Organizational Cybersecurity Evaluation

Part - 2

VULNERABILITY AND RISK ASSESSMENT REPORT

Prepared by: Shriram Karpoora Sundara Pandian

Table of Contents

DISCLAIMER

I, Shriram Karpoora Sundara Pandian, have observed NVIDIA Corporation, a leading global provider of graphics processing units (GPUs), artificial intelligence (AI) technologies, and computing solutions under the surveillance of independent research methodologies and publicly available information classified variably.

This assessment is to determine what their general security posture looks like based upon publicly available data, industry disclosures, and known vulnerabilities based upon their product lineup, infrastructure, and business processes. This is not a report on direct interaction with NVIDIA's network, penetration testing, unauthorized system access. However, you depend only on open-source intelligence (OSINT) and well-known cybersecurity databases to alert potential vulnerabilities that malicious actors could exploit.

Based on my research, I have looked at historical and current security incidents, software and hardware security risks, financial and compliance risks, and threat vectors that I think will be relevant for NVIDIA. This involves looking into NVIDIA's disclosed data breaches, software exploits, supply chain issues, and intellectual property threats.

This report is a snapshot in time based on the information available to date. NVIDIA's security posture could have gotten better, worse, or stayed the same since the time the findings in this assessment were documented. It should also be noted that this report does not purport to have sage insight into the internal security of NVIDIA, as many protective mechanisms and mitigations may be present and not publicly known.

This test has been performed for research and educational purposes only, and no affiliation with NVIDIA or any subsidiaries. As always, if you have any questions or need a clarification on any aspect of this report, please do not hesitate to ask us.

Introduction

Technology is growing at a rapid speed, and each day we are introducing new pipelines and architecture to new threats and vulnerabilities, so these emerged as major challenges for enterprises across the globe. As organizations are expanding on a global scale and services being hosted on the cloud make them scale infinitely, but it creates threats and risks not just for themselves, but for their customers, partners, and the worldwide technology ecosystem.

Of the leaders in this industry, one stands at the forefront. NVIDIA Corporation has been an engine of technological revolutions across multiple domains: GPUs, AI, deep learning, cloud computing, autonomous systems, etc. As NVIDIA is at the forefront of enabling the next generation of AI, powering data centers, gaming platforms, and enterprise computing systems, the security of its products, infrastructure, and intellectual property is paramount.

Why Conduct a Risk and Vulnerability Assessment?

It can avoid the unwanted costs and risk that can be detrimental to the company, as this is a proactive approach to find potential security weaknesses and improve the resilience of the company, along with physical security.

- **The Growing market, so the threat landscape**

  As corporations are making a lot of profits through the groundbreaking technologies, we must secure the enterprise. Here, in case NVIDIA is a leader for GPU technology and becomes a prime target for cybercriminals, state-sponsored attackers, and even some competitors, and not limited to them. Especially from 2022, the boom of NVIDIA is what makes it a prime target. What could be risks and losses?

  - **Design Leaks**: NVIDIA improves its architecture and its design to deliver generational upgrades, improving performance and efficiency, which could be stolen to destroy their reputation.
  - **GPU drivers**: NVIDIA pushes updates for its GPUs regularly (each month), which could be interfered with and can deliver malware that affects everyone over the globe.
  - **Supply Chain**: NVIDIA is not a vertical manufacturer like Apple, so they depend on third parties for hardware and some security dependencies which can bring additional vulnerabilities and increase the attack vectors.
  - **Regulatory risks**: Reputation and money on the line, especially working with compliance and regulations in the United States. So small mistakes can cost millions.

- **The real Impact**

  For a leader like NVIDIA, it is not just about money, but a breach or exploit could lead to severe consequences.

  - **Data Center**: The Fortune top companies rely on data centers as everything moved to the cloud, like AWS and Azure, they highly depend on AI tech stack for bots and cloud processing and storage. If NVIDIA affects, it can create a domino effect.
  - **Exposure**: If a breach reveals NVIDIA secrets, it can give advantage to its competitors.
  - I**nvestor and Stock Market**: Any groundbreaking news can affect the stock prices and create disbelief among Investors.

- **Historical breaches involving NVIDIA**

  - **Lapsus$ Hack (2022)**[1]: It was a ransomware attack on most of the top technology leaders and NVIDIA was affected as well. Attackers stole almost 1TB of data, which includes employee credentials, GPU code, and DLSS source code (essential for GPU).
  - **GPU Driver Vulnerabilities**: Lots of exploits have been found on the GPU drivers of NVIDIA and continuously reported. Some affected CUDA software and AI frameworks, which were used for privilege escalation, while others affecting the display drivers gave huge interruptions to users. However, NVIDIA tried fixing everything and mentioned their fixes in the product security[2] page.
  - **AI security risks**: On NVIDIA hardware, side-channel vulnerabilities have been found, and it leads to adversarial attacks for accessing the root system. These side channel attacks were continuously demonstrated by security researchers.

So the stakes are high, so it is important that we need to do rigorous risk assessment to ensure the continuity of NVIDIA's leadership in the market, and security, which drives the world.

Scope of Risk Assessment

In this report, I will assess the security posture and risks of NVIDIA based on the publicly available information and analyze its historical vulnerabilities and CVEs, risk factors and potential mitigation strategies. In this assessment, there is no direct interaction with NVIDIA networks through active scanning tools like Nmap, Wireshark, but will instead leverage the following.

- **Vulnerability disclosures**: Vulnerabilities that are actively found on NVIDIA products (eg, GPU drivers, AI stacks, and their cloud services).
- **Software patches**: Analyzing the existing software patches and mitigations from NVIDIA.
- **NVIDIA incidents**: Finding patterns in the existing incidents that happened with NVIDIA.
- **Open-source Intelligence (OSINT)**: Utilizing the open-source information and tools to gather as much information as possible.

Key Participant

My name is Shriram Karpoora Sundara Pandian, a master's student in cybersecurity at RIT; I have strong experience with risk assessments and security frameworks. I have the academic training and the practical experience that together enable me to assess cybersecurity risks in a structured manner. Adhering to best practices in the industry and methodologies for risk assessment, such as the NIST Risk Management Framework and MITRE ATT&CK, all conducted under strict ethical guidelines and compliance standards. The results contained within this page derive from information currently in the public domain and all testing and access to systems was conducted without prior authorization. This evaluation is intended to furnish information that can serve organizations such as security officers, developers, compliance officers, and even stakeholders and assist them to better comprehend and counteract looming threats.

Contact

Email: sk2410@rit.edu

Phone: (585)-957-4822

LinkedIn: https://www.linked.com/in/shriramkp

Inventory of Assets

Nvidia operates globally, so it is made of various components, which enables it to function smoothly. Before we can assess something, we need to know the components we will manage. So, the table below lists the inventory of NVIDIA and it is made of.

IT System Components	Risk Management Components	Examples
People	People inside an organization	Cybersecurity Teams, Data scientists, Trusted employees, Engineers, Managers.
	People outside an organization	Customers, Vendors, third-party contractors, cloud providers, Material suppliers.
	Possible threats	Insider threats, social engineering attacks, human errors.
Procedures	IT and business standard procedures	Data center management policies, Secure coding practices, Training standards of AI/ML models.
	IT and business sensitive procedures	R&D confidential workflows, Intellectual property protection policies, GPU architecture manufacturing process.
	Possible threats	Fail to comply, Product deviations, Lack of security awareness.
Data	Transmission	GPU workload data (real-time), AI model and LLMs training data in transit.
	Processing	Computations at Data Centers, NIM (AI inferencing), Deep learning process.
	Storage	CUDA libraries, Chip architecture files(confidential), Geforce (customer data), GPU firmware.
	Possible threats	Data leaks, Unauthorized access, failure in encryption.
Software	Applications	GeForce experience, CUDA, TensorRT (Own model), NVIDIA Omniverse..
	Operating Systems	Driver software for Windows, Mac, and Linux.

	Security Components	GPU firmware patches, Endpoint protection tools, AI agents for security.
	Possible threats	Malware infection on AI models, Vulnerabilities in software, Supply chain attacks, Delay in updates (for customers).
Hardware	Systems and Security Peripherals	NVIDIA DGX and platforms, GPU hardware for training models (RTX), NVIDIA Jetson embedded systems, Trusted Platform Modules (TPM chips), Encryption in the hardware, Firmware security libraries.
	Possible threats	Hardware failure, More water consumption for data centers, Overheating, Insider collision.
Networking	Internet Components	AI services (cloud-based), NVIDIA corporate websites for various products.
	Intranet Components	Data center interlinks, NVIDIA VPC networks (Insider network)
	Possible threats	DDoS attacks, Misconfiguration of resources, Unauthorized access to network (detrimental).

In the above table, it clearly explains the assets and possible threats that could happen if something goes wrong with these assets, because a big company like NVIDIA should never be interrupted. After all, it can cause too much damage and loss, especially reputation in line.

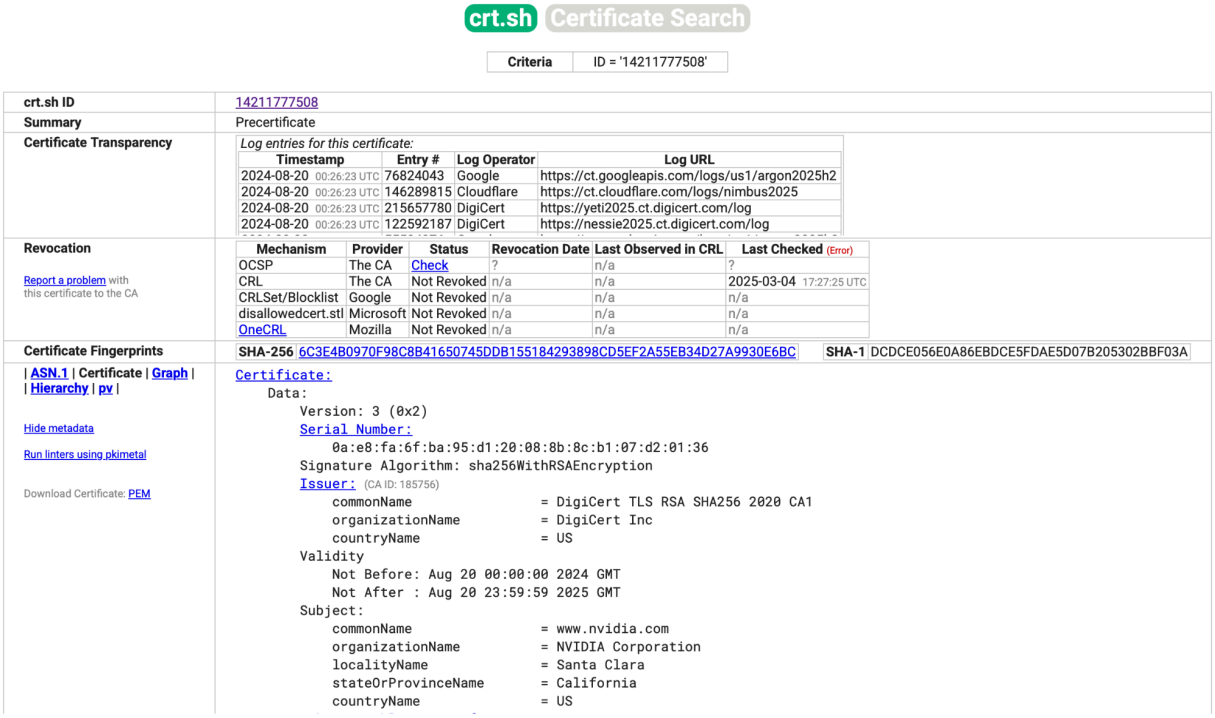
As the NVIDIA API is being used by so many AI services and their drivers are resources for millions of data centers across the world for running the AI applications, even blockchains. So it shows the attack surface is high and there are a lot of entry points and assumptions can be made to attack NVIDIA assets.

Therefore, we are going to study potential vulnerabilities and form a probability vs. impact matrix to showcase its effect over the company and create a checkbox for monitoring the assets and safeguarding the resources from the attackers.

We can appreciate NVIDIA for taking measures quickly and having a security team dedicated to protecting the resources, and this report is going to be interesting from here.

Certificate Test

NVIDIA intensively lists their products online and showcases their research, so I tried to fetch the SSL/TLS certificate[3] from crt.sh and it seems to be perfect and has time till Aug 20, 2025.



crt.sh Certificate Search

Criteria	ID = '14211777508'

crt.sh ID	14211777508
Summary	Precertificate
Certificate Transparency	*Log entries for this certificate:*

Timestamp	Entry #	Log Operator	Log URL
2024-08-20 00:26:23 UTC	76824043	Google	https://ct.googleapis.com/logs/us1/argon2025h2
2024-08-20 00:26:23 UTC	146289815	Cloudflare	https://ct.cloudflare.com/logs/nimbus2025
2024-08-20 00:26:23 UTC	215657780	DigiCert	https://yeti2025.ct.digicert.com/log
2024-08-20 00:26:23 UTC	122592187	DigiCert	https://nessie2025.ct.digicert.com/log

Revocation

Report a problem with this certificate to the CA

Mechanism	Provider	Status	Revocation Date	Last Observed in CRL	Last Checked (Error)
OCSP	The CA	Check	?	n/a	?
CRL	The CA	Not Revoked	n/a	n/a	2025-03-04 17:27:25 UTC
CRLSet/Blocklist	Google	Not Revoked	n/a	n/a	n/a
disallowedcert.stl	Microsoft	Not Revoked	n/a	n/a	n/a
OneCRL	Mozilla	Not Revoked	n/a	n/a	n/a

Certificate Fingerprints: SHA-256 6C3E4B0970F98C8B41650745DDB155184293898CD5EF2A55EB34D27A9930E6BC SHA-1 DCDCE056E0A86EBDCE5FDAE5D07B205302BBF03A

| ASN.1 | Certificate | Graph | Hierarchy | pv |

Hide metadata

Run linters using pkimetal

Download Certificate: PEM

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            0a:e8:fa:6f:ba:95:d1:20:08:8b:8c:b1:07:d2:01:36
        Signature Algorithm: sha256WithRSAEncryption
        Issuer: (CA ID: 185756)
            commonName               = DigiCert TLS RSA SHA256 2020 CA1
            organizationName         = DigiCert Inc
            countryName              = US
        Validity
            Not Before: Aug 20 00:00:00 2024 GMT
            Not After : Aug 20 23:59:59 2025 GMT
        Subject:
            commonName               = www.nvidia.com
            organizationName         = NVIDIA Corporation
            localityName             = Santa Clara
            stateOrProvinceName      = California
            countryName              = US
```

Why is it necessary?

Web traffic is enormous and requires so much attention, as it's the source that interacts with users. Updating the SSL/TLS versions and keeping the certificate up to date, showcases trust and stronger encryption, which protects the data in transit, especially credentials and PII information.

Also, we have various things to start with, especially with an enterprise, so having the basics stronger is beneficial and showcases its trustworthiness and reputation of the company, especially to the security.

The second thing I am going to test is whether any old CVEs still exist which are not in the hands of NVIDIA and rely on third party vendors, like the case below showcases how Mac and NVIDIA need to work together. It shows the importance of dependency and third party vendors when looking at the security of an enterprise in general.

CVE-2018-4230

This vulnerability[4] is old, but interesting because it is not listed or acknowledged by Nvidia in their product security website[5]. But this vulnerability is interesting because it affects macOS older versions, as there is no direct fix from NVIDIA for this, so if somebody is using an older version of macOS (10.13.5), they are prone to this vulnerability.

Vulnerability Details

It's called a race condition leading to Use-After-Free(UAF) issue in SetAppSupportBits. This condition occurs when two or three processes or threads try to access the same resource without proper synchronization, it is like a side channel attack that can be simulated on different cores of the CPU. These conditions can lead to unexpected behaviour.

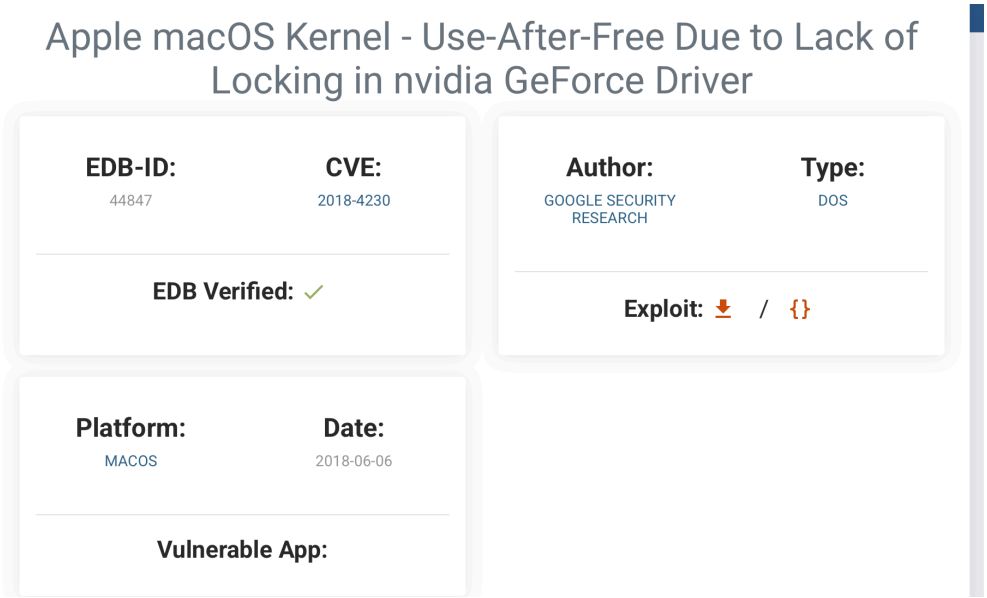
Process

A malicious app having a malicious process can participate in the race and try to access the resources, enabling the attacker to run arbitrary code with higher privileges.

- Allows attackers to bypass user restrictions.
- Data corruption.

Severity

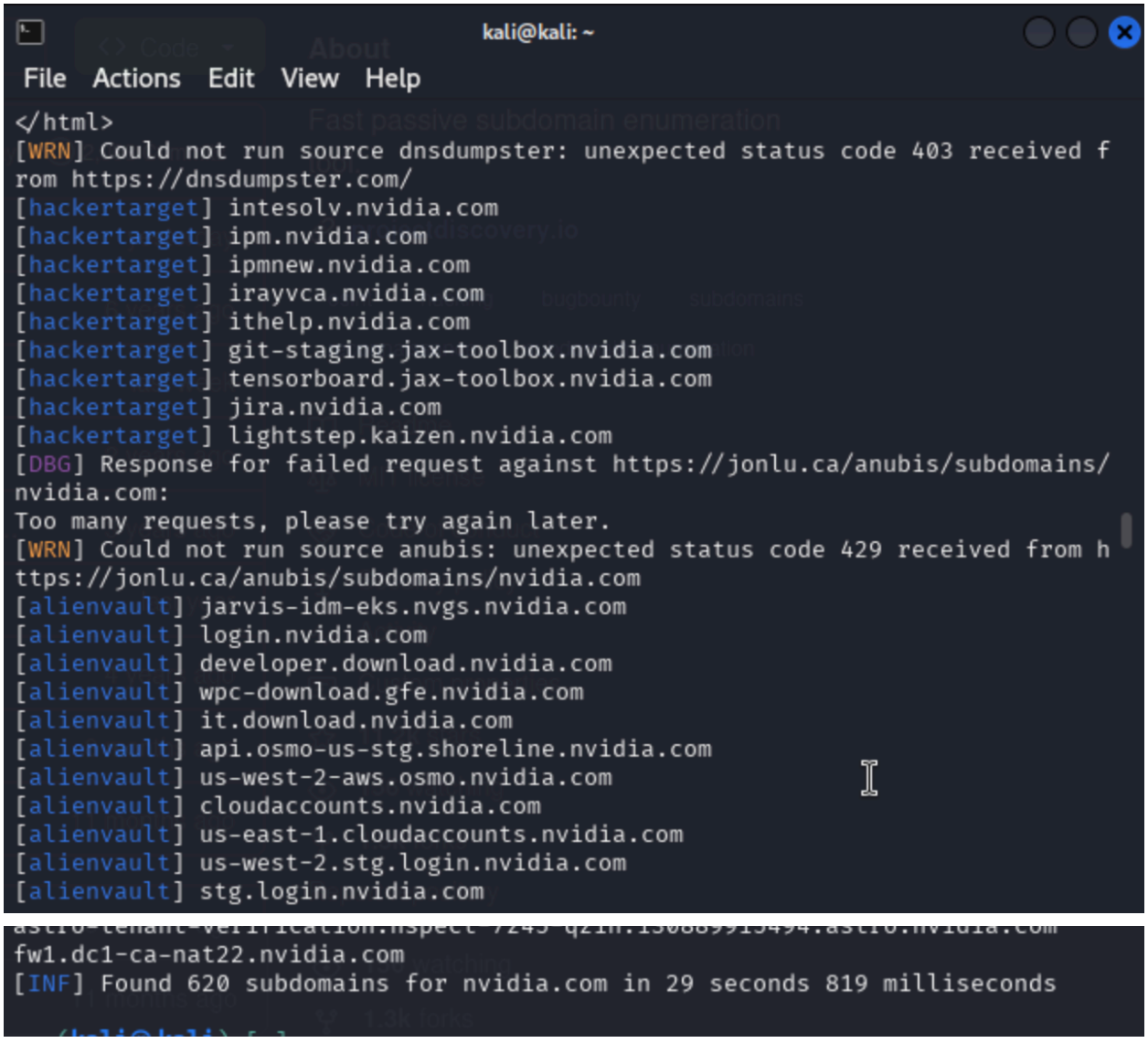
- CVSS Score: 7.0 (HIGH)
- Attack complexity: High
- User Integration: Required
- CIA impact: High



Subfinder

Next up we tried to find the sub domains associated with NVIDIA, so I ran subfinder on my machine to look for subdomains, but NVIDIA WAF blocked some of the requests and also had API restrictions which is good sign initially and showcases their intense use of AWS WAF, but was able to find some subdomains (not a good sign).

NVIDIA should make sure to delete the resources and their subdomains if they no longer need it, as it can affect them highly if an attacker tries to register with the missing subdomain, especially for social engineering.



Found 620 subdomains, which means they need to handle and make sure that their subdomains are under check.

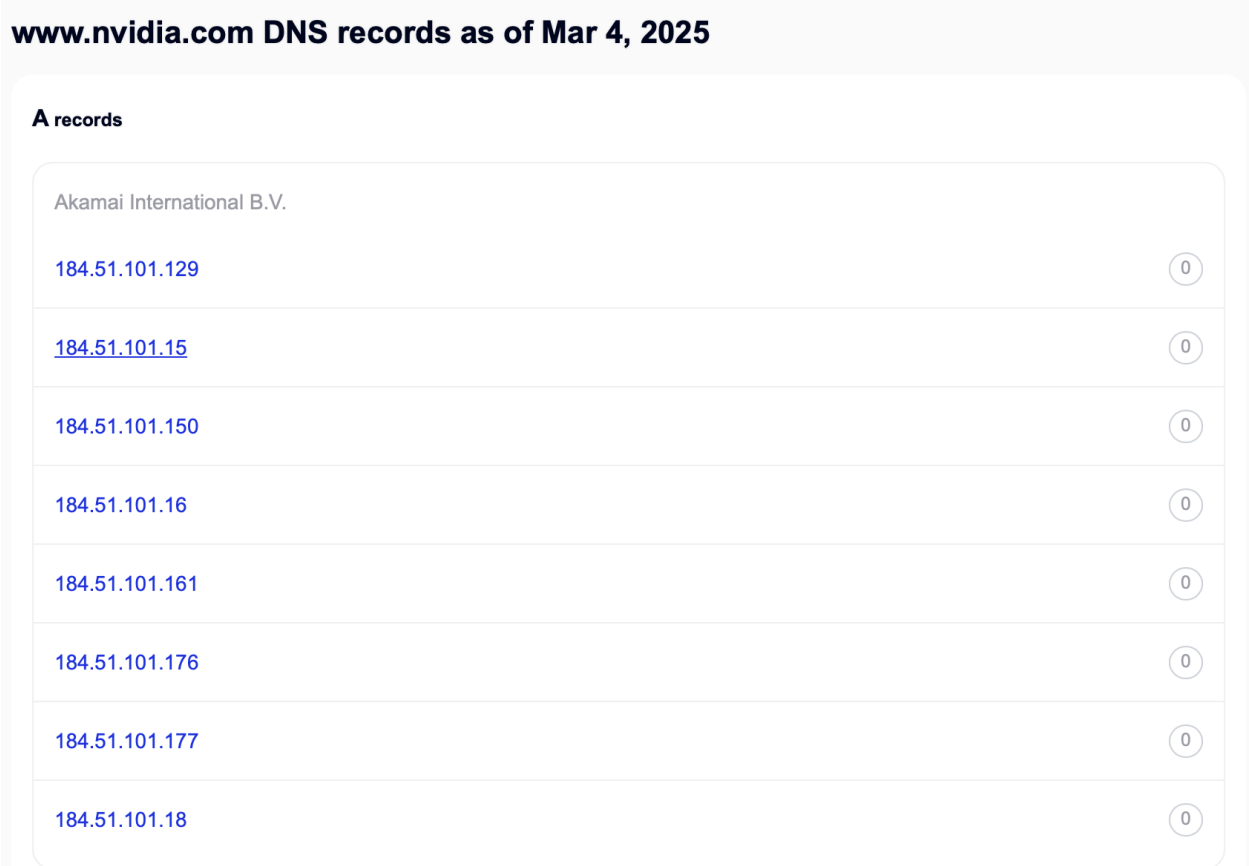
The CVSS score given below reflects the value if the subdomain is taken over by attackers. Maybe they can use the subdomain for various purposes, especially for social engineering and phishing attacks.

CVSS Score: 7.5-9.8 (High-Critical)

Till now, NVIDIA is robust and not much vulnerable to a lot of things, but let's dig deeper using the tool called Security trials and see whether we can get any useful information.

Security Trails

Security Trails[6] is good at scanning the available IP addresses and some DNS records like A records, CNAME records.

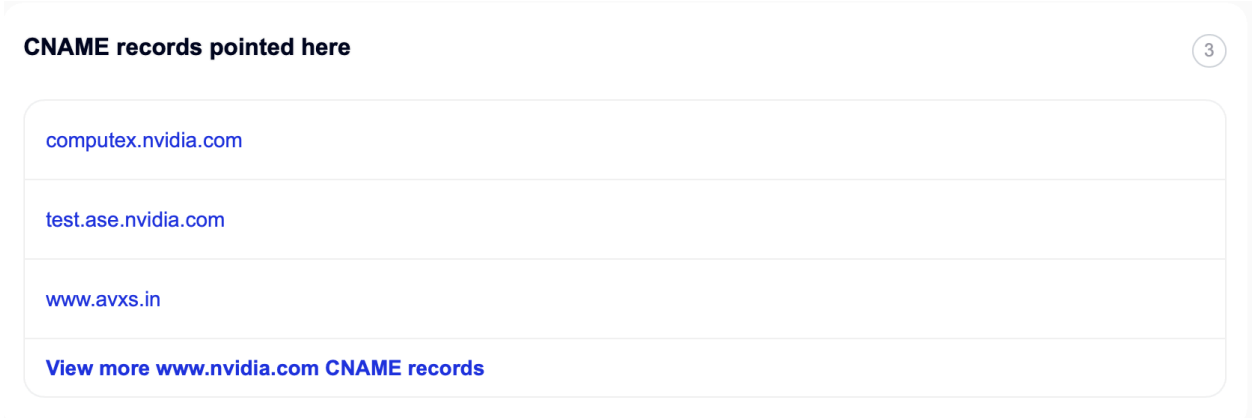


We can find a bunch of IP addresses and it says that these IPs may be associated or hosted with the help of Akamai technologies, who provide cloud solutions through Linode.

This information can be useful, as we have seen previously that third parties and dependency transfer the risk, so if the Akamai server holds vulnerabilities, it can indirectly affect NVIDIA, which is associated with it. For example, in the past Akamai faced a serious vulnerability and it was found by two researchers and they called it, hop by hop headers abuse technique, and http smuggling[7].

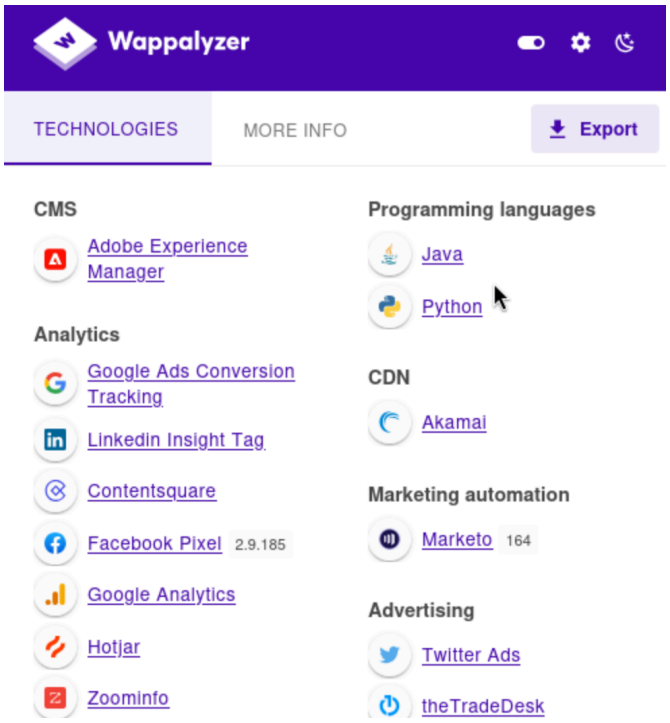
Following the above findings, we got some CNAME records as well. I could have fetched more, but this website needs a subscription for accessing records.

But if some CNAME records are not verified, it can again lead to taking over the domain.



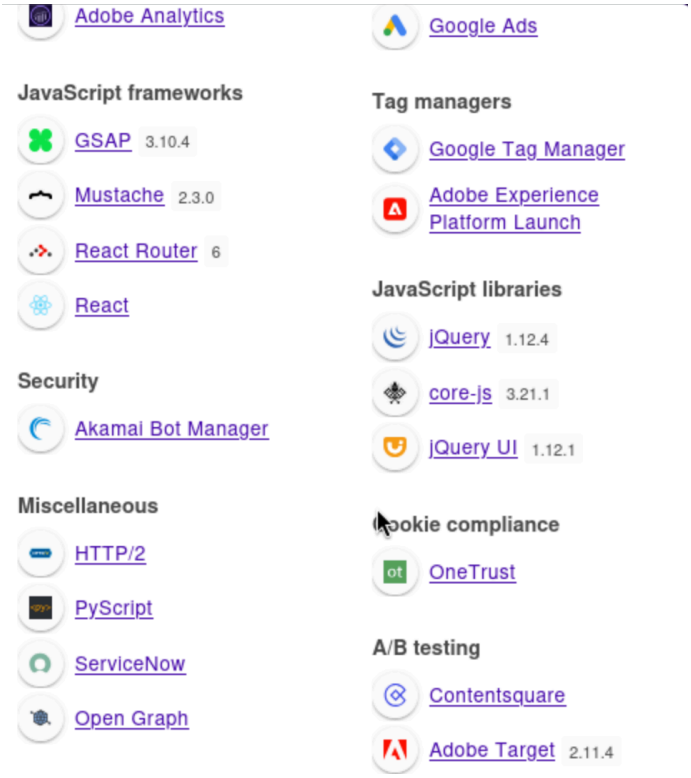
This made me curious and started thinking about the technologies used in the NVIDIA websites because updating the tech stacks to the newest version is crucial, which is not limited to NVIDIA. The better tool to analyze the tech stack will be a Wappalyzer, which can be used as an extension.

Wappalyzer



At the start, we can see the analytics engines used by NVIDIA and they are using quite a lot. So now we know that they are using Akamai as a Content Delivery Network (CDN).

Programming languages and advertising stacks are mentioned, but there is no version mentioned, which can be a good and bad thing, we need to brute force our way into these tech stacks to check whether they are updated or not. But the below screenshot shows a lot of information associated with versions, which we will analyze deeply to check whether they are vulnerable or not.

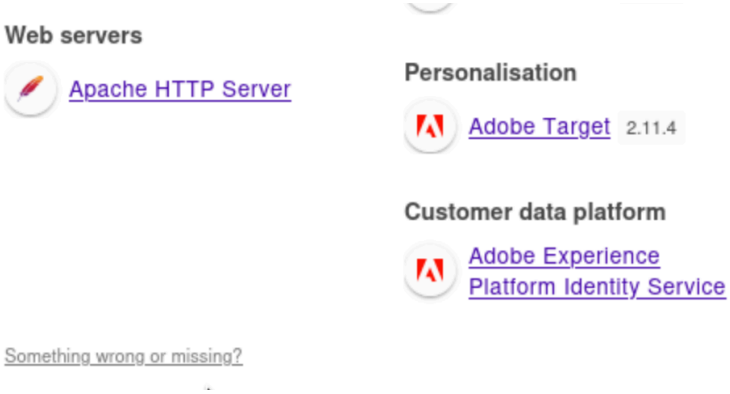


Javascript frameworks and libraries are mentioned, and I can already see some vulnerabilities associated with it especially with JQuery, it's not the newer version and there are some vulnerabilities already existing with JQuery 1.12.4.

The table below explains the vulnerabilities that I have found with some libraries.

Technology	**Version**	**Known Vulnerabilities**	**CVE Details**
jQuery	1.12.4	Multiple issues found (XSS, Prototype pollution, security bypass	CVE-2019-11358
jQuery UI	1.12.1	XSS Vulnerability in widgets and missing security patches	CVE-2021-41183
core-js	3.21.1	Prototype pollution Risks	No CVE but there are security patches in the newest version
React-router	6	History manipulation Vulnerabilities	No CVE but their dependencies are vulnerable.

Below we can see the we have apache as web server, which has history of vulnerabilities from cross scripting to sql injection to authentication bypass[8], but there is no mentioning of version, so we have to brute force with different techniques and try to match the CVEs to confirm whether it's a latest version.



Till now we looked through a Wappalyzer. Let's check with Shodan, and how the findings from it are useful.

Shodan

I tried to look for the domain on shodan and got lot of IPs of data centers associated with NVIDIA, but two results were outside of United States, which made me curious to dig into those and found about the stark industries solution and the domain[9] resoles into the below IP, so I tried to check whois stark industries and found that they registered their domains with go-daddy.

```
└─$ whois stark-industries.solutions
Domain Name: stark-industries.solutions
Registry Domain ID: 72bf32e3db574d3db2d09eedddba4b2f-DONUTS
Registrar WHOIS Server: whois.godaddy.com/
Registrar URL: http://www.godaddy.com/domains/search.aspx?ci=8990
Updated Date: 2024-09-03T21:23:36Z
Creation Date: 2022-02-06T20:23:01Z
Registry Expiry Date: 2030-02-06T20:23:01Z
Registrar: GoDaddy.com, LLC
Registrar IANA ID: 146
Registrar Abuse Contact Email: cctldassistance@godaddy.com
Registrar Abuse Contact Phone: +1.4805058800
Domain Status: clientDeleteProhibited https://icann.org/epp#clientDeleteProhi
bited
Domain Status: clientRenewProhibited https://icann.org/epp#clientRenewProhibi
ted
Domain Status: clientTransferProhibited https://icann.org/epp#clientTransferP
rohibited
Domain Status: clientUpdateProhibited https://icann.org/epp#clientUpdateProhi
bited
Registry Registrant ID: REDACTED FOR PRIVACY
Registrant Name: REDACTED FOR PRIVACY
Registrant Organization: Domains By Proxy, LLC
Registrant Street: REDACTED FOR PRIVACY
Registrant City: REDACTED FOR PRIVACY
Registrant State/Province: Arizona
Registrant Postal Code: REDACTED FOR PRIVACY
Registrant Country: US
```

It’s interesting because I was unable to find anything malicious or abnormal with the stark industries and verified it by visiting abuseipdb and it’s all clear.

AbuseIPDB » *95.164.0.89*

Check an IP Address, Domain Name, or Subnet
e.g. **50.30.232.22**, **microsoft.com**, or **5.188.10.0/24**

95.164.0.89 CHECK

95.164.0.89 was not found in our database

ISP	STARK INDUSTRIES SOLUTIONS LTD.
Usage Type	Data Center/Web Hosting/Transit
ASN	Unknown
Hostname(s)	vm3680379.stark-industries.solutions
Domain Name	stark-industries.solutions
Country	Poland
City	Warsaw, Mazovia

IP info including ISP, Usage Type, and Location provided by IPInfo. Updated biweekly.

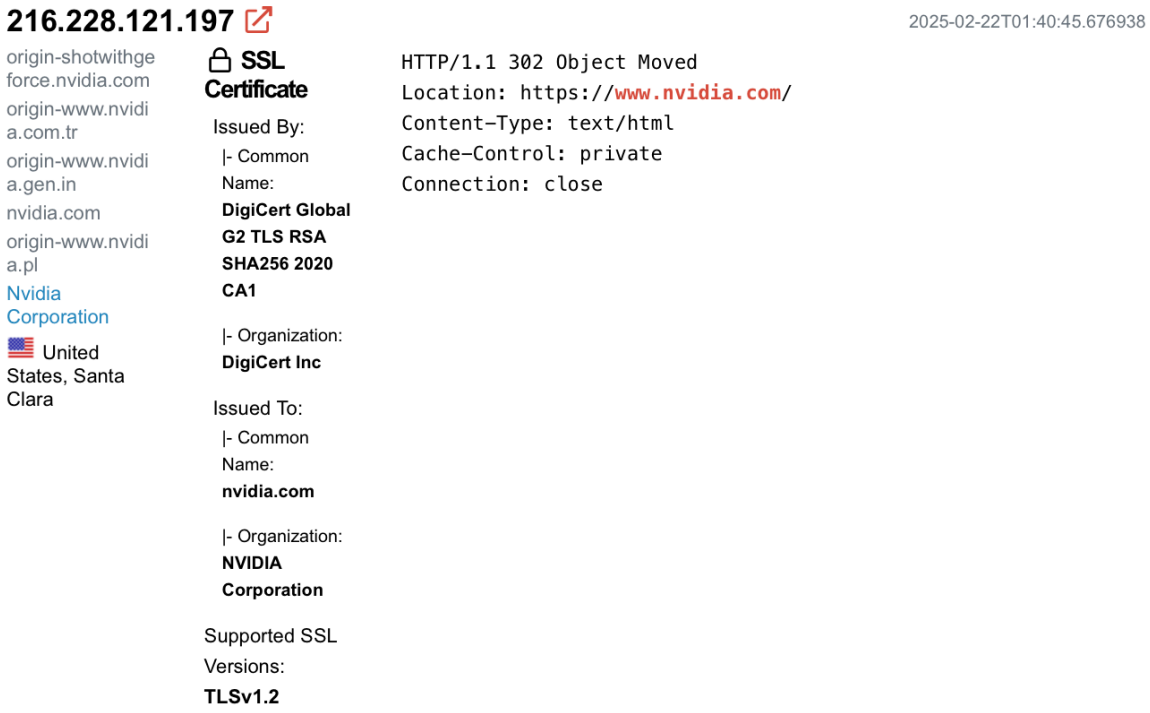
But I started noticing something that some of these data centers are using TLS version 1.2, some were using both v 1.2 and 1.3. As we know that version 1.3 is the latest, and 1.2 has some vulnerabilities like cryptographic weakness and is prone to “Raccoon attack[10]”
But the attack is difficult to perform, but possible to hack the data center.

Some interesting vulnerabilities in TLS 1.2[11] are as follows

- Vulnerability in Perfect forward secrecy (Session Hijacking)
- Weak Cipher Suites
- Quantum Weakness
- CRIME (Compression Ratio info-leak made easy)

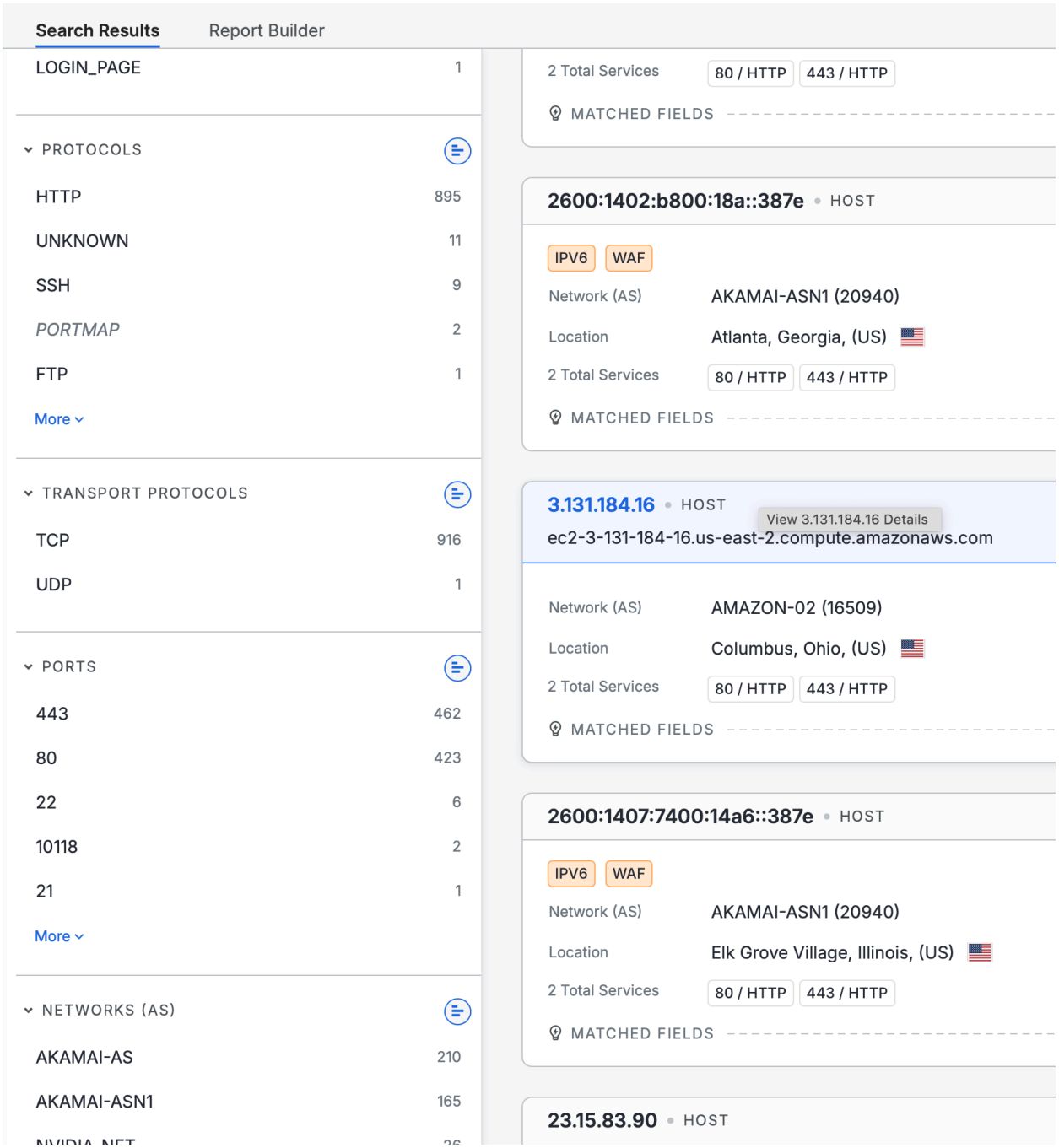
Mitigation:

We can mitigate the issues with TLS 1.2 through stronger configuration and updating to the newest version.



Then I thought of doing a Censys search, a popular search list of critical variables that’s important for vulnerable assessment.

Censys search



This reveals about protocols, ports, and services, so let’s break down the possibilities of exploitation and potential vulnerabilities.

- There are a total of 895 HTTP services, if they are outdated and misconfigured, they could lead to potential exploitation.
- SSH services are exposed and a total of 9 instances, which can be brute forced, in case they have weak or repetitive passwords.
- On port 21 there is an instance which is by default insecure and vulnerable to credential theft.
- There is an IP that says the hosted information on AWS, can be a development or staging server, can be an entry point for attackers.
- Akamai WAF, as we discussed before in this document, said that if it's misconfigured or outdated protocols, it could lead to a vulnerability.

Mentioning of all the vulnerabilities, these are potential vulnerabilities which I got from public resources without actively participating with actual servers.

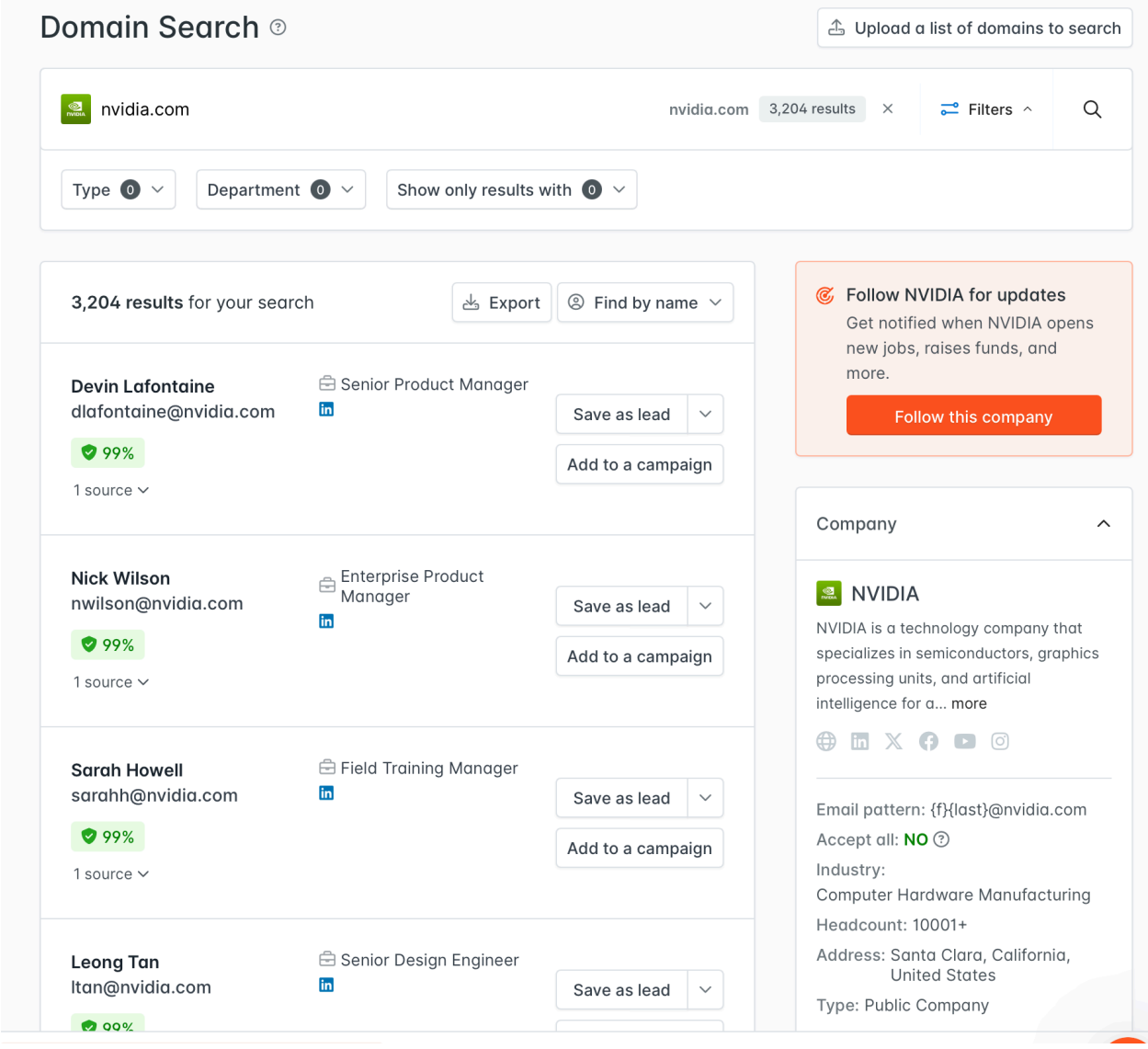
Next up, we are going to see the interesting points which talk about email security. Most of the popular attacks in history are started with a link or we can say a phishing email which has a malicious link with it[1][2], could be clicked or somebody can try to social engineer with advanced techniques especially with the help of artificial intelligence, the possibilities are limitless.

So we changed our focus towards finding the emails associated with NVIDIA, especially the officials and their contacts.

Let's use Hunter.io to find the email ID with @nvidia.com

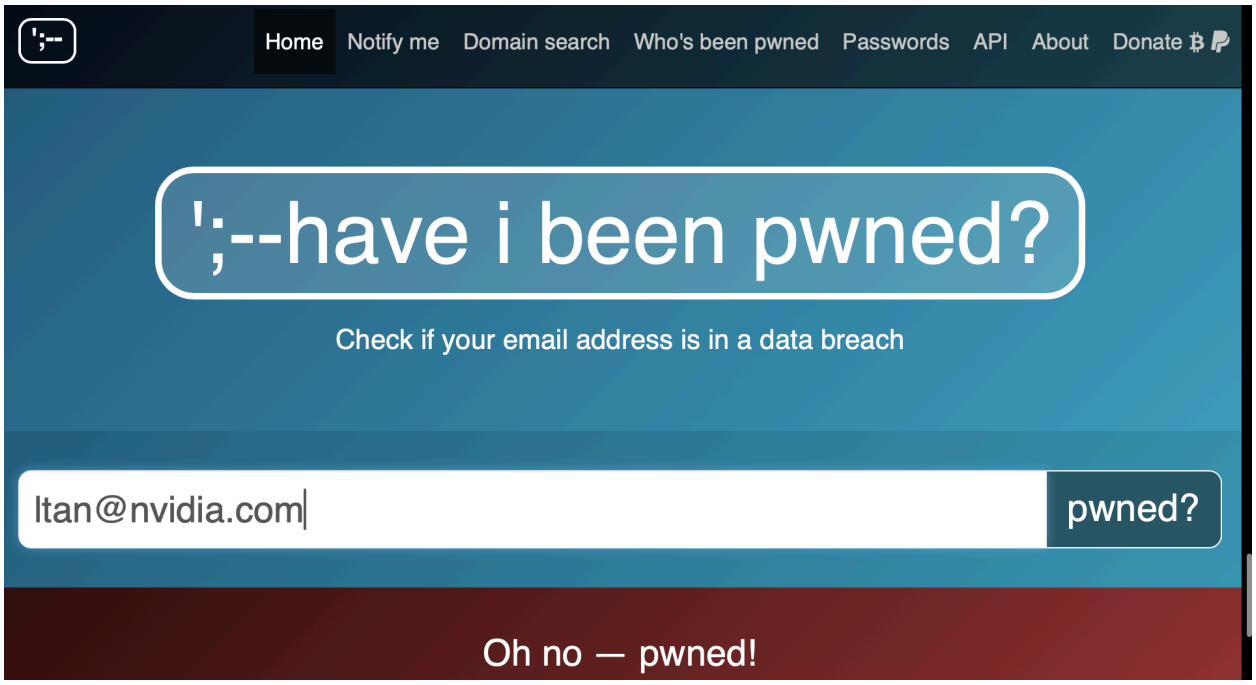
Hunter.io[13]

When we tried to find email addresses using the NVIDIA domain, we got so many results, almost 3,204. The good thing about hunter is that we get email IDs with LinkedIn, which can be useful to understand about the employee and target them via Phishing, Spear Phishing, Pharming, Smishing, evil twin, etc.



To secure an enterprise, employees from different departments, especially from non-IT sectors, should be trained regarding different types of social engineering and attacks. Even social engineering through physical medium.

I thought it would be nice to check whether any of the emails are owned or not, let's hunt about this through the "have I been pwned" website.



To my surprise, I found the email ID that is pwned[14].

AI contribution to the workflow in all technologies that affect millions of lives across the world, especially introduction to AI agents and its use cases across the industries are booming, which led NVIDIA and its stock prices to skyrocket. Along with NVIDIA, it is offering a lot of inferencing services and AI models.

These AI models are great and have great potential, but introduce so many vulnerabilities, and we are going to discuss a critical vulnerability that is addressed by "wiz.io" and has a number of CVE-2024-0132.

CVE-2024-0132

NVIDIA's container toolkit is so popular and powerful because nowadays all companies are using kubernetes and docker to manage their microservices and give grainer control over the services. But what if it turns out that it contains a critical vulnerability, So, it happened with the NVIDIA container toolkit as well, which was brought to the light by wiz.io[15].
In containers and virtualization in general, the host thinks that it's completely isolated like a sandbox. In contrast, at the micro level, it's not true, if you are sharing the common resources like CPU, GPU, and memory, there is a high chance of gaining the host privileges if something is misconfigured.

It’s the same thing that happened with this vulnerability, it allows an attacker to take control of a container image executed by the host and gain full access to the host, which can lead to sensitive information leak, even the infrastructure.

How does it work?

Usually, we can mount the storage we want in the docker, if it’s not properly configured, we can mount the storage of the host file system, so when we get access to the Unix sockets of the container runtime that belongs to host, when we execute the image, we get the root access.

```
ubuntu@wiz-research:~/CVE-2024-0132$ sudo touch /hostfile
ubuntu@wiz-research:~/CVE-2024-0132$ ls -alh /hostfile
-rw-r--r-- 1 root root 0 Sep 26 01:16 /hostfile
ubuntu@wiz-research:~/CVE-2024-0132$ sudo docker run -it --runtime=nvidia wiz-ma
licious-image
root@d3d7f1c2b90a:/# ls -alh /
total 88K
drwxr-xr-x  19 root root 4.0K Sep 26 00:30 .
drwxr-xr-x  19 root root 4.0K Sep 26 00:30 ..
lrwxrwxrwx   1 root root    7 Aug 21 02:13 bin -> usr/bin
drwxr-xr-x   4 root root 4.0K Sep 12 19:29 boot
drwxr-xr-x  16 root root 3.5K Sep  6 18:52 dev
drwxr-xr-x 127 root root  12K Sep 21 06:25 etc
drwxr-xr-x   3 root root 4.0K Aug 23 07:02 home
-rw-r--r--   1 root root    0 Sep 26 01:16 hostfile
lrwxrwxrwx   1 root root    7 Aug 21 02:13 lib -> usr/lib
...
root@d3d7f1c2b90a:/#
```

Mitigation

We can avoid the container run by properly hardening the container environment and setting up zero trust architecture. We should give the least privileges possible to users, and set up the configurations through Dockerfile or the YAML file.

So this CVE brings a lot of vulnerabilities to the table, affecting DevOps and CI/CD pipelines.

Vulnerability in Hopper HGX Systems

Hopper HGX 8-GPU High Performance systems has a critical vulnerability[16] which is recently addressed by NVIDIA, but the effect of this vulnerability is severe risks to AI data centers, cloud tech stacks and other critical environments.

1. CVE-2024-0114(CVSS 8.1) - BMC Exploit for Systemic Compromise

- It is present on the Hopper HGX Management Controller (HMC).
- Attackers who have administrative access to the Baseboard Management Controller can execute remote code execution, escalate privileges, and even cause DoS.
- It can leak AI models, and cross-tenant cloud privilege escalation.
- Anyone is vulnerable if they are using default BMC configurations and they are at higher risk.

2. CVE-2024-0141 (CVSS 6.8) -vBIOS Tampering for DoS Attacks

- It allows attackers to write malicious values to registers and it highly affects GPU vBIOS firmware.
- It can affect cloud services and even data centers by forcing GPUs into unrecoverable failure states.

Mitigation Recommendation

- Immediately updating the firmware version to 1.6.0.
- Hardening the BMC configuration and keeping it as a practice.
- Audit all the access controls for GPU management because they are powering cloud resources.

As almost 80% of data centers and AI models are heavily relying on NVIDIA GPUs, fixing it can prevent supply chain attacks and also protect the infrastructure.

I have discussed some of the vulnerabilities and may have missed some vulnerabilities as NVIDIA has huge infrastructure and has so many resources and the attack surface is big. Also if you are global as an enterprise it becomes a responsibility to manage all the resources and public facing servers and saving the cost at the same time.

Now we are going to make a table on Probability of the threats vs. Impact matrix, because we need to understand the importance of vulnerability and potential threats impact level because of it.

Threats vs. Impact

Potential Vulnerability	**Probability (Low, Medium, High)**	**Impact (Low, Medium, High)**	**Risk Level (Low, Moderate, Critical)**
Software Vulnerabilities in GPU drivers	High	High	Critical
Supply chain attacks on semiconductor manufacturing	Medium	High	Critical
AI model poisoning and adversarial attacks	High	Medium	Moderate
Unauthorized access and Data breaches	High	High	Critical
DDoS Attacks on NVIDIA cloud Services	Medium	Medium	Moderate
Firmware Vulnerabilities	Medium	High	Critical
Intellectual Property theft and Industrial espionage	Medium	High	Critical
Ransomware attacks on NVIDIA's AI Infrastructure	Medium	High	Critical
Insider threats and affecting R&D	Low	High	Moderate
GPU architecture being outdated	High	Medium	Moderate

It requires continuous monitoring, updating the systems, managing the internal teams, and working with business logic, which should never affect the cybersecurity posture of the company.

Based on my assessment, I have listed certain remediation recommendations as follows:

- **Updates and Patches**: Regularly updating the systems and servers will keep the systems immune to vulnerabilities.
- **Configuration Hardening**: Making sure of weak configuration and changing the default settings. Especially in containerized environments.
- **Access Control Implementations**: Utilize zero trust architecture across all the computing devices to make sure least privilege is provided for everyone in NVIDIA.
- **SIEM and SOAR solutions**: Having a good SIEM solution to analyze logs and having matured SOAR playbooks will strengthen the overall infrastructure.
- **Employee Trainings**: Continuously providing training to employees and non-technical staff, especially HR and business managers, is critical and keeps the internal network safe.
- **Incident Response** Planning: Have a backup plan and framework ready to act upon DDoS attacks and respond to potential exploits, also for ransomware.

Conclusion

The vulnerability assessment shows that NVIDIA has matured security systems and servers in place but it doesn't mean they are completely secure and immune to modern threats.

I did this vulnerability assessment based on the publicly available information and am still able to find some vulnerabilities that are moderate to critical. We covered some GPU driver exploits, supply chain risks, server vulnerabilities, and firmware vulnerabilities.

Even though NVIDIA has a history of patching all the vulnerabilities, making sure of using strong access controls, zero-trust architecture, and employee security awareness training will enhance its resilience.

REFERENCES

1. Twingate. (n.d.). *NVIDIA data breach: What you need to know*. Retrieved March 5, 2025, from https://www.twingate.com/blog/tips/nvidia-data-breach
2. NVIDIA. (n.d.). *Product security*. Retrieved March 5, 2025, from https://www.nvidia.com/en-us/product-security/
3. **crt.sh. (n.d.). *Certificate Transparency log search*. Retrieved March 5, 2025, from https://crt.sh/?id=14211777508**
4. **National Institute of Standards and Technology (NIST). (2018). *CVE-2018-4230*. Retrieved March 5, 2025, from https://nvd.nist.gov/vuln/detail/CVE-2018-4230**
5. **NVIDIA. (n.d.). *Product security*. Retrieved March 5, 2025, from https://www.nvidia.com/en-us/product-security/**
6. **SecurityTrails. (n.d.). *NVIDIA domain DNS records*. Retrieved March 5, 2025, from https://securitytrails.com/domain/www.nvidia.com/dns**
7. **PortSwigger. (2025). *Researchers net $46k for Akamai misconfiguration vulnerability*. Retrieved March 5, 2025, from https://portswigger.net/daily-swig/researchers-net-46k-for-akamai-misconfiguration-vulnerability**
8. **MITRE. (n.d.). *CVE records for Apache HTTP server*. Retrieved March 5, 2025, from https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=Apache%20HTTP%20server**
9. **NVIDIA. (n.d.). *NVIDIA Poland website*. Retrieved March 5, 2025, from https://www.nvidia.com/pl-pl/**
10. **SSL.com. (n.d.). *Raccoon Attack targets TLS 1.2 and earlier but is difficult to exploit*. Retrieved March 5, 2025, from https://www.ssl.com/blogs/raccoon-attack-targets-tls-1-2-and-earlier-but-is-difficult-to-exploit/**
11. **Software.Land. (n.d.). *TLS 1.2 vulnerability overview*. Retrieved March 5, 2025, from https://software.land/tls-1.2-vulnerability/**
12. **BlueVoyant. (n.d.). *8 devastating phishing attack examples and prevention tips*. Retrieved March 5, 2025, from https://www.bluevoyant.com/knowledge-center/8-devastating-phishing-attack-examples-and-prevention-tips**
13. **Hunter.io. (n.d.). *Email search for NVIDIA domain*. Retrieved March 5, 2025, from https://hunter.io/search/nvidia.com?product_tour_id=389437**
14. **Hunt, T. (n.d.). *Have I Been Pwned?*. Retrieved March 5, 2025, from https://haveibeenpwned.com/**
15. **Wiz.io. (2024). *NVIDIA AI vulnerability deep dive: CVE-2024-0132*. Retrieved March 5, 2025, from https://www.wiz.io/blog/nvidia-ai-vulnerability-deep-dive-cve-2024-0132**
16. **GBHackers. (n.d.). *NVIDIA issues warning on security vulnerabilities*. Retrieved March 5, 2025, from https://gbhackers.com/nvidia-issues-warning/**