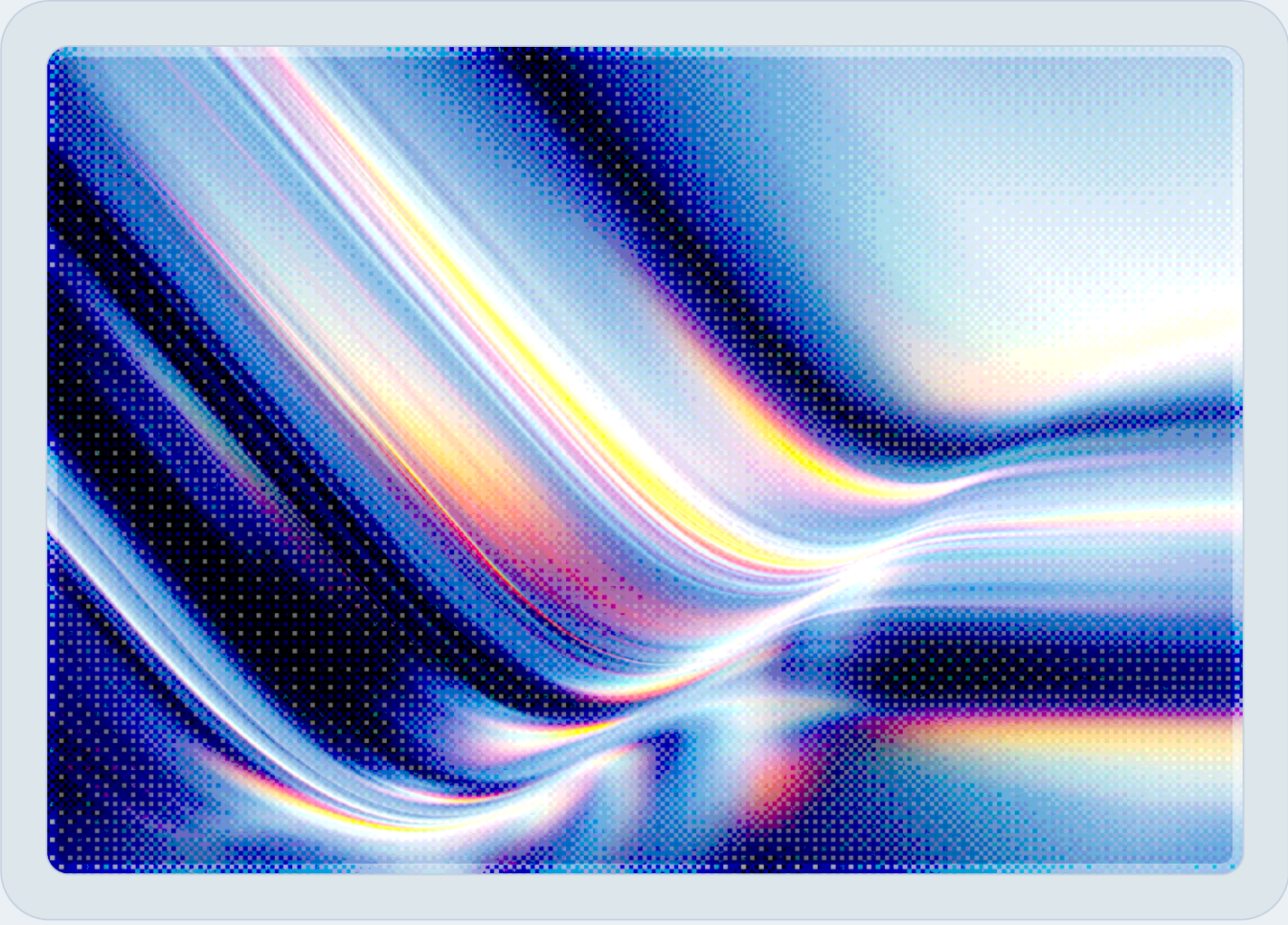
CSAI Foundation | Cloud Security Alliance

NIST and CISA Finalize Guidance on Identity Token Theft

2026-09-23

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

Key Takeaways

On September 15, 2026, NIST published the final version of NIST Interagency Report (IR) 8587, "Protecting Tokens and Assertions from Forgery, Theft, and Misuse: Implementation Recommendations for Agencies and Cloud Service Providers," with CISA jointly announcing the guidance for federal cloud identity systems the same week [1][2][3]. The report gives cloud service providers and federal agencies a shared, outcome-based framework for hardening the digitally signed tokens and assertions that underpin single sign-on, identity federation, API access, and workload-to-workload authentication, replacing prescriptive rules with "MUST" and "SHOULD" requirements that organizations can implement through varied architectures [3][4]. NIST developed the guidance in coordination with CISA's Joint Cyber Defense Collaborative and in support of Executive Order 14306, incorporating roughly 250 public comments on the December 2025 draft and input from a June 2025 technical exchange involving more than fifty experts from providers including Microsoft, Amazon Web Services, Google, Oracle, and Okta [1][3]. The report explicitly cites the 2023 Storm-0558 intrusion, in which a compromised Microsoft consumer signing key was used to forge authentication tokens accepted by enterprise systems, exposing more than 60,000 emails from a single federal agency, as the kind of validation failure its key-management and token-verification requirements are designed to close [4][5]. Core recommendations cover cryptographic key protection and rotation, mandatory token claims and verification steps, shortened token lifetimes, sender-constrained tokens, working revocation, and logging sufficient to reconstruct an incident, while also flagging non-human and AI agent identities as an emerging category that uses the same token mechanisms without dedicated guidance yet in place [3][5][6]. Conformance with NIST IR 8587 is voluntary for organizations outside the federal government unless imposed by policy or contract; this note's assessment is that its detailed, vendor-neutral requirements nonetheless make it a useful reference for any organization securing OAuth 2.0, OpenID Connect, or SAML-based identity infrastructure [6].

Background

Token-based authentication is now widely used as the primary mechanism by which cloud services, federated identity providers, and increasingly autonomous software agents establish trust with one another, alongside or in place of the session cookies and static credentials that once dominated web application security. A signed token or assertion, whether an OAuth 2.0 access token, an OpenID

Connect ID token, or a SAML assertion, allows a relying party to trust a claim about a user's or workload's identity without contacting the issuing identity provider directly, which is precisely what makes the underlying signing keys and verification logic such high-value targets. When those controls fail, the consequences can be severe and difficult to detect: the Storm-0558 campaign that NIST IR 8587 cites by name allowed a nation-state actor to forge authentication tokens for Microsoft enterprise email accounts using a signing key intended only for consumer services, a validation gap that went undetected long enough for the actor to access unclassified email from senior U.S. government officials [4][5]. That incident, alongside the SolarWinds supply chain compromise that NIST IR 8587 also references [3], suggests how a single point of cryptographic failure in identity infrastructure can cascade into widespread, hard-to-remediate access across an organization's cloud environment.

NIST began developing IR 8587 against that backdrop, publishing an initial public draft in December 2025 and soliciting comment through the spring of 2026. The development process included a technical exchange in June 2025 with more than fifty participants from major cloud and identity providers, and the December draft drew close to 250 public comments before NIST finalized the report in September 2026 [1][3]. That engagement shaped a document explicitly built around configurability and interoperability rather than a single prescribed architecture, a design choice NIST's Digital Identity Program Lead, Ryan Galluzzo, framed as applicable well beyond the federal government: "Anyone who is using tokens as part of their access management infrastructure can look to this for insights, whether they are in government or commercial industry" [2]. The guidance also arrives under the authority of Executive Order 14306 and reflects coordination through CISA's Joint Cyber Defense Collaborative, positioning it as part of a broader federal push to harden cloud identity infrastructure following token-based intrusions such as Storm-0558 and the SolarWinds compromise across both government and commercial environments [1].

This research note summarizes the scope and core technical requirements of NIST IR 8587, examines how the guidance addresses the specific failure modes exposed by incidents like Storm-0558, and provides recommendations for security and identity teams, both inside and outside the federal government, that manage token-issuing or token-consuming infrastructure.

Security Analysis

Scope and structure of the guidance

NIST IR 8587 applies to identity and access management systems that rely on digitally signed assertions and tokens based on asymmetric cryptography to make access decisions, a scope that spans traditional human single sign-on and federation as well as API access and authentication for automated workloads

and services [1][6]. The report divides responsibility explicitly between cloud service providers, who issue tokens and must secure the infrastructure and signing keys behind them, and consuming organizations, who configure access policies, apply the security controls providers expose, and monitor for misuse; NIST frames coordinated incident response and revocation across that provider-customer boundary as essential to the guidance working as intended [5]. Rather than mandating a single implementation, the report expresses its requirements as documented, risk-based decisions: organizations must record their token architectures, protocols, lifetimes, validation logic, key-management practices, revocation procedures, session controls, logging, and incident response plans, giving auditors and security teams a checklist against which to evaluate an existing deployment even though the underlying technical choices remain flexible [5].

Key management and signing-key protection

The report's key-management requirements respond directly to the Storm-0558 failure mode, in which a signing key scoped for one trust boundary was mistakenly honored across another. NIST IR 8587 requires that signing and verification keys use NIST-approved algorithms and FIPS 140-validated cryptographic modules, and for moderate-impact systems and above, that keys reside in hardware-based, hardware-backed, or otherwise isolated storage rather than general-purpose application memory [5]. High-impact systems face additional isolation requirements, including separating signing operations from general-purpose compute environments, and the guidance caps the active lifetime of high-impact signing keys at no more than 90 days, while moderate- and low-impact system keys are expected to remain in use for less than a year, with all rotation intervals documented and justified by risk rather than left to informal practice [5][6]. That combination, isolated key storage plus a bounded active lifetime, is designed to narrow the window in which a single compromised key, of the kind at the center of Storm-0558, can be used to forge tokens accepted across unrelated trust boundaries.

Token verification, lifetimes, and sender-constraining

Beyond key protection, NIST IR 8587 specifies what a compliant token must contain and how a resource server must validate it. Tokens are required to carry issuer identification, subject or client identity, an intended audience, issuance time, a validity window, a unique identifier or nonce, authentication time, and a signature, and resource servers must verify the signature, source, integrity, scope, and audience of every token before granting access rather than trusting a token solely because it parses correctly [5]. The report shortens acceptable token lifetimes considerably, calling for access and identity tokens to expire within one hour, with refresh tokens subject to their own expiration, replay protection, secure storage, and revocation requirements [5]. It also recommends sender-constrained tokens, cryptographically binding a token to the specific client or device that requested it so that a stolen token

cannot simply be replayed from a different location, an approach implemented in practice through mechanisms such as Demonstrating Proof of Possession (DPoP) or mutual TLS [5]. Table 1 summarizes how the report's impact-tiered requirements scale across low-, moderate-, and high-impact systems.

Requirement	Low-impact	Moderate-impact	High-impact
Signing key storage	Documented, risk-based	Hardware-based/hardware-backed/isolated	Hardware-based/hardware-backed/isolated, signing isolated from general compute
Active signing key lifetime	Generally < 1 year	Generally < 1 year	≤ 90 days
Access/ID token lifetime	≤ 1 hour	≤ 1 hour	≤ 1 hour
Sender-constraining	Recommended	Recommended	Recommended, higher priority

Emerging gaps: AI agents and post-quantum migration

NIST IR 8587 acknowledges, without fully resolving, two forward-looking concerns that are directly relevant to organizations deploying AI systems. First, it notes that AI agents authenticate and access resources using the same token mechanisms the report otherwise governs, but states that the broader problem of non-human and agentic identity management requires guidance beyond the current document's scope, leaving a gap that organizations building agentic AI systems will need to fill with additional controls [3][5]. Second, the report addresses post-quantum cryptography only at the level of migration planning, urging organizations to prepare for an eventual transition rather than mandating specific post-quantum signing algorithms today, consistent with the broader federal cryptographic transition timeline but leaving implementation specifics to future guidance [3][5].

Recommendations

Immediate Actions

Security and identity teams should inventory every system that issues or consumes signed tokens or assertions, whether for human single sign-on, API access, or workload-to-workload authentication, and compare each system's signing-key storage, key lifetime, and token-lifetime settings against the impact-tiered requirements in NIST IR 8587, prioritizing any environment where a single signing key is shared across multiple trust boundaries in the pattern that enabled Storm-0558 [4][5]. Teams should confirm that resource servers actually validate every required token claim, issuer, subject, audience, issuance time, validity window, nonce, and signature, rather than assuming that token parsing alone constitutes verification, since incomplete validation logic was the specific gap that allowed a consumer-scoped key to be accepted for enterprise tokens [4][5]. Federal agencies and cloud service providers directly in scope of the guidance should begin documenting their token architectures, key-management practices, and revocation procedures now, since NIST IR 8587 frames that documentation as a baseline expectation even though the report does not impose a fixed compliance deadline [5][6].

Short-Term Mitigations

Organizations should evaluate whether their identity infrastructure currently issues access and identity tokens with lifetimes longer than the one hour NIST IR 8587 recommends, and, where longer-lived tokens are in use for operational reasons, assess whether shortening them or layering in refresh-token revocation and replay protection meaningfully reduces exposure without breaking legitimate workflows [5]. Teams operating in cloud environments that support Demonstrating Proof of Possession or mutual TLS should evaluate enabling sender-constrained tokens for their highest-value identity flows, since binding a token to its requesting client closes the specific replay path that allows a stolen bearer token to be used from an attacker-controlled system [5][7]. Security teams should also verify that logging is sufficient to reconstruct a token-based incident after the fact, including which keys signed which tokens, when tokens were issued and revoked, and which resource servers accepted them, since NIST IR 8587 treats that forensic capability as a core requirement rather than an optional enhancement [5].

Strategic Considerations

Organizations building or expanding agentic AI systems should treat NIST IR 8587's acknowledgment of AI agent token usage as an early signal rather than a completed answer: the underlying token-verification, key-management, and sender-constraining requirements in the report apply to agent

authentication today, even though dedicated non-human identity guidance has not yet been finalized, and security teams should not wait for that follow-on guidance before applying the report's existing controls to agent-issued and agent-consumed tokens [3][5]. Because conformance with NIST IR 8587 remains voluntary outside of policy or contractual mandates, commercial organizations should weigh adopting its impact-tiered key-management and token-lifetime requirements voluntarily against the demonstrated impact of token-based intrusions like Storm-0558, particularly for identity providers and cloud services that sit at the center of federated trust relationships with many downstream customers [4][6]. Finally, security and cryptography teams should begin tracking NIST's post-quantum migration guidance as it develops, since the report's current treatment of post-quantum readiness as a planning exercise rather than a mandate suggests more prescriptive requirements are likely to follow as post-quantum signing standards mature [3][5].

CSA Resource Alignment

NIST IR 8587's emphasis on identity-provider architecture, key management, and verification controls aligns closely with CSA's "Zero Trust Principles and Guidance for Identity and Access Management (IAM)," a technology-agnostic implementation guide for applying Zero Trust principles to IAM systems; though the two documents were developed independently, the guide's continuous-verification and least-privilege principles for identity infrastructure map onto NIST IR 8587's requirement that resource servers fully validate every token claim rather than treating a signed token as inherently trustworthy [7]. CSA's "Navigating Identity and Access Management (IAM): Standards and Protocols" complements that alignment with a protocol-level view, walking through how OAuth 2.0, SAML 2.0, OpenID Connect, SCIM, and FIDO2/WebAuthn support authentication, authorization, and provisioning for human, service, and agent identities alike, the same protocol family that NIST IR 8587's token and assertion requirements govern [11]. CSA's "VS Code Zero-Day: One-Click GitHub Token Theft," a rapid-response threat intelligence report on a disclosed vulnerability that allowed theft of broadly scoped OAuth tokens through a webview sandbox escape, offers a concrete, recent illustration of exactly the token-theft risk category NIST IR 8587 addresses, and its exploit-chain analysis underscores why the report's sender-constraining and short token-lifetime recommendations matter even for tokens stolen through client-side vulnerabilities rather than server-side key compromise [8]. CSA's "Agentic AI Identity and Access Management: A New Approach" speaks directly to the gap NIST IR 8587 leaves open around AI agent identity, arguing that conventional protocols such as OAuth 2.1, SAML, and OIDC do not adequately recognize agents as first-class actors and proposing decentralized identifiers, verifiable credentials, and delegation chains that organizations can apply to agent-issued tokens while dedicated federal guidance on non-human identity is still pending [9]. Finally, the AI Controls Matrix (AICM) v1.1, CSA's control framework for cloud and AI environments, provides the IAM domain controls, covering key management,

credential lifecycle, and access verification, that can be used to translate NIST IR 8587's requirements into auditable controls applicable across the model-provider, application-provider, and cloud-service-provider roles [10].

References

[1] Cybersecurity and Infrastructure Security Agency. "CISA and NIST Release Guidelines to Protect Federal Cloud Identity Systems from Token Theft, Forgery, and Misuse." CISA, September 2026.

[2] National Institute of Standards and Technology. "NIST Finalizes Guidelines on Protecting Online Identity and Access Tokens From Misuse." NIST, September 2026.

[3] National Institute of Standards and Technology. "NIST IR 8587: Protecting Tokens and Assertions from Forgery, Theft, and Misuse (Final)." NIST Computer Security Resource Center, September 2026.

[4] Help Net Security. "NIST and CISA finalize playbook to stop token theft and forgery." Help Net Security, September 2026.

[5] Cyber Security News. "CISA and NIST Releases Technical Checklist for Safeguarding the Identity Tokens From Theft and Misuse." Cyber Security News, September 2026.

[6] Biometric Update. "NIST and CISA finalize token security guidance for federal cloud systems." Biometric Update, September 2026.

[7] Cloud Security Alliance. "Zero Trust Principles and Guidance for Identity and Access Management (IAM)." Cloud Security Alliance, 2023.

[8] Cloud Security Alliance. "VS Code Zero-Day: One-Click GitHub Token Theft." Cloud Security Alliance, 2026.

[9] Cloud Security Alliance. "Agentic AI Identity and Access Management: A New Approach." Cloud Security Alliance, 2025.

[10] Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." Cloud Security Alliance, 2026.

[11] Cloud Security Alliance. "Navigating Identity and Access Management (IAM): Standards and Protocols." Cloud Security Alliance, 2026.

