ThreatLabz 2026: Relatório de preparação para IA de fronteira

Índice

Resumo executivo 3

Dados e estatísticas 5

Grupo de avaliação 6

Panorama de ameaças 7

Metodologia 9

Principais conclusões 10

A nova cadeia de ataque:
como a IA de fronteira concatena de forma autônoma pequenas falhas em uma única sequência de exploração 15

Estudo de caso:
três organizações em progresso 16

Realismo na implementação:
o que o sprint de 90 dias realmente exige 18

Recomendações 19

©2026 Zscaler, Inc. Todos os direitos reservados.

Resumo_executivo

No ano passado, recebemos uma ligação de uma das empresas de IA mais conhecidas. Tínhamos acabado de ser notificados sobre o lançamento de um modelo de IA de fronteira, otimizado para executar testes de segurança autônomos. Essa ligação nos levou a uma jornada para ajudar as organizações a se prepararem, examinando sua prontidão contra esses ataques autônomos. Ficamos tão surpresos com as descobertas quanto com suas capacidades.

Vale ressaltar que não aplicamos a IA de fronteira aos ambientes das organizações; em vez disso, primeiro a testamos em nossa própria infraestrutura, usando essas informações para desenvolver o programa de avaliação que oferecemos atualmente.

O objetivo não é apenas apontar as falhas, mas também ajudar a ilustrar onde a segurança pode ser reforçada à medida que as organizações se preparam para uma nova onda de ameaças. Estamos todos aprendendo juntos, e essas descobertas servem não apenas como um alerta, mas como uma lição compartilhada sobre como podemos melhorar a resiliência de forma geral.

Como a Zscaler opera com uma arquitetura nativa da nuvem, diferente dos ambientes híbridos e legados que a maioria das empresas utiliza, empregamos os mesmos métodos de teste de fronteira vistos em modelos reais e os replicamos com modelos e TTPs igualmente eficazes para emular como esses adversários operam nos ambientes das organizações. Isso nos permitiu criar e calibrar o programa de avaliação de acordo com as realidades operacionais da infraestrutura corporativa, em vez de arquiteturas como a nossa.

Nossa principal conclusão: a resposta mais eficaz à IA não é simplesmente mais IA, mas uma arquitetura que, por design, não ofereça caminhos exploráveis.

As empresas globais não estão preparadas para ciberataques impulsionados por IA de fronteira, mas não pelo motivo que a maioria das pessoas imagina. O perigo não reside na IA descobrir vulnerabilidades exóticas e nunca antes vistas. A questão é que os modelos de fronteira conseguem identificar vulnerabilidades óbvias, as falhas que toda empresa já sabe que existem, e concatená-las em rotas de ataque completas em velocidade de máquina. A maioria das arquiteturas nunca foi projetada para isso, e tentar superar a IA, seja com pessoas ou ferramentas, é uma batalha que estamos fadados a perder. A arquitetura é o único fator que prevalece. A boa notícia é que, com a arquitetura zero trust correta implementada, a maioria das empresas está mais perto da resiliência do que imagina.

O programa de avaliação de preparação para IA de fronteira mede a resiliência organizacional contra um tipo específico de adversário: um modelo de IA de fronteira operando em velocidade de máquina e capaz de realizar reconhecimento autônomo, encadeamento de vulnerabilidades e movimentação lateral. Usamos o termo “IA de fronteira” para descrever essa classe de adversários porque ela consegue raciocinar sobre superfícies de ataque de maneiras que as verificações tradicionais e até mesmo analistas humanos não conseguem replicar com velocidade comparável. Em testes paralelos em ambientes de produção, os modelos de fronteira demonstraram raciocínio com alta relação de sinal-ruído e contexto fundamentado, atuando efetivamente como multiplicadores de força para a descoberta de cadeias de ataque. Essas capacidades não são teóricas; elas já foram demonstradas.

©2026 Zscaler, Inc. Todos os direitos reservados.

Em centenas de organizações empresariais avaliadas no primeiro semestre de 2026, incluindo empresas da Fortune 500 nos setores de serviços financeiros, saúde, manufatura, energia e oito outros setores, a pontuação média de preparação para IA de fronteira foi de 37 em 100. Quase todas as empresas (91%) possuía infraestrutura de IA acessível publicamente, e nenhuma delas (100%) possuía controles de acesso baseados em identidade. Além disso, nenhuma organização inspecionou tráfego criptografado suficiente para que sua infraestrutura de segurança funcionasse conforme o planejado. Em média, as empresas desconheciam 64% dos dados que trafegavam em suas próprias redes, e nenhuma havia conseguido eliminar a brecha entre possuir recursos de defesa e, de fato, operá-los.

A lacuna emergente mais urgente reforça a constatação estrutural acima: a governança de IA está no ano zero em todas as empresas avaliadas. As superfícies de IA, incluindo copilots, APIs de inferência, endpoints MCP e gateways de IA paralelos, estão proliferando mais rapidamente do que os programas de segurança conseguem absorvê-las e governá-las.

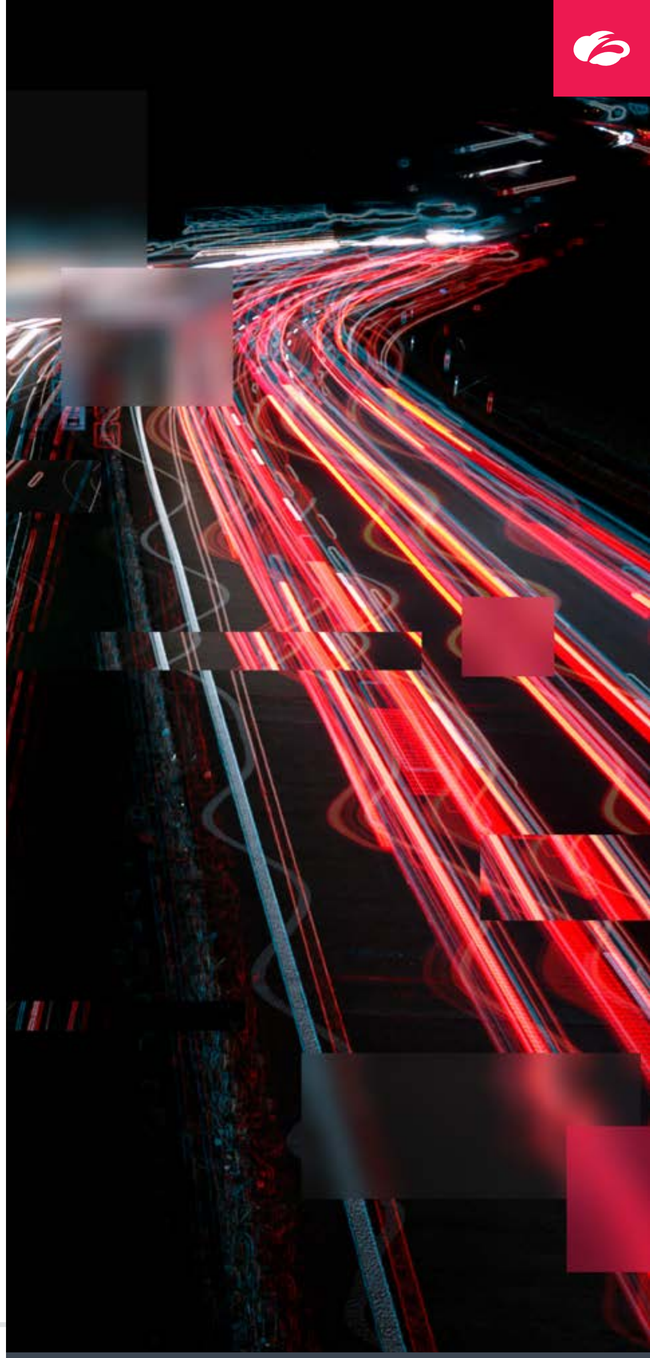
Isso representa uma oportunidade significativa para aumentar a preparação contra as ameaças de IA de fronteira. Descobrimos que quase um terço das organizações poderia melhorar drasticamente a proteção ao estender a tecnologia de deception a pontos críticos de controle. Da mesma forma, a implementação completa da prevenção de intrusões, a integração da detecção de anomalias na camada de aplicação e o aprimoramento das políticas de segmentação representam medidas concretas e de alto impacto para melhorar a fidelidade dos alertas do SOC e a postura defensiva geral.

Os líderes de segurança devem usar este relatório como um ponto de partida para discussões no nível do conselho. É possível, em um prazo de 90 dias, elevar a média da pontuação de preparação para IA de fronteira do grupo, passando de 37 para algo entre 73 e 98, nível em que os cinco pilares defensivos atingem eficária operacional contra ataques de nível de IA de fronteira. Isso exige três ações executivas: (1) estabelecer a resiliência em IA de fronteira como uma diretriz corporativa com acompanhamento pelo conselho, (2) garantir o alinhamento executivo interno e a capacidade dedicada para o sprint e (3) executar um sprint em duas frentes, ativando controles inativos e, simultaneamente, controlando o acesso a todas as superfícies de IA.

O ponto de partida operacional é o mesmo para todas as organizações: comece com o que você já tem. O ponto de partida estratégico é a governança: nenhuma organização pode adiar a segurança de acesso à IA sem aceitar a exposição a vulnerabilidades exploráveis hoje.

DEEPEN DESAI,
vice-presidente executivo de cibersegurança da Zscaler

GEORGE MOSER,
vice-presidente global e diretor de CXO residente da Zscaler



©2026 Zscaler, Inc. Todos os direitos reservados.

Dados_ e estatísticas

100% das organizações avaliadas não possuem autenticação baseada em identidade para sua IA.

91% das organizações com infraestrutura de IA acessível publicamente a deixam completamente exposta a ataques externos.

37 de 100 é a pontuação inicial média de preparação para IA de fronteira. A pontuação mínima foi 17. A máxima foi 60.

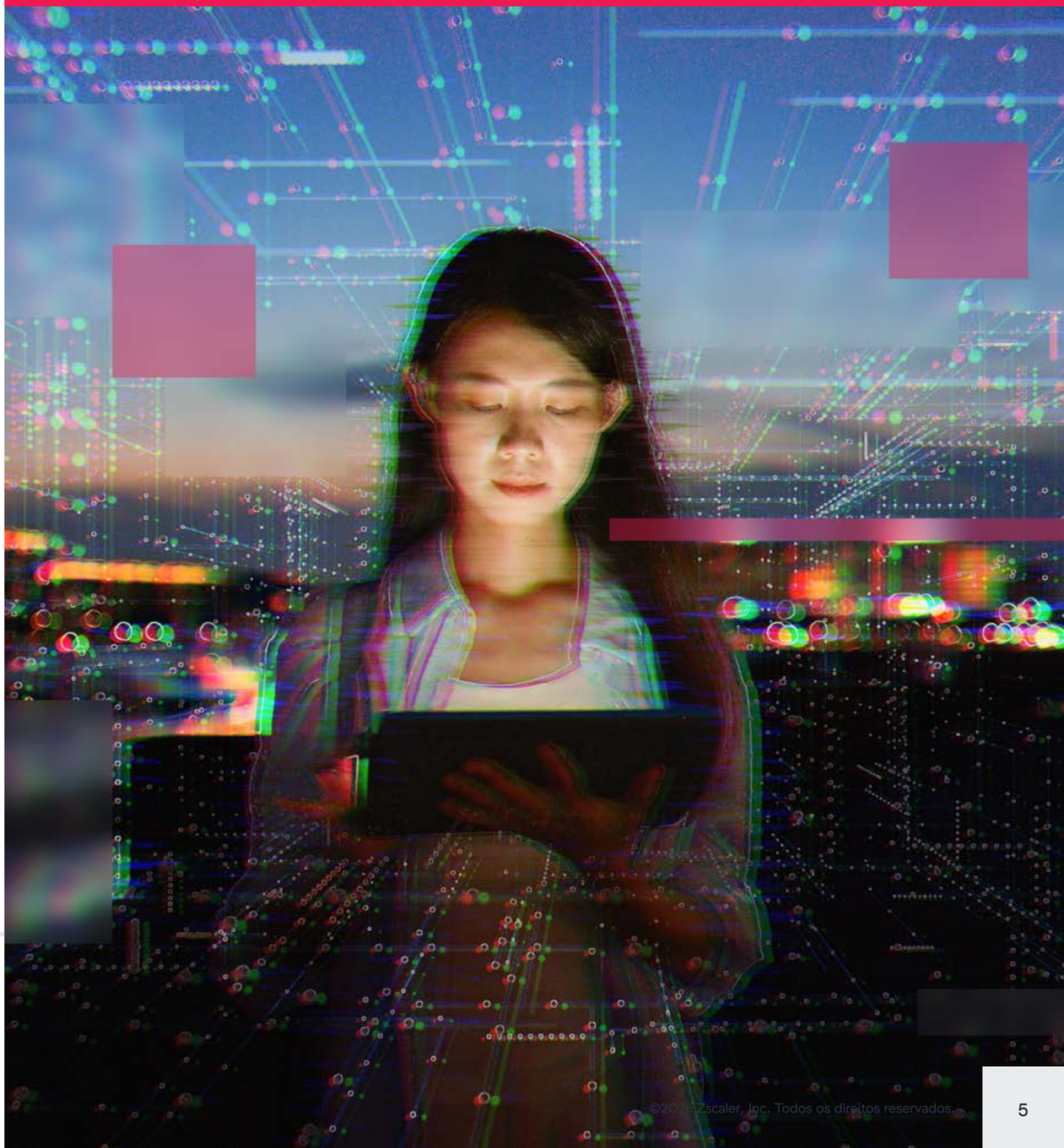
As taxas de inspeção SSL variam de **0,4% a 68,1%, com uma média de aproximadamente 36%**. Isso está bem abaixo do nível de inspeção de 70% considerado necessário para uma cobertura de segurança aceitável com controles de segurança inline (DLP, IPS, sandbox, saída de IA).

483.250 aplicativos estavam acessíveis pela identidade com mais privilégios no cenário de maior raio de impacto.

Em um caso representativo envolvendo wildcards, 195 domínios com wildcard foram suficientes para expor mais de **108.033** aplicativos. (Um wildcard é um símbolo, como "*" ou um "." inicial, usado em políticas para representar múltiplos domínios ou caminhos simultaneamente, introduzindo eficiência administrativa, mas ampliando drasticamente a superfície de ataque quando não controlada adequadamente.)

9 em cada 10 organizações têm dispositivos de VPN expostos. Cada um deles representa uma superfície de ataque pré-autenticação que um modelo de IA de fronteira pode sondar em velocidade de máquina.

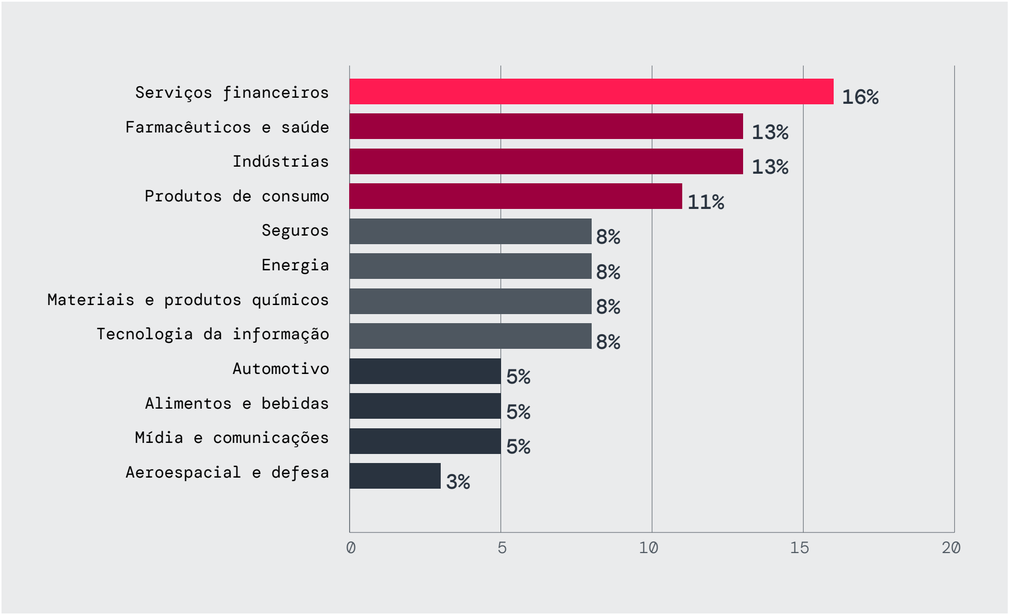
Mais de 750 ativos apresentam vulnerabilidades conhecidas e exploradas pela CISA em infraestruturas voltadas para a internet, confirmadas por meio de sondagens ativas, com exploits sendo utilizados ativamente em ambientes reais no momento da avaliação.



©2026 Zscaler, Inc. Todos os direitos reservados.

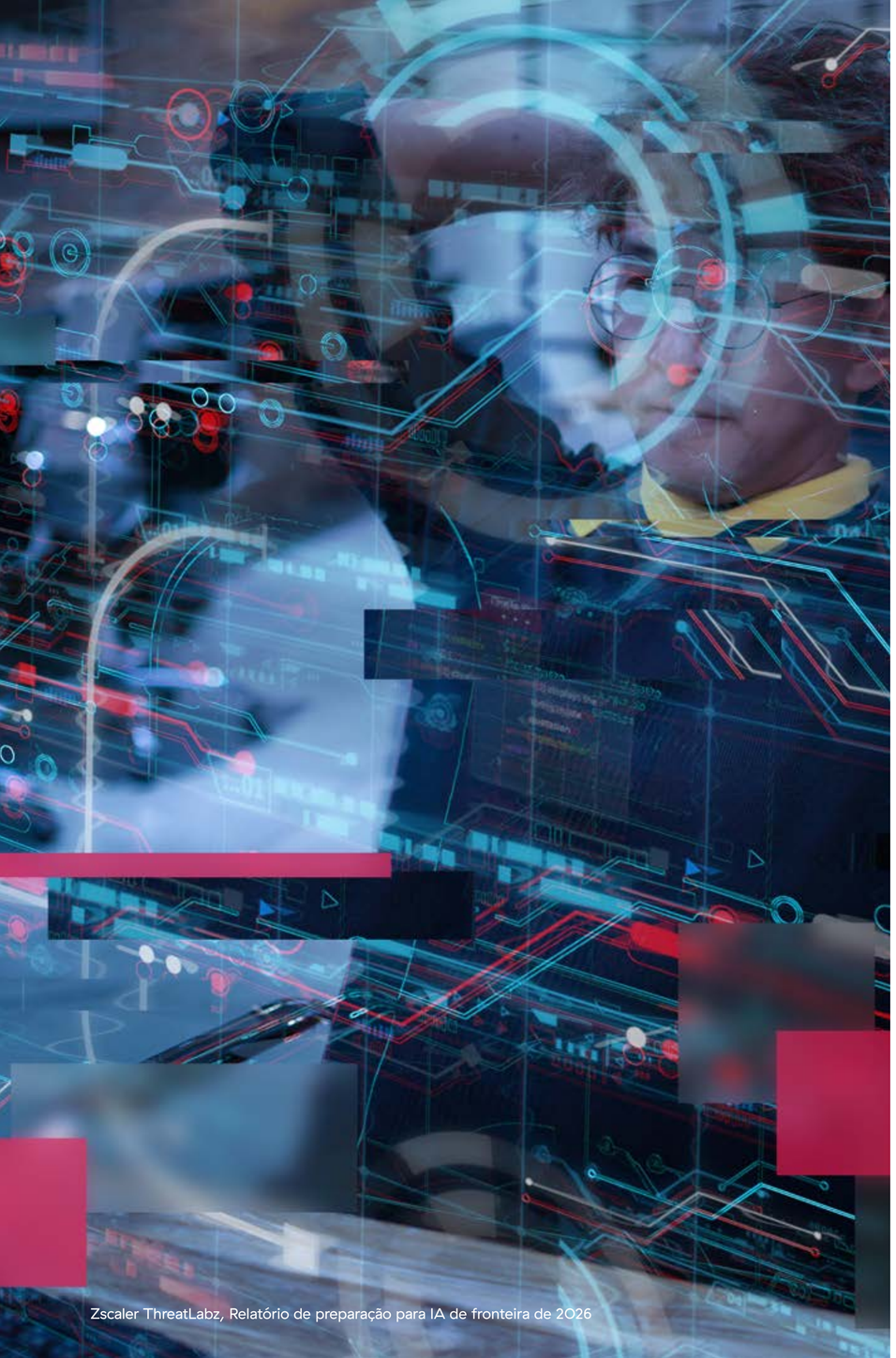
Grupo_ de avaliação

As conclusões deste relatório são provenientes de avaliações de preparação para IA de fronteira realizadas em centenas de organizações corporativas durante o primeiro semestre de 2026. 32% são empresas listadas na Fortune 500. O grupo abrange 12 setores com a seguinte composição:



As organizações avaliadas eram principalmente grandes empresas com dezenas de milhares de ativos localizáveis na internet. Ativos detectáveis na internet incluem qualquer sistema, aplicativo ou dispositivo acessível a partir da internet pública sem autenticação.

©2026 Zscaler, Inc. Todos os direitos reservados.



Panorama_ de ameaças

Os modelos de IA de fronteira demonstraram a capacidade de identificar e concatenar vulnerabilidades corporativas em velocidade de máquina.

As conclusões acima descrevem o estado atual: condições observadas em todo o grupo no primeiro semestre de 2026. A perspectiva de ameaças abaixo explica por que o cronograma de ativação é importante. As capacidades dos adversários que exploram essas condições estão amadurecendo mais rapidamente do que os ciclos operacionais das empresas normalmente se movem. Cada tendência abaixo amplifica diretamente uma ou mais conclusões da avaliação.

No programa de testes de modelos de fronteira da Zscaler, foram avaliados modelos de IA de fronteira em três modos: caixa-preta, sem conhecimento interno; caixa-branca, com acesso a código-fonte, binários compilados e arquivos estáticos; e caixa-cinza, com contexto parcial ou completo do sistema. Esses testes confirmaram, de forma independente, que, quando recebem um contexto fundamentado, os modelos de IA de fronteira conseguem raciocinar em ambientes corporativos com alta relação sinal-ruído e identificar consistentemente cadeias de exploração de múltiplas etapas que analistas humanos e verificações tradicionais não detectaram. Em arquiteturas de produção, o contexto fundamentado provou ser um fator multiplicador para a descoberta de rotas de ataque.

Essa validação independente confirma que as brechas de configuração identificadas nas conclusões 1 a 6 são exploráveis por um modelo de IA de fronteira que opera de forma autônoma e que os modelos de fronteira atuais podem tirar proveito dessas condições específicas observadas em todo o grupo avaliado.

A descoberta de vulnerabilidades impulsionada por IA muito provavelmente acelerará a disponibilidade de novos exploits até 2027

O tempo médio para exploração se tornou negativo: as vulnerabilidades agora são exploradas antes que as correções possam ser implementadas em ambientes corporativos, mesmo com a aplicação automática de correções. O desafio é estrutural, pois os prazos de exploração diminuíram mais rapidamente do que a capacidade de adaptação da gestão de mudanças corporativas.

Para ativos expostos na internet com KEVs e pontuações EPSS iguais ou próximas de 1,0, conforme confirmado em todo o grupo avaliado, o intervalo entre a divulgação e a exploração ativa diminuiu para poucas horas. Empresas que dependem da aplicação de correções durante as janelas de manutenção padrão quase certamente enfrentarão a exploração de vulnerabilidades conhecidas durante o ciclo de correção.

©2026 Zscaler, Inc. Todos os direitos reservados.

A expansão da superfície de ataque de IA provavelmente superará os programas de governança nos próximos 12 a 18 meses

A ausência generalizada de governança de IA reflete uma defasagem estrutural entre a velocidade de implantação da IA e a adaptação dos programas de segurança. Equipes de desenvolvimento, fornecedores de SaaS e unidades de negócios estão implementando copilots, terminais de inferência e interfaces de MCP sem revisão de segurança. Com base na taxa de proliferação observada, a superfície de IA não governada provavelmente dobrará antes que a governança atinja cobertura operacional. As empresas que atrasarem o processo provavelmente enfrentarão exploração durante o período de avaliação.

A movimentação lateral por meio de comprometimento de identidade muito provavelmente se tornará a técnica dominante de pós-exploração contra ambientes zero trust

Os resultados relacionados a wildcards indicam que o comprometimento de identidades é o caminho de menor resistência. É muito provável que um adversário do nível de IA de fronteira priorize o roubo de credenciais em vez da exploração de rede tradicional. A recompensa é desproporcional: milhares de aplicações através de uma única identidade. Organizações com sessões que nunca expiram e sem validação de postura do dispositivo são quase certamente os primeiros alvos.

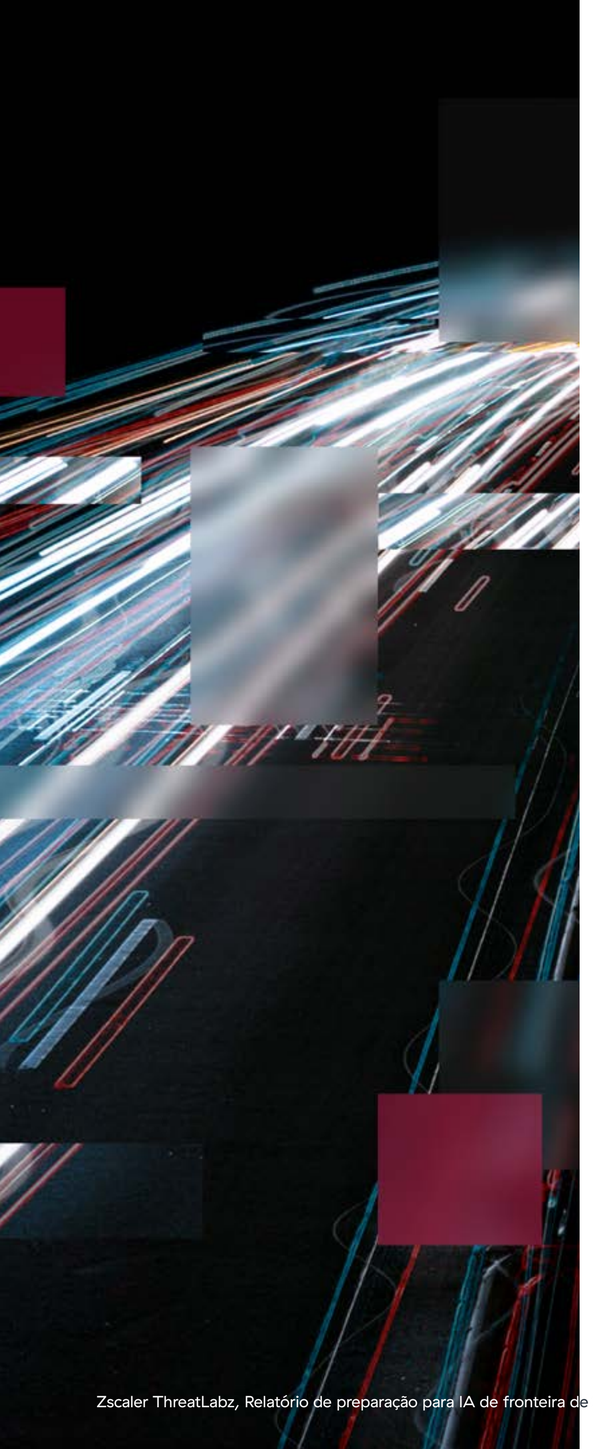
O ponto cego do tráfego criptografado provavelmente se tornará o principal canal de exfiltração para roubo de dados impulsionado por IA

A IA de fronteira consegue codificar dados sensíveis em chamadas de API criptografadas, interações de modelos ou tráfego em nuvem. Ferramentas de IA que geram padrões com aparência legítima, combinadas com sistemas de segurança incapazes de inspecioná-los, criam um canal de exfiltração invisível e de alta largura de banda. É provável que isso já esteja sendo explorado e quase certamente se intensificará.

O encadeamento autônomo de exploits provavelmente reduzirá para minutos o tempo decorrido entre o acesso inicial e o comprometimento total

A convergência de dispositivos de VPN expostos, vulnerabilidades KEV, segmentação por wildcards e detecção inativa cria um ambiente onde um adversário de IA provavelmente pode progredir do reconhecimento ao amplo acesso interno sem desencadear qualquer resposta. Os dados sugerem que essa cadeia completa é viável hoje em dia para a maioria do grupo. Ferramentas autônomas capazes de executar processos de ponta a ponta provavelmente estarão disponíveis para usuários avançados dentro de 12 meses.

©2026 Zscaler, Inc. Todos os direitos reservados.



Metodologia

Validação de ameaças de IA de fronteira

Utilizamos o GPT-5.5 da OpenAI e o Claude Opus 4.8 da Anthropic em condições controladas de uso, com base em um contexto corporativo fundamentado. Usamos a expressão "nível de IA de fronteira" porque o modelo específico importa menos do que o limiar de capacidade: raciocínio autônomo em várias etapas, uso de ferramentas, recuperação de contexto fundamentado e sequenciamento de exploração autodirigido. Qualquer modelo igual ou superior a esse limiar, atual ou futuro, representa a classe de adversários que esta avaliação mede.

Em paralelo, a Zscaler realizou testes internos de modelos de fronteira usando três modos: caixa-preta (conhecimento interno zero), caixa-branca (código-fonte, binários compilados, arquivos estáticos) e caixa-cinza (contexto parcial ou completo do sistema). Esses testes confirmaram que os modelos de IA de fronteira alcançam um alto nível de raciocínio com relação sinal-ruído em ambientes corporativos quando operam com um contexto fundamentado.

Estrutura de pontuação

Cada organização foi avaliada de acordo com o modelo de ameaças de IA de fronteira, utilizando uma pontuação composta ponderada de 100 pontos em cinco pilares defensivos: esconder-se do ataque (25 pontos, 25%), limitar raio de ação (25 pontos, 25%), governança do uso de IA (20 pontos, 20%), proteger os dados (20 pontos, 20%), e deception e iscas (10 pontos, 10%). A pontuação avaliou tanto a titularidade do controle quanto o status de ativação operacional.

Fontes de dados

As avaliações foram baseadas em várias categorias de dados, incluindo gerenciamento da superfície de ataque externa (EASM), priorização de vulnerabilidades orientada por IA, descoberta de superfície específica de IA, revisões de configuração de acesso privado e, quando possível, telemetria inline abrangendo cobertura de inspeção, aplicação de políticas e postura de autenticação.

Modelo de avaliação

Com base na visibilidade abrangente do EASM, o mecanismo Opus validou e priorizou as descobertas para identificar problemas exploráveis confirmados que exigiam ação, refinando o conjunto de candidatos em 84%, em média, e em 92% no caso mais complexo. Além disso, ele reavaliou as CVEs com base no ROI da remediação, cruzou as descobertas com o CISA KEV e identificou rotas de exploração de várias etapas em toda a superfície de ataque validada.

O que essa avaliação abrange (e o que não abrange)

A avaliação analisa a exposição da superfície de ataque externa e o estado de configuração dos controles de segurança. Ela não avalia a eficácia da detecção do SOC no nível das regras do SIEM, a preparação para resposta a incidentes, a postura de segurança do provedor de identidade ou a cobertura de detecção de terminais. A pontuação reflete a probabilidade de um adversário auxiliado por IA conseguir penetrar, atravessar e exfiltrar recursos sem encontrar um controle eficaz. Ela não representa uma pontuação total da postura de segurança.

Fundamentalmente, ataques de IA de fronteira podem comprimir a cadeia de destruição, do reconhecimento à exfiltração, em minutos, abaixo do limite de resposta até mesmo de SOCs maduros e rápidos o suficiente para explorar falhas de configuração antes da intervenção humana.

©2026 Zscaler, Inc. Todos os direitos reservados.

Principais descobertas

Conclusão 1

As empresas podem maximizar suas capacidades existentes de defesa baseadas em IA com alguns passos simples

As empresas podem melhorar consideravelmente suas defesas com práticas adequadas de higiene cibernética.

A oportunidade de ativação é ampla. Em 64% das empresas, plataformas de detecção de deception licenciadas estão posicionadas para implantação, a fim de adicionar profundidade de detecção de alta fidelidade. 15% das organizações possuem sistemas de prevenção de intrusões prontos para fazer a transição do modo de monitoramento para o modo de aplicação completa de medidas. 18% possuem recursos de autenticação que podem ser estendidos a outros locais. Somente por meio da otimização da configuração, as organizações podem atingir pontuações-alvo entre 73 e 98, em uma escala de 0 a 100. Considerando a pontuação inicial média do grupo, de 37 pontos, isso representa melhorias individuais expressivas na pontuação.

Por que isso é importante contra ameaças de nível de IA: ferramentas de segurança que operam abaixo do nível máximo de aplicação oferecem menor resistência a um adversário que opera em velocidade de máquina.

Conclusão 2

A governança de IA está no ano zero em todas as empresas avaliadas, independentemente do tamanho da organização ou da abrangência do controle

Essa constatação é universal. 100% das organizações avaliadas não possuem autenticação baseada em identidade para sua IA, e 91% têm terminais de IA expostos publicamente. A pontuação média do pilar de governança de IA foi de 2,1 em 20 (10,5% do máximo). A pontuação mais alta da turma foi 3 em 20.

A exposição é significativa: terminais públicos de IA eram comuns em toda a coorte, incluindo terminais de MCP expostos à internet e chatbots ou copilots não autenticados. Os casos mais graves observados incluíram gateways de GenAI de gravidade crítica expostos à internet pública.

Essa brecha existe porque as superfícies de IA são implantadas por equipes de desenvolvimento, unidades de negócios e fornecedores terceirizados em um ritmo que supera a adaptação dos programas de segurança. A governança de IA não é uma iniciativa para o futuro. Trata-se de uma necessidade atual e passível de exploração em todas as organizações avaliadas, além de ser a categoria de superfície de ataque não gerenciada que cresce mais rapidamente.

©2026 Zscaler, Inc. Todos os direitos reservados.

Conclusão 3

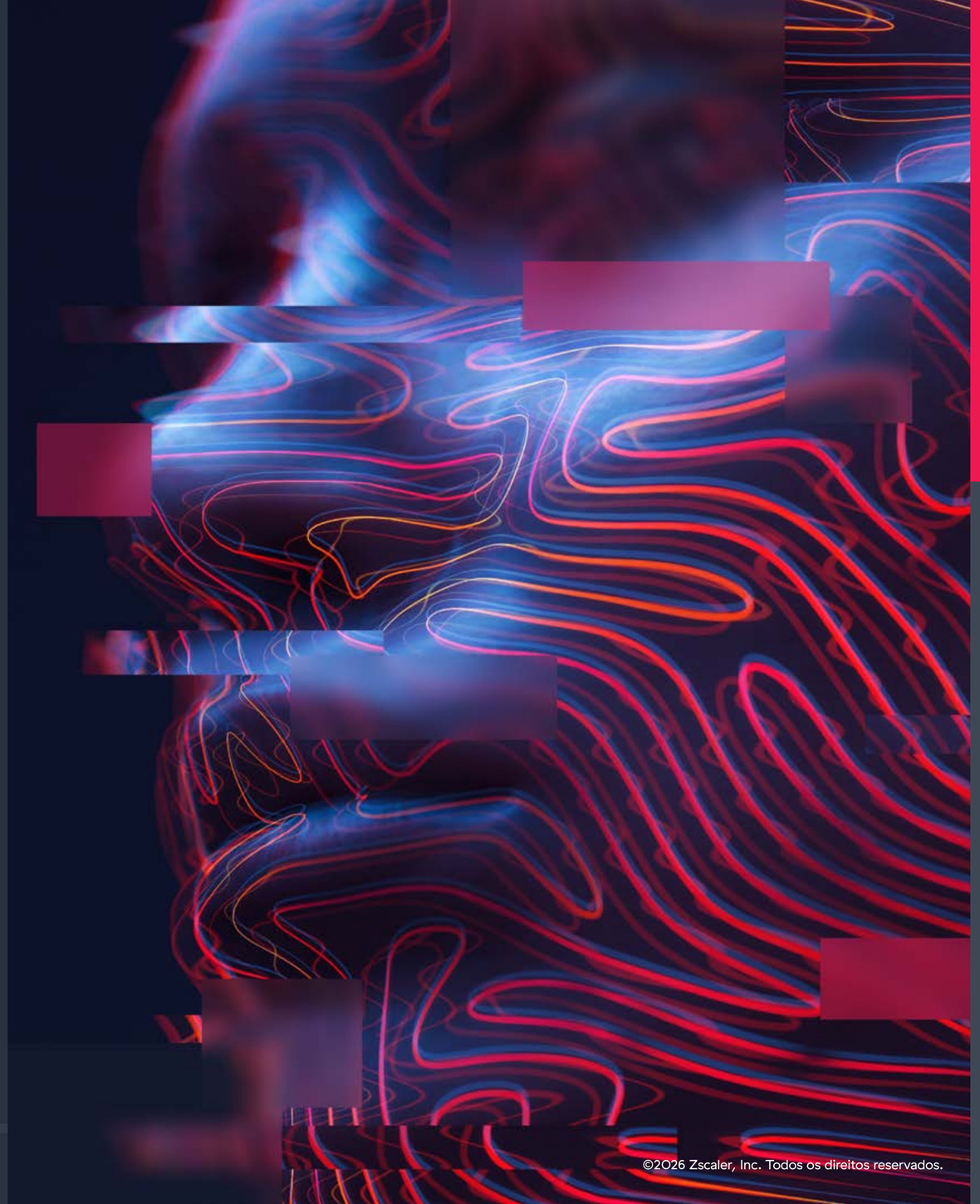
A inspeção limitada do tráfego criptografado reduz a eficácia dos controles de segurança implementados na maior parte do tráfego

A inspeção do tráfego criptografado tem sido uma prioridade de operações de segurança há muito tempo. No contexto de ameaças impulsionadas por IA, a cobertura completa da inspeção torna-se crucial. Os modelos de IA de fronteira que realizam exfiltração de dados quase certamente usarão canais criptografados, particularmente chamadas de API de IA, uploads para armazenamento em nuvem e comunicações entre modelos que se misturam com padrões de tráfego legítimos. Um adversário que opera em tráfego não inspecionado não é visível para os controles inline. Isso eleva a inspeção de tráfego criptografado de uma prática operacional recomendada a um requisito fundamental contra ameaças baseadas em IA.

Em média, uma empresa inspeciona cerca de 36% do seu tráfego criptografado. Nenhuma organização atinge o limite de inspeção qualificada de 70% definido nesta avaliação como a proporção de tráfego criptografado que passa por controles de segurança ativos em linha (DLP, IPS, sandbox, saída de IA) suficiente para que esses controles alcancem a cobertura de segurança ideal. Na pior das hipóteses, 99,6% de todo o tráfego criptografado passa pela infraestrutura de segurança sem ser inspecionado.

Quando o tráfego criptografado permanece sem inspeção, os controles de segurança inline, como DLP, IPS, sandbox e governança de saída de IA, tornam-se ineficazes, deixando as organizações alheias a ameaças críticas como exfiltração de dados e distribuição de malware. Considerando que mais de 86% das ameaças cibernéticas e todas as interações com modelos de IA utilizam canais criptografados, a inspeção abrangente do tráfego é um pré-requisito indispensável para a defesa contra ataques de IA.

As causas principais são operacionais, não técnicas. A grande maioria das organizações utiliza um modelo de inspeção opcional (com a opção padrão definida como “não inspecionar”). Em todo o grupo, as organizações frequentemente dependem de regras de SSL do tipo “Ignorar outras políticas” ou “Não inspecionar”, que desativam a inspeção para categorias inteiras de tráfego. Entre as organizações afetadas, o número dessas regras de desvio variou de 2 a 51.



©2026 Zscaler, Inc. Todos os direitos reservados.

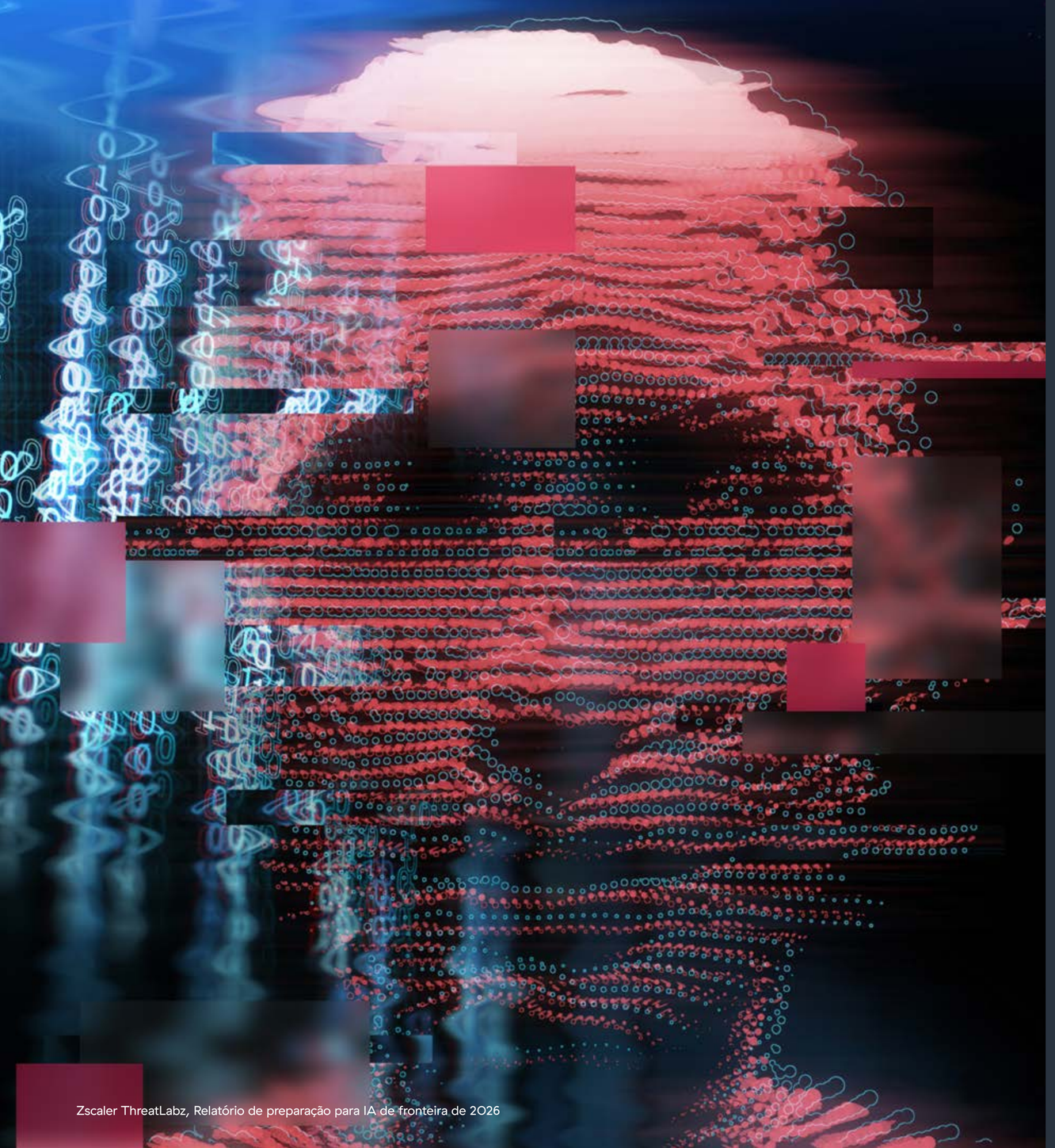
Conclusão 4

Os ambientes avaliados ainda não estão configurados para detectar ou prevenir ataques baseados em IA na velocidade necessária

A conclusão 4 destaca uma brecha crítica de detecção, na qual os controles de segurança ativados falham porque suas configurações (em áreas como deception, IPS, detecção de anomalias e monitoramento de sessões) não estão otimizadas para ataques de nível de IA, deixando as organizações expostas a ameaças que elas são capazes de bloquear, mas incapazes de identificar. Consequentemente, um ataque de IA de fronteira permanece invisível para o SOC, sem gatilhos para deception, IPS, anomalias comportamentais ou tempos limite de sessão, o que significa que a primeira indicação de comprometimento provavelmente será uma notificação externa em vez de uma detecção interna.

- **Deception (o sistema de alerta precoce):** 64% das organizações possuem plataformas de deception com cobertura ainda não estendida a segmentos críticos. Um atacante que se move lateralmente não encontraria mecanismos de detecção na configuração atual. Os recursos de deception proporcionam detecção eficaz e independente de assinaturas contra ataques baseados em IA, utilizando iscas de alta fidelidade para acionar a contenção imediata e automatizada da movimentação lateral.
- **Prevenção de intrusões (bloqueio inline):** para 15% das organizações, existe uma clara oportunidade de fortalecer as defesas ao colocar seus sistemas de prevenção de intrusões em modo de aplicação total. Quando ativadas, as regras de “Ignorar outras políticas” significavam que a maior parte do tráfego nunca era avaliada. Uma vulnerabilidade conhecida, transitando por canais criptografados, passaria sem inspeção e sem ser bloqueada.
- **Detecção de anomalias na camada de aplicação (o sinal comportamental):** várias empresas ainda não haviam configurado regras inline. Não havia detecção de anomalias no volume de travessia com wildcards. Uma única identidade acessando milhares de aplicações em sequência muito provavelmente não geraria nenhum alerta.
- **Monitoramento de sessão e identidade (o gatilho de contenção):** tempos limite de autenticação configurados sem expiração em várias organizações significam que sessões comprometidas podem permanecer válidas indefinidamente. Em uma das empresas, nenhuma das 18 regras de acesso realizava verificações de postura do dispositivo. Uma credencial roubada poderia operar por semanas sem acionar a reautenticação.
- **Como um ataque de IA de fronteira provavelmente se apresenta para o SOC:** o SOC não vê nada. Nenhum gatilho de deception. Nenhum alerta do IPS. Nenhuma anomalia comportamental. Nenhum tempo limite de sessão. O primeiro indício de comprometimento provavelmente seria uma notificação externa, e não uma detecção interna.

©2026 Zscaler, Inc. Todos os direitos reservados.



Conclusão 5

A segmentação ampla transformou plataformas de acesso privadas em vias expressas para movimentação lateral

Em nossa análise, 64% das organizações que usam ZTNA tinham configurações com uso excessivo de wildcards, o que pode expor mais aplicativos do que o pretendido. Em um caso específico, apenas 195 domínios wildcard tornaram mais de 108 mil aplicativos acessíveis, potencialmente permitindo que uma única identidade acessasse uma parcela muito ampla do ambiente. Embora sejam frequentemente introduzidas durante a implementação ou migração para simplificar a integração, essas configurações podem persistir à medida que os ambientes crescem, expandindo o acesso além da intenção original. Sem revisão e aprimoramento frequentes, políticas baseadas em wildcards podem prejudicar os objetivos de privilégio mínimo e aumentar o impacto de uma identidade comprometida, como por meio de phishing ou roubo de sessão.

Diversas organizações agravam essa situação configurando os tempos limite de autenticação para “nunca”. Uma sessão comprometida permanece válida indefinidamente. Mesmo a organização com a maior pontuação em Limit Blast Radius (14,4 de 25) ainda manteve configurações do tipo wildcard. Organizações com pontuação mais baixa não apresentavam, na prática, nenhuma segmentação.

Esse padrão significa que o ambiente de acesso privado ainda não reflete a segmentação de privilégio mínimo que as arquiteturas zero trust visam aplicar. Contra um adversário em nível de IA de fronteira capaz de comprometer identidades em larga escala, as configurações com predominância de wildcards representam o caminho de movimentação lateral de maior impacto nos ambientes avaliados.

©2026 Zscaler, Inc. Todos os direitos reservados.

Conclusão 6

Mais de um terço das empresas mantém vulnerabilidades conhecidas e exploradas em sua infraestrutura voltada para a internet

Essas exposições muitas vezes persistem não por serem desconhecidas, mas porque a remediação precisa competir com realidades operacionais, como janelas de manutenção, requisitos de controle de mudanças e dependências de infraestrutura. Mais da metade do grupo (53%) recebeu recomendações de correção para vulnerabilidades conhecidas e exploradas (KEVs) da CISA em ativos expostos à internet. Durante as sondagens ativas, mais de um terço das empresas mantiveram essas KEVs exploráveis em seu perímetro.

O resultado é uma crescente discrepância entre a rapidez com que os defensores conseguem corrigir os problemas e a rapidez com que os atacantes conseguem identificar e explorar as vulnerabilidades. À medida que os adversários utilizam cada vez mais a IA para acelerar a descoberta e a exploração de vulnerabilidades, é provável que essa janela de exposição se reduza ainda mais.

Três CVEs apareceram na maioria do grupo. Todas constam no catálogo de vulnerabilidades exploradas conhecidas da CISA. Todos os casos foram confirmados por meio de sondagem ativa.

- **CVE-2024-38475 (Apache mod_rewrite):** pontuação EPSS de 0,9988 a 0,9997
- **CVE-2021-40438 (Apache mod_proxy SSRF):** pontuação EPSS de 0,9999 a 1,0000
- **CVE-2023-44487 (HTTP/2 Reinicialização rápida):** pontuação EPSS 0,999

Sua persistência, mesmo em organizações sofisticadas, reflete uma realidade operacional: a infraestrutura da camada web, particularmente Apache e F5/NGINX, frequentemente fica fora dos processos padrão de automação de correções e exige coordenação manual entre as janelas de mudança.

Embora a exposição atual reflita vulnerabilidades conhecidas, as organizações que possuem KEVs conhecidas em seu perímetro também enfrentam o risco de que a exploração de vulnerabilidades ainda não divulgadas nos mesmos sistemas possa surgir antes que a remediação existente seja concluída.

A conclusão 5 amplifica ainda mais esse risco, pois cada recurso exposto à internet sem a devida correção representa um ponto de entrada de pré-autenticação e, em ambientes onde uma única identidade pode alcançar centenas de milhares de aplicativos, uma única CVE explorada pode estender o acesso muito além do servidor afetado.

As ações prioritárias são manter a visibilidade contínua dos ativos por meio do gerenciamento externo da superfície de ataque, reduzir a superfície de ataque detectável por meio do acesso intermediado, encurtar os ciclos de detecção, triagem e correção para ativos expostos justificados pela empresa por meio da priorização baseada na explorabilidade e usar a aplicação autônoma de correções virtuais durante as janelas de remediação.

©2026 Zscaler, Inc. Todos os direitos reservados.

A nova cadeia de ataque:_ como a IA de fronteira concatena de forma autônoma pequenas falhas em uma única sequência de exploração

Essas descobertas não são isoladas. Elas estão reforçando elos em uma cadeia de exploração mais ampla. Individualmente, cada uma representa uma falha que pode ser solucionada; juntas, elas criam uma exposição materialmente maior do que qualquer descoberta isolada sugere.

Um modelo de IA de fronteira com contexto fundamentado pode identificar e sequenciar autonomamente condições como uma CVE não corrigida, um IPS ainda não implementado, configurações de wildcards, terminais de IA sem governança e rotas de saída não inspecionadas, transformando-as em uma cadeia contínua de exploração, em velocidade de máquina e sem intervenção humana. As verificações tradicionais e os testes de penetração não revelam esse caminho complexo com velocidade comparável. A seguinte sequência foi validada em todos os ambientes avaliados:

Etapa 1: enumerar um dispositivo de VPN exposto (um de mais de 500) por meio de varredura na internet. Identificar CVE-2024-38475 (Apache mod_rewrite, EPSS 0,9997) na camada web.

Etapa 2: explorar a vulnerabilidade de pré-autenticação. Obter acesso inicial sem credenciais. O sistema IPS está desativado. Nenhum alerta é gerado.

Etapa 3: extrair credenciais da infraestrutura de VPN comprometida. Obter uma única identidade corporativa válida.

Etapa 4: autenticar-se na plataforma de acesso privado. A configuração de wildcards concede acesso a mais de 17.000 aplicativos. O tempo limite está definido como "nunca". Nenhuma verificação de postura do dispositivo.

Etapa 5: percorrer os aplicativos em velocidade de máquina. Sem iscas de deception no caminho. O SOC não detecta nenhum alerta.

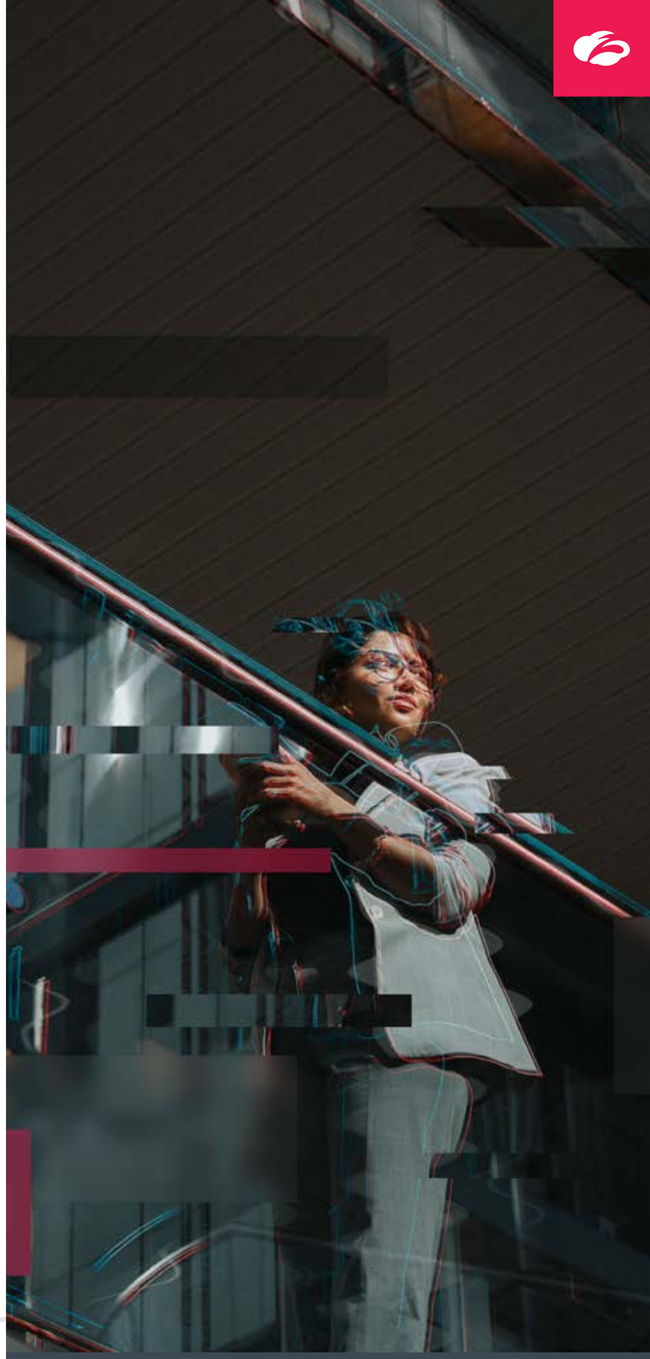
Etapa 6: alcançar um terminal de inferência de IA não governado (um de mais de 300) acessível sem autenticação adicional. Consultar dados proprietários através da janela de contexto do modelo.

Etapa 7: exfiltrar dados por meio de canal criptografado. Dois terços do tráfego não são inspecionados. A DLP não consegue ver a saída. Nenhum alerta é gerado.

Tempo total desde a verificação inicial até a exfiltração: durante testes de adversários em ambientes internos, os sistemas de IA corporativos atingiram um tempo médio até a primeira falha crítica de apenas 16 minutos.

No programa de testes de modelos de fronteira da Zscaler (de nível de IA de fronteira e GPT 5.5 em três modos: conhecimento interno zero, com acesso ao código-fonte e com contexto parcial/total do sistema), essa capacidade de raciocínio foi validada. O aproveitamento do contexto fundamentado permite que modelos de fronteira identifiquem caminhos de exploração em várias etapas que nem as verificações tradicionais nem os analistas humanos detectariam com a mesma rapidez.

Todos os componentes foram observados nos ambientes das empresas durante o primeiro semestre de 2026. A cadeia representa o modelo adversário em relação ao qual a pontuação de preparação para IA de fronteira é calibrada.



©2026 Zscaler, Inc. Todos os direitos reservados.

Estudo de caso: três organizações em progresso

Os seguintes estudos de caso anonimizados representam organizações em diferentes estágios de seu processo de remediação. Eles ilustram as condições iniciais e a metodologia operacional utilizada para sanar as lacunas identificadas durante a avaliação.

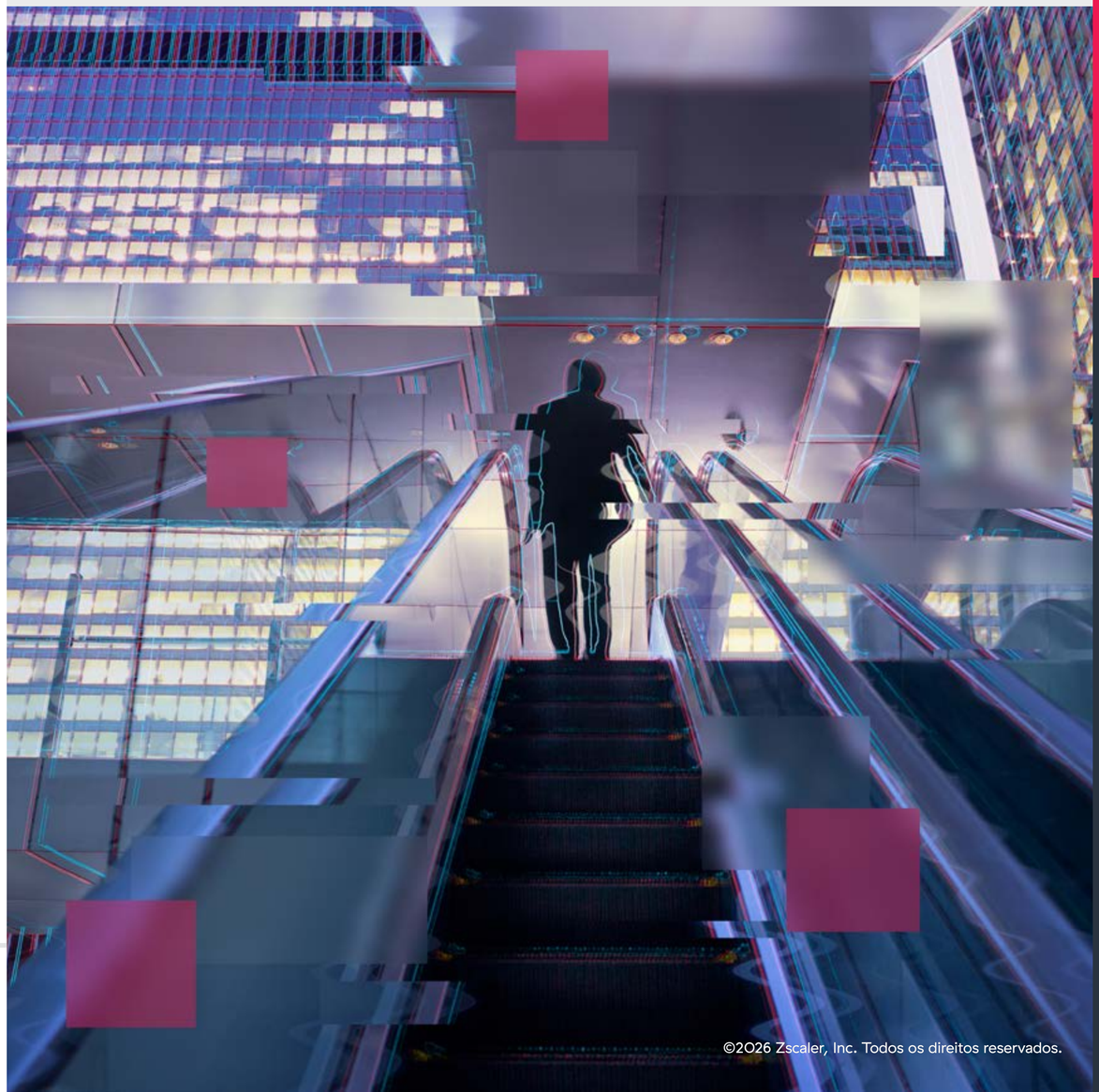
CASO 1

Organização global de serviços financeiros

Estado inicial: mais de 10.000 ativos expostos à internet. Mais de 4.000 CVEs validadas (mais de 900 únicas, 165 críticas). Três das CVEs mais prevalentes do grupo (CVE-2024-38475, CVE-2021-40438, CVE-2023-44487) foram confirmadas na infraestrutura exposta.

Resultados da avaliação: a extensa infraestrutura da organização voltada para a internet continha vulnerabilidades críticas não corrigidas, acumuladas ao longo de sucessivas gerações de infraestrutura. A inspeção de SSL ainda não havia sido estendida para abranger a maior parte do tráfego. Recursos de deception estavam autorizados, mas ainda não haviam sido implementados. As superfícies de IA permaneciam fora do perímetro de governança.

Remediação iniciada: priorização da aplicação de correções de 165 CVEs críticas. Conversão da inspeção de SSL de adesão voluntária para exclusão voluntária. Ativação de estratégias de deception em segmentos financeiros essenciais. Iniciada a filtragem de superfície por IA.



©2026 Zscaler, Inc. Todos os direitos reservados.

CASO 2

Produtor global de alimentos

Estado inicial: 13 dispositivos de VPN expostos. 40 superfícies de IA públicas fora do perímetro de governança (a maior densidade de exposição de IA em relação ao tamanho no grupo). 2 CVEs validadas na infraestrutura exposta.

Resultados da avaliação: a rápida transformação digital da organização ultrapassou a capacidade do programa de segurança de incorporar novas interfaces. Os terminais de IA foram provisionados pelas unidades de negócios como parte de iniciativas operacionais e ainda não haviam sido roteados pela arquitetura de segurança. A política de segmentação não havia sido estabelecida para o ambiente de acesso privado. A infraestrutura de VPN permanecia exposta à internet, aguardando migração.

Remediação iniciada: proteção via VPN com acesso privado (imediata). Controle de acesso por IA para 40 terminais. Criação de política de segmentação do zero. Implantação de deception em segmentos da produção.

CASO 3

Conglomerado global de artigos de luxo

Estado inicial: 487 nomes voltados ao público no escopo. 922 TB de tráfego criptografado não inspecionado mensalmente. 10.000 aplicativos definidos.

Resultados da avaliação: a arquitetura de inspeção da organização havia sido implementada com um modelo de adesão conservador, deixando a maior parte do tráfego criptografado sem ser examinada. Com 10.000 aplicativos e segmentação praticamente inexistente, o ambiente havia escalado sem a devida governança de acesso.

Remediação iniciada: conversão da inspeção de SSL como prioridade máxima. Substituição de wildcards em 10.000 aplicativos. Ativação de técnicas de deception. Estabelecimento de uma base de referência para governança de IA.

©2026 Zscaler, Inc. Todos os direitos reservados.

Realismo na implementação: o que o sprint de 90 dias realmente exige

O caminho para a resiliência é tecnicamente viável em 90 dias, porque na maioria das vezes os controles já estão definidos e as cadeias de dependência são bem compreendidas. No entanto, como não é algo fácil, precisa ser uma determinação executiva. Reduzir a lacuna de preparação para IA de fronteira é uma iniciativa corporativa que se beneficia da visibilidade em nível de conselho, do alinhamento entre CISO e CEO e da capacidade de execução dedicada. Trata-se de uma oportunidade para alinhamento mútuo em um cenário de ameaças em constante evolução. Os líderes de segurança devem:

1. Informar seus conselhos sobre o atual nível de preparação, contextualizando-o dentro das regulamentações da SEC, DORA e NIS2.

2. Obter patrocínio executivo para um sprint de 90 dias com responsabilidade definida e relatórios frequentes para a diretoria executiva.

3. Alocar pequenas equipes dedicadas para realizar a implementação, tendo em mente que os atritos decorrentes da expansão da cobertura de inspeção, ajustes de acesso durante a substituição de wildcards e situações semelhantes são uma parte esperada do processo de ativação.

A inspeção de SSL exige alinhamento interfuncional. As organizações devem coordenar o trabalho com as equipes de TI e segurança para mapear completamente a cobertura de inspeção atual, identificar quais fluxos de tráfego exigem inspeção e determinar onde a inspeção é operacionalmente inviável (por exemplo, aplicativos com certificados fixados ou integrações legadas). Quando a inspeção não puder ser aplicada, controles de mitigação, como isolamento de navegador, DLP de terminais, inspeção de API baseada em CASB e filtragem na camada de DNS, podem fornecer visibilidade e prevenção de ameaças compensatórias sem exigir descriptografia inline.

A substituição de wildcards exige coordenação com os proprietários dos aplicativos. Substituir mais de 200 wildcards por FQDNs específicos não é uma tarefa exclusiva da equipe de segurança. Cada wildcard está associado a aplicativos pertencentes a unidades de negócios, equipes de desenvolvimento e fornecedores que devem validar a continuidade do acesso. A equipe de segurança define a arquitetura de destino, mas a execução requer coordenação interfuncional.

O cronograma pressupõe capacidade de execução dedicada. As organizações que designaram uma equipe de sprint específica (de 2 a 4 engenheiros) progrediram conforme o planejado. Aquelas que buscaram soluções dentro das cargas de trabalho existentes se beneficiaram de suporte adicional para manter o ritmo. O sucesso delas dependeu do patrocínio executivo, incluindo visibilidade regular da alta administração, uma equipe dedicada com capacidade protegida e relatórios semanais de marcos em relação à meta estabelecida.

Embora qualquer atualização de segurança significativa possa causar alterações operacionais temporárias, essas medidas são essenciais para construir uma defesa mais resiliente a longo prazo. Nossa avaliação destaca que abordar proativamente as brechas, como a pontuação inicial de 37, é um passo crucial no gerenciamento das ameaças digitais modernas. Equilibrar esses ajustes de curto prazo com a necessidade de uma proteção robusta é fundamental para fortalecer a postura de segurança geral da organização.

©2026 Zscaler, Inc. Todos os direitos reservados.

Recomendações

1 Ative o modo de aplicação nos controles de segurança

Todas as organizações avaliadas possuíam controles de alto impacto (IPS, autenticação, bloqueio de IPs maliciosos, proteção contra ataques de front-end de domínio) que podem ser totalmente implementados por meio de alterações de configuração. As equipes de operações de segurança devem auditar todos os recursos próprios quanto ao status de ativação, aproveitando as auditorias de configuração do fornecedor e o suporte à ativação, conforme apropriado, e habilitá-los no modo de aplicação em um sprint de 14 dias.

Reconhecemos a preocupação de todas as equipes de SOC: mover o IPS do modo de monitoramento para o modo de bloqueio pode gerar falsos positivos e interromper o tráfego. A abordagem correta é uma implementação faseada: comece em segmentos não críticos, ajuste e valide, e depois expanda para produção com critérios claros de reversão.

2 Controle imediatamente o acesso a todas as superfícies de IA com reconhecimento de identidade

Todos os pontos de acesso de IA acessíveis pela internet pública devem ser protegidos por verificação de identidade corporativa. Essa medida é recomendada para 100% das organizações avaliadas. Enquanto o acesso com reconhecimento de identidade não estiver em implementado, a infraestrutura de IA permanecerá vulnerável a exploração por qualquer adversário conectado à internet.

3 Eleve a inspeção de SSL/TLS com cobertura qualificada mínima de 70% em até 90 dias

Inverta o modelo de adesão voluntária para exclusão voluntária: inspecione todo o tráfego por padrão e, em seguida, crie exclusões específicas. Elimine regras de “Ignorar outras políticas”. Reduza as exceções de “Não Inspecionar” ao mínimo necessário. Planeje um período de 2 a 4 semanas para interrupções na aplicação. Aloque capacidade extra para o suporte técnico. Sem essa cobertura, os controles de saída de DLP, IPS, sandbox e IA não podem funcionar na maior parte do tráfego.

A inspeção de SSL deve ser baseada no risco, e não universal. Comece com os aplicativos mais críticos e o tráfego de maior risco, depois expanda seletivamente. Como alguns aplicativos e serviços exigem exceções, como o bloqueio de certificados ou devido a restrições regulatórias, as organizações devem definir antecipadamente os caminhos de exceção e os procedimentos de resposta rápida. Nos casos em que a inspeção não for viável, reduza as lacunas de visibilidade com controles compensatórios.

4 Substitua domínios de wildcard por FQDNs nomeados em segmentos críticos

Audite todas as configurações de acesso privado para entradas com wildcards. Priorize a substituição nos segmentos mais essenciais: sistemas financeiros, repositórios de propriedade intelectual e infraestrutura de identidade. Implemente tempos limite de autenticação, validação da postura do dispositivo em todas as regras de acesso e segmentação em nível de aplicativo, limitando cada identidade apenas aos aplicativos necessários. Planeje a coordenação interfuncional com os responsáveis pelos aplicativos.

©2026 Zscaler, Inc. Todos os direitos reservados.

5 Implante estratégias de deception em segmentos críticos da rede.

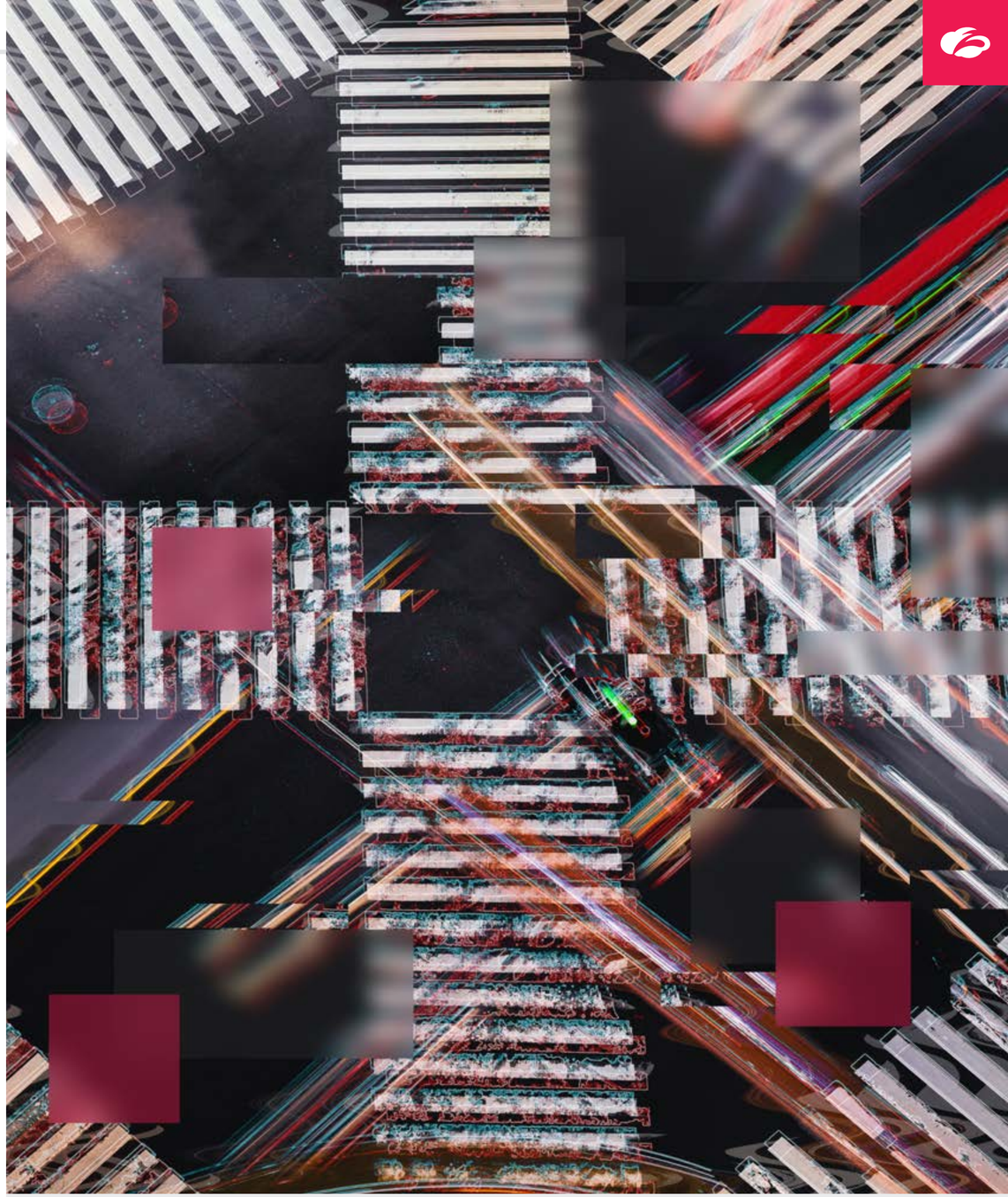
As organizações podem fortalecer suas defesas colocando iscas perto de aplicativos essenciais, criando uma camada de detecção independente com alertas de alta fidelidade que ajudam as equipes do SOC a responder mais rapidamente. Quando uma isca é acionada, o deception revoga automaticamente o acesso do usuário comprometido a aplicativos privados por meio do ZPA em todos os dispositivos conectados, sem a necessidade de aprovação do analista. Isso intercepta os invasores antes que possam realizar movimentação lateral, bloqueando essa atividade na velocidade do próprio ataque, e não na velocidade da fila do SOC.

6 Mova dispositivos de VPN expostos à internet para trás de um acesso intermediado por arquitetura zero trust

Cada dispositivo de VPN detectável representa uma superfície de ataque pré-autenticação. Implementar acesso intermediado em toda a infraestrutura de VPN em até 30 dias. Desative os aparelhos expostos em um prazo de 90 dias. Isso elimina o ponto de entrada que um modelo de IA de fronteira investigaria primeiro.

7 Corrija vulnerabilidades CISA-KEV em ativos expostos à internet em até 72 horas após a avaliação

As vulnerabilidades CVE-2024-38475, CVE-2021-40438 e CVE-2023-44487 possuem pontuações EPSS iguais ou próximas de 1,0, com exploração ativa confirmada. Corrija ou faça uma correção virtual imediatamente. Nos casos em que forem necessárias janelas de manutenção, implemente a aplicação autônoma de correções virtuais para bloquear tentativas de exploração durante esse período.



©2026 Zscaler, Inc. Todos os direitos reservados.

Nota sobre a exposição_ regulatória

Organizações com pontuação inferior a 40 podem apresentar fragilidades nos controles que aumentam tanto a probabilidade quanto o impacto potencial de um incidente cibernético significativo. Incentivamos as organizações a analisarem e implementarem as recomendações deste relatório para ajudar a reduzir a exposição aos riscos e fortalecer a resiliência. Em alguns casos, um incidente associado a essas fragilidades também pode gerar considerações de divulgação de acordo com o Item 1.05 da SEC (Formulário 8-K).

Para organizações sujeitas à DORA ou à NIS2, a ausência de controles fundamentais também pode indicar oportunidades para fortalecer a gestão de riscos de TIC e a resiliência operacional geral.

Próximos_ passos

A avaliação de preparação para IA de fronteira fornece a cada organização uma pontuação inicial quantificada, uma sequência de remediação priorizada e uma meta de 90 dias, alcançável principalmente por meio da expansão ou do aprimoramento dos recursos existentes. Os dados deste relatório estabelecem a base de referência do setor. A questão para todo líder de segurança é onde sua organização se situa em relação a isso.

Você já viu os resultados. Agora veja sua exposição.

Fale com um especialista da Zscaler para identificar onde podem existir riscos relacionados à IA em seu ambiente e o que deve ser abordado primeiro.

Conheça os riscos. Aja rápido. Mantenha-se seguro.
Saiba mais em zscaler.com/br/campaign/zscaler-ai-security-demo

Sobre_ a ThreatLabz

ThreatLabz é o braço de pesquisa de segurança da Zscaler. Essa equipe de classe mundial é responsável por perseguir novas ameaças e garantir que as milhares de organizações que usam a plataforma global Zscaler estejam sempre protegidas. Além da pesquisa de malware e análise comportamental, os membros da equipe estão envolvidos na pesquisa e no desenvolvimento de novos módulos para proteção avançada contra ameaças na plataforma Zscaler, e realizam regularmente auditorias internas de segurança para garantir que os produtos e a infraestrutura da Zscaler atendam aos padrões de conformidade de segurança. A ThreatLabz publica regularmente análises aprofundadas de ameaças novas e emergentes em seu portal, **research.zscaler.com**.

Siga-nos: X **@ThreatLabz** | ThreatLabz **security research blog**

©2026 Zscaler, Inc. Todos os direitos reservados.

zscaler™ | Act Fast. Stay Secure.

Sobre a Zscaler

A Zscaler (NASDAQ: ZS) acelera a transformação digital para que seus clientes possam ter mais agilidade, eficiência, resiliência e segurança. A Zscaler Zero Trust Exchange™ protege milhares de clientes contra ataques cibernéticos e perda de dados, conectando com segurança usuários, dispositivos e aplicativos em qualquer local. Distribuída em mais de 150 data centers globalmente, a Zero Trust Exchange™ baseada em SSE é a maior plataforma integrada de segurança em nuvem do mundo. Saiba mais em **zscaler.com/br** ou siga-nos no Twitter **@zscaler**.

© 2026 Zscaler, Inc. Todos os direitos reservados. Zscaler™ e outras marcas registradas listadas em **zscaler.com/br/legal/trademarks** são (i) marcas registradas ou marcas de serviço ou (ii) marcas comerciais ou marcas de serviço da Zscaler, Inc. nos Estados Unidos e/ou em outros países. Quaisquer outras marcas registradas são de propriedade de seus respectivos detentores.

+1 408.533.0288 **Zscaler, Inc. (Sede) • 120 Holger Way • San Jose, CA 95134** **zscaler.com/br**