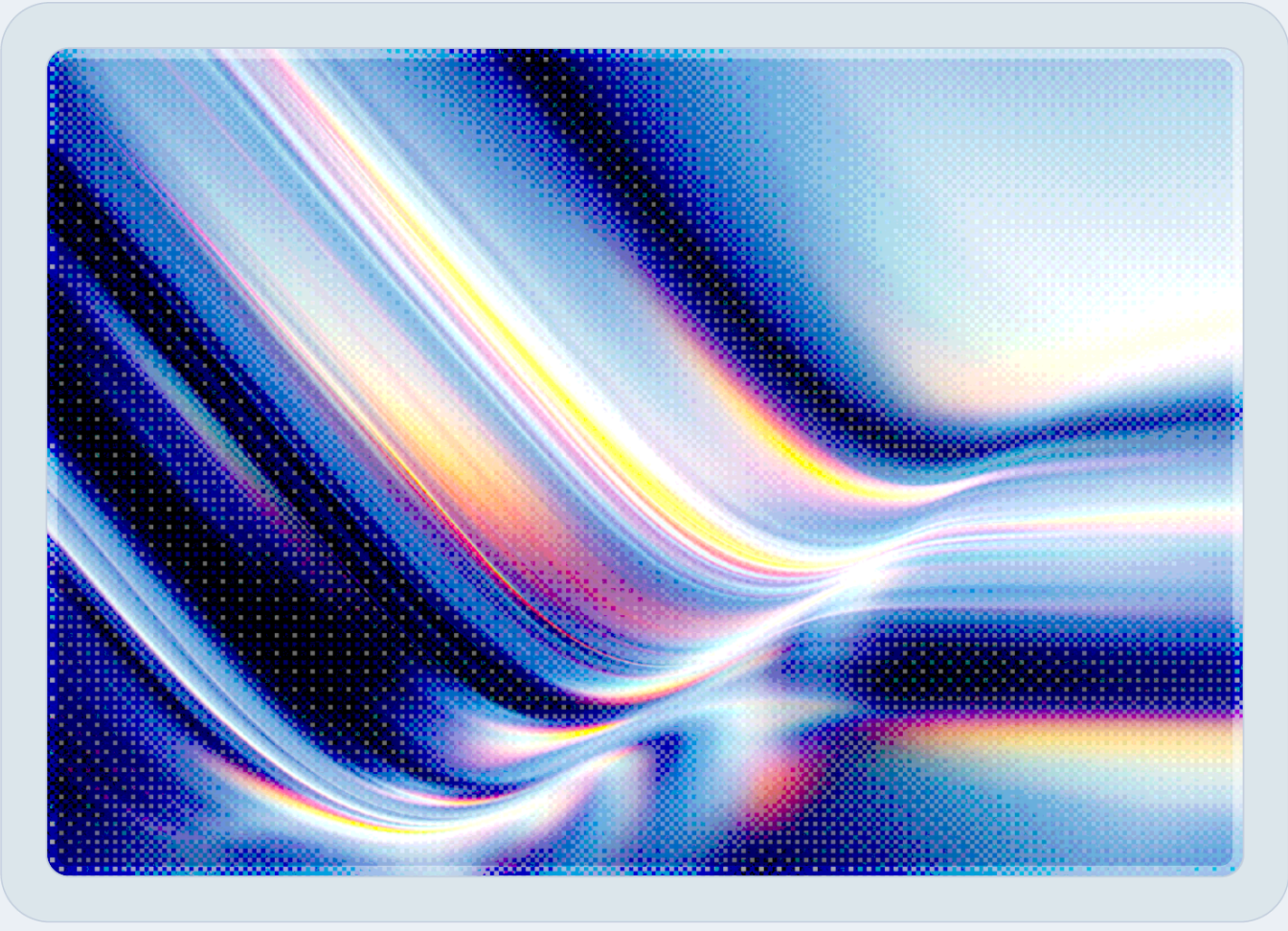
CSAI Foundation | Cloud Security Alliance

Black October: Concentration Risk in the AI Build-and-Host Layer

AI Developer Platforms as Credential Reservoirs

2026-11-05

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

© 2026 Cloud Security Alliance. Some rights reserved.

You may download, store, display, view, print, redistribute, and link to this document in its original, unmodified form, provided that attribution to the Cloud Security Alliance is maintained and all trademark and copyright notices remain intact.

This document may not be modified or altered. You may quote portions of the document as permitted by the Fair Use provisions of the United States Copyright Act, provided that attribution is given to the Cloud Security Alliance.

This document may be shared on professional and social media platforms in its original form with attribution.

This document was generated with AI assistance and has not undergone official CSA review and approval processes.

Key Takeaways

During October and early November 2026, twelve AI developer, hosting, and inference platforms disclosed security incidents, a cluster Wiz has named "Black October" [1]. Earlier CSA notes looked at individual incidents and at how stolen keys are abused afterward [2][3]. This note covers the structural risk those incidents expose, which comes down to five findings:

- **The platforms are credential reservoirs.** Build, hosting, evaluation, and coding-agent platforms hold large numbers of customer secrets. AI provider keys are among the most attractive of these, because they are bearer credentials that bill the victim and have an active resale market [3]. Wiz reports that attackers went after stored environment variables, especially ones that grant AI access [1].
- **Spreading AI keys across more vendors increases exposure.** Using several vendors spreads availability risk, but every extra platform that holds a copy of a secret is one more place it can be stolen. Organizations that use many AI developer tools are exposed through each of them.
- **The failures came together, across many platforms.** The root causes reported so far differ, and no single vulnerability links the incidents [1]. Even so, the platforms were hit in the same weeks by attackers chasing the same type of asset, a pattern that looks like a market sweeping a class of target rather than a one-off breach.
- **Notice arrives far behind exploitation.** In the two downstream abuse cases Wiz investigated, the first unauthorized model call came within 30 minutes of the reported credential access [1]. Where notice timelines are public, they ran far longer: Fireworks began notifying customers about 22 hours after detection [4], and Brev activity that began October 4 was reported only on November 2 [1].
- **For AI provider keys, the model provider is the one point every stolen key passes through.** Provider-side scoping, usage alerts, and proactive revocation apply whichever platform leaked the key. Non-AI secrets held in the same reservoirs, such as SSH keys, registry credentials, and wallet keys, have no such chokepoint and depend on inventory and rotation.

Taken together, these findings shift the question from how secure any one platform is to how many platforms hold an organization's credentials and how quickly those credentials can be revoked when several platforms fail at once.

Background

The Black October Incident Cluster

On October 30, 2026, Wiz published an analysis that grouped a series of incidents at AI developer platforms under the name "Black October" [1]. The original count was at least nine incidents disclosed between October 6 and October 30, and Wiz later added three more incidents reported on November 2 and 3 [1]. The affected organizations range from inference providers to general-purpose platform-as-a-service (PaaS) hosts, AI coding tools, and GPU development environments. The table below summarizes the incidents using the platforms' own disclosures where available and Wiz's tally otherwise. Date ranges in the table span from the first reported activity or disclosure to the latest public update. Most of these organizations have not published root causes, and investigations are still under way.

Platform	Category	Disclosure (2026)	Reported cause or exposure
Fireworks AI	Inference and evaluation	Oct 6–10	Internal credentials used to reach environment variables in evaluation jobs for a limited number of customer accounts. API keys and tokens stored as secrets are believed captured [4].
Andromeda	AI infrastructure	Oct 9	Described by the company as part of an "industry-wide attack" [1].
Zeabur	PaaS	Oct 14	Credential access. The company is shutting down [1].
Lightning AI	AI development platform	Oct 15	Credential access. The platform had earlier been affected by the Mini Shai-Hulud supply chain campaign, which Wiz does not link to the October incident [1].
OpenCode Console	AI coding tool (hosted console)	Oct 16–17	Intrusion into a legacy system. Third-party AI provider keys saved by about 1.4% of Console users may have been exposed [5][6].

Platform	Category	Disclosure (2026)	Reported cause or exposure
Railway	PaaS	Oct 20–22	An unauthorized party may have accessed encrypted environment variables and, for some, obtained the decrypted contents [7]. Preliminary cause cited by Wiz: privilege escalation in build infrastructure [1].
Warp	AI terminal and agent	Oct 23	Attributed to a third-party vendor dependency [1].
CommandCode	AI coding tool	Oct 28	Attributed by the company to "distillation" [1].
Northflank	PaaS	Oct 30	Customers told to "assume all environment variables have been decrypted." Customers later reported unauthorized use of stored AI keys [8].
NVIDIA Brev	GPU development environments	Activity Oct 4–20, reported Nov 2	Listed by Wiz without technical details [1]. Secondary reports describe possible exposure of SSH keys used to access customer virtual machines, which NVIDIA rotated, stating it found no evidence of misuse [12].
Trigger.dev	Background-job platform	Nov 1, reported Nov 2	No technical details shared [1].
Chainstack	Blockchain and node infrastructure	Oct 29–31, reported Nov 3	No technical details shared [1].

Some affected organizations have pointed to a shared adversary. OpenCode and CommandCode linked their incidents to "distillation," and Fireworks, OpenCode, and Warp reported signs of an "agent swarm" attack, according to Wiz [1]. Railway said its incident was part of a series of attacks on AI inference

providers and that it had told those providers to revoke their keys [7]. None of these links has been publicly confirmed, and some skeptics have called the distillation framing in OpenCode's case unproven [5]. This note therefore treats the cluster as a pattern of attackers seeking the same kind of asset, not as a confirmed single campaign.

What Happened to the Stolen Keys

Wiz investigated two enterprise cases in which Amazon Bedrock credentials were stolen through one of these platform breaches [1]. Each key received one test model call within 30 minutes of the reported vendor credential access. A low-rate verification job then kept checking the key for several days, followed by enumeration of quotas and model catalogs. Bulk abuse came next: in one case it ran for about thirty hours and made about 800,000 Bedrock calls, and in the other it ran for about thirteen hours and made about 1.7 million calls. At list price the two cases together likely exceeded $100,000 in usage, with most of the traffic going to the most expensive models [1]. Infrastructure linked to discount resellers of frontier-model access later handled both keys, though Wiz could not definitively tie the resellers to the bulk abuse [1].

CSA's November 2 note covers this resale pipeline in detail [2]. This note focuses instead on a prior question: why so many keys were stored in places where they could be stolen in bulk.

Security Analysis

The Build-and-Host Layer as a Credential Reservoir

A modern AI application has to authenticate somewhere at runtime. The simplest and most widely documented way is to store the model provider's API key as an environment variable on the platform that builds or hosts the application. Platforms encrypt these variables at rest, but they also have to decrypt them to inject them into builds, jobs, and containers. As CSA's Northflank note observed, encryption at rest therefore does not stop an attacker who reaches the build system, the orchestration layer, or the key-management path [2]. The Railway and Fireworks disclosures match this pattern. One involved a reported escalation within build infrastructure [1][7], and the other involved internal credentials that could read secrets made available to evaluation jobs [4].

The result is a reservoir. A single PaaS, inference, or coding-agent platform may hold secrets for thousands of tenants, and the platform's own administrative credentials and build systems sit in front of all of them. Compromising the platform yields many customers' secrets at once, without breaching any customer. AI provider keys are among the most attractive secrets in such a store, because they are bearer

credentials that work from anywhere, they bill the victim, and resale markets will pay for them [3]. Railway's scale shows how large these stores can be: BleepingComputer reports that the platform has about 4.8 million users hosting more than 16 million services [7].

The reservoirs also hold more than AI keys. Railway urged users to rotate private registry credentials as well as API keys [7]. Wiz cites unverified social media reports that about $371,000 in cryptocurrency was stolen using wallet keys stored in Railway [1], and secondary reports on Brev describe possible exposure of SSH keys for customer virtual machines [12]. AI keys may be what attackers are mainly after, but anything else stored alongside them is exposed at the same moment.

Why Diversification Does Not Help Here

Concentration risk usually brings to mind availability, and for availability the standard answer is to spread workloads across vendors. CSA's earlier concentration work makes this argument for model providers and cloud regions [9][10], though [9] already cautioned that diversification offers limited benefit when providers share critical dependencies. Secrets are an extreme case of that caveat, because they work the other way. Every platform that holds a copy of a given credential is one more place to lose it, and a theft from any of them compromises the key everywhere. A team whose Anthropic or Bedrock key sits in a hosting platform, an evaluation service, a coding-agent console, and a workflow tool has four separate chances of compromise for the same key. For secrets, more vendors means more exposure.

This changes how third-party risk teams should think about the AI build-and-host layer. In our experience, platforms like these are often adopted bottom-up by individual developers, on self-service plans, with no procurement review. The cluster included large platforms such as Railway and NVIDIA's Brev, but it also included small, young companies, and Zeabur's announcement that it is shutting down after its incident shows that a vendor may not be around to support remediation [1]. Organizations therefore tend to accumulate a long tail of platforms that each hold high-value credentials. Few of them were risk-assessed, and many were never recorded as holding secrets at all.

Correlated Failure Across the Layer

The second concentration issue is that failures occur together. The incidents differ in their technical root causes, but they share an attacker objective, a time window, and in at least one case reported infrastructure. Wiz observed that an IP address Zeabur named in its compromise later appeared checking enterprise keys that Wiz attributes to another of the vendor breaches [1]. Correlation across platforms means risk at one platform does not simply add to risk at another. When attackers appear to be sweeping AI developer platforms for keys, as the October timing and shared infrastructure suggest [1], an organization that uses five of them should not assume that only one could be affected at a time. Incident

response plans that expect a single vendor notice and a single rotation event did not fit October 2026. Teams instead faced a run of notices, each requiring a separate inventory of which keys sat on which platform.

This pattern is consistent with CSA's earlier observation that the AI stack consolidated rapidly and that existing risk frameworks were not designed for the resulting shared chokepoints, which carry industry-wide blast radius [9][10]. Black October shows the same dynamic at a layer that earlier analyses covered less closely: the commodity platforms that build, host, and evaluate AI applications, as opposed to the model providers and hyperscalers above them.

The Provider as the Common Point

Every AI provider key stolen from these platforms is used against one of a few model providers. Whichever platform leaked it, the abuse shows up at the provider's API, which makes the provider the one place where controls cover every theft path for that class of credential at once. Providers acted on this in the OpenCode case: Anthropic and Google blocked keys they believed were exposed, and OpenAI worked with OpenCode on restoring access and refunding credits [5]. Wiz's cases also show the limits of this kind of control. One stolen key was limited to US regional inference profiles; throttling blocked a large share of its roughly 1.7 million calls, yet more than a million succeeded [1]. Scoping limited the damage without stopping it.

For customers, the point is that controls on the AI key itself act at the one place every theft path for that key leads to. Spend caps, model allowlists, source restrictions, short key lifetimes, and fast revocation at the provider all apply no matter which platform leaks the key, while platform-side controls, however good, protect only the copy that platform holds. The same logic does not extend to the other secrets in the reservoir. SSH keys, registry credentials, and wallet keys have no common provider-side chokepoint, so their protection depends on knowing where each copy lives and rotating it quickly.

Timing

The 30-minute validation window Wiz observed [1] is, in our assessment, the key planning figure for incident response, although it comes from two investigated cases. Fireworks detected activity on October 6 and began notifying customers the next morning [4]. By the standards of vendor disclosure that is fast, but it is still many times slower than an attacker validating a key, and Brev's report on November 2 of activity dating from October 4 shows how much longer the gap can run [1]. Northflank's guidance to treat every environment variable as decrypted [8] is the realistic stance, because customers

usually cannot tell from a notice whether their own keys were used. Containment therefore depends on how quickly a customer can find and revoke every key that a named platform held. Better vendor notification helps, but it cannot close that gap on its own.

Recommendations

The recommendations below add to the key-level controls in CSA's earlier Black October notes [2][3]. They focus on the concentration problem, meaning how many reservoirs hold an organization's credentials and how the organization responds when several fail together. They are ordered by time horizon, moving from immediate containment to longer-term changes in vendor management and credential architecture.

Immediate Actions

Organizations should first confirm they have rotated credentials held on every platform in the Background table that they use. Old keys must be revoked at the issuing provider, because deleting a stored secret from a platform does not invalidate the credential [4]. Billing and provider-side usage logs should be reviewed from October 1 onward, using the indicators Wiz published [1].

Security teams should then build a credential reservoir map: a list of every external build, hosting, inference, evaluation, coding-agent, and workflow platform that holds any organizational secret, with the specific secrets each one holds. The map should cover self-service accounts set up by individual developers, not only contracted vendors, and expense reports and SSO logs are often the quickest way to find them. Without this map, responding to the next notice will again depend on searching by hand while attackers are already using the keys.

Short-Term Mitigations

Over the next 30 to 90 days, organizations should reduce the number of places each credential is stored. Each high-value credential should live in as few external platforms as possible, and AI keys in particular should be issued per platform and per environment, so that a theft from one reservoir can be revoked without breaking others. Where a platform can retrieve secrets at runtime from an organization-controlled secrets manager, or federate workload identity to the model provider, organizations should prefer that to storing static keys in the platform's environment-variable store. This moves where the trust sits rather than removing the need for trust, but it reduces how many tenants' secrets a single compromise exposes.

Incident runbooks should also cover several vendors failing at once. They should set a target time from any platform breach notice to completed revocation at the provider, measured in hours, and rehearse revoking keys across several platforms simultaneously. Provider-side alerting on new source networks, model-catalog enumeration, quota queries, and spend anomalies should go to an on-call function, not a billing inbox [2].

Strategic Considerations

Third-party risk programs should treat AI build-and-host platforms as a distinct tier, classified by the value of secrets they hold rather than by contract size. Due diligence for this tier should ask how secrets are encrypted, who and what can decrypt them, how build infrastructure is isolated from secret stores, and how fast the vendor notifies customers. It should also check that the vendor can revoke or help revoke credentials when it shuts down. Contracts should require notification on the scale of hours, given how quickly stolen keys are used.

Organizations should also decide deliberately where AI credentials live. Running traffic through an organization-controlled AI gateway, with platforms holding only short-lived or gateway-scoped tokens, combines provider keys into a single store that the organization defends itself. That too is a concentration point, so the trade-off is worth making only when the organization can secure the gateway better than the platforms it replaces. Finally, organizations should press model providers for finer-grained scoping, workload identity federation, and proactive revocation of keys known to be exposed. As argued above, for AI provider keys the provider is the one point every stolen key passes through.

CSA Resource Alignment

This note continues CSA's coverage of the Black October cluster. **"Black October Reaches PaaS: Northflank, Secret Stores, and AI Key Resale"** [2] analyzed the Northflank incident, the decryption problem facing PaaS secret stores, and the validate-hold-flood pattern of key abuse. Its key-level controls (scoping, spend caps, and non-human identity governance) underpin the short-term mitigations above. **"Harvesting Keys for Distillation: Attacks on AI Credential Platforms"** [3] set out the demand side: distillation operators, resellers, and the bearer-token nature of AI keys that makes platform-side fixes insufficient. This note builds on both by examining the platforms together as a concentrated, correlated risk.

CSA's concentration research supplies the framing. **"AI Developer Ecosystem Concentration: Critical Infrastructure's Hidden Risk"** [9] argues that concentration in the AI stack has passed the point where it should be treated as a critical infrastructure issue. It recommends explicit dependency

thresholds and short contractual incident-notification windows, which the strategic considerations above apply to the build-and-host tier. **"AI Development Stack Concentration Risk"** [10] identified shared chokepoints across the AI development stack and recommended mapping every AI system's vendor dependencies. The credential reservoir map recommended here extends that mapping to the secrets each dependency holds.

For control structure, the **CSA AI Controls Matrix (AICM) v1.1** [11] provides the assessment basis. Its Identity and Access Management and Cryptography, Encryption and Key Management domains cover credential scoping, rotation, and secret storage. Its Supply Chain Management, Transparency and Accountability domain covers third-party tiering and due diligence for AI platform vendors. Its Security Incident Management, E-Discovery, and Cloud Forensics domain covers response when several vendors fail at once.

References

[1] Wiz (B. Read, R. McCarthy). "“Black October” for AI Developer Platforms." Wiz Blog, October 30, 2026 (updated November 2–3, 2026).

[2] Cloud Security Alliance AI Safety Initiative. "Black October Reaches PaaS: Northflank, Secret Stores, and AI Key Resale." CSA Labs, November 2, 2026.

[3] Cloud Security Alliance AI Safety Initiative. "Harvesting Keys for Distillation: Attacks on AI Credential Platforms." CSA Labs, October 27, 2026.

[4] Fireworks AI. "October 2026 Security Incident." Fireworks AI Security Bulletin, October 9–10, 2026.

[5] A. Mehta. "Hackers hit AI coding tool OpenCode, running up charges on users' AI accounts." TechCrunch, October 19, 2026.

[6] S. Khandelwal. "OpenCode Console Breach May Have Exposed OpenAI, Anthropic, and Google API Keys." The Hacker News, October 19, 2026.

[7] B. Toulas. "Railway warns devs to rotate secrets after potential breach." BleepingComputer, October 22, 2026.

[8] L. Abrams. "Northflank security breach leads to abuse of customers' AI API keys." BleepingComputer, October 31, 2026.

[9] Cloud Security Alliance AI Safety Initiative. "AI Developer Ecosystem Concentration: Critical Infrastructure's Hidden Risk." CSA Labs, May 22, 2026.

[10] Cloud Security Alliance AI Safety Initiative. "AI Development Stack Concentration Risk." CSA Labs, May 3, 2026.

[11] Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." Cloud Security Alliance, 2025.

[12] DiarioBitcoin. "NVIDIA rota claves SSH tras detectar actividad no autorizada en Brev durante 16 días." DiarioBitcoin, November 2, 2026.

© 2026 Cloud Security Alliance. AI-assisted Rapid Research.