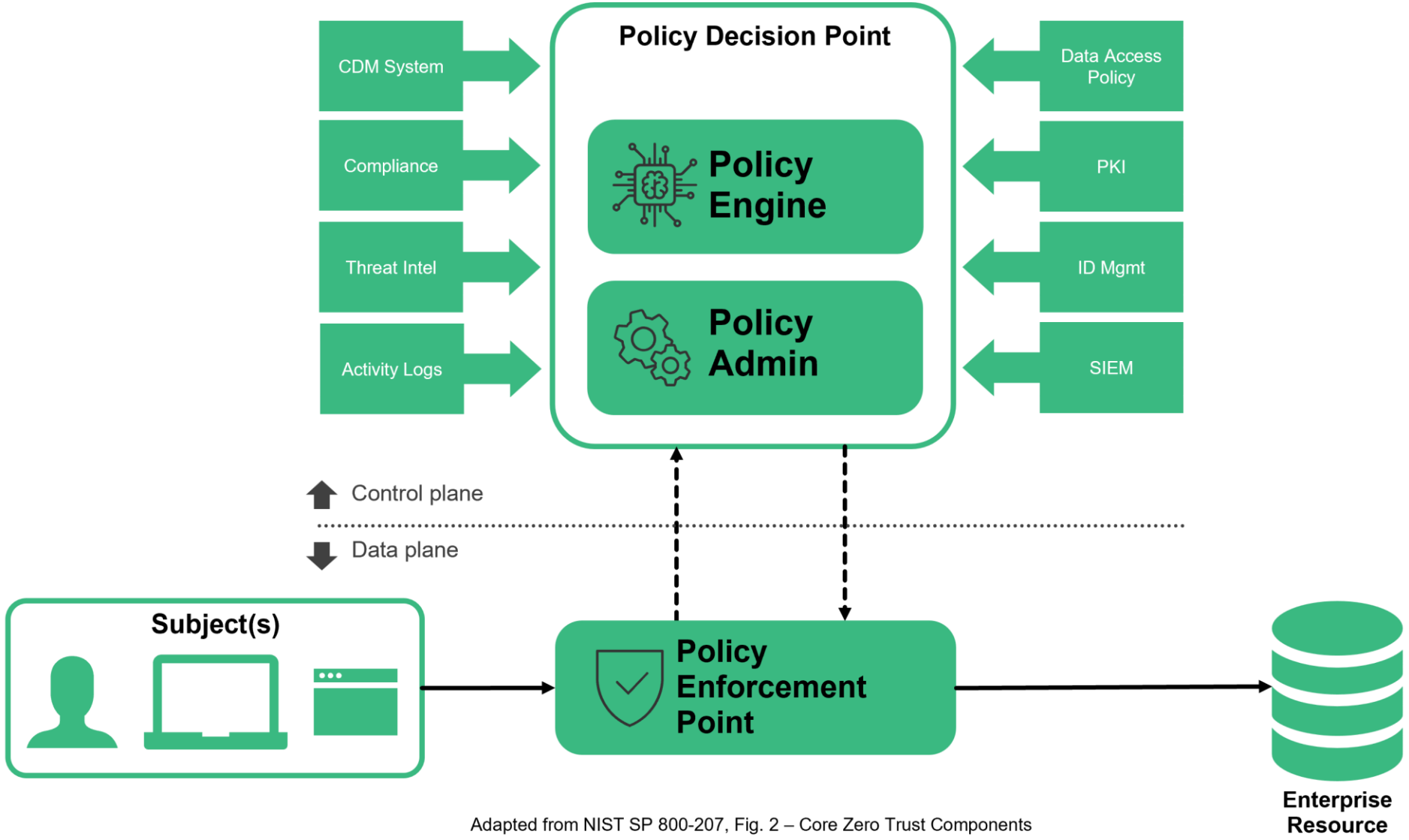
Actor	• Human • Non-human
Device	• Identity/type • Config/posture
Client app	• Identity • Composition/config/posture
Channel/flow	• Protocol • Method
Action	• Type • Sensitivity
Resource	• Identity/type/sensitivity • Posture

What's getting in the way of Zero Trust?

"What operational aspects of Zero Trust are viewed as the most challenging?"

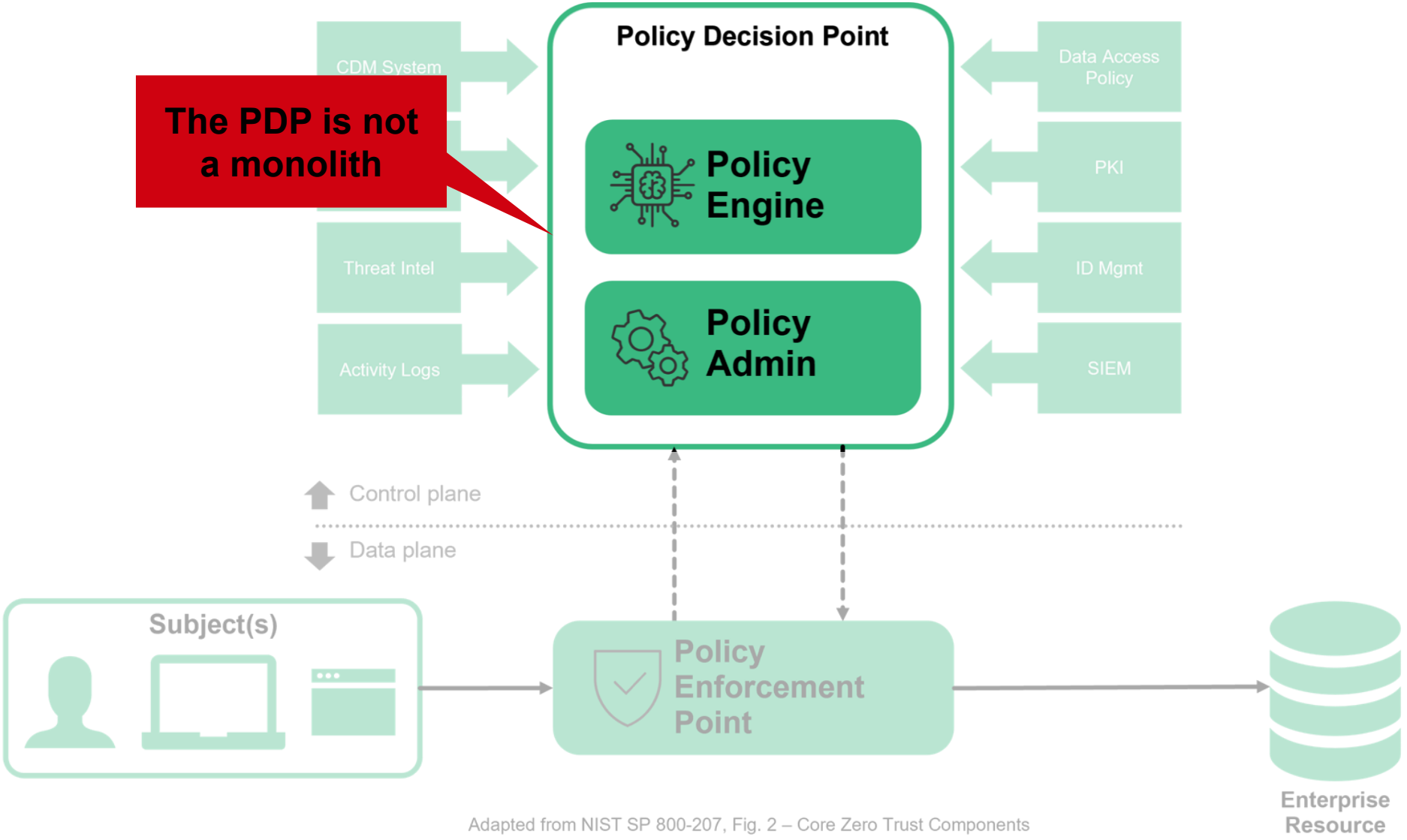
Source: Forrester's Security Survey, 2025
Base: 488 security decision makers who are a major influencer or final decision maker in Zero Trust for their organization

The myth of core Zero Trust components



Adapted from NIST SP 800-207, Fig. 2 – Core Zero Trust Components

The myth of core Zero Trust components



Adapted from NIST SP 800-207, Fig. 2 – Core Zero Trust Components

Implementing Zero Trust takes a “village”

Which means leveraging platforms and their ecosystems.

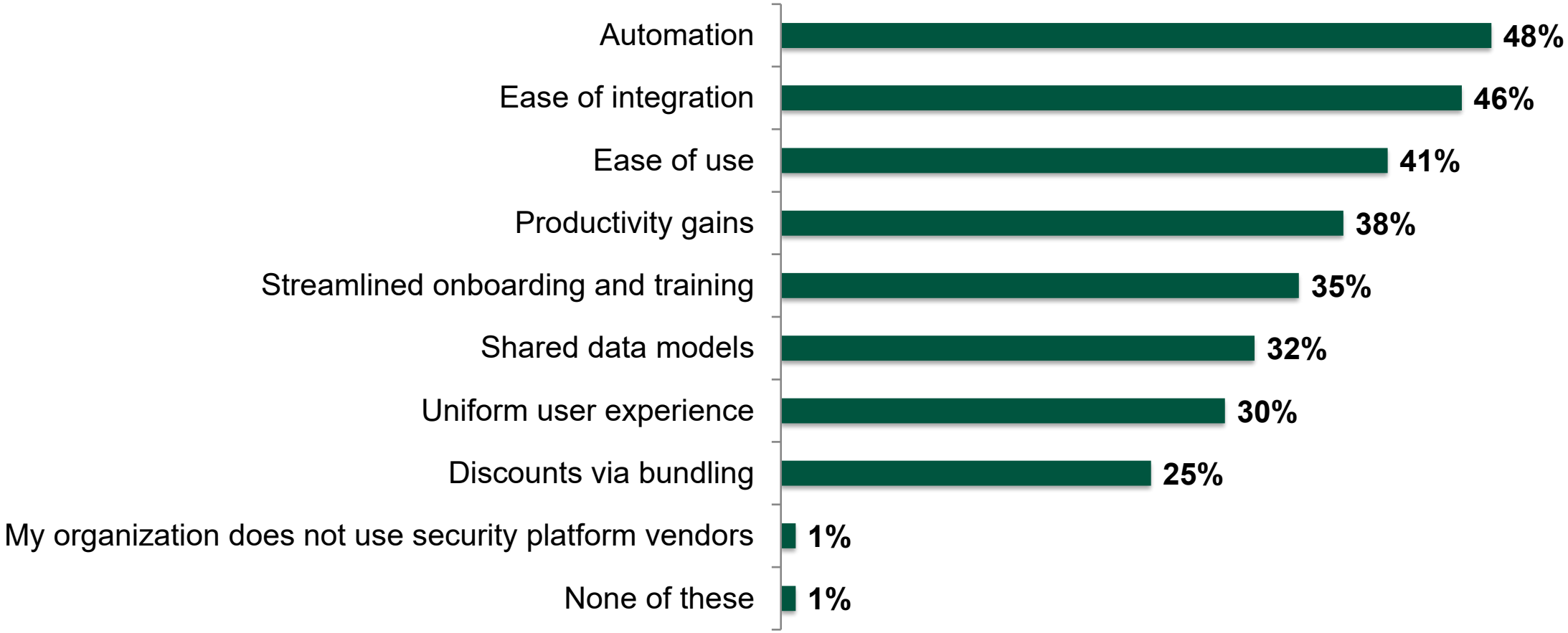
1%

organizations that report they do not use security platform vendors

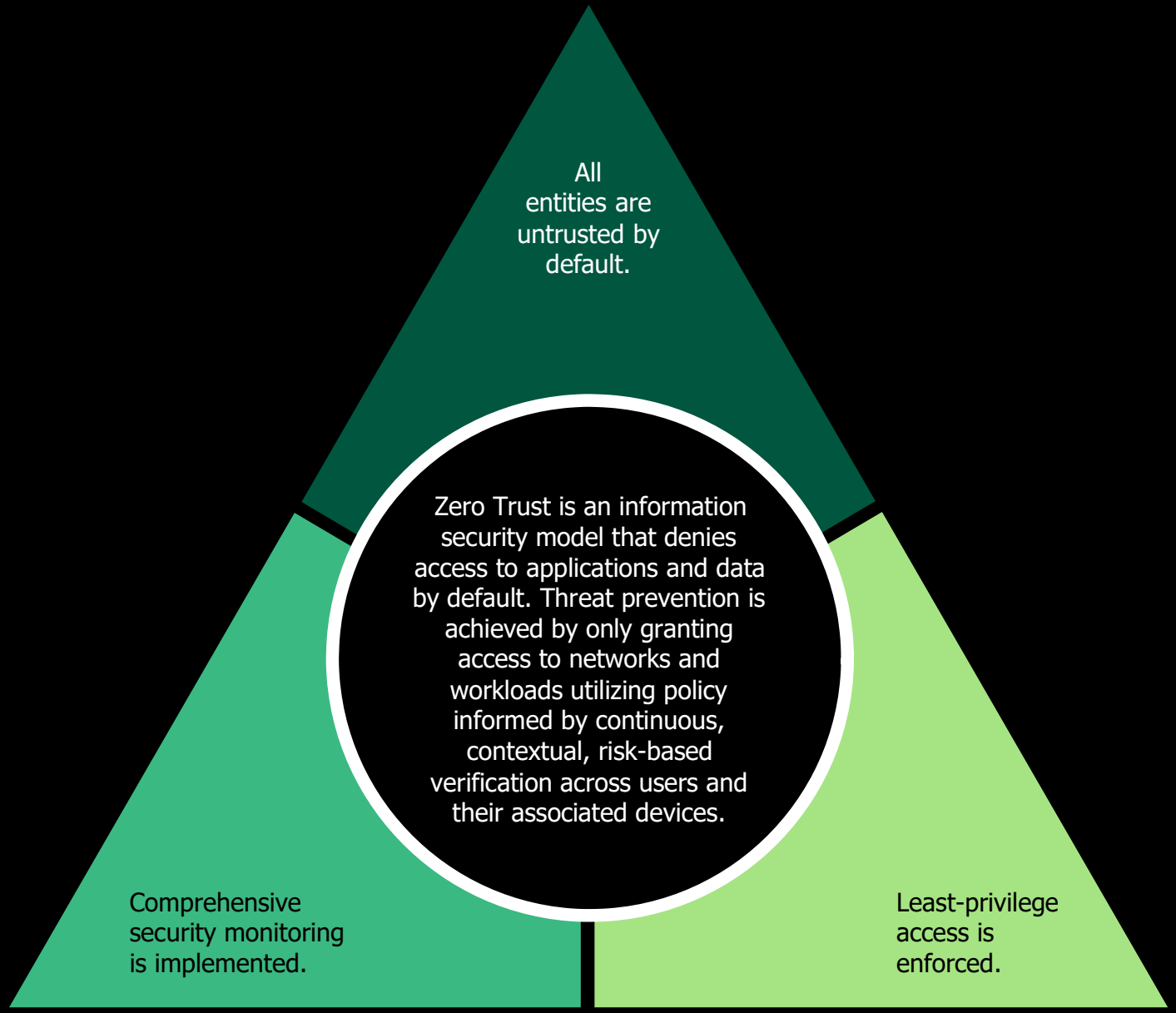
Source: Forrester's Q4 2025 Tech Pulse Survey
Base: 155 IT/technology decision makers

The benefits of platforms

"What most influenced your organizations decision to use security platform vendors in your security program?"



INTERNAL




Thank you!