NETWITNESS

REPORT

When Trust Becomes the Attack Surface: Tokens, Secrets, and the Identity Supply Chain

Series: 2025 Threat Landscape

Signals, Shifts, and Lessons Learned

Author: Rajas Save
NetWitness FirstWatch Threat Research & Intelligence

REPORT

Executive Summary

In 2025, significant intrusions primarily resulted from the exploitation of trusted access rather than novel exploits. Compromised OAuth tokens (digital credentials for authorization), leaked CI/CD (Continuous Integration/Continuous Deployment) secrets, and stolen software supply chain identities enabled adversaries to authenticate as legitimate users, integrate seamlessly into standard workflows, and propagate across interconnected systems.

To illustrate this strategic shift, the following section presents three operational examples demonstrating how attackers leveraged trusted access to bypass traditional defenses in 2025. These cases offer concrete evidence of evolving adversary tactics.

- **SaaS token abuse at scale:** Google Threat Intelligence Group (GTIG) reported UNC6395 activity in Aug 2025, using compromised OAuth tokens (authorization credentials for third-party apps) associated with Salesloft Drift to access and exfiltrate Salesforce customer data.

- **CI/CD as a credential spillway:** The compromise of tj-actions/changed-files involved retroactive tag manipulation and exposed secrets via workflow logs, impacting 23,000+ repositories.

- **Wormable open-source supply chain:** The Shai-Hulud npm (Node Package Manager, a package repository for JavaScript) campaign began with malicious package versions published on Sep 15, 2025, and used stolen tokens to propagate across additional packages and repositories.

In all these cases, intrusion relied on reusing trusted assets rather than bypassing controls. Tokens, grants, build identities, and maintainer credentials became primary intrusion vectors,

Table of Contents

Chapter	Page No.
Trust as the New Attack Surface	4
Drift OAuth Token Compromise Enables Salesforce Data Theft (UNC6395)	4
GitHub Actions Supply Chain Attack: Compromising tj-actions/changed-files	5
Shai-Hulud: Wormable npm Supply Chain Compromise	6
The Connective Tissue: How Modern Trust Gets Abused	7
Response Priorities: Minimizing the Blast Radius	7
NetWitness FirstWatch Insights: Why Correlation Wins	8
Looking Ahead	9
References	9

Trust as the New Attack Surface

A key takeaway from 2025 is that authorization materials, including tokens, sessions, application grants, CI identities, and registry tokens, became more valuable to adversaries than malware. This shift facilitated unobtrusive access and rapid lateral movement. Security teams increasingly recognize identity as the new control plane.

This development presents a fundamental paradox for defenders:

- Modern environments are optimized for speed and integration (SaaS↔SaaS, CI↔registry, CI↔cloud)
- Each integration creates a **trusted path**: scopes (permissions specifying what actions an app can perform), refresh tokens (credentials that renew access tokens), service accounts (automated user accounts for apps), and automation identities.
- If an attacker compromises any node, lateral movement can occur through **legitimate interfaces** (APIs, workflows, and grants), resulting in fewer detectable malware-related signals.

GTIG's Drift/Salesforce reporting captures this model end-to-end: compromised OAuth tokens enabled systematic data access, and the actor then searched for stolen credentials and tokens to expand access further.

Drift OAuth Token Compromise Enables Salesforce Data Theft (UNC6395)

GTIG described a widespread campaign attributed to UNC6395, operating as early **as Aug 8, 2025, through at least Aug 18, 2025**, targeting Salesforce customer instances via **compromised OAuth tokens** associated with the Salesloft Drift third-party application.

A notable consideration for defenders is the apparent normalcy of the access path:

- **Not a Salesforce platform vulnerability:** GTIG frames access as stemming from compromised third-party OAuth tokens rather than a core Salesforce flaw.
- **Bulk collection through legitimate query mechanisms:** GTIG details systematic querying and export activity consistent with automation and bulk retrieval.
- **Data theft as a means to harvest additional access:** GTIG observed post-exfiltration searches for "secrets," including strings consistent with **AWS access keys** and **Snowflake-related access tokens**, effectively turning CRM data into a credential-discovery surface. According to Google Cloud, the Logs Explorer lets users specify and view selected subsets of log entries for a project using Boolean expressions, enabling operational visibility and investigation even if certain query jobs are deleted. A single compromised integration token can serve as an access multiplier across multiple tenants and workflows.

NetWitness FirstWatch Insights: Hunting Priorities

Token-driven SaaS intrusions become apparent when defenders establish baselines for typical integration use, revealing clear and actionable patterns.

- **SaaS audit + query telemetry:** Baseline and alert on abnormal query volume, wide object enumeration, and high-limit query patterns from integration identities.
- **Connected-app authentication anomalies**: Investigate unusual authentication paths tied to connected apps (IP reputation, geography, time-of-day drift, unexpected user agents).
- **Downstream credential exposure:** Treat the "secret mining" behavior as a pivot for response. Prioritize rapid credential rotation and investigation in systems referenced in stolen records

GitHub Actions Supply Chain Attack: Compromising tj-actions/changed-files

In March 2025, multiple sources documented the compromise of tj-actions/changed-files, a widely used GitHub Action. The mechanics of this incident are significant because they exemplify a repeatable CI/CD failure mode.

StepSecurity reported that adversaries **modified the action's code and retroactively updated multiple version tags** (labels used to identify code versions) to reference a malicious commit (an unauthorized code change). As a result, workflows pinned only to tags could silently begin executing attacker-controlled content. The compromised action printed CI/CD secrets into GitHub Actions build logs (records generated during automated tasks), posing an elevated risk when those logs are publicly accessible.

CISA's alert corroborated the severity and scope of impact, describing information disclosure of secrets, including valid access keys, GitHub personal access tokens, npm tokens, and private RSA keys.

GitHub's advisory likewise summarizes the incident and its impact ("over 23,000 repositories") and notes that the compromised action executed a malicious Python script (a file that carries out automated tasks) as part of the intrusion chain.

Wiz further reported linkage and follow-on exposure involving reviewdog/action-setup, reinforcing that the immediate risk for many victims was not continued exploitation, but **persistent exposure** of secrets already printed into historical logs.

Prevalence of CI/CD in intrusions: CI systems are positioned on the most direct path to production, with deployment credentials, signing keys, registry tokens, and cloud access commonly present in runners and workflows. When secrets are exposed, adversaries can pivot directly into production control planes with minimal reliance on endpoint malware.

NetWitness FirstWatch Insights: Hunting Priorities

- **Tag movement and release integrity**: Treat unexpected tag changes/retagging as alertable activity. The incident demonstrates why "pinning to a tag" is not the same as pinning to immutable code.
- **Historical workflow log exposure as IR scope:** Assume compromise may have created durable exposure in logs; IR should include workflow history review and comprehensive secret rotation decisions.
- **Runner behavior and unexpected scripts:** Prioritize detection of anomalous runner execution paths and unexpected script stages during CI runs, especially when secrets are involved.[5]

Shai-Hulud: Wormable npm Supply Chain Compromise

In September 2025, Wiz documented a campaign commonly referred to as Shai-Hulud, beginning with malicious versions of popular npm packages published on September 15, 2025. Wiz describes a post-install script that harvested sensitive data and exfiltrated it to attacker-created public GitHub repositories, and highlights worm-like behavior: when malware encountered additional npm tokens, it could publish malicious versions of other accessible packages.

Unit 42 characterized Shai-Hulud as a **self-replicating worm** that compromises hundreds of packages, emphasizing that stolen tokens enable attackers (or automated payloads) to identify and republish compromised packages at scale.

JFrog's continued tracking underscored the persistence of this threat class, reporting a later wave with more advanced tactics and **an additional 796 new malicious packages.**

Distinction from typical dependency risk: The Shai-Hulud campaign combines credential theft with automated propagation. Containment efforts must address not only the removal of malicious versions, but also the likelihood that environment tokens have been harvested and may already be used to facilitate further compromise. (Intelligence, 2023)

NetWitness FirstWatch Insights: Hunting Priorities

- **Install-time execution signals**: Monitor for unexpected lifecycle script execution, unusual child processes during installs, and spikes in outbound connections during dependency resolution.
- **Exfil patterns to code hosting:** Where audit data exists, monitor for suspicious repository creation and for unusual automated commits/repositories associated with compromised tokens.
- **Token hygiene and rapid rotation:** Propagation depends on available credentials; rapid token rotation and privilege minimization directly reduce worm "fuel."

The Connective Tissue: How Modern Trust Gets Abused

Although these incidents appear distinct (SaaS token abuse, CI action compromise, and npm worm), they share a common pattern of intrusion logic:

Obtain trust material	Operate through legitimate interfaces	Blend into expected behavior	Harvest more secrets to expand
(tokens/keys/ maintainer identity)	(APIs, workflows, registries)	(automation looks like automation)	(CRM to cloud tokens; CI logs to prod keys; npm tokens to more packages)

It is essential for defenders to recognize that in 2025, secrets hygiene and integration governance were not merely best-practice checklists. They constituted the primary security controls.

Response Priorities: Minimizing the Blast Radius

Rather than relying on extensive checklists, the incidents of 2025 underscore three priorities that consistently influence security outcomes.

Treat tokens as incident artifacts, not implementation details

- Inventory connected apps and integration grants (who has what scopes, where refresh tokens live).
- Build a token incident runbook that includes **revocation**, **rotation**, and downstream hunting as first-hour actions.
- Use SaaS audit telemetry to hunt for high-volume export patterns and anomalous API activity.

Assume CI/CD is production and harden accordingly

- Prefer immutable pinning for dependencies (avoid trusting tags alone).
- Reduce and compartmentalize secret exposure: least-privilege tokens, short-lived credentials, strict log access controls.
- After CI supply chain events, scope IR to include workflow history and historical log exposure (not just current workflow state).

Defend the software supply chain against propagation

- Pin dependencies (version + integrity) and monitor for unexpected install-time execution.
- Enforce strong maintainer authentication and registry token governance; monitor for anomalous publishing activity.
- When compromise is suspected, rotate credentials broadly. Worms spread by harvesting what's already present.

NetWitness FirstWatch Insights: Why Correlation Wins

As intrusions increasingly mimic legitimate user and system behavior, effective Security Operations Centers (SOCs) must correlate signals across identity, API activity, build telemetry, and network context rather than relying solely on malware detection.

NetWitness FirstWatch delivers regularly updated threat logic, detection rules, and machine-readable threat intelligence mapped to the MITRE ATT&CK framework. This content is continuously informed by the latest research and intelligence from diverse sources and distributed to customers via NetWitness Live as detection rules, parsers, and threat feeds. Integrating NetWitness into an environment provides organizations with a continuous stream of actionable intelligence, operationalizing new detections and IOCs directly into detection and response workflows to help security teams keep pace with evolving adversary tactics.

FirstWatch's mission is to enable detection-driven intelligence and make it directly consumable within the NetWitness platform. The NetWitness approach centralizes analytics, detection, and incident response and is designed to support complex investigations spanning SaaS audit trails, CI/CD systems, and developer ecosystems. This approach reduces gaps as threat actors move laterally across modern environments.

Looking ahead »

While 2025 demonstrated the critical role of trust primitives, 2026 is expected to see accelerated and increasingly automated exploitation of these vulnerabilities, particularly at the intersection of SaaS integrations, developer tools, and CI pipelines.

The next installment in this series (Part 2), 'Exploitation at Machine Speed: Edge Breakouts and the Rise of Agentic Operations,' will examine how 2025 became the year of large-scale exploitation campaigns targeting edge infrastructure and enterprise platforms at unprecedented speed.

References

[1] Google Threat Intelligence Group (GTIG): Widespread Data Theft Targets Salesforce Instances via Salesloft Drift
[2] Step Security: tj-actions/changed-files action is compromised
[3] Wiz: Shai-Hulud npm Supply Chain Attack
[4] Wiz: Shai-Hulud 2.0 – Ongoing Supply Chain Attack
[5] CISA: Supply Chain Compromise of Third-Party tj-actions/changed-files (CVE-2025-30066) and reviewdog/action-setup
[6] GitHub Advisory Database: CVE-2025-30066 / GHSA-mrrh-fwg8-r2c3
[7] Unit 42: "Shai-Hulud" Worm Compromises npm Ecosystem in Supply Chain Attack
[8] JFrog: Shai-Hulud npm supply chain attack – new compromised packages detected
[9] JFrog Research: Shai-Hulud "The Second Coming"
[10] Microsoft Security Blog: Automatic disruption of human-operated attacks through containment of compromised user accounts
[11] Cybersecurity Insiders: 2025 Pulse of the AI SOC Report: From Alert Fatigue to Actionable Intelligence
[12] NetWitness: FirstWatch maps threat intelligence content to the MITRE ATT&CK Framework
[13] NetWitness: Threat Detection & Response platform overview

Contact Information

Email for customer Service
support@netwitness.com

Website
www.netwitness.com

Follow us for regular updates

NETWITNESS