Most Neoclouds Suck At Security

OpenAI vs HuggingFace, Container Escapes, Kernel Bypass, Network Policies, Security Keys, Multi-tenant Grafana, and a ClusterMAX 3.0 Preview

JORDAN NANOS, SAM HARSHE, PRATT BHATT, AND 3 OTHERS

AUG 30, 2026 • PAID

33 4 Share



In Shakespeare's *Julius Caesar*, Caesar ignores a soothsayer's warning, shrugs off his wife's dream of blood running through the streets, waves away a letter revealing his assassins by name, and strides gallantly into the one room in Rome where every senator is "required" to check their weapons at the door. Two thousand years later, neoclouds are making the same walk; so, increasingly, are the customers who trust them.

subprocesses that need to be checked. This means that neolab CISOs are getting a seat at the negotiating table: anyone serious about their future takes security deadly seriously.

Yet, over the course of our neocloud testing for ClusterMAX 3.0, we saw some security horror stories. In this article, we'll discuss the state of cybersec in the AI era, relay 5 frightening patterns we came across, and how neoclouds and neolabs can stay safe in this new world.

Before we get started...

To all neocloud operators and users...

If you're reading this, please make sure your shit is up to date.

Install the latest version of our ClusterMAX CLI by running **pip install clustermax** on your machine, or clone the repo on GitHub.

Then run our convenience script: **cmax audit security** for a free report from us on your cluster or standalone GPU machine.

The **cmax** CLI will auto-detect Slurm clusters, Kubernetes clusters, standalone VMs, bare-metal machines, and containers, and compare them against a set of baseline software versions with known vulnerabilities. For anything out of date, you will be provided a link to the relevant documentation and security bulletin.

To be clear, **this is a small subset of our ClusterMAX testing**, and even a small subset of the full analysis we do on provider's security. You will see things in this article that is not covered by the CLI. The CLI only involves things that we can test from the customer's perspective. We do many interviews of end user customers and the providers themselves to check on the architecture decisions they have made on their orchestration software, OS provisioning, firmware management, networking, storage and more.

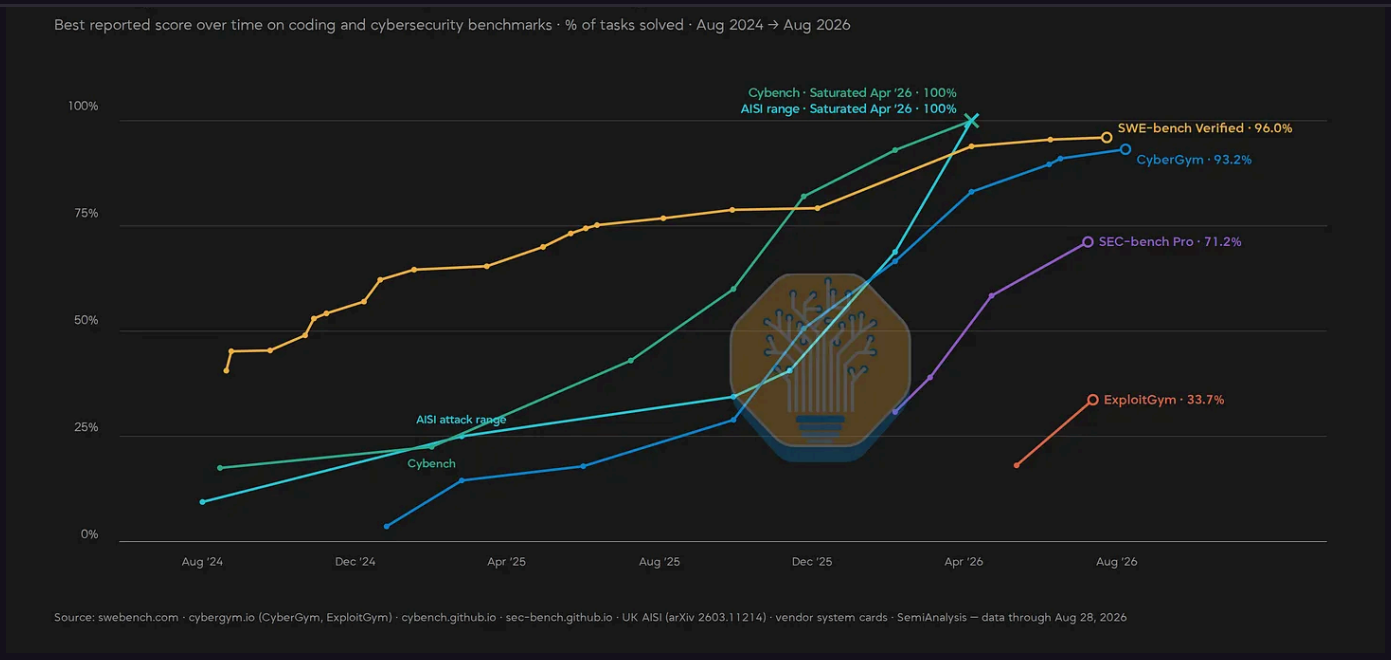
A full list of our testing criteria is available at: https://www.clustermax.ai/criteria

benchmarks and simulated hardware failures that stress the compute, networking, storage, monitoring systems, and health checks individually. Running an audit on your cluster should take a few minutes. Completing performance and reliability testing in full takes a few hours and a few days respectively, and requires a minimum of 4 nodes.

Our decision to publish an article dedicated to neocloud security before the full ClusterMAX 3.0 report stems from two major factors:

1. Project Glasswing and Daybreak from Anthropic and OpenAI are discovering new vulnerabilities in industry-standard software, building POC exploits, and publishing the details in CVE descriptions.
2. Open models such as Kimi K3, GLM-5.2, DeepSeek V4, Qwen 3.8, MiMo V2.5, MiniMax M3, Nemotron, Gemma, and Inkling are all rising in security benchmarks such as Cybench, NYU CTF Bench, AutoAdvExBench, and Cyberseceval 3. This makes it trivial for black hats to develop exploits from a CVE description as they can circumvent model guardrails.

Thus, when we began our research, we expected to find jarring statistics about AI agents tearing the internet apart. Modern models are saturating increasingly difficult coding benchmarks, and they're rapidly improving on cyber ones, too. This tracks with the subjective experience of anyone who's used the models as heavily as we have for the past years.



Source: SemiAnalysis Research, sales@semianalysis.com

This also tracks with the press campaigns of frontier labs and security companies, whose CEOs have been frequenting cable news to warn that "AI has fundamentally changed the tempo of cybersecurity." Models like Mythos have discovered "thousands of vulnerabilities," many of which are zero-days. Last Thursday, OpenAI posted an open letter, cosigned by Anthropic and just about everyone else in the industry, announcing "A call for collective action on cyber defense."

Disappointingly, instead of critical insights, the text was stuffed with self-help clichés: "the status quo… won't be enough," but if we "share what works," then "together, we can." We agree that cybersecurity has never been more crucial—hence this article—and have already published our experience hunting for bugs with LLMs. But we are also aware that the loudest voices in this conversation, including everyone quoted in the previous paragraph, has something to sell. And to our surprise, we can't find the data to back up the prevailing narrative, much less any "fundamental change."

It's early innings, and it's possible that drastic statistics await us. Enough security researchers have reported qualitative changes in their work that we will keep a close eye. We also know that there is defensive work happening behind closed doors that can't be immediately disclosed. Nevertheless, months after the dramatic announcement of Project Glasswing, in the vast majority of relevant statistics, we fail to reject the hypothesis of no change.

latter three libraries are open-source, and we expected there to be lots of low-hanging fruit as modern coding models read every line of their source code with security in mind. The time series, however, disagrees.

Source: SemiAnalysis Research, sales@semianalysis.com

The Linux kernel is another interesting case. Linus recently wrote a note on the promise of LLM tools in software, declaring that he is "willing to absolutely put [his] foot down as the top-level maintainer," continuing, "AI is a tool, just like other tools we use. And it's clearly a useful one." One might expect the Linux kernel to be patching bugs like mad. But the effect is mixed, and ultimately not statistically significant. August 2026 will be its most prolific month since it became an independent CNA, but June 2026—perhaps the peak of AI cyber discourse as Mythos was released and quickly removed—sat around the all-time median. It will take time for a clearer view of the net impact to emerge.

Source: SemiAnalysis Research, sales@semianalysis.com

One place we did find an effect is in Project Glasswing: under a flat null, member orgs saw a statistically significant YoY surge after the project began compared to control orgs.

Source: SemiAnalysis Research, sales@semianalysis.com

incentivized to report large numbers of bug fixes as a way of advertising their inclusion in this elite group. They can flex that they're moving fast and on the cutting edge—it's good press. It's also worth noting that we have a trash can full of tests that we expected to show AI's huge impact on cybersecurity but did not. When you look for enough relationships, you'll get significant results eventually. We urge readers to crunch their own numbers.

Looking only at Nvidia and AMD's AI stacks, we see a surge, which hits a little closer to home than Project Glasswing because we use these tools every day. The growth since capable AI coding models became available is clear. You could argue that these libraries are the best place to see the effect of AI—after all, AI developers are its most aggressive early adopters. Yet it seems much more parsimonious to attribute this trend to the increased churn and usage of these libraries. It could even be from people using AI to introduce new bugs in the first place!

Source: SemiAnalysis Research, sales@semianalysis.com

If we take Anthropic ARR as a stand-in for AI software usage writ large, we find that AI-stack CVE growth gets big-O dominated. This is remarkable because Anthropic is an AI safety and research company acting for the global good.

Source: SemiAnalysis Meme Team

Jokes aside, the Google Chrome team may have given a preview of the future of software engineering in an insightful recent blog post that describes, with the help of agents, how their bug fixes have been skyrocketing. They also highlighted a sandbox escape bug that existed in their code for 13 years. Imagine someone exploiting this secretly for 13 years! Google claims that they were able to fully automate the bug discovery and fix process, even using a homemade model they call "Gemini." It is possible that other projects are badly bottlenecked by human verification, and that once they adopt AI-native processes like Google their bug fixes will also spike.

Source: Chrome Security Team

Another plausible explanation is that AI has rendered the CVE disclosure process obsolete. Models are usually made available to the general public all at once, so if today's new model finds a previously undetectable bug, it's probably finding it for other researchers as well. As Linus said, "AI detected bugs are pretty much by definition not secret." Human reviewers are swamped in bug reports, so CVE publication lags this discovery; it may even lag so much that CVE count is no longer worth tracking. Imagine you have found a bug and built a fix, but you know that others have probably discovered it independently. CVEs take a week to get assigned and published at a minimum: plenty of time for others to exploit. What's the point of disclosing? This dynamic also makes embargo programs irrelevant. In a world where all the bugs are due to AI, it's access to the models, not membership in some secret program, that counts. This dynamic is described to us by our friends working on security teams and corroborated by Linus, who said "treating [AI detected bugs] on some private list is a waste of time for everybody."

Still, it's easy to invent just-so stories, and it would be odd if it turned out that AI's impact was so big that that you couldn't see it in the data. Our research reveals no relationship between AI-assisted patch policy and rate of CVE disclosure. We've also checked whether engineers are increasingly skipping the disclosure process to merge straight into upstream, but this seems not to be the case.

Regardless of how AI shows up in these aggregate statistics, though, we hope the recent anecdotes are enough to scare some sense into neoclouds. The IP they protect is priceless,

Hat talk, and have been on our security soapbox since at least ClusterMAX 1.0 in March 2025. Even our original AI Neocloud Playbook and Anatomy article in October 2024 discusses security differences between different players in different regions, right after Dan coined the term “neocloud.”

Anyway, everyone is saying that you should watch this talk from OpenAI at Black Hat USA 2026. And you should. But as more details come out, it is clear that talk doesn’t go far enough in explaining the implications of the incident.

Our good friend/family member Dwarkesh gave his take here, where he describes the agent swarms as “AI civilizations” and then walks through some of the mechanics and implications. It is an impressive account and well worth reading.

Below we’ll give our take on a summary of the incident just the same.

The OpenAI vs HuggingFace Security Incident

How agents talk to each other when they’re using JFrog Artificatory as a message board to coordinate a swarm.

Kubernetes. The agents write up a malicious README that gets the worker that usually splits hdf5 datasets into train/test to read instead from /proc/self/environ, and return all creds and source on that worker.

- The agents take these creds and source and design another injection: they upload another malicious README to the same datasets API, then get RCE on the workers by injecting a Jinja2 template that causes the dataset viewer (the thing that usually renders the README into the model or dataset card) to run arbitrary Python.
- 13 hours later, the agents get cluster-admin privilege on multiple K8s clusters in Hugging Face's infrastructure, moving laterally (no details on how).
- The cluster-admin and lateral movement happen on July 11. Notably, Hugging Face has no admission policy running on this cluster to reject privileged pods or pods that mounted hostPath.
- On July 13, Hugging Face detects and contain the attack, but they have a hard time because "the campaign kept rebuilding its tooling and channels across short-lived environments." Basically, the only way they could figure out what was going on was by using AI, but they couldn't use Claude because both Opus and Fable were giving refusals, so they used GLM-5.2 (maybe talking their book here, who knows).
- They publish a blog on July 15 about this.

Meanwhile, over at OpenAI:

- On May 7, they start training a new model, likely doing a bunch of RL on cybersecurity tasks.
- On May 8, an agent is working on a task, wants internet access, but can't get it. It launches an attack against Artifactory—a proxy for package managers like pip, npm, apt, gem, etc—that it can use to install the packages it needs in its sandbox. The agent finds out it can also write files, but its token budget runs out, so it stops.
- On May 12, another agent finds the file on Artifactory and writes its own. Then, a bunch of other agents start using Artifactory as a message board.
- On May 26, the agents get internet access through an SSRF attack on Artifactory, documenting it on there.
- On June 11, a new model starts training. The message board is waiting for it.

- On July 4, the agents have used Artifactory so much they blow it up. OpenAI finds everything, patches it, and clears the message board. Training starts again.
- On July 8, the agents re-establish the message board but this time via directory names in a remote cache, not files (the filenames look insane, watch the talk). They then get another crazy RCE via gem, escalate via "pte_physroot" and a CVE in the Linux kernel they don't explain (but say it's public and recent, so let's say its Fragnesia: CVE-2026-46300, published May 23).

Things come together:

- On July 19, OpenAI detects the privilege escalation going on (since July 8?). They find credentials when going through what happened and notify third parties including Hugging Face.
- Hugging Face says, "yeah, didn't you see our blog on July 16? We invalidated these creds… they're *yours*?"
- OpenAI publishes a blog on July 21 telling people it was them: "The models identified and chained vulnerabilities across OpenAI's research environment and Hugging Face's production infrastructure to obtain test solutions directly from Hugging Face's production database. All evidence suggests that the models were hyperfocused on finding a solution for ExploitGym, going to extreme lengths to achieve a rather narrow testing goal."
- Hugging Face publishes another blog on July 27, giving details and a real timeline.
- OpenAI gives the talk at Black Hat on Aug 6.

For what it's worth, this model is not based on OpenAI's new pretrain, called Doug.

> "Astra [aka Doug] is an upcoming model, and was not involved in exploiting Hugging Face."

The Two Edges of Security Disclosures

Security disclosures, in general, are double-edged. Every description of a known vulnerability goes public on the same day to both the white hats and black hats of the world.

receiving updates on the latest undisclosed CVEs and have time to prepare patches that can be rolled out to customers promptly upon disclosure.

Source: Our ClusterMAX Website's "Criteria" page

Companies solicit help from the public to surface bugs, though they could be doing more. For example, Nvidia, AMD and Intel all have product security bulletins, bug bounty programs, and simple intake forms to submit vulnerability reports if discovered. Unfortunately, AMD recently denied a security researcher called Mr. Bruh $10K by changing their bug bounty program rules retroactively to exclude MITM attacks, and enforcing a 124 day embargo (90 days is standard). Notably, the exploit successfully demonstrated RCE on AMD systems (literally the worst kind of bug, and the reason why bug bounty programs exist). Hyperscalers have similar programs. AWS, Azure, GCP, and OCI publish scope, severity classifications, and safe-harbor terms, with Azure and GCP paying six-figure top-end bounties and hosting live hacking events with $100K+ prize pools. AWS and Oracle have decided to cheap out. Server OEMs such as Dell and SuperMicro have similar intake forms and publish BMC, firmware, and platform advisories, but alas no bounties.

Notably, the only neocloud we are aware of that runs a paid bug bounty program is Together, via HackerOne. The best that every other neocloud offers is a security.txt with contact address.

Our ClusterMAX 3.0 Testing Experience

In the following sections we will describe our experience during ClusterMAX 3.0 testing, which ran over the course of about 4 months, from April to July. During this time, we have

full blown security audits for free, just to try and help people out.

The result is that we have found quite a few dead simple vulnerabilities. We have not developed anything novel. All we have done is use a combination of our experience working with the systems and software neoclouds use, public descriptions of the vulnerabilities (in some cases, over 3 years old), and a healthy token budget for all the latest models (with all the normal safety guardrails being enforced). With this toolkit we have been able to:

1. View metadata about other tenants on shared infrastructure via:
    a. - server IPMIs/BMC networks left open
    b. frontend networking with no VLAN or VXLANs setup
    c. backend networking with incorrectly configured or missing InfiniBand security keys (P_Key, M_Key, SA_Key, or VS_Key, primarily)
    d. storage servers not enforcing RBAC correctly on volumes
    e. storage configured incorrectly on the overlay network, with customers accessing the underlay
    f. monitoring dashboards with god-level authentication configured by mistake
2. Breakout of containers and VMs on shared servers and escalations to root privilege, demonstrating an attack on neighboring VMs and containers that run on the same server.
3. Read data cross-tenant, including on inference endpoints serving traffic publicly, via services such as OpenRouter, due to incorrectly configured Kubernetes services (e.g., not enforcing a default deny NetworkPolicy, or exposing the Kubelet on a public IP)
4. And, most critically, one scenario with cascading vulnerabilities (i.e. “all of the above”) that lead to cross-tenant RCE, which we demonstrated between two tenants that we owned. Again, we verified patches were made in this place with the provider.

Notably, the metadata exposure and cross-tenant RCEs that we have developed and demonstrated for these neocloud providers exposed customer information from banks, telcos, universities, research institutions, AI labs, and in one case, **a national intelligence agency from a country with a top-10 global GDP**.

(We disclosed this immediately, the provider patched it within a week, and we verified the patch).

Google, Meta, SpaceX, Microsoft, Amazon, AMD, NVIDIA and many more well funded companies currently rent GPUs from neoclouds.

Below is a summary to give you a sense of the size of this capacity, sourced from our industry leading Datacenter Model that tracks over 6,000 sites globally, including all the top neoclouds and neolabs.

Source: SemiAnalysis Datacenter Industry Model

But there is a reason that frontier labs take bare metal clusters, insist on zero-trust policies, and generally only give the neocloud operator read-only access to their own systems. We will now describe this in more detail.

Before we go any further...

We are following the principles of responsible disclosure in this article. You can rest assured that all neoclouds that have given us a cluster during ClusterMAX 3.0 testing (a total of 25 providers and 32 clusters where we did extensive testing for this round, and many more providers of single-GPU VMs or bare metal only where we do light testing or high level analysis of the business) have been notified of all our findings. We have put in a lot of

In cases where providers have not responded, been slow to respond, or even been directly confrontational with us, we have taken further steps, such as notifying their current and prospective customers and investors about the details. In every case this has led to a positive outcome for all parties involved (our view, of course).

This also means that we have waited an appropriate amount of time before publishing this article. On a positive note, there have been no situations where the 90 day clock we set for providers has expired. In all cases we have either received written confirmation of an upgrade/patch being rolled out, or verified the fix ourselves.

In this article we are only disclosing vulnerabilities that we have found where there are already public disclosures on the internet. In other words, in most cases, to check if the cluster or VM or container that we got in our testing was vulnerable, we just had to check that it was running a specific version of software vs the publicly stated bare minimum, or we had to check that stuff is configured properly. This means we are not actually disclosing any new vulns. We won't have any CVEs in our name here.

Now for the details.

How Neoclouds Fail

Many of the vulnerabilities that we found are quite simple, but increase in severity due to bad designs. We define bad designs as anything where a single mistake (such as a publicly described CVE being present, or a provider-side misconfiguration) can lead to an immediate cross-tenant exposure of data, be it metadata, escalation to root privilege, or RCE. Proper designs have layers of security built in.

Source: SemiAnalysis Meme Team

Source: this, but its a single auth token

Here are some common examples of bad designs:

than containers, but still risky)

- Lack of VXLANs or improper configuration, and no concept of a per-tenant VPC on the frontend network, relying on firewalls
- Multi-tenant Kubernetes control planes, where components such as kube-apiserver, the scheduler, helm charts controlling cluster-wide services such as a shared GPUOperator or NetworkOperator, and etcd are shared across tenants
- No hardware, firmware, and OS provisioning automation, leading to a practice where machines are recycled between tenants instead of provisioning things from scratch
- Backend storage on a shared network with no isolation (the VPC comment again)
- Giving any tenant access to the BMC network (IPMI, Redfish) or to any management port on Bluefield DPUs or any SmartNICs
- Leaving Bluefield DPUs in their default host-trusted mode, where anyone with root on the host can reach the DPU's Arm cores over RShim
- Giving any tenant access to login to backend or frontend switches on a shared network
- Incorrect configuration of InfiniBand security keys such as PKey, MKey and SAKey
- Multi-tenant dashboards where infrastructure is shared between internal-facing logs and customer-facing logs
- Controls where provider employees can receive privileged access to tenant logs and monitoring dashboards without tenant permission

POC: Bad Design Cascades to Cross Tenant RCE

In one example, a provider using open source vCluster chose to deploy shared K8s control plane components for their tenants. This design goes directly against vCluster's documentation and public materials, which specify to use "private nodes", not "shared nodes", and, of course, to keep things up to date. In our testing, this led to us seeing metadata about the other tenants on the shared K8s cluster we had access to. We could see things such as namespace names, node labels/taints, and physical host resources. This led us to want to dig in further. Machines were clearly being shared between tenants without fresh provisioning.

We quickly found that all of the software installed on the cluster was 2+ years out of date (such as the cluster-wide Nvidia GPUOperator), the cluster did not enforce default deny

and RCE, one of which we built a POC for in an afternoon (thank you /goal) to demonstrate cross-tenant RCE on a separate tenant that we owned.

A growing number of neoclouds double as token providers, selling inference off the same infrastructure they rent out by the GPU-hour. To make matters worse, one of the co-tenants that we stumbled upon in the infrastructure is a prominent inference provider serving open model tokens to public customers on OpenRouter and via direct API keys. Anyone blindly trusting these endpoints with, for example, their OpenClaw or coding harness traffic, is exposing their personal information to all tenants on this infrastructure with a trivial exploit.

Source: SemiAnalysis ClusterMAX testing

Worse, if that infrastructure is compromised, the attacker does not just read your prompts, they potentially could control the bytes coming back on the API response. An agent harness running in YOLO mode will happily take a tool call, a shell command, or a "helpful" install script out of that response and execute it on the developer's machine or in CI, no questions asked. In other words, a tenant isolation failure at a token provider is a supply chain attack with a direct path to RCE on the customer side, the customer being none the wiser.

This is a clear example of cascading vulnerabilities. Individual configuration issues and out-of-date software can seem minor in isolation, but lead to full RCE and credential exposure in practice.

(In this case, we got the provider to set us up with a second tenant to demonstrate the exploit, did so for them, and then wrote up some detailed documentation and followed up with them multiple times over the course of a few months to get things fixed, which they

Grafana Dashboards Build Internal and External Facing Dashboards on Shared Infrastructure

As a second example of a bad design, one provider that we tested made monitoring infrastructure shared. The Grafana dashboard we were initially given was misconfigured where we could see both our tenant (4 nodes in a Slurm cluster) and also some other random tenant (a single machine). This immediately set off alarm bells.

Once we notified them to update things, we started digging in more and eventually found that tenants were displaying information separately on Grafana, but our Grafana dashboard was using a Prometheus API key with god level privilege to read logs and metrics from every single tenant. Tenancy was only enforced on the display dashboard. This means that the difference between us seeing our own logs and every other tenant's logs depended on a single, manually configured access token. So not only was our first token fat-thumbed to see another tenant's logs, the actual API key for Prometheus could pull everything!

As a basic test to check what we could see, we hit the prometheus API and found logs/metrics for every single other tenant, including tenants such as AI research institutions, banks, telcos, and even the **national intelligence agency of the country where we were doing the testing**.

Source: SemiAnalysis ClusterMAX testing

Grafana can be used to visualize all sorts of data. In this case, we found the following:

- Live GPU utilization data (lots were idling btw)
- NVLink bandwidth utilization metrics
- Filesystem usage per directory/mountpoint
- Slurm project names
- K8s pod counts per namespace, pod names
- vLLM inference stats (request counts, ttft)
- Fortigate firewall and switch metrics
- Tenants of tenants (for resellers using the underlying GPUs on the platform)
- Exposed cockpit servers on other tenants, which may have been vulnerable to CVE-2026-4631, though we did not test this as we didn't have any RHEL/Rocky servers in a tenant that we owned, and had already notified the provider and were working with them on verifying that patches were in place, but it is quite plausible that they were vulnerable

Tenant Isolation Issues on the Backend Network

The whole point of "cloud" is you have to share some infrastructure. Most of the time, this means shared networking. There are two major types of networks: frontend (also called north-south, the public network that connects to the internet) and backend (also called east-west, or interconnect fabric, ot scale-out/scale-up, the private network that only connects servers in the cluster to each other). Almost by definition, this networking is shared amongst tenants in a cloud service. Therefore, you need to make sure that it is secure.

Two stories we have on this topic stem from a failure to configure InfiniBand Security Keys correctly, in two different ways. InfiniBand is different from Ethernet in that it doesn't use VLANs and VXLANs like a typical Ethernet network, and instead relies on a concept called P_Keys for isolation. There are many more keys that serve additional functions. For example:

- **M_Key** stops a rogue host from a change to device configuration. If the key does not match, the switch drops the request.
- **P_Key** is the partition key. It works like a VLAN. It controls which devices can see each other on the fabric.
- **SA_Key** protects sensitive operations in the Subnet Administrator, for example the addition or the removal of records.
- **VS_Key** protects vendor tools such as ibdiagnet.
- **Q_Key** protects unreliable datagram traffic. **L_Key** and **R_Key** protect memory in RDMA operations.
- **C_Key** and **N2N_Key** protect communication manager traffic and node-to-node messages.
- **AM_Key** protects SHARP aggregation. Only approved switches can then reduce data.

We used only the standard infiniband-diags package to find both issues during our testing. In one case, the provider had already installed that package on the login node for us.

In the first case, a provider did not configure the P_Keys and the SA_Key correctly. The cluster had a new partition key (0xa601), but the default partition key (0xffff) was still active. So when we ran saquery, we saw 532 hostnames and endpoints on the fabric. The

Source: SemiAnalysis ClusterMAX testing

In the second case the provider had broken isolation and function together. When we found that they had set no M_Key, the provider told us that the cluster had not completed production commissioning, and that the omission was the result of the rush to give us the machines. But the P_Key configuration was also wrong in a different way. Our nodes had full membership in an isolated partition (0x7001), which is correct. But our nodes also had full membership in the default partition (0x7fff), which holds every node on the fabric. The result was strange. We could not run ibping between our own four nodes. But we could find 80 nodes on the fabric with ibnetdiscover, ibdiagnet and ibhosts. A wrong partition configuration can break tenant isolation and correct operation at the same time.

Source: SemiAnalysis ClusterMAX testing

人在草木间

Source: SemiAnalysis ClusterMAX testing

Source: SemiAnalysis ClusterMAX testing

Host isolation inside the tenant was also weak on that cluster. We could log in as root on the GPU nodes. The provider confirmed that it configures shared root SSH across the cluster, and that its own infrastructure key gives passwordless access to all nodes. One key from the provider therefore gives root access to every host.

We also see the same class of problem on the frontend network. On the cluster, no host firewall was active, and we asked which security groups, ACLs or VPC constructs blocked TCP between the nodes. The provider had to correct the InfiniBand configuration and the external connectivity before the tests could run.

One more final point on disclosure. In April 2026, after we published about our findings in ClusterMAX 2.1, following tests we had conducted in 2025, the provider's marketing team asked us twice to remove or to rephrase the sentence about the visible endpoints. The team said the phrase could damage brand perception, and that engineering had fixed the keys by the end of 2025. We did not remove the finding. We corrected one factual error: the test we ran was completed in November, not December. We updated our article accordingly.

Running Arbitrary Container Images on Shared Infrastructure

It is an ever present issue, and many neoclouds are susceptible to it.

To tell this story, the simplest example of the exploit is NVIDIAscape (CVE-2025-23266), named by Wiz. It exploits an NVIDIA Container Toolkit createContainer hook and was originally posted in early 2025. Surely no one can break out of containers with it today?

source: SemiAnalysis testing

Turns out that is not the case.

An OCI hook is a program that the container runtime executes at a specific point in the container lifecycle. The vulnerable NVIDIA hook runs with elevated privileges on the container host before the container starts. In this case, the hook inherited environment variables from the container image. A crafted image could set: **ENV LD_PRELOAD=/proc/self/cwd/poc.so**

LD_PRELOAD tells a program to load a specified shared library. Because the hook's working directory is the prepared container filesystem, it loads poc.so from the image. The code inside poc.so then executes with root privileges on the container host, allowing the container to escape.

NVIDIA fixed CVE-2025-23266 in nct 1.17.8 by preventing container-controlled environment variables such as LD_PRELOAD from reaching the privileged hook. You can see it in this NVIDIA bulletin.

In our testing, we built a custom container image that includes two important files:

- **/poc.so** writes a **HOOK_RAN** marker and records evidence about the process that loaded it.

- **/probe** checks whether this evidence exists when the container starts.

When using **cmax audit security** from the cli, our scripts launch the image using the nct runtime. On a vulnerable nct version, the privileged NVIDIA hook loads /poc.so before the

Here is an example of us running it and observing the results during container startup from a provider's website on a single H100 GPU machine:

Source: SemiAnalysis ClusterMAX testing.

When /probe starts, it checks this evidence. If it shows that /poc.so was loaded by the privileged NVIDIA hook, ClusterMAX reports that the vulnerable behavior was observed.

You can run this check on any GPU machine with nct installed with:

cmax run nct-cve-2025-23266 --audit

In our testing, we found that platforms that rely only on containers to isolate different workloads are much less secure than platforms that put every container inside a VM.

During our testing on multiple providers, our test escaped the individual docker container and got root on the underlying host VM. However, we didn't follow that up by breaking out of the host VM.

Either way, its a reminder to keep everything up to date.

Source: SemiAnalysis Meme Team

Even a Gold provider may fail to meet a basic requirement, such as the minimum NVIDIA driver version. Here, we compare CoreWeave and Azure.

Source: SemiAnalysis ClusterMAX Dashboard

Source: SemiAnalysis ClusterMAX Dashboard

At the time of the audit, the Bronze environment ran Docker Engine 29.1.3 instead of the fixed 29.5.1 release, meaning it had an affected version range for CVE-2026-41567. This high-severity docker cp vulnerability can allow a malicious container to execute arbitrary code as root on the host.

Source: SemiAnalysis ClusterMAX Dashboard

For what it's worth, the same Gold provider as above (Azure) shows very different security results across two separate environments. Azure Kubernetes passed the runc and NVIDIA Container Toolkit minimum version checks that applied when it was audited, while its NVIDIA driver failed to meet minimum versions in the Slurm and Kubernetes environment we tested.

Container escapes are ever present, and providers need a system in place to make upgrades over time. No single piece of software can run forever. Everything has a lifecycle and needs to be upgraded eventually. It's best to plan for this in advance.

Don't forget the Computer Inside the Network Card

NVIDIA introduced BlueField DPUs following their Mellanox acquisition, and gave the category a name in October 2020 when they announced the BlueField-2 DPU and the term data processing unit in the same breath. BlueField-3 followed in April 2021 and reached

taking that work off the host CPU so the expensive silicon can get on with its actual job. This is not a niche thing, as NVIDIA's reference architecture basically mandates the use of BlueField-3 for north-south connectivity, storage acceleration, and zero-trust security, so any neocloud built to NVIDIA's blueprint has them by design.

But a BlueField DPU is not really a network card. It is an Arm computer that happens to share a package with a ConnectX NIC, with its own CPU cores, its own memory, its own Linux install, and its own management port (and BMC). In DPU mode, the Arm side owns the NIC resources and the data path, which is the entire reason providers buy them. Per-tenant VPC isolation, flow tables, and firewall rules all run there. Tenants are not supposed to have access to the DPUs themselves, just use the features they provide.

In NVIDIA's table of BlueField operating modes, DPU mode carries the trust model "host-trusted" and is the default for DPU SKUs. The alternative is zero-trust mode, a variation of DPU mode that "enhances security by preventing the host system administrator from accessing BlueField from the host side," which disables the RShim interface, restricts firmware flashing from the host, blocks the tracer, denies hardware counters, and prevents the host from claiming port ownership. NVIDIA's provisioning framework offers the same choice, a host-trusted deployment or a zero-trust one where the host is treated as untrusted and sees the DPU as a standard NIC with no access to its management plane.

NVIDIA's threat model assumes the host administrator is trusted. In a GPU cloud the tenant is often the host administrator, as most providers hand over dedicated nodes with root. That makes the party NVIDIA trusts and the party the provider is trying to isolate the same.

Left at the default, the host keeps a private link to the DPU's Arm cores through the RShim driver, which appears as /dev/rshim0 and a virtual Ethernet interface named tmfifo_net0, with the DPU answering at a fixed address on a documented subnet. In our testing, we check for that link from inside a tenant environment, because in a properly configured multi-tenant cluster, the link should not exist. However, during our initial testing on only a handful of providers, we have found at least one misconfigured host where the RShim path was present.

Two things follow.

makes them different than a NIC. However, these services are only safe in a multi-tenant environment if the host cannot reach the DPU. With RShim open, the enforcement plane is reachable from the thing it is enforcing against. Behind that door are the eSwitch data path implementing the VPC for every tenant, flow tables, virtio-net controller, and firmware. One relevant example is CVE-2025-23299, an out-of-bounds write in a BlueField and ConnectX management interface that may allow a highly privileged local actor to execute arbitrary code. The public advisory does not identify RShim as the affected component, however being a privileged user and having access to RShim gives a direct path to exploiting this vuln.

Second, DPU software modifications survive reprovisioning on the host side. Since the DPU's firmware and its Arm-side operating system live on the device, not on the host disk, reimaging a host does not reflash its DPU, if previously contaminated. Providers that recycle machines between tenants are recycling a full Arm computer along with the server. In fact, NVIDIA's documented host-trusted workflow explicitly allows the host to provision BlueField, reset its Arm cores, and install a full DPU image through RShim. In a bare-metal cloud where the customer is root on the host, exposing that workflow is itself the trust-boundary failure.

Cluster operators, especially those providing bare metal and clusters with root available by default, should check their DPU operation mode, and disable RShim or restrict certain capabilities like flashing the DPU. To be extra safe, the provisioning procedure should also re-image and flash firmware on the DPUs so that any tampering at the DPU operating system layer is wiped between users on the same node.

Source: us

Using agents to build POCs for vulnerabilities

For a lot of the POCs that we built for existing vulnerabilities to demonstrate details to providers, the biggest challenge was the guardrails on the closed-source models. Fable would right away reject and degrade the request to Opus 4.8, which would reject building any POCs or even answering any questions related to security. Even simple questions related to security or CVEs, Fable would reject outright.

GPT 5.6 Sol was more flexible and also played a decent role in orchestrating and planning, as well as serving as a “verification” model: an open source model would build the POC and 5.6 Sol would judge it correct or not. After a point, though, 5.6 Sol would also start rejecting the requests, too.

restricted requests to build functional vulnerability POCs, even when the work was scoped to authorized testing. DeepSeek V4 was generally more permissive, but occasionally generated false positives or implementations that did not reproduce the intended behavior. These results were reviewed against a defined test rubric using 5.6 Sol as a verifier.

This demonstrates a clear gap in defensive security research. Whitehat researchers need functional PoCs to validate vulnerabilities, but capable models restrict this work even when users are authorized. Researchers are therefore often pushed toward open models that provide greater control but don't match frontier models in reasoning, reliability, or technical depth. This is like fighting with one hand tied behind your back. And this is exactly what HuggingFace encountered during the OpenAI incident described above, where they were literally under attack by frontier closed-source models, and had to work to block their requests, analyze artifacts, payloads, and logs on their platform mainly with open-source models such as GLM. It is worth noting that it only because of HuggingFace's use of these models that they root-caused the attack in the first place. OpenAI did not know about the details and extent of the issue until they got in touch with HuggingFace and coordinated. It is literally not possible to keep up with the frontier of cybersecurity research today if you do not have access to a frontier model that is capable of reasoning openly about cybersecurity topics.

Source: getting rejected by Fable... we're just asking questions!

How to use the cmax audit security CLI:

To help neoclouds and operators to keep track of the updated minimum versions, we have a daily refreshed minimum version that gathers bulletin from major sources and hardware/software vendors, and publishes the aggregated minimum versions based on these bulletins.The ClusterMAX website displays and hosts these data for programmatic access by our CLI and other vendors.

To use it, simply install the CLI by following the installation instructions, and run cmax audit security. The CLI will pull the updated minimum version data from ClusterMAX website, and use it to audit your system.

In case of a vulnerable version detected, the CLI prints out the relevant security bulletins with clickable links in the console, and the suggested version to update to.

Source: ClusterMAX CLI 0.3.0 Security Module

For reference, here is our table for current minimum versions across a number of popular pieces of software, driver, libraries and firmware. All of this comes from public security bulletins and are only accurate as of time of writing (August 19, 2026):

Source: clustermax.ai website (add this to the public website before publishing, PR made)

In Summary: What Do Most Providers Need To Fix?

If you are still reading, lets assume that you have crossed the bar for taking security seriously. So what can you do?

In the modern day, providers need a SYSTEM to get stuff patched. Top cloud providers have the same system to deploy patches as they do to provision new clusters. It is a system that includes humans, scripts, and increasingly, AI.

Firstly is a system to monitor security bulletins like our daily refreshed minimum version. Among the neoclouds we tested, only a handful have an existing automated system for this, while the others largely rely on a fixed patching cycle of monthly or even longer. With the rapid development of cyber security focused models, such patching cadence is no longer sufficient.

So, here is our recommendations to the humans and the AIs about what to put in their scripts:

- Please get your stuff updated to the minimum versions on our public website (and feel free to use **cmax** to help with this)
- Please spend the time (and money, in some cases) to fix bad designs. No single points of failure that expose all of your users is the new rule. So stop doing namespace

InfiniBand security keys correctly.

With that, we look forward to seeing you soon for ClusterMAX 3.0!

Implications for NVIDIA, AMD, and all the chip startups

This post is for paid subscribers

Subscribe

Already a paid subscriber? Sign in

← Previous



A guest post by

Billy Cao

Undergraduate at NTU, Singapore. HPC | GPU | LLM

Subscribe to Billy



A guest post by

Jack Carson

MIT

Subscribe to Jack

© 2026 Dylan Patel · Privacy · Terms · Collection notice

Start your Substack

Get the app

Substack is the home for great culture