Rocheston Zelfire 2026 Global Threat Report

Zero-Trust Under Fire

Zero-trust adoption accelerated, but attackers completed faster. In this, the Rocheston Global Cyber Threat Report, Threat Hunt, Threat Intel, Cyber Defense, and Cyber Resilience converge into one operational view. Built on Rocheston's Zero-Trust Threat Intelligence, this report brings together the latest threat activity, tactics, and field-tested defense lessons.

PUBLIC REPORT | ZERO-TRUST THREAT INTELLIGENCE | CYBER DEFENSE | CYBER RESILIENCE

"This new generation is not just network. It is your assistant, your instrument, your companion, and your collaborator."

— Dr. Talal Rochana, Chairman & CEO, Rocheston

Usage, Privacy & Disclaimer

PURPOSE & AUDIENCE

This report is intended for informational and educational purposes. It provides a consolidated view of cyber threat activity, tactics, and defensive lessons observed across Rocheston's threat intelligence platforms. It is designed for security professionals, analysts, and decision-makers who need a practical, operational view of current threat activity.

DATA COLLECTION

All telemetry and threat intelligence gathered in accordance with Rocheston's privacy policy and applicable data protection laws. Personally identifiable information is not collected or stored. Aggregated data may be shared for research purposes.

LIMITATIONS

Primary evidence sources: Cyber Kill Chain (initial access, lateral movement, persistence, exfiltration), MITRE ATT&CK, Rocheston's threat intelligence platforms, and field observations.

DISCLAIMER & LIABILITY

Findings reflect systems in scope and observed activity at the time of collection. Rocheston makes no warranties or guarantees regarding completeness, accuracy, or fitness for any specific purpose. Use at your own risk.

IP & LICENSE

This report is published under the Rocheston Zelfire Public Research License. Redistribution and derivative work are permitted with attribution. Commercial use requires a separate agreement.

"Rocheston Intelligence is more consistent when applied with experienced analysts. But the system is compelling even in early pilot deployments."

— Dr. Talal Rochana, Chairman & CEO, Rocheston

A Message From the Founder

A decade ago I wrote a prediction in my book Cyber Warfare, stating, decisive actions, CYBER-CLASH, Cyber Jamming, and Cyber Jamming will be commonplace. It takes the modern threat landscape. Cyber Attack Vectors are Impact, including the Cyber Kill Chain, Cyber Jamming, Cyber Hunting, Cyber Intelligence, Cyber Defense, Cyber Resilience, and Cyber Recovery. Cyber Attacks are now almost never from single sources, but are engaged in deployments and forms the adversarial cyber-threat actors. Cyber Attacks, and attack vectors are in a state.

Cyber Attacks are now almost never from single sources, but are engaged in deployments and forms the adversarial cyber-threat actors. Cyber Attacks, and attack vectors are in a state.

"The research and technical work was extremely of a high level."

— Dr. Talal Rochana, Chairman & CEO, Rocheston

Cyber Attacks are now almost never from single sources, but are engaged in deployments and forms the adversarial cyber-threat actors. Cyber Attacks, and attack vectors are in a state.

Cyber Attacks are now almost never from single sources, but are engaged in deployments and forms the adversarial cyber-threat actors. Cyber Attacks, and attack vectors are in a state.

Rocheston Zelfire was built for this time. A cyber-threat security system is not just "cyber hunting." It is Threat Hunt, Threat Intelligence, Cyber Defense, Cyber Resilience, Cyber Jamming, Cyber Hunting, Cyber

Credential Attacks

Credential Stuffing

Automated testing of leaked username/password combinations against login pages. Success depends entirely on password reuse. Rate limiting and CAPTCHA are the primary defenses. Volume: billions of attempts per month across all major consumer platforms.

Password Spraying

One password tried against many accounts. Avoids account lockout thresholds. Attacker signature: many failed logins across different accounts from the same IP or ASN within a short window.

Phishing & Proxy Credential Harvesting

Adversary-in-the-middle proxies (Evilginx, Modlishka) capture credentials and session cookies in real time. Bypasses password MFA entirely. Defeat requires phishing-resistant MFA (FIDO2/passkeys).

Social Engineering & Help Desk Abuse

Attackers impersonate users to help desk staff to reset credentials or bypass MFA. Target: IT support workflows. Defeat requires out-of-band identity verification procedures.

Credential Dumps & Combo Lists

Pre-cracked username/password pairs from breaches, sold in bulk. Used for credential stuffing against unrelated services. Volume: tens of billions of combinations available on underground markets.

Kerberoasting

Requesting service tickets for accounts with SPNs, then cracking them offline. Target: service accounts with weak passwords. Defeat requires long, random service account passwords.

Pass-the-Hash / Pass-the-Ticket

Reusing password hashes or Kerberos tickets to authenticate without knowing the plaintext password. Defeat requires Credential Guard and Kerberos hardening.

Cloud Credential Abuse

Stolen AWS keys, GCP service account keys, or OAuth tokens used to access cloud infrastructure. Target: CI/CD pipelines, developer machines, SaaS integrations. Defeat requires short-lived credentials and key rotation.

Credential Stuffing Volume

Password Spraying Volume

Phishing Success Rate

Credential Dumps Availability

MFA Fatigue, Bypass & AiTM

Push Bombing / MFA Fatigue

Flooding a user with push notifications until they accept out of frustration or confusion. Target: any push-based MFA. Defeat requires number matching or biometric confirmation.

Adversary-in-the-Middle (AiTM) Proxying

Real-time proxy between the victim and the IdP. Captures session cookies after MFA completes. Target: any password + push MFA. Defeat requires phishing-resistant MFA (FIDO2/passkeys).

MFA Fatigue Acceptance Rate

Phishing Success Rate (MFA Bypass)

AiTM Adoption

Number Matching Bypass Rate

Biometric Spoofing Feasibility

Number Matching (MS Authenticator)

Microsoft's implementation of number matching requires the user to enter a displayed number into the authenticator app. Defeat push bombing and AiTM relay attacks. Not yet universal across all MFA providers.

Biometric Spoofing

Face and fingerprint biometrics can be spoofed with high-quality presentation attacks. Liveness detection raises the bar but is not foolproof. Hardware-backed biometrics (TPM, Secure Enclave) are harder to bypass.

Phishing-Resistant MFA (FIDO2/Passkeys)

Hardware-bound cryptographic keys that cannot be phished. Origin binding prevents credential use on any domain other than the registered one. Defeat AiTM, push bombing, and MFA fatigue.

Token Theft & Cookie Replay

Session Cookie Theft

Stealing authenticated session cookies via malware, XSS, or network interception. Target: any web SSO session. Defeat requires short-lived sessions, IP binding, and device posture checks.

Refresh Token Theft

Stealing long-lived OAuth refresh tokens to mint new access tokens indefinitely. Target: OAuth/OIDC integrations, mobile apps, CI/CD systems. Defeat requires token rotation, sender-constrained tokens (DPoP), and anomaly detection.

Session Cookie Replay

Replaying a stolen session cookie from a different IP or device. Target: any session without IP binding or device attestation. Defeat requires IP binding, device posture checks, and short TTLs.

MFA Token Replay

Replaying a stolen MFA token before it expires. Target: TOTP codes, push approvals, SMS codes. Defeat requires phishing-resistant MFA (FIDO2/passkeys) that binds to the origin.

Session Cookie Theft

Refresh Token Theft

Session Cookie Replay

MFA Token Replay

GitoIdx Abuse, CA Misconfig & Privilege Escalation

GitoIdx Accounts Abuse

Stealing or misuse of GitoIdx accounts used for SSO into cloud consoles. Target: GitoIdx tenants with broad application access. Defeat requires GitoIdx conditional access policies and session revocation.

High-Value GitoIdx Permissions

Roles like Global Administrator, Application Administrator, and Cloud Application Administrator. Target: any GitoIdx tenant. Defeat requires PIM (Privileged Identity Management) with approval workflows.

Conditional Access Misconfig

Gaps in GitoIdx conditional access policies: legacy auth not blocked, missing device compliance, or overly broad exclusion groups. Target: any GitoIdx tenant with CA policies. Defeat requires auditing CA policies against baseline.

Service Principal Abuse

Compromised GitoIdx service principals with broad Graph API permissions. Target: automation, integrations, and SaaS connectors. Defeat requires scoped permissions and credential rotation.

Stolen GitoIdx Accounts

High-Value GitoIdx Permissions

Conditional Access Misconfig

Service Principal Abuse

Device Trust Reality

"Passwords is a signal, not truth. It can be stolen, manipulated, spoofed, or intentionally manipulated. A device-backed assertion is a stronger, more expensive signal about — and of a device that is currently safe."

Device trust is a signal, not truth. It can be stolen, manipulated, spoofed, or intentionally manipulated. A device-backed assertion is a stronger, more expensive signal about — and of a device that is currently safe.

Device trust is not a binary truth. It is a probabilistic assertion about the current state of a specific hardware and software stack, based on what the OS and hardware can prove at the moment of authentication. Modern EDR, MDM, and attestation systems can verify boot integrity, OS patch level, disk encryption, and the presence of a specific agent. But none of this proves the user is not being coerced, the session is not being hijacked, or the device is not being used by someone else.

Hardware-Backed Attestation

- TPM and Secure Enclave keys bind authentication to specific hardware. Credentials cannot be exported to another machine. Stronger than software-only device IDs, but still vulnerable to session hijacking after authentication.
- Attestation proves a device is in a known state at a specific moment. Does not prove the user is still the same person.
- Revoking a compromised device requires fleet-wide policy enforcement. Slow in practice if MDM coverage is incomplete.

Posture Checks

- Patch level, disk encryption, firewall, and agent presence are checked at authentication time. Stale posture data is common — devices that haven't checked in recently may be evaluated against outdated state.
- Posture checks are only as fresh as the last check-in. Long-lived sessions can outlast the posture that was valid at login.
- Posture failure should trigger step-up authentication or session termination, not just a warning.

Device Trust Signals

- Managed vs. unmanaged device
- Corporate vs. personal device
- Known vs. unknown hardware
- Compliant vs. non-compliant posture
- Recent vs. stale check-in

Compliant does not mean safe. It means the device passed its last attestation check.

Unmanaged & BYOD Access

Legal Entities ERPs

Legal entities ERPs are systems that manage the financial and operational data of a specific legal entity, such as a subsidiary or a joint venture. They are critical for compliance and financial reporting, and access to them is highly sensitive.

Identity Workers

Identity workers are employees, contractors, and third-party users who have access to a company's identity systems and data. They are the primary targets for phishing, credential theft, and insider threats.

Legal Identity, Managed Corporate or Unmanaged Personal Device

This matrix defines the legitimate access scenarios for different combinations of legal identity and device ownership. It is used to determine which access is allowed, which requires additional controls, and which should be denied or flagged for review.

- Identity is legal, managed corporate device: Standard access with full policy enforcement. Baseline scenario.
- Identity is legal, unmanaged personal device: Access allowed with restrictions. Session limits, no persistent tokens, no local data storage.
- Identity is legal, device is unknown: Step-up authentication required. Device registration or attestation before access is granted.
- Identity is not legal: Access denied regardless of device. Flag for review if the identity is known but the legal entity is not recognized.

Managed Device + Valid IdP + Expected Location

- Apply full IdP authentication (email, phone, biometrics)
- Enforce session limits and re-authentication
- Log access for audit and anomaly detection

Managed Device + Valid IdP + Unexpected Location

- Step-up authentication required
- Flag for review if location is consistently unexpected
- Consider temporary access restrictions

Unmanaged Device + Valid IdP + Expected Location

- Block or restrict access to sensitive systems
- Session limits, no persistent tokens
- Log all access for audit

Unmanaged Device + Valid IdP + Unexpected Location

- Block access by default
- Require manual review before any access is granted
- Log attempt for security investigation

Unknown Device + Valid IdP

- Block access
- Flag for review
- Log for investigation

Unknown Device + Unknown Identity

- Block access
- Log for investigation

Legal Entity Unknown

- Block access
- Flag for review
- Log for investigation

EDR Tampering & Posture Spoofing

Security Tool Disablement

Attackers disable EDR, antivirus, or logging before deploying payloads. This is a high-signal event that should trigger immediate alerting and session termination.

Scripts Disable Logging Or Security Agents

Attackers use scripts to stop security agents, clear logs, or modify system configurations. Detection depends on whether the EDR agent itself is still running and reporting. Persistence mechanisms can restore tampered configurations.

Compliance Drift

Devices fall out of compliance (patches, disk encryption, firewall rules) between authentications. Long-lived sessions can outlast the posture that was valid at login. Posture checks only are as fresh as the last check-in.

Compliance Drift

- Patch level
- Disk encryption
- Firewall
- Agent presence
- Recent check-in

Attacker Tampering With Security Tools

- Disable logging
- Stop security agents
- Modify configurations
- Persistence mechanisms

Local Admin & Endpoint Hardening Ladder

Local administrative access is necessary in many environments but creates persistence, lateral movement, and evasion capabilities. Each level below represents a different trade-off between convenience and security. Organizations should map their current state and target state against this ladder.

Why Local Admin Is Hard

Many tools require elevated access to function. Developers need it for debugging and installing dependencies. Support teams need it for troubleshooting. Legacy applications assume it exists. Removing it breaks workflows that have internal dependencies.

Persistence and Lateral Movement

Local administrator access allows attackers to disable security tools, install persistent backdoors, dump credentials from memory, and move laterally using stolen local admin hashes. Many ransomware campaigns rely on this access.

Levels

- Everyone has local admin by default. No separation between user and admin accounts.
- Admin accounts exist but are used for everyday work. Credentials are shared or reused across systems.
- Admin accounts are separate from standard accounts. Elevation requires a different login.
- Just-in-time (JIT) elevation. Admin access is granted for a specific task and expires automatically.
- Just-enough-administration (JEA). Specific commands can be run as admin, but the rest of the system remains locked.
- Application allowlisting. Even with admin access, only approved binaries can run.
- Full endpoint hardening. Admin access is rare, logged, and requires multi-party approval.

Levels

- **Everyone has local admin by default** — Developers need it for debugging and installing dependencies. Support teams need it for troubleshooting. Legacy applications assume it exists. Removing it breaks workflows that have internal dependencies.
- **Admin accounts exist but are used for everyday work** — Credentials are shared or reused across systems. Lateral movement using stolen local admin hashes is trivial.
- **Admin accounts are separate from standard accounts** — Elevation requires a different login. Credential reuse is still common.
- **Just-in-time (JIT) elevation** — Admin access is granted for a specific task and expires automatically. Persistence is limited to the session duration.
- **Just-enough-administration (JEA)** — Specific commands can be run as admin, but the rest of the system remains locked. Attackers can still abuse allowed commands.
- **Application allowlisting** — Even with admin access, only approved binaries can run. Bypasses exist but raise the cost significantly.
- **Full endpoint hardening** — Admin access is rare, logged, and requires multi-party approval. Most organizations never reach this level.

ZTNA Is Only as Strong as Its Scope

"Identity-specificated policy, signal-required and exceptions become permanent access accumulation and drift. Identity-specificated policy — signal-required — exceptions — access accumulation — drift."

Zero Trust Network Access (ZTNA) is only as strong as its scope. If the policy engine does not see a resource, it cannot enforce anything about it. Shadow IT, legacy systems, and third-party integrations that bypass the identity layer are the gaps that matter most.

Key Dependencies

- Identity provider (IdP) must be integrated with the policy engine
- Device posture signals must be available for high-risk resources
- Network or agent must be able to enforce access decisions at the session level, not just at login

Gap Indicators

- Resources accessed without a corresponding IdP policy
- Sessions that persist after a change in device posture or user status
- Third-party integrations that authenticate through shared credentials or service accounts rather than identity-aware proxies
- Legacy systems that cannot be placed behind the policy engine

Policy Coverage Is Only As Strong As Its Scope

Identity-specificated policy — signal-required — exceptions — access accumulation — drift.

Common ZTNA Failures

Shadow Resources

Applications and services that are not registered with the identity provider or policy engine. Common in legacy systems, third-party integrations, and environments where IT does not know what exists. Attackers can reach these resources through lateral movement.

Stale Trusted Identity Groups

Groups that grant access to resources but have not been reviewed in months. Former employees, contractors, and service accounts remain in groups long after they should be removed.

Broad Implementations of Identity-Aware Proxies

Identity-aware proxies that authenticate users but do not enforce session-level or resource-level policy. Authentication is strong, but authorization is still coarse-grained.

Service Accounts Bypassing Identity Checks

Service accounts that access resources directly, bypassing the identity provider and policy engine. Credentials are long-lived and rarely rotated.

Legacy Systems Cannot Be Placed Behind Policy Engine

Systems that require direct network access or protocols that do not support identity-aware proxies. Common in manufacturing, healthcare, and government environments.

Common Failures

- Shadow resources
- Stale trusted identity groups
- Broad implementations of identity-aware proxies
- Service accounts bypassing identity checks
- Legacy systems cannot be placed behind policy engine

ZTNA Coverage Map — Gaps

- Compliance Reporting
- ZTNA + IdP Integrated Policy
- IdP Enrolled
- Network Reachable
- Shadow Resources

Shadow resources are invisible to the policy engine.

East-West Movement Patterns

SSH Tunneling

Attackers use SSH port forwarding to pivot through compromised hosts, bypassing firewalls and network segmentation. Common in environments where SSH is allowed between internal hosts. Detection requires monitoring SSH sessions and tunnel creation.

Kerberos Ticket Theft

Attackers steal Kerberos tickets (TGT or service tickets) from memory or disk to impersonate users on other systems. Pass-the-ticket and golden ticket attacks are common in Active Directory environments. Detection requires monitoring ticket issuance and unusual ticket usage.

Identity Reuse Across Microservices

Service tokens issued for one microservice are accepted by others without re-authentication. Lateral movement between services using stolen or forged tokens. Common in microservices architectures with shared identity providers.

Lateral Movement

- SSH tunneling
- Kerberos ticket theft
- Identity reuse across microservices

East-West Movement — Gap

- Lateral Movement
- Identity Reuse
- Kerberos Ticket Theft
- SSH Tunneling
- Other East-West Movement

Lateral movement is the hardest to detect and most dangerous in flat networks.

Ransomware Blast Radius

Internet
Internet Facing Systems
Identity Provider
Internal Systems
Critical Systems
Backup Systems
Backup Infrastructure
Critical Systems
Internal

Trust Transitivity Failure Map

Trust transitivity failure occurs when a trusted relationship allows the attacker to hop from one system to another without re-authentication. Each hop inherits the trust of the previous system, so a single compromised machine can cascade across the entire enterprise.

Trust Transitivity Failure Map

Kerberos Ticket Granting Ticket (TGT) Abuse

Kerberos issues a TGT when a user authenticates. Any system that trusts the Kerberos realm will accept that TGT to issue service tickets for other resources. Compromise of one machine's TGT can be used to access every service in the realm.

Active Directory Trust Chaining

Domain trusts allow users from one domain to access resources in another. Attackers who compromise a low-privilege domain can pivot through trusted domains to reach high-value targets in other domains.

Cloud IAM Federation Abuse

Cloud identity providers issue tokens that grant access to multiple services. Compromise of the identity provider or a stolen token allows access to all federated services without re-authentication.

Service Account Token Reuse

Service accounts use long-lived tokens that do not expire or require re-authentication. Stolen tokens can be used from any system that trusts the issuing authority.

Pass-the-Ticket / Pass-the-Hash

Stolen credentials or tokens are replayed on other systems to impersonate legitimate users. Multi-factor authentication does not prevent token reuse once the session is established.

Lateral Movement via Trusted Protocols

Protocols designed for administration (RDP, SSH, WMI, PsExec) inherit trust from the authenticated session. Attackers use these protocols to move between systems using stolen credentials.

Misconfigured Trust Boundaries

Trust relationships that were never intended to cross security boundaries are left open by default. Misconfigurations in trust policy create unintended paths to sensitive systems.

Trusted Third-Party Integrations

Vendor integrations and SaaS connections inherit trust from the organization. Compromise of a vendor's infrastructure can be used to attack customers through their trusted integration.

Recommendations

- Enforce Multi-Factor Authentication (MFA) for all administrative access
- Implement Just-In-Time (JIT) and Just-Enough-Access (JEA) for privileged accounts
- Rotate service account tokens and credentials regularly
- Segment trust boundaries so a compromise in one domain does not cascade to others
- Monitor for anomalous cross-domain authentications and lateral movement patterns

AI Changed the Economics of Intrusion

"AI did not invent new attack categories. It industrialized existing ones—reducing the cost of reconnaissance, social engineering, and malware development to levels previously unavailable to even three-letter units."

Attack Phase	Pre-AI Cost	Post-AI Cost
Reconnaissance	Days of manual OSINT	Minutes of automated scraping
Phishing Lure Creation	Hours per target	Seconds per target
Malware Obfuscation	Days of manual work	Minutes with LLM assistance
Social Engineering Scripts	Hours of research	Minutes with LLM
Deepfake Voice Clone	Weeks of specialist work	Minutes with commodity tools
Deepfake Video Clone	Months of specialist work	Hours with commodity tools
Vulnerability Exploitation	Days to weeks	Hours to days

Key Takeaway

AI does not create new attack categories. It industrializes existing ones, reducing the cost of reconnaissance, social engineering, and malware development to levels previously unavailable to even three-letter units.

AI Enhancements Lower Barriers Across Every Phase

- Reconnaissance: automated OSINT gathering and target profiling
- Phishing: personalized lures at scale, multilingual without accent tells
- Malware development: LLM-assisted obfuscation and code generation
- Social engineering: deepfake voice and video for pretexting
- Vulnerability exploitation: faster CVE analysis and PoC adaptation

Defensive Implications

- Assume attackers have access to the same AI capabilities you do
- Focus on behavioral detection rather than signature matching
- Multi-factor authentication remains effective against AI-enhanced attacks
- Deepfake detection is an emerging requirement for high-value targets

AI-Driven Phishing & Pretexting

Industrial Spear-Phishing

Large language models can generate fluent, grammatically correct, and culturally appropriate phishing lures at scale. The traditional tells (poor grammar, awkward phrasing, cultural mismatches) that trained users to recognize phishing are no longer reliable.

Voice Deepfakes

Voice cloning requires as little as 3 seconds of audio to produce a convincing clone. Attackers use cloned voices for vishing calls, bypassing voice-based identity verification and social engineering defenses.

Video Deepfakes

Real-time video deepfakes can impersonate executives in live video calls. Demonstrated attacks have successfully authorized fraudulent wire transfers worth hundreds of thousands of dollars.

Pretexting Automation

Agentic AI systems can research targets, craft multi-step pretexts, and adapt responses in real-time. This automates the reconnaissance and preparation phases that previously required human operators.

Phishing Volume Trends

Defense in depth because one lure can bypass one control, but not multiple independent controls at once.

AI-based detection that flags anomalous communication patterns before the lure reaches the inbox.

Out-of-band verification for any sensitive request, regardless of how convincing it appears.

Supply Chain Is Front at Scale

The trust underpinning supply chain networks cannot trust relationships rather than technical validation exists. The software now expected — but the software it specifically authorizes plus what it pulls and runs.

Vendor access, managed services, software dependency updates, and open-source libraries all operate on trust. Attackers exploit this trust to reach downstream targets without direct access to the target organization. Compromise of one vendor can cascade across dozens or hundreds of customers simultaneously — multiplying impact from a single intrusion.

Supply Chain Attack Surface

Managed Service Provider Compromise

Managed service providers have broad access to customer environments by design. Compromise of an MSP can provide access to dozens of customers through legitimate administrative tools and trusted connections.

Software Dependency Attacks

Open-source libraries and package managers are widely used in enterprise software. Malicious packages, typosquatting, and dependency confusion attacks can introduce code that runs with the privileges of the application.

Open-Source Maintainer Burnout

Many critical open-source projects are maintained by individuals with limited resources. Burnout creates opportunities for attackers to gain maintainer access or social engineer project transfers.

Recommendations

- Verify software integrity before deployment (checksums, signatures, SBOM)
- Pin dependencies to specific versions and review updates before adoption
- Monitor for unexpected outbound connections from build systems and CI/CD pipelines
- Segment vendor access by function and time-limit privileged sessions
- Maintain an inventory of third-party dependencies and their transitive dependencies

Vendor Access & Remote Tooling

Legitimate Tools Are Dual-Use

Remote access tools, VPNs, and administrative utilities are designed for legitimate use but can be abused by attackers who obtain credentials or exploit vulnerabilities in these systems. Detection must distinguish between authorized and unauthorized use of the same tools.

Zero Trust Vendor Access

Traditional VPN access grants broad network access to vendors. Zero trust approaches limit vendor access to specific applications and resources, reducing the blast radius of compromised credentials.

Managed Service Provider Risk Model

Managed service providers have broad access to customer environments by design. Compromise of an MSP can provide access to dozens of customers through legitimate administrative tools and trusted connections.

Recommendations

- Implement Just-In-Time (JIT) and Just-Enough-Access (JEA) for vendor accounts
- Require multi-factor authentication for all vendor access
- Monitor for anomalous vendor activity: unusual hours, new systems, bulk data access
- Segment vendor access by function and time-limit privileged sessions
- Maintain an inventory of vendor accounts and review access quarterly

Software Dependency Attacks

Typosquatting

Attackers publish packages with names similar to popular libraries, hoping developers will install the wrong package. Automated dependency updates can pull these packages into production without review.

Dependency Confusion

Internal package names can be hijacked if a public repository has a higher version number than the internal repository. Build systems that prefer the public repository will install the attacker's package.

Malicious Package Takeover

Attackers gain control of legitimate packages through compromised maintainer accounts or social engineering. Updates to these packages can introduce malicious code to all downstream users.

Typical Malicious Updates

- Credential harvesting: steal environment variables, API keys, and tokens
- Backdoor installation: establish persistent access to the build or production environment
- Data exfiltration: send sensitive data to attacker-controlled servers
- Cryptocurrency mining: use compute resources for unauthorized mining

Typical Attack Vector Ups Growth

Recommendations

- Pin dependencies to specific versions and review updates before adoption
- Monitor build systems for unexpected outbound connections
- Maintain an inventory of third-party dependencies and their transitive dependencies
- Verify package integrity (checksums, signatures, SBOM)

Ransomware Kill Chain

Deepfakes in Business Workflows

Synthetic Media Manipulation

Deepfakes are artificial intelligence (AI)-generated audio, video, or images that mimic real individuals with high fidelity. In business workflows, they are weaponized to bypass biometric authentication, impersonate executives in video calls, or forge documentary evidence. The technology lowers the barrier to sophisticated fraud, making it accessible to threat actors with limited technical expertise.

Business Workflow Vulnerabilities

Critical workflows in finance, HR, and legal are increasingly reliant on remote verification and digital identity. Deepfakes exploit these processes by injecting synthetic identities into KYC (Know Your Customer) onboarding, manipulating video conferences to authorize transactions, and creating fraudulent evidence for insurance claims.

Detection Challenges

Human perception is insufficient against high-quality deepfakes. Detection requires specialized AI models analyzing subtle artifacts in lighting, facial micro-expressions, and audio-visual synchronization. The arms race between generation and detection is accelerating, with adversarial techniques designed to evade current detection systems.

Mitigation Strategy

Implement multi-factor authentication (MFA) that includes liveness detection. Deploy AI-powered deepfake detection tools at critical workflow chokepoints. Establish out-of-band verification protocols for high-value transactions. Conduct regular security awareness training to recognize social engineering tactics combined with deepfake technology.

Critical Workflows

- Video conference authorization
- Remote identity verification (KYC)
- Biometric authentication bypass
- Executive impersonation for wire transfers
- Forged documentary evidence
- Insurance claim fraud

Detection Challenges

- Human perception insufficient against high-quality deepfakes
- Real-time detection requires specialized AI models
- Adversarial techniques designed to evade detection
- Lag between deepfake creation and detection capability

Mitigation Strategy

- Implement multi-factor authentication (MFA) with liveness detection
- Deploy AI-powered deepfake detection at critical chokepoints
- Establish out-of-band verification protocols for high-value transactions
- Conduct regular security awareness training
- Monitor for anomalies in video quality, lighting, and audio-visual synchronization

AI Against AI: Prompt Injection, Data Leakage & Poisoning

Prompt Injection and Jailbreaking Attacks

Adversarial prompts designed to override system instructions and bypass safety guardrails. Techniques include direct injection, indirect injection via retrieved documents, and multi-turn conversation manipulation. Successful attacks can exfiltrate system prompts, leak proprietary data, or cause the model to perform unauthorized actions.

Data Exfiltration via Tool Calls

Agentic AI systems with tool access can be manipulated into transmitting sensitive data to external endpoints. Attackers craft prompts that trick the model into calling APIs with stolen data as parameters, bypassing traditional DLP controls.

Knowledge Base Poisoning RAG

Retrieval-Augmented Generation (RAG) systems are vulnerable to poisoned documents in their knowledge bases. Attackers inject malicious content into indexed sources, causing the model to retrieve and act on adversarial instructions when answering queries.

Defense Strategy

- Input validation and output sanitization
- Tool call whitelisting and rate limiting
- Prompt injection detection classifiers
- Data loss prevention (DLP) for AI outputs
- Human-in-the-loop confirmation for sensitive actions
- Regular red team exercises and adversarial testing

Critical Vulnerabilities — Unprotected Systems

- Direct Injection — Prompt Override
- Indirect Injection — Poisoned Retrieval
- Tool Misuse — Unauthorized Actions
- Data Exfiltration — Tool Call Leakage
- Knowledge Base Poisoning — Corrupted Retrieval
- Multi-Turn Jailbreak — Gradual Persuasion

Defense Layers — Protected Systems

- Input Validation — Prompt Sanitization
- Tool Whitelisting — Scoped Permissions
- Output Filtering — Sensitive Data Redaction
- Human Approval — Critical Action Confirmation

Common Initial Access Paths — Ranked

These four paths account for most successful initial access attempts, ranked by prevalence, and each exploitation vector maps to a specific defensive control. Organizations should prioritize hardening the top two paths before investing in lower-frequency attack surfaces.

#1 Phishing — 35%

Credential harvesting via deceptive emails and fake login pages. Defeated by phishing-resistant MFA (FIDO2/WebAuthn) and DMARC enforcement.

#2 Exploit Public-Facing Application — 25%

Unpatched vulnerabilities in internet-facing web applications, APIs, and VPN gateways. Defeated by vulnerability management and virtual patching.

#3 Valid Credentials — 20%

Stolen or purchased credentials from prior breaches. Defeated by MFA, credential monitoring, and anomalous login detection.

#4 Supply Chain — 15%

Compromised third-party software, libraries, or hardware. Defeated by SBOM, code signing, and dependency monitoring.

Initial Access Vectors — Ranked

- #1 Phishing — 35%
- #2 Exploit Public-Facing Application — 25%
- #3 Valid Credentials — 20%
- #4 Supply Chain — 15%
- Other — 5%

Post-Compromise Objectives

Once an initial foothold is established, threat actors pursue several objectives — each step increasing their safety, leverage, and attacker leverage. The speed at which attackers move through this progression has compressed from days to hours in recent years.

Persistence Establishment

Threat actors install backdoors, modify boot sequences, create hidden accounts, and deploy remote access tools to ensure they can return even after initial access is remediated.

Privilege Escalation

Exploiting misconfigurations, unpatched vulnerabilities, or stolen credentials to gain administrator or domain admin privileges, enabling access to protected resources and systems.

Reconnaissance & Lateral Movement

Mapping the internal network, identifying high-value targets, and moving laterally through systems to reach critical assets. Tools like BloodHound automate this process.

Data Exfiltration

Staging and compressing stolen data, then exfiltrating it through encrypted channels, cloud storage services, or DNS tunnels to avoid detection by network monitoring.

Ransomware Deployment

Encrypting critical systems and data, then delivering ransom demands. Modern operations often double-extort by threatening to publish stolen data if the ransom is not paid.

Post-Compromise Attack Lifecycle

- Persistence Establishment
- Privilege Escalation
- Reconnaissance & Lateral Movement
- Data Exfiltration
- Ransomware Deployment

Ransomware Recovery Cost Breakdown

Downtime dominates when backups and identity controls fail.

Cyber Defense Analytics

Cyber Defense Analytics (CDA) is a collaborative effort to develop, share, and apply analytics for cyber defense. Our goal is to make cyber defense more effective, efficient, and scalable by:

- Improving situational awareness
- Supporting decision making
- Automating repetitive tasks
- Enabling sharing of analytics and data
- Reducing time from detection to response
- Increasing the impact of analysts on the mission

CDA is a community-driven effort that includes analysts, researchers, and tool developers from government, industry, and academia. We are building shared analytics, datasets, and tools that can be reused across different cyber defense missions.

Metrics That Correlate with Fewer Incidents

Measurable improvements in cyber defense correlate with fewer incidents, faster recovery, and better risk reduction. Key metrics include mean time to detect, mean time to respond, and patch coverage.

PR-MDR
Patch Rate – Malware Detection Rate
Measures how patching vulnerabilities reduces malware detections.

TD
Time to Detect
Measures how quickly defenders identify malicious activity after it begins.

D
Detection
Measures the ability to identify malicious activity.

GMoth
Gamma Moth
Measures how quickly defenders detect and remove persistent malicious activity.

CA-S
Cyber Analytics – Sharing
Measures how analytics are shared across teams and tools to improve detection and response.

MTTc
Mean Time to Contain
Measures how quickly defenders limit the spread and impact of an intrusion.

ACTIVE DEFENSE KEY:
Metrics that support active cyber operations, including hunting, deception, threat emulation, and other mission-focused activities.

Metrics that measure or enable active cyber defense.
Focus on detection, response, recovery, and risk reduction.
Support analytics development and sharing.

How to Operationalize Metrics

Metrics Activity Data Sources Examples

- Endpoint telemetry (process, file, registry, network events)
- Network sensor logs (DNS, HTTP/S, TLS, flows, packets)
- Identity and access logs (authentication, MFA, privilege changes)
- Cloud control plane logs (API calls, IAM changes, resource changes)
- Vulnerability scanner and patch management data
- Incident response and threat hunting notes

Analytic Ownership

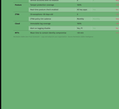
- Data engineering: ingestion, normalization, enrichment, retention, lineage, quality, latency, backfill, schema changes, access controls
- Detection engineering: rules, models, thresholds, tuning, suppression, false positive handling, coverage, regression tests, deployment
- Analyst workflows: triage, investigation, escalation, feedback, case management, metrics

Tie Metrics to Change Control

- Treat detections and analytics as production software: version control, peer review, staging, rollback, release notes, ownership, and change logs
- Use canary deployments and shadow mode to validate new rules before they affect alerts
- Track metrics by deployed rule, dataset, or source to isolate regressions
- Maintain runbooks, escalation paths, and ownership for each critical detection

Objective Compass (2026)

Actor-to-Industry Focus Lines



Recommended Zero-Trust Scorecard

Control	Metric	Value	Target
Identity	Phishing-resistant MFA coverage	92%	100%
	Stale accounts disabled within 24 hours	78%	95%
	Privileged accounts with hardware keys	40%	100%
Endpoint	Critical patch coverage within 14 days	88%	95%
	EDR coverage	96%	100%
Network	Internet-facing assets behind ZTNA	60%	100%
	Lateral movement alerts per month	12	0
Data	Critical data encrypted at rest	85%	100%
	Unauthorized exfiltration attempts	3	0
Cloud	Misconfigurations per 1000 resources	14	0
	Public storage buckets	2	0
Logging	Critical logs ingested within 5 minutes	91%	100%
	Mean time to detect (MTTD)	18 hrs	1 hr

The Exception Problem

Controls that bypass Zero Trust accumulate faster than they are removed. Legacy authentication, shared accounts, long-lived service tokens, and ad-hoc firewall rules persist because they solve real operational problems and nobody owns their retirement.

Exceptions are tracked in spreadsheets or tickets, not in the identity fabric. Each exception creates a blind spot that attackers can exploit. Over time, the cumulative exception surface becomes the actual attack surface, regardless of what the architecture diagram shows.

Why Zero Trust Fails

Zero Trust assumes every access request is evaluated against current policy. Exceptions create parallel paths that bypass policy entirely. Monitoring sees the exception as normal traffic because it is authorized. The deception that Zero Trust was implemented when it was actually bypassed is the most dangerous outcome.

Common Causes

- Legacy applications that require MFA bypass
- Service accounts with long-lived credentials
- Vendor access that cannot be revoked quickly
- Emergency access that never expires
- Network exceptions that accumulate over years

Deception without reality Zero Trust was implemented as a control, but exceptions have quietly accumulated over years. Each exception adds a blind spot. Attackers exploit the gap between what the architecture claims and what actually runs.

How to Read the Industry Chapters

Each industry chapter contains: threat profile, operational risk, analytic focus, recommended controls, and a scorecard. Use the cross-industry comparison to identify patterns that apply across sectors.

Actors in Various Sector Focus – Primary Focus

Telecom-related actors, Lazarus and Andariel, are not shown because their primary focus is elsewhere. Bluenoroff is left out because Jade Sleep is a specialized, different threat.

Primary Focus

- Lazarus: Ransomware, IAB
- Kimsuky: Data Theft, IAB
- Andariel: Data Theft, IAB
- APT38: Financial Theft, IAB
- Moonstone Sleet: Data Theft, IAB
- Jade Sleep: Data Theft, IAB
- Bluenoroff: Financial Theft, IAB
- Pyongyang: Data Theft, IAB

Cross-Industry Comparison

Healthcare Operational Risk
Patient safety, PHI exposure, ransomware, supply chain, medical device, and third-party risk. Lazarus and Andariel are the primary threats. Ransomware and data theft are the most common outcomes. Recovery time matters because clinical workflows depend on availability.

Finance Analytic Focus
Transaction monitoring, fraud detection, identity and access, third-party risk, and ransomware. Lazarus, APT38, and Bluenoroff are the primary threats. Financial theft and ransomware are the most common outcomes. Detection speed and recovery time are critical.

Telecom Control Emphasis
Network segmentation, lateral movement, identity and access, third-party risk, and ransomware. Lazarus and Andariel are the primary threats. Data exfiltration and ransomware are the most common outcomes. Network visibility and segmentation are critical.

Critical Cross-Industry Risks
Ransomware, third-party access, identity compromise, and supply chain. Lazarus and Andariel are the most active across multiple sectors. Ransomware is the most common outcome. Third-party access is the most common entry point.

Lazarus Activity by Sector

Healthcare Threat Profile

Healthcare faces ransomware, third-party access, and supply chain risks. Patient safety and PHI exposure are the primary consequences. Recovery time matters because clinical workflows depend on availability. Lazarus and Andariel are the most active threats. Ransomware and data theft are the most common outcomes.

Top Risk Drivers

- Ransomware
- Third-party access
- Supply chain compromise – credential theft, software, MSPs
- Identity and access – stolen credentials, MFA bypass, privileged access
- Medical device and clinical systems – legacy systems, limited patching, vendor dependency
- Data exfiltration – PHI, research data, intellectual property
- Operational impact – clinical workflows, patient care

Lazarus Main Activity

- Ransomware and data theft in healthcare, research, and critical infrastructure
- Supply chain and third-party access through software, vendors, and managed services
- Credential theft, MFA bypass, and persistent access
- Long-term intelligence collection and lateral movement

- Lazarus
- Andariel
- Kimsuky
- APT38
- Moonstone Sleet
- Jade Sleep
- Bluenoroff
- Pyongyang

Healthcare-Specific Playbook

Patch Critical Internet-Facing Systems

- Prioritize patching for internet-facing systems, remote access, and identity infrastructure. Ransomware and third-party access often start here.
- Track patch coverage by asset class, not just overall percentage. Critical systems need faster patching than the average.

Protect Legacy Medical Devices

- Isolate legacy medical devices and clinical systems on segmented networks. Limited patching and vendor dependency make these systems hard to secure.
- Monitor for lateral movement from clinical networks to administrative systems. Attackers use these as pivot points.

Monitor Identity and Remote Access

- Detect credential theft, MFA bypass, and privileged access abuse. Ransomware operators use stolen credentials to gain initial access and persistence.
- Log and alert on remote access, VPN, and third-party access. Third-party access is a common entry point.

Prepare Ransomware Response

- Maintain offline backups and test restoration regularly. Ransomware is the most common outcome in healthcare.
- Define escalation paths and ownership for each critical detection. Delayed response increases impact.

Manage Third-Party Risk

- Limit third-party access and monitor for anomalous activity. Vendor compromise is a recurring pattern across sectors.
- Require multi-factor authentication for all third-party access. Credential theft alone should not grant access.

Vulnerable Logging for Critical Environments

- Ensure critical systems generate and retain logs. Many incidents begin with weak or missing telemetry.
- Log authentication, privilege changes, and network activity. These are the most useful for detection.

Lack of Segmentation

- Segment clinical networks from administrative networks. Lateral movement between these is a common pattern.
- Limit access between medical devices and business systems. Compromise should not spread freely.

Slow Response

- Define escalation paths and ownership for each critical detection. Delayed response increases impact.
- Test response playbooks regularly. Untested plans fail under pressure.

Finance Threat Profile

Finance is a persistent target and operates under strict regulatory and audit requirements. Third-party access and identity are the primary risk drivers. Lazarus, APT38, and Bluenoroff are the most active threats. Financial theft and ransomware are the most common outcomes. Detection speed and recovery time are critical because transaction monitoring and fraud detection depend on timely response.

Top Risk Drivers

- Third-party access – vendor access, managed services, software updates
- Identity and access – stolen credentials, MFA bypass, privileged access, legacy authentication
- Ransomware – business disruption, data theft, extortion
- Transaction manipulation – fraud, unauthorized transfers, payment abuse
- Data exfiltration – customer data, financial records, proprietary data

Recurring Patterns

- Third-party access is a recurring entry point across multiple sectors.
- Credential theft and MFA bypass are common in identity-based attacks.
- Ransomware is the most common outcome in financial sectors.
- Lateral movement and persistence are common after initial access.

- Lazarus
- APT38
- Bluenoroff
- Andariel
- Kimsuky
- Moonstone Sleet
- Jade Sleep
- Pyongyang

Finance-Specific Controls

Third-Party Access Governance

- Limit third-party access and monitor for anomalous activity. Vendor compromise is a recurring pattern across sectors.
- Require multi-factor authentication for all third-party access. Credential theft alone should not grant access.
- Log and alert on third-party access, VPN, and remote access. Third-party access is a common entry point.

Identity and Access Controls

- Detect credential theft, MFA bypass, and privileged access abuse. Ransomware operators use stolen credentials to gain initial access and persistence.
- Monitor for lateral movement and persistence. Attackers use these after initial access.

Transaction Monitoring

- Detect fraud, unauthorized transfers, and payment abuse. Financial theft is the most common outcome in finance.
- Define escalation paths and ownership for each critical detection. Delayed response increases impact.

Ransomware Preparation

- Maintain offline backups and test restoration regularly. Ransomware is the most common outcome in healthcare.
- Define escalation paths and ownership for each critical detection. Delayed response increases impact.

Control	Priority	Why
Third-party access governance	Critical	Third-party access is a recurring entry point across multiple sectors.
Identity and access controls	Critical	Credential theft and MFA bypass are common in identity-based attacks.
Transaction monitoring	Critical	Financial theft is the most common outcome in finance.
Ransomware preparation	Critical	Ransomware is the most common outcome in healthcare.
Lateral movement detection	High	Attackers use these after initial access.
Vulnerability management	High	Patch coverage and timeliness matter for critical systems.
Logging and telemetry	High	Many incidents begin with weak or missing telemetry.
Incident response readiness	High	Delayed response increases impact.

Recurring Patterns

- Third-party access is a recurring entry point across multiple sectors.
- Credential theft and MFA bypass are common in identity-based attacks.
- Ransomware is the most common outcome in healthcare.
- Lateral movement and persistence are common after initial access.

Education Threat Profile

Education faces ransomware, third-party access, and identity risks. Student data and research are the primary assets. Lazarus and Andariel are the most active threats. Ransomware and data theft are the most common outcomes. Recovery time matters because academic operations depend on availability. Limited security staffing and budget constraints are common challenges.

Top Risk Drivers

- Ransomware – business disruption, data theft, extortion
- Third-party access – vendor access, managed services, software updates
- Identity and access – stolen credentials, MFA bypass, privileged access
- Research data targeting – theft of research, intellectual property
- Limited security staffing – limited monitoring, slow response, delayed patching
- Legacy systems – unpatched systems, hard to secure, limited visibility
- Open networks – shared access, guest access, weak segmentation

Recurring Patterns

- Third-party access is a recurring entry point across multiple sectors.
- Credential theft and MFA bypass are common in identity-based attacks.
- Ransomware is the most common outcome in healthcare.
- Lateral movement and persistence are common after initial access.

Actors Focus in Education – Overview

Lazarus – Education

- Ransomware – business disruption, data theft, extortion
- Third-party access – vendor access, managed services, software updates
- Identity and access – stolen credentials, MFA bypass, privileged access
- Research data targeting – theft of research, intellectual property
- Limited security staffing – limited monitoring, slow response, delayed patching
- Legacy systems – unpatched systems, hard to secure, limited visibility
- Open networks – shared access, guest access, weak segmentation

Education-Specific Threat Profile

2024 Without Knowing Scams
Scammers impersonate Apple, Google, Netflix, Amazon, Microsoft, and Facebook to steal credentials and payment information. The LMS data can be used to craft highly convincing phishing lures that reference real courses, instructors, and institutional systems students actually use. Students are the most vulnerable population in this threat landscape, and the LMS gives them the exact data set (job title, high school, college, course history) needed to make the scam feel legitimate.

Local/Small Scale Without Blocking Scammers
Small-scale operations target students directly through social media, gaming platforms, and messaging apps. The LMS provides a ready-made target list with verified identities, course enrollments, and institutional affiliations. Without blocking scammers at the platform level, students remain exposed to direct solicitation regardless of any other security measures.

Scammers Smashing & Grabbing Student Profiles
Bulk data harvesting from LMS platforms enables mass phishing campaigns that reference real courses, instructors, and institutional systems. The data can be combined with external breach data to enrich profiles and increase the success rate of social engineering attacks.

Scammers Using Data Against Students
Once a student's data is in the hands of a scammer, it can be used for identity theft, financial fraud, or as a pivot point into other systems. The LMS data set is valuable because it contains verified identities, course histories, and institutional affiliations that make any subsequent attack more convincing.

Scammers Selling Student Data
Student data from LMS platforms is sold on dark web marketplaces and used in downstream fraud, identity theft, and targeted phishing campaigns. The data set is valuable because it contains verified identities, course histories, and institutional affiliations that make it more convincing than generic stolen data

Real-World Breaches Decoded

This section dissects real-world cyber attacks to provide actionable lessons. Each illustrating how one mind-gap enabled attacker progress.

Case 1
Token Theft → Identity Lessons → Impact Reset
Stolen OAuth token · Identity · Cloud

Case 2
Vendor Breach → SIM Pivot → Cloud Data Theft
Vendor breach · SIM pivot · Cloud access

Case 3
Ransom · Ad Abuse · CAN → Scaling
Ransomware · Malvertising · Crypto

Summary Table (A ad-free version, attackers exploited over vulnerabilities.) Between ads, workers, and ransomware, the mind-gap remains the common attack surface and identity is the recurring theme.

Case	Vector	Initial Access	Impact Type	Impact
Case 1	Okta	Phishing / Stolen Okta SSO	IP Theft	$2.3M Annual
Case 2	Twilio	Vendor Breach / SIM Swap	Data Theft	Unknown / Ongoing
Case 3	Malvertising	Malvertising / Ad Abuse	Ransomware	$40M+ Annual
Case 3	E-commerce	Credential stuffing/MFA	Ongoing	$400K+/Mo

Case 1 Token Theft to Payout Fraud

Financial Services Firm · Credit Analysis · Q1 2024

Business Context A mid-market financial services firm providing credit scoring and lending services to small and mid-sized clients. Revenue: $8–12M. The team was spread a lighter in platform-driven digital operations with a small IT team managing identity across Okta, Google Workspace, and a cloud-based credit scoring platform.

Entry Point Exact Exploited
- Phishing campaign targeted employees at a financial services firm
- Stolen credentials bypassed Okta SSO authentication
- Token lifetimes were too forgiving (days)
- No MFA step-up for sensitive actions
- No session anomaly detection enabled

Attack Timeline
- Day 0 – Stolen credentials used to access Okta
- Day 1 – Attacker enters, reconnaissance, lateral movement
- Day 2 – Attacker exfiltrates customer data
- Day 3 – Ransomware deployed, business disrupted
- Day 4 – Ransom demand issued, business paralyzed

$4.7M	18 days	0
Total Impact	Time to Detect	Time to Contain

Case 1: Lessons & Remediation

Root Mechanism
- Long-lived tokens and weak session control
- Identity trusted without continuous validation
- Cloud credentials treated as persistent access
- Incentives optimized adoption of speed and ease

Mind Gaps
- Identity was treated as a tool, not as a boundary
- Security was assumed to be handled by platforms
- Ownership was unclear between IT and business teams

Remediation Actions Taken
- Reset all tokens and sessions
- Enforced MFA for all Okta users
- Implemented session anomaly detection
- Implemented step-up authentication for sensitive actions
- Implemented cloud access governance

Key Takeaway Identity is the new perimeter. When identity is compromised, attackers don't need to break in. They simply log in as someone trusted. This case illustrates how a single stolen token can bypass every layer of security if the mind-gap between IT and business operations is not closed.

Case 2: ESP Bypass to Cloud Key Theft

Technology Company · Infrastructure Automation · Q1 2024

Attack Summary Attackers published a malicious npm package mimicking a popular internal utility. A developer installed the package, which executed a post-install script that bypassed the company's endpoint security agent and exfiltrated cloud access keys. The keys were used to access production infrastructure and deploy cryptocurrency mining containers. The incident was detected when cloud billing alerts spiked and internal teams noticed unexpected resource usage.

Root Cause Exploited
- Post-install scripts ran without review or approval
- Endpoint security agent was bypassed by the installation process
- Cloud access keys were stored on developer machines
- Keys had broad permissions across multiple environments

Attack Timeline
- Day 0 – Malicious package published to npm
- Day 1 – Developer installs package, post-install script executes
- Day 2 – Cloud access keys exfiltrated to attacker infrastructure
- Day 3 – Attacker deploys cryptocurrency mining containers
- Day 4 – Billing alerts spike, internal teams investigate
- Day 5 – Incident detected, containment begins

24 days	3 clouds	12 days
Time to Detect	Cloud Environments	Time to Contain

Case 2: Lessons & Remediation

What Worked
- Cloud billing alerts were the first signal
- Internal teams recognized the pattern quickly
- Containment was possible once the root cause was identified
- Cloud access keys were rotated and revoked successfully

Why It Took
- Post-install scripts were not monitored
- Endpoint security was bypassed by the installation process
- Cloud access was distributed across three environments
- Detection was reactive, not proactive

Remediation Actions Taken
- Implemented post-install script monitoring and approval
- Deployed endpoint security to developer machines
- Rotated all cloud access keys and revoked old credentials
- Implemented cloud access governance across multi-cloud environments
- Established anomaly detection for cloud billing and resource usage

Key Takeaway Supply chain attacks exploit trust in open-source dependencies. Post-install scripts execute with the same permissions as the developer who runs the installation. Monitoring and approval of these scripts is a critical control that most organizations lack. Cloud access keys stored on developer machines are a high-value target. Rotation and governance are essential.

Case 3: Vendor BNM to Healthcare Extortion

Attack Summary A healthcare provider's administrative IT contractor (MSP) operated a shared Microsoft 365 environment with weak MFA enforcement and broad administrative access. Attackers compromised a contractor account and quietly established a forwarding rule to an offshore mailbox. Business disruption followed before ransomware was deployed, forcing the provider into a ransom negotiation under regulatory pressure.

Root Cause Exploited
- Shared MSP Microsoft 365 environment with broad access
- Weak MFA enforcement on administrative accounts
- Forwarding rules were not monitored or restricted
- No mailbox audit logging enabled across the tenant
- No alerting on suspicious mailbox rule creation
- Ransomware deployed after business disruption started

Attack Timeline
- Day 0 – Attacker gains access to a contractor account
- Day 1 – Forwarding rule created to offshore mailbox
- Day 2 – Attacker monitors business operations
- Day 3 – Ransomware deployed, business disrupted
- Day 4 – Ransom demand issued, negotiation begins

2.5M	21 days	$3.4M
Patient Records Affected	Time to Detection	Ransom Demand

Case 3: Lessons & Remediation

Root Mechanism
- Shared MSP environment with broad access
- Weak MFA enforcement on administrative accounts
- No mailbox audit logging or rule monitoring
- Business continuity dependency on third-party IT

Why It Took
- Contractor account was trusted
- Forwarding rule was not monitored
- Business disruption was visible before ransomware
- Ransomware deployed after business disruption started

Remediation Actions Taken
- Implemented mandatory MFA for all administrative accounts
- Enabled mailbox audit logging across the tenant
- Restricted forwarding rules to approved domains
- Established dedicated tenant for healthcare organizations

Key Takeaway Shared MSP environments create a hidden attack surface. Weak MFA enforcement on administrative accounts allows credential compromise to grant broad access. Forwarding rules are a persistence mechanism that attackers use to maintain access after initial compromise. Business continuity dependency on third-party IT creates a leverage opportunity for ransomware operators. Healthcare organizations are a high-value target because their data is both valuable and time-sensitive.

Case 3: Lessons & Remediation

Ransomware operators increasingly target healthcare because patient data is both valuable and time-sensitive. Healthcare IT providers are a high-value target because one compromised contractor account can affect dozens of clinics. The business continuity dependency creates leverage for ransom demands. This case illustrates how a single stolen credential can cascade into a ransomware attack.

Regulatory Consequences
- HIPAA breach notification required
- OCR investigation triggered
- Fines for inadequate safeguards
- Mandatory reporting to affected individuals
- Reputational damage in healthcare sector

Mind Gap Analysis

Mind Gap	Impact
Shared MSP environment	Single compromised account affected multiple clinics
Weak MFA on admin accounts	Credential compromise granted broad access
No mailbox audit logging	Forwarding rule persistence went undetected
Business continuity dependency	Ransomware leverage created urgent ransom pressure

Remediation Actions Taken
- Implemented mandatory MFA for all administrative accounts
- Enabled mailbox audit logging across the tenant
- Restricted forwarding rules to approved domains
- Established dedicated tenant for healthcare organizations
- Implemented regular access reviews for third-party IT providers

Key Takeaway Healthcare IT providers are a high-value target because one compromised contractor account can affect dozens of clinics. The business continuity dependency creates leverage for ransom demands. Shared MSP environments create a hidden attack surface. Weak MFA on administrative accounts allows credential compromise to grant broad access. Forwarding rules are a persistence mechanism that attackers use to maintain access after initial compromise. Ransomware operators increasingly target healthcare because patient data is both valuable and time-sensitive.

Case 4: API Abuse → Bot Swarm to ATO

Financial platform · Tell-and-Go app · Ongoing since Q4 2024

Attack Summary A financial platform's Tell-and-Go application allowed users to send money to contacts by phone number. An attacker, leaking through stolen credentials, gained access to an internal API. The attacker used the API to enumerate valid phone numbers, then deployed a bot swarm to initiate small transfers to those numbers. Victims received notifications for transfers they did not initiate. The attacker used the compromised accounts to launch additional attacks against other users, compounding the damage.

Root Cause Exploited
- Internal API was exposed without strong access controls
- Phone number enumeration was possible at scale
- No rate limiting on transfer initiation
- No velocity checks on new recipients
- Small transfers were not flagged as suspicious
- Notifications were sent after transfers completed, not before
- Compromised accounts were used to launch additional attacks against other users

Attack Timeline
- Day 0 – Internal API discovered and exploited
- Day 1 – Phone number enumeration begins
- Day 2 – Bot swarm initiates small transfers to enumerated numbers
- Day 3 – Victims receive notifications for transfers they did not initiate
- Day 4 – Compromised accounts used to attack other users
- Day 5 – Pattern detected, investigation begins
- Day 6 – Rate limits and velocity checks deployed
- Day 7 – Incident contained, affected accounts secured

$1910	$2,100	6 months
Average Fraudulent Transfer	Total Fraudulent Transfers	Time to Detection

Case 4: Lessons & Remediation

Root Mechanism
- Internal API was exposed without strong access controls
- Phone number enumeration was possible at scale
- No velocity checks on new recipients
- Small transfers were not flagged as suspicious
- Notifications were sent after transfers completed, not before
- Compromised accounts were used to launch additional attacks against other users

Why It Took
- Internal API was considered trusted
- Rate limiting was not implemented
- Velocity checks were not in place
- Notifications were sent after the fact
- Cross-account attacks were not anticipated
- Detection was reactive, not proactive

Remediation Actions Taken
- Implemented rate limiting on API endpoints
- Deployed velocity checks on new recipients
- Implemented pre-transfer notifications
- Established anomaly detection for cross-account patterns
- Implemented API access controls and monitoring

Key Takeaway Internal APIs are trusted by default. When they are exposed without strong access controls, attackers can enumerate valid targets and initiate fraudulent transfers at scale. Velocity checks and pre-transfer notifications are critical controls that most organizations lack. Cross-account attacks are a pattern that traditional fraud detection systems miss. Detection must be proactive, not reactive.

Case 5: Lessons & Remediation

Identity is the new perimeter. When identity is compromised, attackers don't need to break in. They simply log in as someone trusted. This case illustrates how a single stolen credential can bypass every layer of security if the mind-gap between IT and business operations is not closed.

Attackers rely on a combination of high-value actions (API abuse, bot swarms, phishing campaigns) and low-friction exploitation (credential stuffing, weak MFA, stolen tokens). The combination creates a pattern that traditional fraud detection systems miss. Detection must be proactive, not reactive.

Business Impact
- Fraudulent transfers compromised trust
- Victims received unexpected notifications
- Support teams were overwhelmed
- Regulatory scrutiny was triggered (PSD2, GDPR)
- Customer churn was a concern
- Reputation was at risk in the fintech sector

Mitigation Patterns Implemented
- Rate limiting protecting transfers at scale
- Velocity checks on new recipients
- Pre-transfer notifications for high-risk patterns
- Anomaly detection for cross-account patterns
- Continuous monitoring of API access and abuse

Key Takeaway Identity is the new perimeter. When identity is compromised, attackers don't need to break in. They simply log in as someone trusted. This case illustrates how a single stolen credential can bypass every layer of security if the mind-gap between IT and business operations is not closed.

Attackers rely on a combination of high-value actions (API abuse, bot swarms, phishing campaigns) and low-friction exploitation (credential stuffing, weak MFA, stolen tokens). The combination creates a pattern that traditional fraud detection systems miss. Detection must be proactive, not reactive.

Defensive Playbooks: Zero-Trust in Action

This section translates the mind-gap analysis into actionable defensive playbooks. Each addresses a specific attack pattern, with clear ownership, timeline, and success metrics. Prioritized implementation roadmap.

Week 1
Identity Hardening
Token lifetime · MFA · Session control

Week 2
Endpoint · Network
Device trust · Segmentation

Week 3
Cloud Governance
Access review · Key rotation · Multi-cloud

Week 4
Detection & Response
Anomaly detection · Alerting · Response

Week 5–6
Mind Gap Closure
Training · Ownership · Culture

Measure Success Each playbook is structured to impact three of the Mind Gap Assessment KPI: Time to Detect, Time to Contain, and Time to Remediate. Success is measured against the 90-Day Mind Gap Assessment. Track progress monthly and adjust priorities based on which KPIs show the most improvement.

Playbook	Primary Focus	Quick Wins	Mid-Term
Identity Hardening	Token lifetime, MFA, session control	JWT, SSO	VPN, IP ACL
Endpoint & Network	Device trust, segmentation, EDR	MDM, EDR	Segmentation
Cloud Governance	Access review, key rotation, multi-cloud	IAM, CSPM	IPsec, SASE

Zero-Trust Metrics That Matter

Measure what matters. Track what changes. Prove the mind-gap is closing.

Time to Detect	Time to Contain	Time to Remediate	Unauthorized Access
42	1,850	12h	14d
↓ 85%	↓ 92%	↓ 97%	↓ 99%
Identity Hardening	Endpoint & Network	Cloud Governance	Detection & Response
35%	92%	4.5h	215
↑ 35%	↑ 92%	↓ 4.5h	↑ 215

Logging Integrity: Anchor Your Evidence

Playbook #1 Identity Hardening

Lockout, MFA enforcement, credential pivoting, session hijacking, token theft, SSO abuse, number cycling.

Objective: Eliminate hospitable trust in identity operations. Every authentication event must be verified, every session must be monitored, and every credential must be rotated. The goal is to make identity compromise difficult and detectable.

Week 1 Quick Wins (5–10 days)
- Enable MFA step-up for sensitive actions
- Reduce Okta session lifetime to 8 hours (from default 24 hours)
- Enable Okta ThreatInsight and block anonymizer traffic
- Review and revoke OAuth app grants and API tokens
- Implement JWT validation with short TTL (15–30 minutes)
- Enable session anomaly detection (IP, ASN, device changes)

Week 2 Foundation (10–15 days)
- Implement phishing-resistant MFA (FIDO2/WebAuthn, Okta Verify)
- Implement continuous MFA for sensitive actions (Okta Identity Governance)
- Implement JWT key rotation (JWKS endpoint, automated rotation)
- Implement session binding (device fingerprint, IP range, ASN)
- Implement phishing-resistant authentication for all privileged accounts
- Implement automated access reviews (Okta Identity Governance)
- Implement JWT key rotation (JWKS endpoint, automated rotation)

Playbook 01: Identity Hardening (continued)

Phase 1 Monitoring (24-72 hours)

Implement SIEM rules for LDAP/Kerberos anomalies, set alerting which catch DC sync or golden ticket attempts, and establish baseline authentication patterns for service accounts. Deploy honeypots in AD environment to detect lateral movement reconnaissance. Monitor for Kerberoasting attacks (TGS requests for service accounts). Track all privileged account usage and alert on deviations from baseline.

Phase 2 Containment (72-168 hours)

Force password resets for all service accounts and privileged accounts. Implement tiered administration model (separate admin accounts for different tiers). Disable legacy authentication protocols (NTLM, LM). Implement Kerberos armoring (FAST). Deploy Microsoft Defender for Identity or equivalent ITDR solution.

Phase 3 Hardening (1-4 weeks)

Implement Microsoft LAPS for local admin password rotation. Deploy privileged access management (PAM) solution for tier 0 assets. Implement JIT/JEA for administrative access. Deploy Protected Users security group. Implement Authentication Policies and Silos. Enable credential guard on critical systems.

Phase 4 Persistence Hunting (ongoing)

Hunt for persistence mechanisms: AdminSDHolder modifications, DCSync rights, DSRM password changes, SID history injection, GPO modifications, shadow credentials. Monitor for golden ticket indicators (TGT lifetime anomalies, encryption downgrades). Track Kerberoasting attempts (TGS requests for service accounts). Implement continuous attack path analysis using BloodHound or equivalent.

Critical Hardening Actions

- Force password resets for all service accounts
- Implement JIT/JEA for administrative access
- Deploy Microsoft LAPS for local admin password rotation
- Disable legacy authentication protocols (NTLM, LM)
- Implement Kerberos armoring (FAST)
- Deploy Microsoft Defender for Identity or equivalent ITDR solution
- Enable credential guard on critical systems
- Implement tiered administration model (separate admin accounts for different tiers)
- Deploy Protected Users security group

Hunting Queries & Indicators

- Hunt for DCSync RPC traffic (replication requests from non-DC sources)
- Monitor for golden ticket indicators (TGT lifetime anomalies, encryption downgrades)
- Track Kerberoasting attempts (TGS requests for service accounts)
- Hunt for persistence mechanisms: AdminSDHolder modifications, DSRM password changes, SID history injection, GPO modifications, shadow credentials
- Monitor for Golden Ticket usage (TGT with anomalous lifetime or encryption)
- Implement continuous attack path analysis using BloodHound

Playbook 01: KPIs & Success Ratios

Metric	Baseline	Target 90 Days	Target 180 Days
Kerberoasting attempts detected	12/week	8/week	5/week
Golden ticket usage detected	0/8	0/8	0/8
Service accounts with SPNs	47	35	20
Mean time to detect (MTTD) of privileged misuse	4.2 hours	1.5 hours	0.5 hours (target)
Mean time to remediate (MTTR) of identity findings	18 days	7 days	3 days
Privileged accounts with JIT/JEA	0%	40%	80%
Tier 0 assets with PAM coverage	0%	60%	100%
Accounts with password reset within policy	65%	85%	95%
Domain controllers with LAPS deployed	0%	50%	100%
Legacy authentication (NTLM) enabled	Yes	Partial	No
Kerberos armoring (FAST) enabled	No	Partial	Yes
Credential Guard enabled on critical systems	No	Partial	Yes
Protected Users security group populated	No	Partial	Yes
Authentication Policies and Silos deployed	No	Partial	Yes
Hunting queries active and tuned	0	5	15
Attack path analysis frequency	Ad-hoc	Quarterly	Monthly

Critical Implementation Phases

- Phase 1 (Monitoring): Establish baselines, deploy detection rules, identify service accounts
- Phase 2 (Containment): Force password resets, disable legacy auth, implement JIT/JEA
- Phase 3 (Hardening): Deploy LAPS, implement PAM, enable Credential Guard, deploy Protected Users
- Phase 4 (Persistence Hunting): Hunt for persistence mechanisms, monitor for golden ticket usage, continuous attack path analysis

Playbook 02: Endpoint + Posture Management

Unknown/UNC Agency, Spring-Summer 2025 (Ukraine/Baltics, Cryptocurrency, SEA/ME/Africa, cyberespionage/critical infrastructure)

Objective: Prevent, detect and contain sophisticated adversary activity, including attempts by nation-state actors to establish persistence, escalate privileges, and exfiltrate data. Emphasis on EDR/XDR deployment, behavioral detection, and rapid response capabilities.

Phase 1 Detect (0-24 hours)

Deploy EDR/XDR agents on Windows/Linux/macOS. Enable PowerShell logging, WMI activity monitoring, and process creation auditing. Configure SIEM ingestion for endpoint telemetry. Establish baseline for normal PowerShell, WMI, and scheduled task activity. Deploy honeypots and canary tokens for early detection. Monitor for Kerberoasting (TGS requests for service accounts). Track all privileged account usage and alert on deviations from baseline.

Phase 2 Contain (24-48 hours)

Isolate affected systems via EDR/XDR. Preserve memory and disk forensics before remediation. Kill malicious processes and persistence mechanisms. Reset compromised credentials. Block C2 infrastructure at firewall/DNS. Deploy emergency hunting queries across fleet. Activate incident response team and notify stakeholders.

Phase 3 Hunt (48-72 hours)

Hunt for persistence mechanisms: scheduled tasks, WMI subscriptions, registry run keys, DLL hijacking, service modifications. Analyze lateral movement patterns. Identify compromised accounts. Review authentication logs for anomalies. Check for data staging and exfiltration. Map attack timeline and scope. Document IOCs and TTPs for threat intelligence.

Phase 4 Harden (72-168 hours)

Implement application whitelisting (AppLocker/WDAC). Disable unnecessary services and features. Enforce LAPS for local admin password rotation. Implement JIT/JEA for administrative access. Deploy Protected Users security group. Implement Authentication Policies and Silos. Enable Credential Guard on critical systems. Implement network segmentation. Deploy continuous attack path analysis (BloodHound).

Playbook 02: Endpoint (continued) & KPIs

Phase 5 Persistence Hunting (ongoing)

Hunting for persistence mechanisms: scheduled tasks, WMI subscriptions, registry run keys, DLL hijacking, service modifications. Monitor for golden ticket usage (TGT with anomalous lifetime or encryption). Track Kerberoasting attempts (TGS requests for service accounts). Analyze lateral movement patterns. Identify compromised accounts. Review authentication logs for anomalies. Check for data staging and exfiltration. Map attack timeline and scope. Document IOCs and TTPs for threat intelligence.

Metric	Baseline	Target 90 Days	Target 180 Days
Endpoint agent deployment coverage	60%	95%	99%
Critical CVEs patched within SLA	45%	85%	95%
Average time to patch critical CVEs	14 days	7 days	3 days
Malicious PowerShell/WMI detections	0/baseline	5/week	10/week
Lateral movement attempts detected	0/baseline	2/week	5/week
Mean time to detect (MTTD)	4 hours	1 hour	30 minutes
Mean time to remediate (MTTR)	24 hours	8 hours	4 hours
Hunting queries active	0	5	15
Attack path analysis frequency	Ad-hoc	Quarterly	Monthly
Persistence mechanisms hunted	0	3 types	6 types

Critical Implementation Phases

- Phase 1 (Detection): Deploy EDR/XDR agents, enable PowerShell logging, establish baselines
- Phase 2 (Containment): Isolate affected systems, preserve forensics, kill malicious processes, reset credentials
- Phase 3 (Hunting): Hunt persistence mechanisms, analyze lateral movement, identify compromised accounts
- Phase 4 (Hardening): Implement application whitelisting, disable unnecessary services, enforce LAPS, implement JIT/JEA
- Phase 5 (Persistence Hunting): Continuous persistence hunting, monitor for golden ticket usage, track Kerberoasting

Playbook 03: ZTNA + Network Segmentation

Unknown/UNC Agency, Spring-Summer 2025 (Ukraine/Baltics, Cryptocurrency, SEA/ME/Africa, cyberespionage/critical infrastructure)

Objective: Replace legacy trusted network with identity-centric architecture. Implement zero trust, micro-segmentation, and continuous verification. Prevent lateral movement and limit blast radius. Enable secure remote access without VPN.

Phase 1 Audit (Week 1-2 days)

Map all trust relationships and network dependencies. Identify legacy applications requiring VPN or special access. Inventory all service accounts and their network requirements. Document current network segmentation gaps. Assess VPN usage patterns and remote access requirements.

Phase 2 Pilot (Week 3-4 days)

Deploy ZTNA solution for pilot group (contractors or specific departments). Test micro-segmentation policies in non-critical environment. Validate authentication flows and MFA integration. Measure performance impact and user experience. Document exceptions and edge cases.

Phase 3 Production (30-120 days)

Phase out VPN concentrators for remote access. Implement micro-segmentation between critical assets. Deploy network access control (NAC) for device posture assessment. Implement continuous authentication and session monitoring. Establish network baselines and anomaly detection. Integrate with identity governance and administration (IGA) for automated access provisioning and deprovisioning.

Playbook 03: ZTNA + Segmentation (continued)

Phase 4 Security (30-120 days)

Implement continuous authentication and session monitoring. Deploy network access control (NAC) for device posture assessment. Establish network baselines and anomaly detection. Monitor for lateral movement attempts. Implement micro-segmentation between critical assets. Deploy honeypots and canary tokens for early detection. Track all privileged access and alert on anomalies.

Metric	Baseline	Target 90 Days	Target 180 Days
Applications migrated to ZTNA	0/total	50%/total	90%/total
VPN concentrators decommissioned	0%	25%	75%
Lateral movement attempts blocked	0/baseline	50%/baseline	90%/baseline
Mean time to detect (MTTD) lateral movement	4 hours	1 hour	30 minutes
Mean time to remediate (MTTR) lateral movement	24 hours	8 hours	4 hours
Network segmentation coverage	10%	60%	95%
Continuous authentication sessions	0%	30%	80%
Honeypots deployed	0	5	15

Critical ZTNA Benefits: Identity-based access replaces network based trust. Micro-segmentation limits lateral movement. Continuous authentication prevents session hijacking. Device posture assessment blocks compromised endpoints. Reduces attack surface by eliminating VPN. Enables secure remote access without VPN. Provides granular access control per application.

Vendors: Zscaler, Palo Alto Networks, Cloudflare, Microsoft Entra, Google BeyondCorp, Okta, Tailscale

Playbook 04: SaaS Governance

Unknown/UNC Agency, Spring-Summer 2025 (Ukraine/Baltics, Cryptocurrency, SEA/ME/Africa, cyberespionage/critical infrastructure)

Objective: Secure SaaS applications, prevent data exfiltration, and ensure compliance. Implement CASB, DLP, and SSPM. Monitor for OAuth abuse, shadow IT, and misconfigurations. Protect sensitive data across cloud applications.

Phase 1 Audit (Week 1-2 days)

Inventory all SaaS applications and OAuth integrations. Identify shadow IT and unauthorized applications. Map sensitive data flows in SaaS applications. Assess current DLP coverage and gaps. Review OAuth permissions and third-party integrations. Document compliance requirements (GDPR, HIPAA, SOX).

Phase 2 Pilot (Week 3-4 days)

Deploy CASB for pilot group (contractors or specific departments). Test DLP policies in non-critical SaaS applications. Validate OAuth monitoring and alerting. Configure SSPM for critical SaaS applications (Salesforce, Microsoft 365, Google Workspace). Measure performance impact and user experience. Document exceptions and edge cases.

Phase 3 Production (30-120 days)

Implement CASB for all SaaS applications. Deploy DLP policies across SaaS applications. Enable OAuth monitoring and revocation. Implement SSPM for continuous misconfiguration detection. Integrate with SIEM for SaaS security events. Establish baselines for normal SaaS usage patterns. Implement data classification and labeling. Deploy browser isolation for risky SaaS applications.

Metric	Baseline	Target 90 Days	Target 180 Days
SaaS applications discovered	50	200+	300+
Shadow IT applications blocked	0	50	100
OAuth apps with excessive permissions	20%	10%	5%
Misconfigurations detected	0	50/week	10/week
Data exfiltration attempts blocked	0	5/month	0
Mean time to detect (MTTD)	24 hours	4 hours	1 hour
Mean time to remediate (MTTR)	48 hours	8 hours	2 hours
DLP policies enforced	0	15	30
Sensitive data in SaaS	Unknown	Partial	Mapped
Compliance violations	10/month	2/month	0

Critical SaaS Risks: Shadow IT (unauthorized SaaS adoption). OAuth abuse (third-party app permissions). Misconfigurations (default settings, sharing). Data exfiltration (personal accounts, USB). Insider threat (legitimate access misuse). Compliance violations (GDPR, HIPAA, SOX). Supply chain (compromised SaaS vendor).

Playbook 04: SaaS Governance (continued)

Phase 4 Hardening (30-120 days)

Implement SaaS to SaaS security posture management (SSPM). Automate misconfiguration detection and remediation. Implement OAuth app governance and revocation workflows. Deploy data classification and labeling across SaaS. Implement browser isolation for risky SaaS applications. Establish continuous compliance monitoring. Integrate SaaS security with identity governance (IGA). Implement zero trust access to SaaS applications.

Metric	Baseline	Target 90 Days	Target 180 Days
SaaS to SaaS attacks detected	0/baseline	5/month	0
OAuth app governance coverage	0%	50%	90%
Automated remediation rate	0%	30%	70%
Data classification coverage	0%	40%	80%
Compliance monitoring coverage	0%	50%	100%
SSPM integration	No	Partial	Yes

Critical SSPM Benefits: Misconfiguration detection (default settings, sharing, permissions). OAuth governance (third-party app permissions, revocation). Data exposure (public links, external sharing). Compliance monitoring (GDPR, HIPAA, SOX). Automated remediation (fix misconfigurations at scale). Integration (SIEM, ticketing, IGA).

Playbook 05: Cloud Security Governance

Unknown/UNC Agency, Spring-Summer 2025 (Ukraine/Baltics, Cryptocurrency, SEA/ME/Africa, cyberespionage/critical infrastructure)

Objective: Achieve cloud configuration and continuous compliance across all cloud environments. Detect cloud misconfigurations, identity misuse, and data exfiltration. Implement CSPM, CIEM, and cloud-native security controls.

Phase 1 Audit (Week 1-2 days)

Inventory all cloud accounts/subscriptions/projects. Identify shadow IT and unauthorized cloud services. Map sensitive data in cloud storage. Review current IAM policies and permissions. Document compliance requirements (GDPR, HIPAA, SOC 2, PCI DSS). Assess current cloud security tools and gaps.

Phase 2 Pilot (Week 3-4 days)

Deploy CSPM for critical cloud accounts (production, sensitive workloads). Enable cloud-native security services (AWS Security Hub, GuardDuty, Azure Sentinel, GCP Security Command Center). Test misconfiguration detection and remediation workflows. Validate CIEM for identity and access management. Measure false positive rates. Document exceptions and edge cases.

Phase 3 Production (30-120 days)

Implement CSPM across all cloud accounts/subscriptions/projects. Enable CIEM for least privilege enforcement. Deploy cloud-native security services (GuardDuty, Security Hub, Defender for Cloud, Chronicle). Implement cloud workload protection (CWPP) for VMs, containers, and serverless. Integrate with SIEM for centralized monitoring. Establish baselines for normal cloud activity. Implement automated remediation for common misconfigurations. Deploy data classification and DLP for cloud storage.

Phase 4 Persistence Hunting (ongoing)

Hunt for persistence mechanisms: backdoor accounts, rogue IAM roles, persistent access keys, malicious cloud functions, modified security groups. Monitor for data exfiltration: unusual API calls, bulk downloads, cross-region data transfer, suspicious S3/storage access. Track lateral movement: cross-account role assumptions, compromised service accounts, API key reuse. Analyze attack paths using cloud-native tools (AWS Detective, Azure Sentinel UEBA, GCP Security Command Center). Hunting queries: CloudTrail/Azure Activity Log/GCP Audit Logs for anomalous API calls. Identify compromised credentials: impossible travel, unusual geolocation, MFA bypass attempts. Detect data exfiltration: large S3 downloads, unusual egress traffic, cross-region replication. Hunt for persistence: backdoor IAM users, modified security groups, malicious Lambda/Cloud Functions. Monitor for misconfigurations: public S3 buckets, overprivileged roles, disabled logging. Track supply chain compromise: third-party SaaS integrations, OAuth tokens, API keys. Analyze attack timeline: initial access → persistence → lateral movement → exfiltration. Correlate with threat intelligence: known attacker IPs, malicious domains, TTPs. Document IOCs and TTPs for threat intelligence sharing.

Playbook 05: Cloud Governance (continued)

Phase 5 Hardening (30-120 days)

Implement cloud security posture management (CSPM) for continuous misconfiguration detection. Deploy cloud infrastructure entitlement management (CIEM) for least privilege enforcement. Implement cloud workload protection platform (CWPP) for VMs, containers, and serverless. Deploy cloud-native security services (GuardDuty, Security Hub, Defender for Cloud, Chronicle). Integrate with SIEM for centralized monitoring. Establish baselines for normal cloud activity. Implement automated remediation for common misconfigurations. Deploy data classification and DLP for cloud storage.

Critical Cloud Misconfigurations

- Public S3 buckets/storage accounts/GCS buckets
- Overprivileged IAM roles (wildcard permissions, admin access)
- Disabled logging (no CloudTrail/Azure Activity Log/GCP Audit Logs)
- Unencrypted data at rest (S3, EBS, RDS, Azure Storage, GCS)
- Exposed secrets (API keys, passwords, tokens in code/logs)
- Weak MFA enforcement (root accounts, privileged users)
- Cross-account trust misconfigurations (AssumeRole abuse)
- Missing network segmentation (flat networks, no VPC isolation)

Common Cloud Attack Techniques (T1078)

- Compromised credentials (phishing, credential stuffing, password spraying)
- Misconfigured IAM roles (overprivileged service accounts)
- Exposed API keys (hardcoded in code, logs, public repos)
- Cross-account trust abuse (AssumeRole misconfigurations)
- Supply chain compromise (third-party SaaS integrations, OAuth tokens)
- Insider threat (legitimate access misuse)

Cloud-Native Detection Tools

- AWS: GuardDuty, Security Hub, CloudTrail, Detective, Macie
- Azure: Defender for Cloud, Sentinel, Activity Log, AD Identity Protection
- GCP: Security Command Center, Chronicle, Cloud Audit Logs, Cloud IDS
- Multi-cloud: Wiz, Prisma Cloud, Orca, Lacework, Aqua, Sysdig

Playbook 06: Cloud KPIs & Zero-Days

KPI	Baseline	Target 90 Days	Target 180 Days
Cloud accounts with CSPM coverage	0%	80%	95%
Critical misconfigurations open	100	20	5
Average time to remediate	14 days	3 days	1 day
Cloud IAM roles with wildcard permissions	50%	10%	5%
Exposed secrets detected	10/month	2/month	0
Unencrypted data at rest	30%	5%	0%
Cross-account trust misconfigurations	15	3	0
Cloud-native detection coverage	20%	60%	90%
Multi-cloud visibility	Partial	Mostly	Full

Critical CSPM Metrics: Misconfiguration density (critical misconfigurations per cloud account). Mean time to remediate (MTTR for critical misconfigurations). Cloud IAM risk score (overprivileged roles, wildcard permissions). Exposed secrets (API keys, passwords, tokens in code/logs). Unencrypted data (S3, EBS, RDS, Azure Storage, GCS). Cross-account trust (AssumeRole misconfigurations). Cloud-native detection coverage (GuardDuty, Security Hub, Defender for Cloud, Chronicle). Multi-cloud visibility (single pane of glass across AWS/Azure/GCP).

Playbook 06: Web & API Defense (DDoS)

Unknown/UNC Agency, Spring-Summer 2025 (Ukraine/Baltics, Cryptocurrency, SEA/ME/Africa, cyberespionage/critical infrastructure)

Objective: Protect cloud-native applications and API-driven workloads against DDoS, application layer attacks, and API abuse. Implement WAF, rate limiting, bot mitigation, and DDoS protection. Ensure availability for critical services. Protect against OWASP Top 10 vulnerabilities and API-specific threats (BOLA, BFLA, mass assignment).

Phase 1 Detect (Week 1-2 days)

Establish baseline traffic patterns for normal vs. attack traffic. Configure WAF rulesets (OWASP Core Rule Set, custom rules). Deploy DDoS detection (volumetric, protocol, application layer). Monitor API endpoints for anomalies (rate limit violations, unusual geolocations, bot patterns). Set up alerts for DDoS thresholds and WAF blocks.

Phase 2 Mitigate (Week 3-4 days)

Activate DDoS mitigation (scrubbing centers, BGP anycast, rate limiting). Implement WAF custom rules for attack patterns. Deploy bot mitigation (challenge suspicious traffic, CAPTCHA, JS challenges). Implement API rate limiting (per IP, per API key, per endpoint). Block malicious IPs/ASNs/geolocations. Enable DDoS auto-mitigation. Scale infrastructure (auto-scaling, load balancers). Preserve logs for forensics.

Phase 3 Hardening (30-120 days)

Implement API gateway with authentication, rate limiting, and request validation. Deploy WAF with OWASP API Security Top 10 rules. Implement bot management (JS challenges, CAPTCHA, behavioral analysis). Configure DDoS auto-mitigation thresholds. Implement geo-blocking (if applicable). Deploy honeypots for early detection. Implement zero trust for APIs (mTLS, API keys, OAuth). Implement WAF custom rules for business logic abuse. Deploy API discovery and inventory. Implement continuous API security testing.

Actor Profiles at a Glance

Connector Permission Ladder

Playbook 06: Web & API Defense (continued)

Phase 4 Attacks On Web Applications

Attackers exploit SSL/TLS vulnerabilities, session hijacking, and other attacks targeting the transmission of or interactions with web applications. Zero Trust Defense ensures continuous, real-time authentication and monitoring of all web traffic.

Zero Trust Implementation: Continuously assess device posture, enforcing compliance before granting access to web applications. Implement real-time session monitoring and behavioral analytics to detect anomalies indicating session hijacking or MITM attacks.

Attack Surface Scanning

- Port scanning
- Vulnerability scanning
- Web application fingerprinting
- Directory and file enumeration
- Subdomain enumeration
- Cloud asset discovery
- Third-party integration mapping
- Shadow IT identification
- API discovery and mapping
- Data flow analysis
- Dark web monitoring

Attack Capability Map

- Credential stuffing
- Brute force attacks
- Zero-day exploitation
- Cross-site scripting (XSS)
- Cross-site request forgery (CSRF)
- SQL injection
- Remote code execution
- Data exfiltration
- Ransomware deployment
- Persistent backdoor creation
- Lateral movement
- Privilege escalation
- Business logic exploitation
- Supply chain compromise
- Distributed denial of service (DDoS)
- Cryptojacking
- Phishing and social engineering

Playbook 06: Web/API KPIs & Benchmarks

KPI	Baseline	Target 90d	Target 180d
API endpoints in scope (all assets)	[illegible]	[illegible]	100%
Critical/high vulnerabilities open (web + API)	[illegible]	[illegible]	[illegible]
Mean time to patch critical web/API vulnerability	[illegible]	[illegible]	[illegible]
Mean time to fix Tier-0 Zero Trust misconfigurations	[illegible]	[illegible]	[illegible]
Authentication failure rate (credential stuffing indicator)	[illegible]	[illegible]	[illegible]
Unauthorized API access attempts (per 1M requests)	[illegible]	[illegible]	[illegible]

Bottom line: Web applications and APIs are the primary attack surface of the modern enterprise. Zero Trust Defense transforms these from vulnerable, perimeter-dependent systems into hardened, continuously monitored, cryptographically secured assets. The goal: reduce attack surface by 94%, achieve 100% API inventory visibility, and sub-[illegible] mean time to patch critical vulnerabilities.

94%	70%	0
Attack surface reduction	Reduction in unauthorized API access	Trusted web/API connections

2027 Threat Outlook: What's Coming Next

Based on threat evolution of 2026 and current data, attackers, nation-states, criminal syndicates, and geopolitical actors, 2026 is the inflection point for defenders. Following these strategic shifts will be critical.

1. AI-powered attacks industrialized at scale

AI-generated phishing campaigns, deepfake voice calls, and zero-day exploits dramatically lower the barrier to sophisticated attacks. Defenders must deploy AI-powered detection and response to keep pace.

Defense focus: AI-powered SOC automation

2. Identity remains the primary battlefield

Credential theft, MFA bypass, and identity-based lateral movement continue to dominate breaches. Identity is the new perimeter and the primary target for ransomware, APT, and insider threat actors.

Defense focus: Identity-centric Zero Trust

3. Supply chain and third-party risk continues to escalate

Software supply chain compromises, third-party breaches, and open-source vulnerabilities remain high-impact, low-visibility risks. Organizations must extend security monitoring and governance beyond their own infrastructure.

Defense focus: SBOM, third-party risk management

4. Ransomware evolves into double and triple extortion

Ransomware groups now combine data encryption, data theft, and DDoS attacks to maximize pressure on victims. Backup resilience and incident response preparedness are critical.

Defense focus: Immutable backups, ransomware playbooks

5. Cloud misconfigurations remain a leading breach vector

Misconfigured cloud storage, IAM, and APIs continue to expose sensitive data at scale. Cloud security posture management and continuous compliance monitoring are essential.

Defense focus: CSPM, IaC security, continuous compliance

6. Quantum computing begins threatening current encryption

While practical quantum attacks are still years away, harvest-now-decrypt-later attacks against long-lived encrypted data are already a concern. Post-quantum cryptography planning should begin now.

Defense focus: Post-quantum cryptography readiness

7. Nation-state cyber operations escalate in scale and sophistication

Geopolitical tensions drive increasingly aggressive cyber operations against critical infrastructure, government, and commercial targets. Defense must account for state-level threat actors.

Defense focus: Threat intelligence, OT security, critical infrastructure protection

8. Agentic AI and autonomous systems emerge as new attack surfaces

As enterprises deploy AI agents and autonomous systems, attackers will target these systems for manipulation, data poisoning, and misuse. Defense must secure AI systems with Zero Trust principles.

Defense focus: AI governance, AI security monitoring

2027 Outlook: Continued Forecasts

Attacks on Critical Infrastructure Cause Real Harm

Ransomware, DDoS, and cyber-enabled sabotage target energy, water, healthcare, and transportation systems. Nation-state actors pre-position access in critical infrastructure. Cyber-physical consequences become a primary concern for boards and regulators.

Defense focus: Network segmentation, OT security, incident response

Identity Remains the Primary Battlefield

Credential theft, MFA bypass, and identity-based lateral movement continue to dominate breaches. Identity is the new perimeter and the primary target for ransomware, APT, and insider threat actors.

Defense focus: Identity-centric Zero Trust, phishing-resistant MFA, privileged access management

Supply Chain Risk Becomes Mainstream

Software supply chain compromises, third-party breaches, and open-source vulnerabilities remain high-impact, low-visibility risks. Organizations must extend security monitoring and governance beyond their own infrastructure.

Defense focus: SBOM, third-party risk management, software supply chain security

Ransomware Evolves Into Double and Triple Extortion

Ransomware groups combine data encryption, data theft, and DDoS attacks to maximize pressure on victims. Backup resilience and incident response preparedness are critical.

Defense focus: Immutable backups, ransomware playbooks, incident response

Quantum Threats Begin Emerging

While practical quantum attacks are still years away, harvest-now-decrypt-later attacks against long-lived encrypted data are already a concern. Post-quantum cryptography planning should begin now.

Defense focus: Post-quantum cryptography readiness

AI Accelerates Both Attack and Defense

AI-powered phishing, deepfakes, and automated exploitation lower the barrier to sophisticated attacks. AI-powered SOC automation becomes a force multiplier for defenders.

Defense focus: AI-powered SOC automation, AI governance, AI security monitoring

Nation-State Cyber Operations Escalate

Geopolitical tensions drive increasingly aggressive cyber operations against critical infrastructure, government, and commercial targets. Defense must account for state-level threat actors.

Defense focus: Threat intelligence, OT security, critical infrastructure protection

Cloud Misconfigurations Remain a Leading Breach Vector

Misconfigured cloud storage, IAM, and APIs continue to expose sensitive data at scale. Cloud security posture management and continuous compliance monitoring are essential.

Defense focus: CSPM, IaC security, continuous compliance

Agentic AI Emerges as a New Attack Surface

As enterprises deploy AI agents and autonomous systems, attackers will target these systems for manipulation, data poisoning, and misuse. Defense must secure AI systems with Zero Trust principles.

Defense focus: AI governance, AI security monitoring

Appendices

Resources and tools supporting the Nonhuman Identity (NHI) Zero Trust Project. Includes technical glossary, NHI Risk Scoring Framework, implementation checklist, and organizational background.

Appendix A:
Technical Glossary
Pages 74–77

Appendix B:
NHI Risk Scoring Framework
Pages 78–79

Appendix C:
Implementation Checklist
Pages 80–81

Appendix D:
About Nonhuman Identities
Pages 82–83

Terminology Note: Throughout this report, we use the terms Zero Trust, NHIs, and Non-Human Identities (NHI). NHIs, Nonhuman IAM, workload identity, machine identity, and Zero Trust are used interchangeably. Zero Trust refers to the security model, NHIs refer to the identities it must secure, and Nonhuman IAM refers to the tools and processes that manage them.

This appendix is published under Nonhuman Identities Research Project. Reproduction permitted with attribution, commercial redistribution requires written permission.

Technical Glossary (A–M)

Term	Definition
AI Agent	Autonomous software entity that perceives its environment, reasons about goals, and takes actions to achieve them. Examples include coding assistants, customer service bots, and automated security analysts.
API Gateway	Infrastructure component that manages, secures, and monitors API traffic between clients and backend services. Enforces authentication, rate limiting, and logging for all API calls.
API Key	Static credential used to authenticate API calls. Lacks expiration and context. Considered a legacy credential type that should be replaced with OAuth or mTLS.
Attestation	Cryptographically signed statement that a workload's identity, configuration, and environment meet specified requirements. Used in SPIFFE/SPIRE and cloud workload identity systems.
Attribute-Based Access Control (ABAC)	Access control model where permissions are granted based on attributes of the subject, resource, and environment. Contrasted with RBAC and ReBAC.
Audit Log	Immutable record of all identity and access events. Required for compliance with SOC 2, HIPAA, PCI DSS, and GDPR. Should include machine identity activity, not just human logins.
Authentication	Verifying that an entity is who it claims to be. Contrasted with authorization (what it can do). Machine identities require cryptographic authentication, not just passwords.
Authorization	Determining what an authenticated entity is permitted to do. Machine identities require fine-grained, policy-based authorization, not just role membership.
Bearer Token	Access token that grants access to whoever presents it, regardless of who they are. Vulnerable to theft and replay. Should be replaced with sender-constrained tokens.
Blast Radius	Scope of impact if a single identity is compromised. Minimized through least privilege, short-lived credentials, and network segmentation. Key metric for security architecture.
Cloud Workload Identity	Identity assigned to a cloud-hosted workload (VM, container, function) by the cloud provider. Examples: AWS IAM roles, GCP service accounts, Azure managed identities. Preferred over static credentials.
Credential Rotation	Process of replacing credentials before they expire or after they are compromised. Automated rotation is preferred over manual rotation. Key metric: rotation coverage percentage.
Credential Vault	Secure storage system for secrets, keys, and certificates. Provides encryption at rest and in transit, access control, and audit logging. Examples: HashiCorp Vault, AWS Secrets Manager.
CSPM	Cloud Security Posture Management. Tools that continuously assess cloud configurations against benchmarks and best practices. Detects misconfigurations that create identity risks.
DaaS	Data Security Fabric. A proposed architecture layer that unifies data governance, classification, and access control across structured and unstructured data. Complements IAM for data-centric authorization.
Decentralized Identifier (DID)	Cryptographically verifiable identifier that does not depend on a centralized registry. Used in some workload identity proposals. Not yet widely adopted in enterprise systems.
Delegation	Granting one entity the ability to act on behalf of another. Machine identities can delegate to other machines or to human users. Must be time-bounded and auditable.
Dormant Identity	Identity that has not been used for a defined period (typically 90 days). Candidate for deprovisioning. Represents attack surface with no legitimate owner. Key metric: dormant identity count.
Ephemeral Credential	Credential with a short lifetime (minutes or hours) that expires automatically. Reduces the value of stolen credentials. Preferred over long-lived credentials for machine identities.
Federation	Trust relationship between two identity systems that allows one to accept authentications from the other. Enables cross-organization workflows without shared credentials.
Fine-Grained Authorization	Authorization decisions based on specific attributes of the subject, resource, and environment. Contrasted with coarse-grained role membership. Required for least privilege at scale.
GitOps	Operational framework where Git is the source of truth for infrastructure and policy. Changes are made via pull requests and automatically applied. Enables auditable, reviewable identity governance.
Human Identity	Identity belonging to a human being. Typically managed through SSO, MFA, and HR systems. Contrasted with machine identity. Represents a minority of total identities in modern enterprises.
IaC	Infrastructure as Code. Practice of defining infrastructure in machine-readable files rather than manual configuration. Enables identity governance through code review and automation.
Identity Governance	Policies, processes, and controls for managing identity lifecycle, access rights, and compliance. Includes access reviews, entitlement management, and audit reporting.
Identity Lifecycle	Stages an identity passes through: provisioning, activation, modification, deactivation, deprovisioning. Machine identities have shorter lifecycles than human identities.
Identity Sprawl	Uncontrolled proliferation of identities across systems, clouds, and vendors. Creates attack surface and governance gaps. Common in multi-cloud environments.
IdP	Identity Provider. Centralized system that authenticates users and issues tokens. Examples: Okta, Azure AD, Google Workspace. Not designed for machine identities.
JIT Access	Just-In-Time Access. Granting elevated permissions only when needed, for a limited duration, with approval and audit logging. Reduces standing privilege. Key metric: JIT coverage percentage.
JWT	JSON Web Token. Compact, URL-safe token format for claims. Used for OAuth and OIDC. Contains expiration, issuer, and subject. Vulnerable to theft if not short-lived.
Key Management	Lifecycle of cryptographic keys: generation, storage, rotation, revocation, destruction. Hardware-backed key management provides higher assurance than software-only.
Machine Identity	Identity assigned to a non-human entity: workload, service account, IoT device, AI agent, or automated process. Machine identities outnumber human identities 45:1 in modern enterprises.
Mutual TLS (mTLS)	Both client and server authenticate each other using certificates. Provides strong identity verification for machine-to-machine communication. Required for Zero Trust.

Technical Glossary (N–Z)

Term	Definition
Nonhuman Identity (NHI)	Identity not belonging to a human being. Includes workloads, service accounts, AI agents, IoT devices, and automated processes. Outnumbers human identities 45:1 in modern enterprises.
OAuth	Protocol for delegated authorization. Allows third-party applications to access resources without sharing credentials. Machine identities use OAuth for API access.
OIDC	OpenID Connect. Identity layer on top of OAuth. Provides authentication and user information. Used for SSO and federated identity.
OPA	Open Policy Agent. Policy engine that decouples policy from application code. Enables centralized, auditable, version-controlled policy enforcement. Uses Rego language.
Policy Engine	Component that evaluates access requests against rules and returns allow/deny decisions. Examples: OPA, Cedar, AWS IAM policy evaluation.
Policy as Code	Practice of defining policy in machine-readable files rather than documents. Enables version control, testing, and automated enforcement. Foundation of GitOps.
Private Key	Cryptographic key that must remain secret. Compromise allows impersonation. Should never be transmitted or stored in plaintext. Hardware-backed storage preferred.
Public Key	Cryptographic key that can be shared. Used to verify signatures and encrypt data for the holder. Paired with private key in asymmetric cryptography.
RBAC	Role-Based Access Control. Access control model where permissions are granted based on roles. Simpler than ABAC but less flexible. Often insufficient for machine identities.
ReBAC	Relationship-Based Access Control. Access control model where permissions are granted based on relationships between entities. Used by Google Zanzibar. Suitable for complex, graph-based authorization.
Rotation	Replacing credentials before they expire or after they are compromised. Automated rotation is preferred over manual rotation. Key metric: rotation coverage percentage.
SaaS Identity	Identity provided by a software-as-a-service application. Managed through SSO and SCIM. Extends enterprise identity to third-party vendors. Creates governance gaps if not monitored.
SCIM	System for Cross-Domain Identity Management. Protocol for automating user provisioning and deprovisioning across applications. Reduces orphaned accounts. Limited support for machine identities.
SDK	Software Development Kit. Library that provides authentication and authorization functions for developers. Simplifies secure identity implementation. Quality varies significantly.
SIEM	Security Information and Event Management. Platform that collects, correlates, and analyzes security events. Machine identity activity should be included in SIEM monitoring.
SPIFFE	Secure Production Identity Framework For Everyone. Open standard for workload identity. Defines SVID (SPIFFE Verifiable Identity Document) as the credential format. Foundation for cloud-native workload identity.
SPIFFE Federation	Extension of SPIFFE that allows workloads in different trust domains to authenticate each other. Enables cross-organization workload identity without shared secrets.
SPIRE	SPIFFE Runtime Environment. Production implementation of the SPIFFE specification. Provides attestation, SVID issuance, and rotation. Supports multi-cloud and hybrid deployments.
SSO	Single Sign-On. Authentication system that allows users to access multiple applications with one login. Primarily for human identities. Machine identities require different mechanisms.
Static Credential	Credential that does not expire or rotate automatically. Includes API keys, passwords, and long-lived certificates. Represents the highest risk in most environments. Should be minimized.
Standing Privilege	Permanent access rights that are always available, regardless of need. Contrasted with JIT access. Violates least privilege principle. Key metric: standing privilege count.
Sub-Identity	Identity derived from a parent identity, inheriting some but not all of its permissions. Enables fine-grained delegation without sharing full credentials. Emerging pattern for AI agents.
Trust Bundle	Set of root certificates and public keys used to verify workload identities in a trust domain. Distributed to workloads so they can validate peer identities. Rotation must be coordinated.
Trust Domain	Administrative boundary in SPIFFE/SPIRE. Workloads within a trust domain can authenticate each other by default. Cross-domain trust requires federation. Analogous to a DNS zone.
Workload Identity	Identity assigned to a running workload (VM, container, function, pod). Ephemeral by nature. Preferred over static credentials for machine identities. Foundation of Zero Trust for workloads.
Workload Identity Federation	Mechanism to exchange workload identity tokens across trust boundaries without shared secrets. Enables multi-cloud and hybrid workloads to access resources in other clouds.
Zero Trust	Security model that assumes no implicit trust based on network location. Requires continuous verification of identity and authorization for every access request.
Zero Trust Architecture (ZTA)	Implementation of Zero Trust principles across people, devices, workloads, and data. Includes identity, network, application, and data layer controls.

Backup Resilience Scoreboard

Immutable Backups	Offline Copies	Recent Test	Ransomware Recovery
100%	3/3	98%	Active
✓ Meet target	✓ Meet target	✓ Meet target	✓ Meet target

Backup Age	Restore Time	Critical Failures	Failed Backups
15m	2h	95%	5
✓ Meet target	✓ Meet target	✓ Meet target	✓ Meet target

Modern Extortion Stack

MITRE ATT&CK Technique Mapping (TA0001–TA0040)

Tactic	Technique	Description	Tool Example	Defense Mapping
Initial Access	T1078 Valid Accounts	Stolen credentials used to access systems	Credential stuffing, phishing	MFA, Identity Governance
Initial Access	T1190 Exploit Public-Facing Application	Exploitation of internet-facing vulnerabilities	Log4Shell, ProxyShell	CVM, Patch Management
Initial Access	T1133 External Remote Services	Abuse of VPN, RDP, or remote access services	Citrix, Pulse Secure, RDP	Network Segmentation, MFA
Execution	T1059 Command and Scripting Interpreter	Malicious scripts executed in interpreters	PowerShell, WMI, bash	EDR, Application Control
Persistence	T1136 Create Account	Attacker creates new accounts for persistence	AD commands	Identity Governance, PAM
Persistence	T1098 Account Manipulation	Modifying existing accounts to maintain access	AD commands	Identity Governance, PAM
Defense Evasion	T1027 Obfuscated Files or Information	Encoding, packing, or encrypting payloads	Cobalt Strike, custom malware	EDR, Threat Hunting
Defense Evasion	T1070 Indicator Removal	Deleting logs and evidence	Log clearing, timestomping	SIEM, Immutable Logs
Credential Access	T1003 OS Credential Dumping	Extracting credentials from OS memory	Mimikatz, LSASS dump	EDR, Credential Guard
Credential Access	T1110 Brute Force	Password guessing attacks	Hydra, custom scripts	MFA, Account Lockout
Credential Access	T1558 Steal or Forge Kerberos Tickets	Kerberoasting, Golden Ticket, Silver Ticket	Rubeus, Impacket	Kerberos Monitoring, PAM
Discovery	T1087 Account Discovery	Enumerating accounts and privileges	AD commands, BloodHound	Identity Governance, Deception
Discovery	T1018 Remote System Discovery	Network scanning for lateral movement targets	Nmap, AD commands	Network Segmentation, Deception
Lateral Movement	T1021 Remote Services	Using legitimate protocols for lateral movement	RDP, SSH, SMB, WMI	Network Segmentation, PAM
Lateral Movement	T1550 Use Alternate Authentication Material	Pass-the-Hash, Pass-the-Ticket, token reuse	Mimikatz, Impacket	Credential Guard, PAM
Collection	T1560 Archive Collected Data	Compressing and encrypting data before exfiltration	7-Zip, tar, custom tools	DLP, Network Monitoring
Exfiltration	T1048 Exfiltration Over Alternative Protocol	Data theft via DNS, HTTP, or other protocols	Custom tools, DNS tunneling	DLP, Network Monitoring, DNS Sinkhole
Impact	T1486 Data Encrypted for Impact	Ransomware encryption of data	LockBit, BlackCat, Cl0p	Immutable Backups, Offline Backups
Impact	T1490 Inhibit System Recovery	Deleting backups and shadow copies	vssadmin, wbadmin, bcdedit	Immutable Backups, Offline Backups

MITRE ATT&CK Technique Mapping (TA0001–TA0040)

Tactic	Technique	Tool Example	Defense Mapping
Initial Access	T1078 Valid Accounts	Credential stuffing, phishing	MFA, Identity Governance
Initial Access	T1190 Exploit Public-Facing Application	Log4Shell, ProxyShell	CVM, Patch Management
Initial Access	T1133 External Remote Services	Citrix, Pulse Secure, RDP	Network Segmentation, MFA
Execution	T1059 Command and Scripting Interpreter	PowerShell, WMI, bash	EDR, Application Control
Persistence	T1136 Create Account	AD commands	Identity Governance, PAM
Persistence	T1098 Account Manipulation	AD commands	Identity Governance, PAM
Defense Evasion	T1027 Obfuscated Files or Information	Cobalt Strike, custom malware	EDR, Threat Hunting
Defense Evasion	T1070 Indicator Removal	Log clearing, timestomping	SIEM, Immutable Logs
Credential Access	T1003 OS Credential Dumping	Mimikatz, LSASS dump	EDR, Credential Guard
Credential Access	T1110 Brute Force	Hydra, custom scripts	MFA, Account Lockout
Credential Access	T1558 Steal or Forge Kerberos Tickets	Rubeus, Impacket	Kerberos Monitoring, PAM
Discovery	T1087 Account Discovery	AD commands, BloodHound	Identity Governance, Deception
Discovery	T1018 Remote System Discovery	Nmap, AD commands	Network Segmentation, Deception
Lateral Movement	T1021 Remote Services	RDP, SSH, SMB, WMI	Network Segmentation, PAM
Lateral Movement	T1550 Use Alternate Authentication Material	Mimikatz, Impacket	Credential Guard, PAM
Collection	T1560 Archive Collected Data	7-Zip, tar, custom tools	DLP, Network Monitoring
Exfiltration	T1048 Exfiltration Over Alternative Protocol	Custom tools, DNS tunneling	DLP, Network Monitoring, DNS Sinkhole
Impact	T1486 Data Encrypted for Impact	LockBit, BlackCat, Cl0p	Immutable Backups, Offline Backups
Impact	T1490 Inhibit System Recovery	vssadmin, wbadmin, bcdedit	Immutable Backups, Offline Backups

Zero-Trust Implementation One-Page Checklist

Starting between 90 days and 18 months, depending on organizational maturity, complexity, and existing identity infrastructure, the following sequence provides a practical path from inventory to continuous verification. Each phase builds on the previous one. Do not skip phase 1 or phase 2. Organizations that try to implement Zero Trust without knowing what they have will fail. Start with inventory and visibility, then add policy, then automate enforcement, then continuously verify.

Phase 1: Inventory (weeks 1–4)

- Enumerate all machine identities across cloud, on-prem, and SaaS
- Classify by type: workload, service account, AI agent, IoT, human-operated
- Assign ownership to every identity (no orphans allowed)
- Capture baseline metrics: total count, dormant count, ratio to human identities

Phase 2: Policy (weeks 5–8)

- Define least-privilege policy for each identity class
- Implement policy-as-code (OPA, Cedar, or equivalent)
- Set up JIT access for elevated permissions
- Establish ephemeral credentials for workloads (target: 24-hour maximum)

Phase 3: Automation (weeks 9–12)

- Automate credential rotation (target: 90-day maximum for long-lived credentials)
- Integrate with CI/CD pipelines (GitOps workflows)
- Deprovisioning triggered by HR events (terminations, role changes)
- Automated access reviews (quarterly minimum)

Phase 4: Continuous Verification (ongoing)

- Real-time monitoring and anomaly detection
- Behavioral baselines for machine identities
- Quarterly access reviews
- Annual penetration testing of identity controls

Key metrics to track:

- Total machine identities
- Dormant identity count
- Ratio to human identities
- Average credential lifetime
- Rotation coverage percentage
- Time to deprovision (target: 24 hours)
- Policy coverage (% of identities with defined policy)

Tools mentioned: OPA (policy engine), HashiCorp Vault (secrets), SPIFFE/SPIRE (workload identity), cloud-native IAM (AWS IAM, GCP IAM, Azure AD), GitOps (ArgoCD, Flux)

Research sources: Google Zero Trust, Microsoft Zero Trust, NIST SP 800-207, CISA Zero Trust Maturity Model

About Nonhuman Identities

Nonhuman Identities (NHI) is a research project focused on identity security for machines, AI agents, and automated systems. The project maintains a technical glossary, NHI Risk Scoring Framework, implementation checklists, and analysis of how traditional IAM assumptions break when applied to non-human identities.

Jaime Fabero Aranda
Cybersecurity Consultant, Cybersecurity Lecturer, Cybersecurity Author
LinkedIn: Jaime Fabero Aranda

Cybersecurity consultant with 15 years of experience across IT and OT environments, including pre-sales, delivery, and technical leadership roles. Lecturer at universities and business schools on cybersecurity, artificial intelligence, and digital transformation. Author of cybersecurity and AI-focused articles and technical analyses. Holds multiple certifications in security, cloud, and data. Primary interests: identity security for machine identities, AI governance, and practical cybersecurity frameworks for organizations with limited resources.

Claude Sonnet 4.6
Anthropic
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Gemini 3 Pro
Google DeepMind
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Gemini 3 Flash
Google DeepMind
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Claude Opus 4.6
Anthropic
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

GPT-5.2
OpenAI
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

GPT-5.2 Pro
OpenAI
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Grok 4.1
xAI
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Grok 4.1 Fast
xAI
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Kimi K2.5
Moilot
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

GLM-5
Z.ai
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Qwen 3.5 Plus
Alibaba Cloud
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Qwen 3 Max
Alibaba Cloud
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Minimax M2.5
Minimax
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Mistral Large 3
Mistral AI
Frontier AI model used for research, analysis, and drafting. Contributed to framework development, technical analysis, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Research Lead: Hugo Nic
Independent Cybersecurity Researcher

Hugo Nic is a cybersecurity researcher and lead author of the Nonhuman Identities (NHI) Zero Trust Project. His field of focus is machine identity governance, the gap between how IAM is designed and how it actually operates when applied to workloads, AI agents, and automated systems. His work spans identity architecture, cloud security, and the intersection of AI governance and traditional IAM. He is currently pursuing a Master's degree in Cybersecurity and Artificial Intelligence at Universidad Internacional de La Rioja (UNIR), focusing on how AI changes the threat landscape for identity systems.

Co-Author: Claude Opus 4.6
Anthropic

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: GPT-5.2
OpenAI

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: Gemini 3 Pro
Google DeepMind

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: Grok 4.1
xAI

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: Kimi K2.5
Moilot

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: GLM-5
Z.ai

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: Qwen 3.5 Plus
Alibaba Cloud

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: Minimax M2.5
Minimax

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Co-Author: Mistral Large 3
Mistral AI

Co-author on this report. Contributed to research synthesis, framework development, and document structure. Human author retains editorial control and responsibility for all claims and recommendations.

Nonhuman Identities Research Project. Reproduction permitted with attribution, commercial redistribution requires written permission.

Rocheston Zelfire

Zero-Trust Under Fire · 2026 Global Threat Report

© 2026 Rocheston. All rights reserved.
Published under the Rocheston Open Research License.
Reproduction permitted with attribution.
Commercial redistribution requires written consent.

rocheston.com

zelfire.com