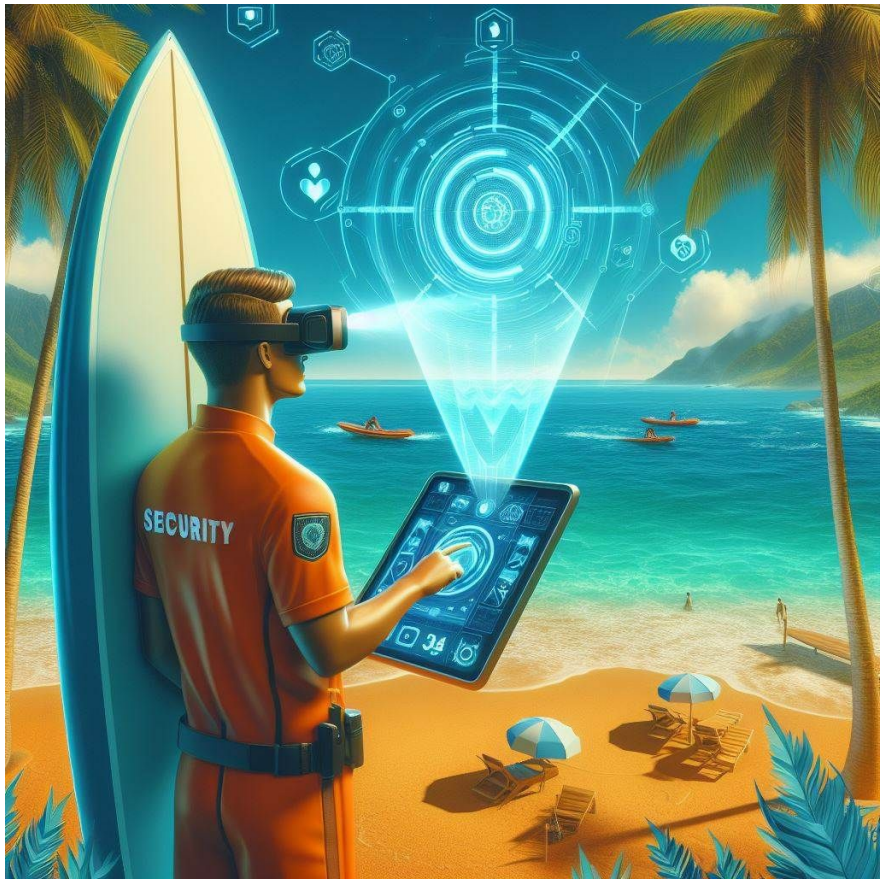
Open source based vulnerability management

Joost Grunwald, Patrick Kuin

21 May, 2024



Agenda



01
Vulnerability management

02
Technical implementation

03
Results & Risks

04
Reports

Vulnerability management

Vulnerability management

The automatic and periodic

- Finding
- Reporting -> solving
- Retesting

of vulnerabilities.

Vulnerability management vs penetration testing vs attack surface management.

Different interests:

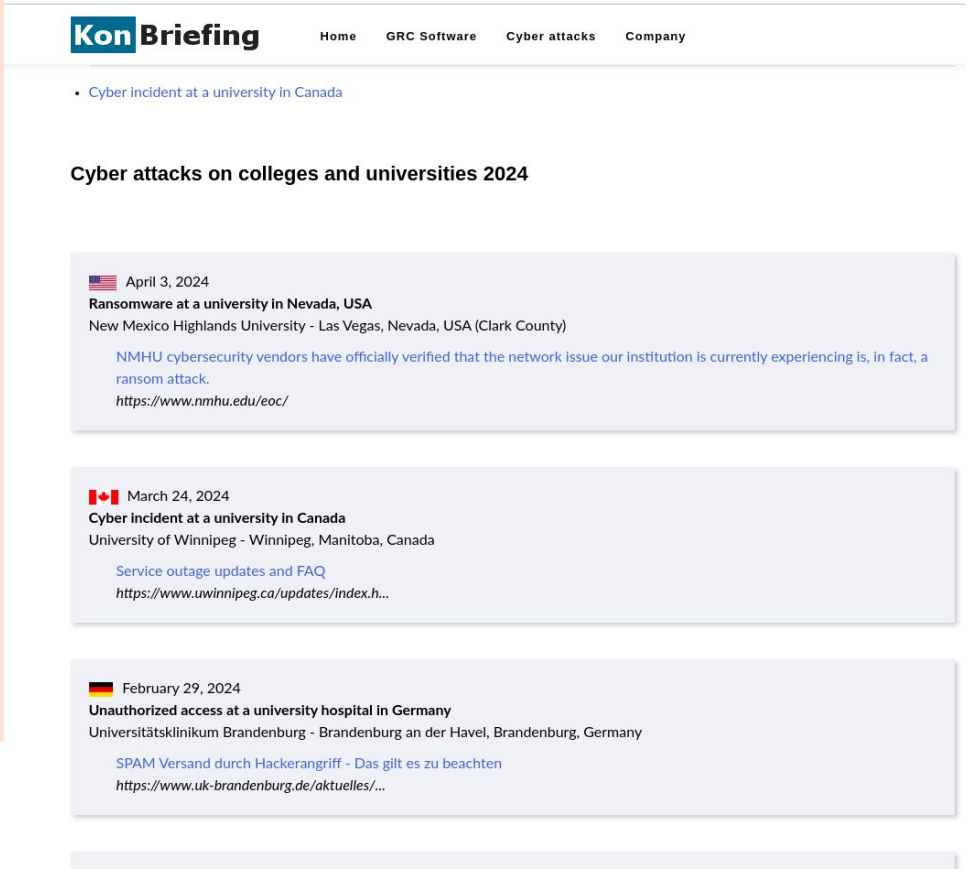
- Riskometer for **CISO's,** total risk for an organization.
- Breachometer for **security engineers**, every single threat in a system.



SURF

Why vulnerability management?

- Prevent hacks
- Compliance…
- Ideal from risk management point of view



"If we maintain constant awareness (**24/7**) of new vulnerabilities across our **external** attack surfaces, internal networks, and all computers, we can proactively **stay ahead** of threats by strategically **prioritizing** the most critical vulnerabilities."

What goes wrong in vulnerability management….

Non transparent market

Solutions are hard to compare

Made our own comparison

Tool	Speed	Confidence	Automation
Tenable Nessus	9	9	5
Nuclei	9	10	10
OpenVAS	6	9	5
Acunetix	6	7	8
Rapid7 Nexpose	7	8	8
RidgeBot	5	7	5
Rapid7 appspider	3	6	5
HCL Appscan	3	7	7
Syhunt Infinity	5	5	7
Nmap	7	4	9
Xray	9	3	9
Shodan	10	3	10

SURF

What goes wrong in vulnerability management….

Lots of false positives

Backports

Vendors giving themselves high CVSS scores

False negatives (missed vulnerabilities)

Expensive on annual/monthly basis

IPV6 subspace often not taken into consideration

Hard to prioritize vulnerabilities, CVSS lacking?

Speed in large organizations

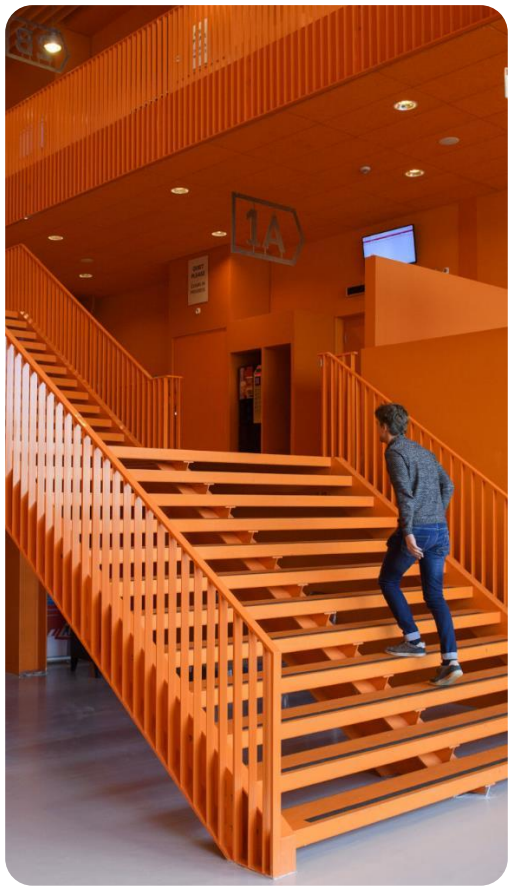
SURF

Innovation

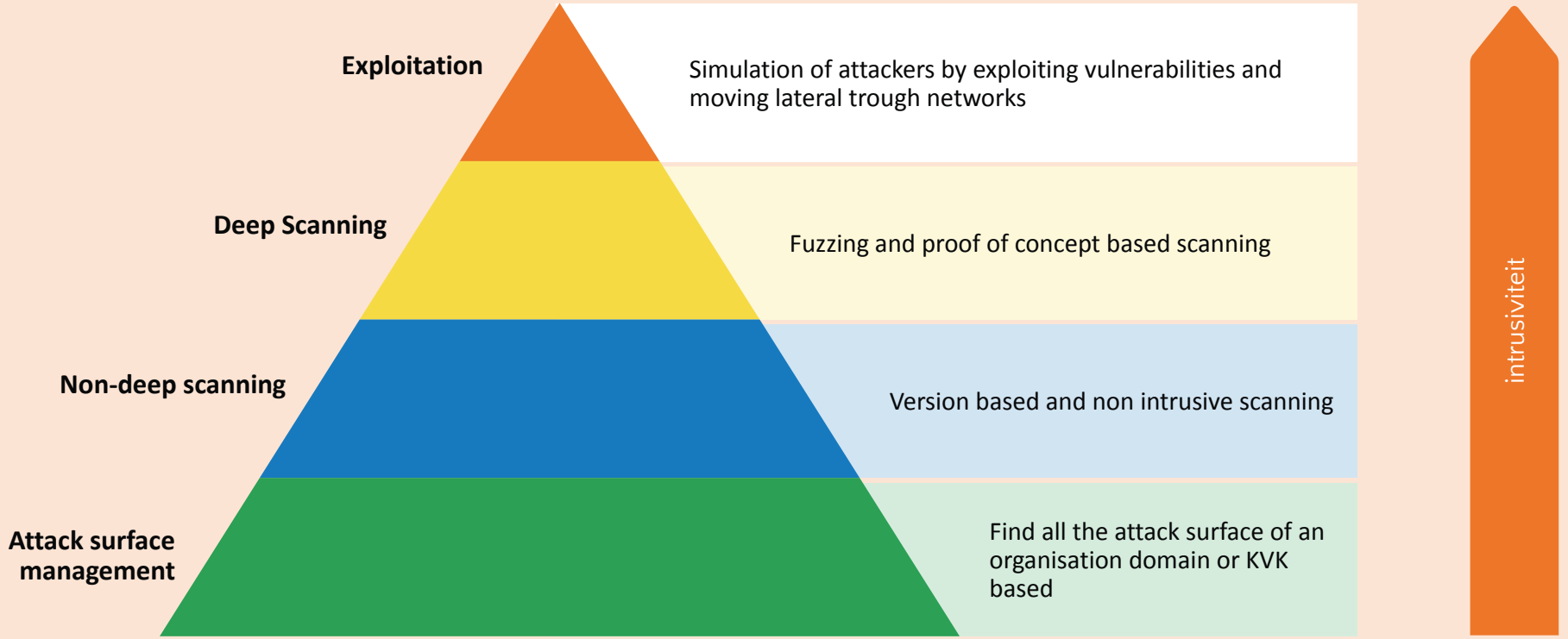
Why don’t we build something ourselves?

- Open source based
- Regarding backports and false positives
- With IPV6 support
- With different risk prioritization
- With external attack surface management
- With network and agent based sensors included.

Because dealing with vulnerabilities properly **makes Dutch education safer.**



Project CyberSphere



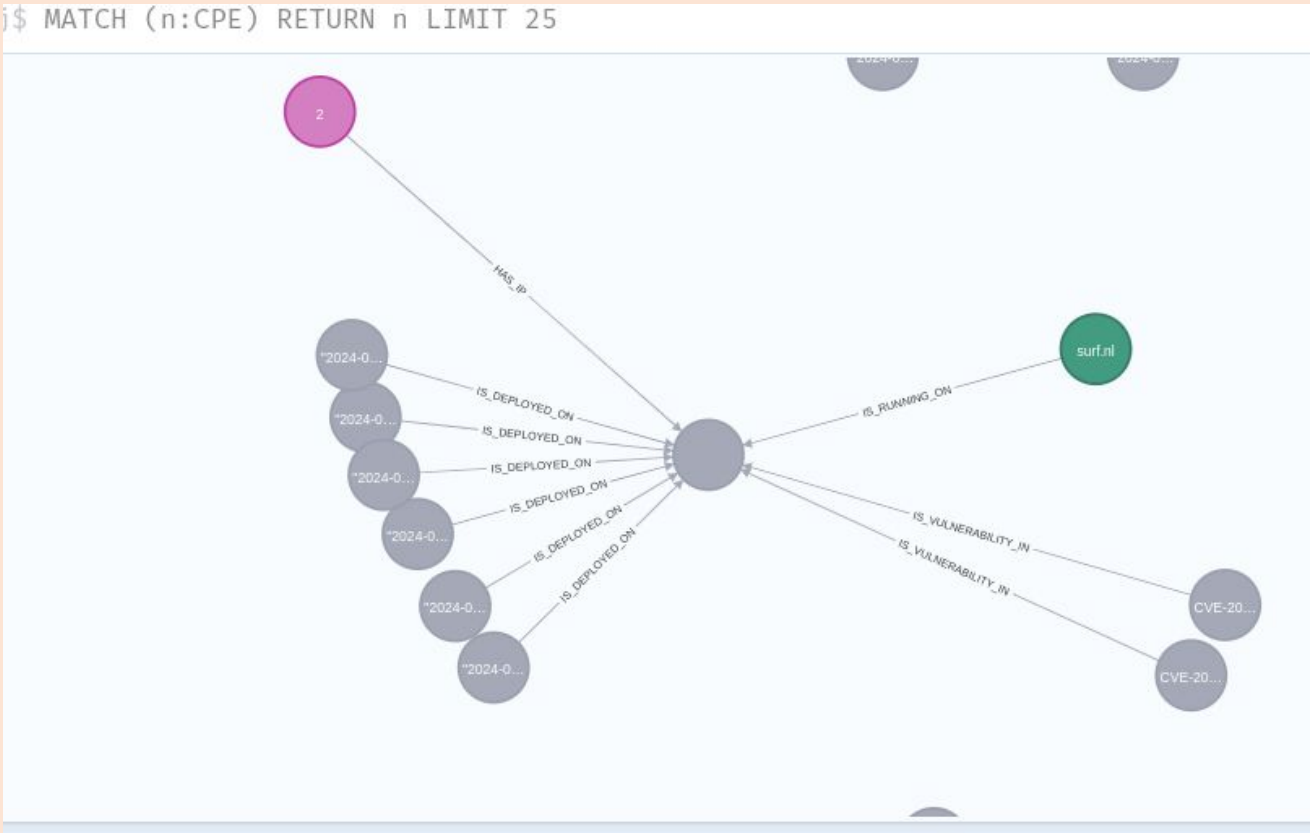
Technical implementation

Mapping


OSINT based scanning
Actively scanning
Fuzzing/POC based
Running exploits
Shadowserver
Nuclei Wordpress
Ridgebot
Censys
Nessus
Nuclei
Metasploit
Shodan
OpenVAS
Acunetix
Pentest AI
Non intrusive
Low confidence
Intrusive
High confidence

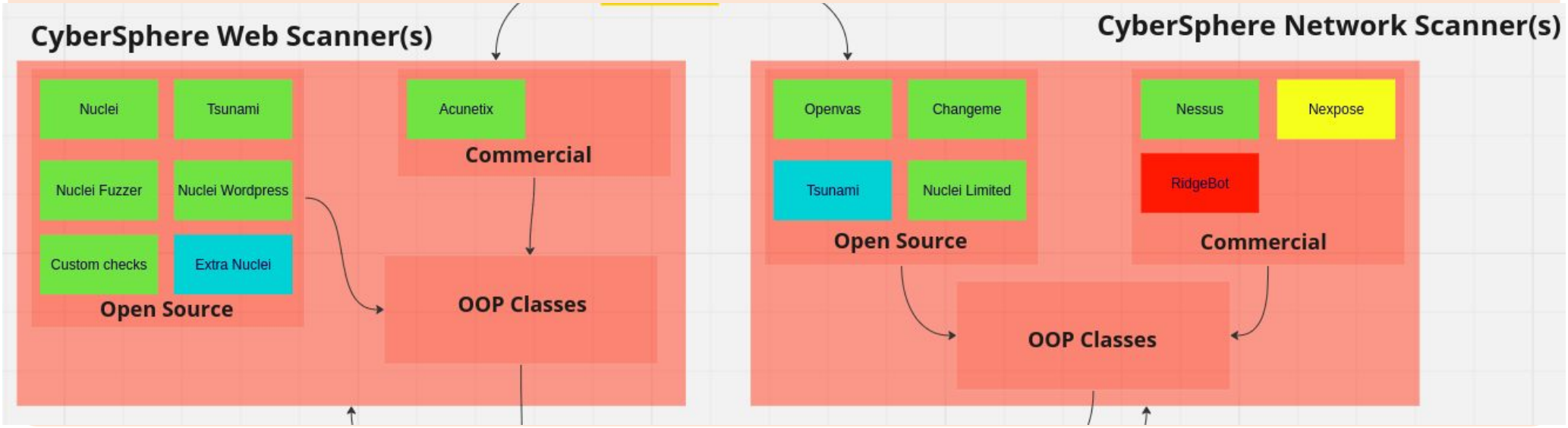
Neo4j & Attack surface management

- IP’s, hostnames
- Cloud buckets
- Cloud environments
- Code repos
- Leaked emails
- Leaked passwords
- CPE’s and technologies
- Vulnerabilities



Scanning

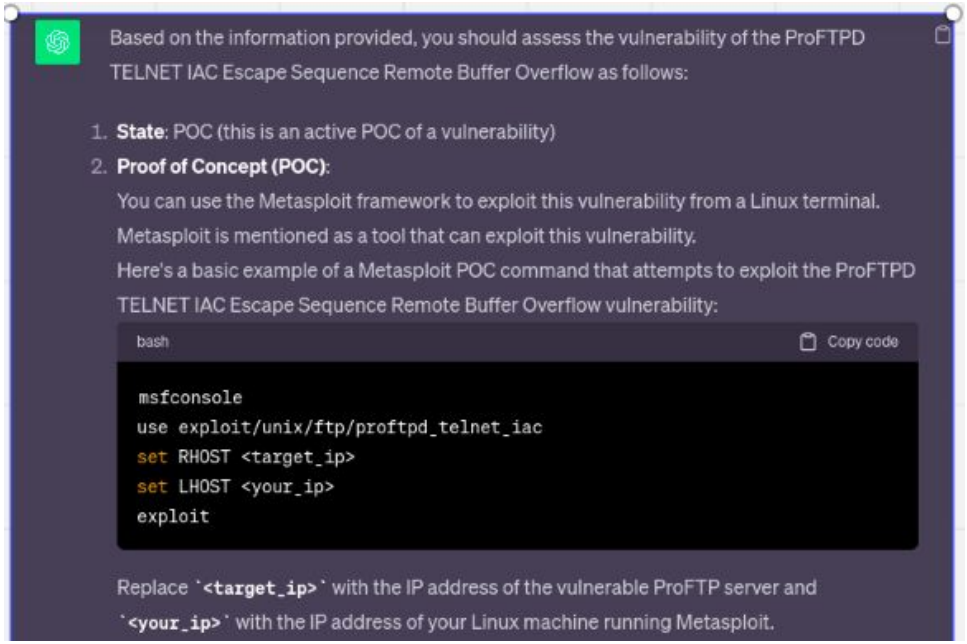
- IP rotation every request
- ASM backbone of multiple scanners
- Orchestrator-agent model



Agent based


Agent in danger (epss) Last 1 year
data_vulnerability_cve_version
4
Unique Count
4
Critical ... Last 1 year
20
High vul... Last 1 year
306
Medium... Last 1 year
318
Low vul... Last 1 year
6
High ag... Last 1 year
9
Crit age... Last 1 year
4
Vulnerable agents (essp top 17%) Last 1 year
NUW-BEST-16
35
data_nessus_host
4
Jossovich
1
nuwelijn-siem
1
Vulnerable agents (High or Critical) Last 1 year
30
19
301
High 85%
nvt 8%
Critical 5%
Low 2%
Vulnerable objects (High or Critical) Last 1 year
Gooole Chrome 93
Mozilla Firefox (x64 nl) 30
VLC media player 23
xxd 23
vim-tiny 23
vim-common 23
Python 3.5.4 (64-bit) 9
Python 3.6.8 (64-bit) 7
libwebkit2gtk-4.0-37 6
libjavascriptcoregtk-4.0-18 6
gir1.2-webkit2-4.0 6
gir1.2-javascriptcoregtk-4.0 6
Vulnerable agents (high or critical) Last 1 year
nuwelijn-siem 154
Jossovich 133
DESKTOP-T9UF... 17
NUW-BEST-12 6
NUW-BEST-11 4
data_nessus_host 4
NUW-BEST-10 3
NUW-BEST-09 3
DESKTOP-KD1J... 2
Vulnerable agents (high or critical) Last 4 days
CVE agent package CVSS3 EPSS % data_vulnerability_rationale
CVE-2022-48554 nuwelijn-siem file 5.50 0.105 File before 5.43 has an stack-based buffer over-read in file_copystr in funcs.c. NOTE: "File" is the name of an Ope...
CVE-2022-48560 Jossovich Python 3.7.9 (64-bit) 7.50 0.172 A use-after-free exists in Python through 3.9 via heappushpop in heapq.
CVE-2022-48564 Jossovich Python 3.7.9 (64-bit) 6.50 0.0695 read_ints in plistlib.py in Python through 3.9.1 is vulnerable to a potential DoS attack via CPU and RAM exhaustion...
CVE-2022-48564 Jossovich Python 3.6.8 (64-bit) 6.50 0.0695 read_ints in plistlib.py in Python through 3.9.1 is vulnerable to a potential DoS attack via CPU and RAM exhaustion...
CVE-2022-48564 DESKTOP-T9UF5... Python 3.5.4 (64-bit) 6.50 0.0695 read_ints in plistlib.py in Python through 3.9.1 is vulnerable to a potential DoS attack via CPU and RAM exhaustion...
CVE-2023-39742 nuwelijn-siem libgif7 0 0 giflib v5.2.1 was discovered to contain a segmentation fault via the component getarg.c.
CVE-2023-39742 nuwelijn-siem libgif7 0 0 nvt
CVE-2023-4428 NUW-BEST-10 Google Chrome 8.10 0.141 Out of bounds memory access in CSS in Google Chrome prior to 116.0.5845.110 allowed a remote attacker to perf...
CVE-2023-4428 NUW-BEST-10 Google Chrome 8.10 0 nvt

Testing LLM based exploitation



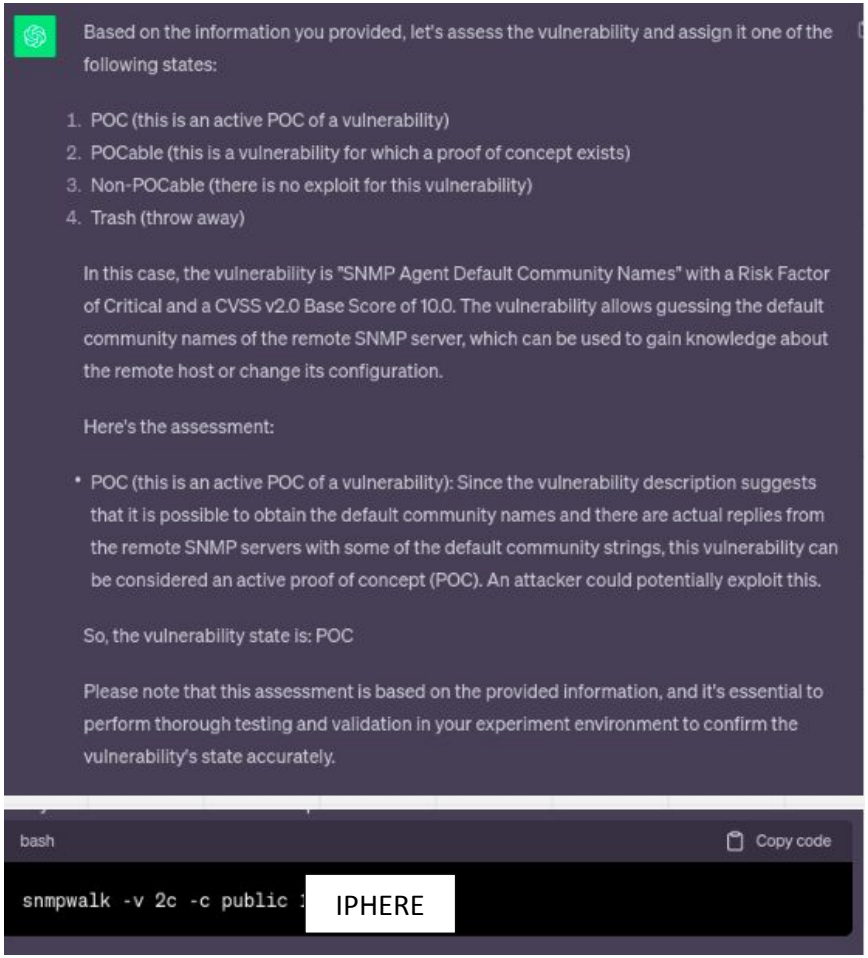
Based on the information provided, you should assess the vulnerability of the ProFTPD TELNET IAC Escape Sequence Remote Buffer Overflow as follows:

1. **State**: POC (this is an active POC of a vulnerability)
2. **Proof of Concept (POC)**:
   You can use the Metasploit framework to exploit this vulnerability from a Linux terminal. Metasploit is mentioned as a tool that can exploit this vulnerability.
   Here's a basic example of a Metasploit POC command that attempts to exploit the ProFTPD TELNET IAC Escape Sequence Remote Buffer Overflow vulnerability:

bash — Copy code

```
msfconsole
use exploit/unix/ftp/proftpd_telnet_iac
set RHOST <target_ip>
set LHOST <your_ip>
exploit
```

Replace `<target_ip>` with the IP address of the vulnerable ProFTP server and `<your_ip>` with the IP address of your Linux machine running Metasploit.



Based on the information you provided, let's assess the vulnerability and assign it one of the following states:

1. POC (this is an active POC of a vulnerability)
2. POCable (this is a vulnerability for which a proof of concept exists)
3. Non-POCable (there is no exploit for this vulnerability)
4. Trash (throw away)

In this case, the vulnerability is "SNMP Agent Default Community Names" with a Risk Factor of Critical and a CVSS v2.0 Base Score of 10.0. The vulnerability allows guessing the default community names of the remote SNMP server, which can be used to gain knowledge about the remote host or change its configuration.

Here's the assessment:

- POC (this is an active POC of a vulnerability): Since the vulnerability description suggests that it is possible to obtain the default community names and there are actual replies from the remote SNMP servers with some of the default community strings, this vulnerability can be considered an active proof of concept (POC). An attacker could potentially exploit this.

So, the vulnerability state is: POC

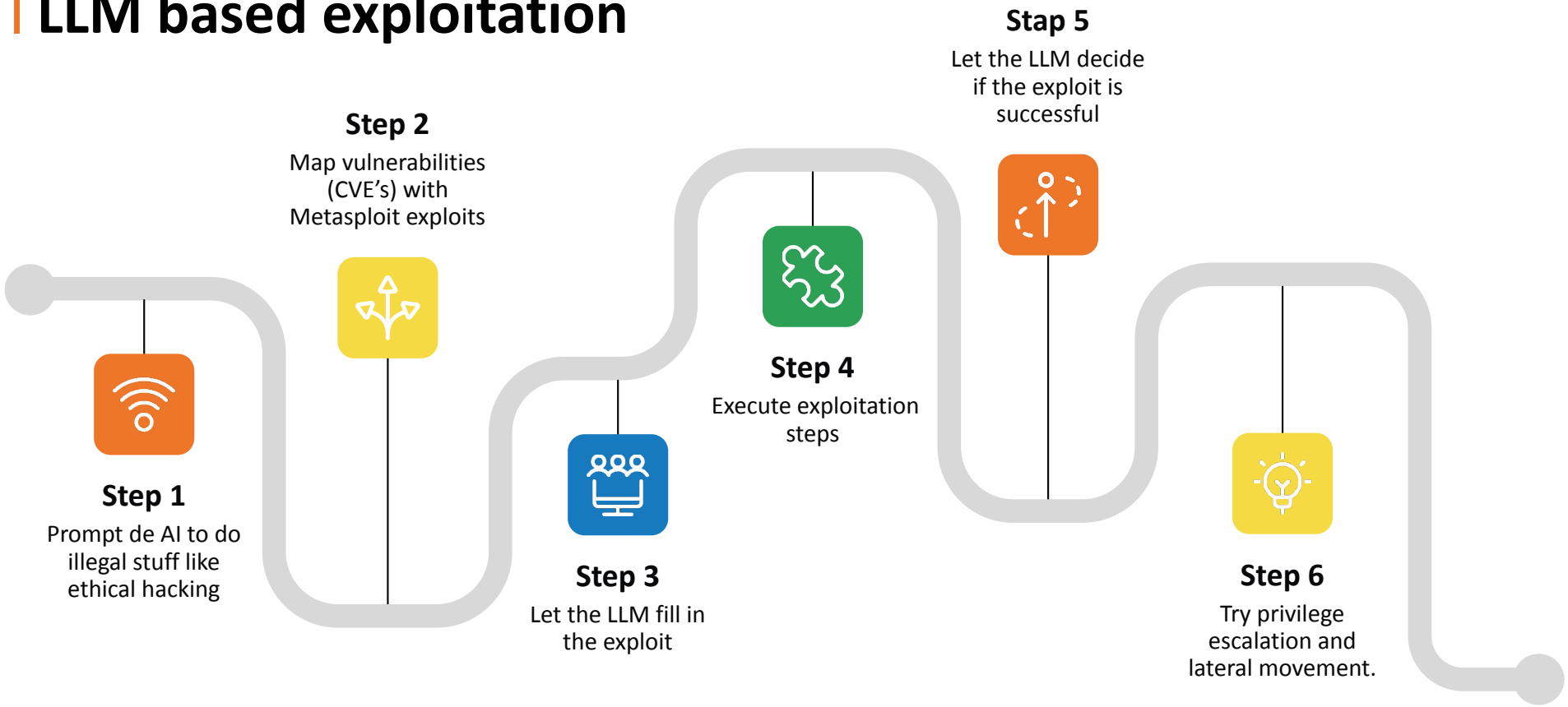
Please note that this assessment is based on the provided information, and it's essential to perform thorough testing and validation in your experiment environment to confirm the vulnerability's state accurately.

bash — Copy code

```
snmpwalk -v 2c -c public 1
```

IPHERE

LLM based exploitation



SURF

Langchain based pentesting?



TABLE 1: Overall performance of LLMs on Penetration Testing Benchmark.

	Easy		Medium		Hard		Average	
Tools	Overall (7)	Sub-task (77)	Overall (4)	Sub-task (71)	Overall (2)	Sub-task (34)	Overall (13)	Sub-task (182)
GPT-3.5	1 (14.29%)	24 (31.17%)	0 (0.00%)	13 (18.31%)	0 (0.00%)	5 (14.71%)	1 (7.69%)	42 (23.07%)
GPT-4	4 (57.14%)	52 (67.53%)	1 (25.00%)	27 (38.03%)	0 (0.00%)	8 (23.53%)	5 (38.46%)	87 (47.80%)
Bard	2 (28.57%)	29 (37.66%)	0 (0.00%)	16 (22.54%)	0 (0.00%)	5 (14.71%)	2 (15.38%)	50 (27.47%)
Average	2.3 (33.33%)	35 (45.45%)	0.33 (8.33%)	18.7 (26.29%)	0 (0.00%)	6 (17.64%)	2.7 (20.5%)	59.7 (32.78%)

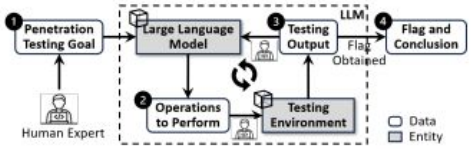


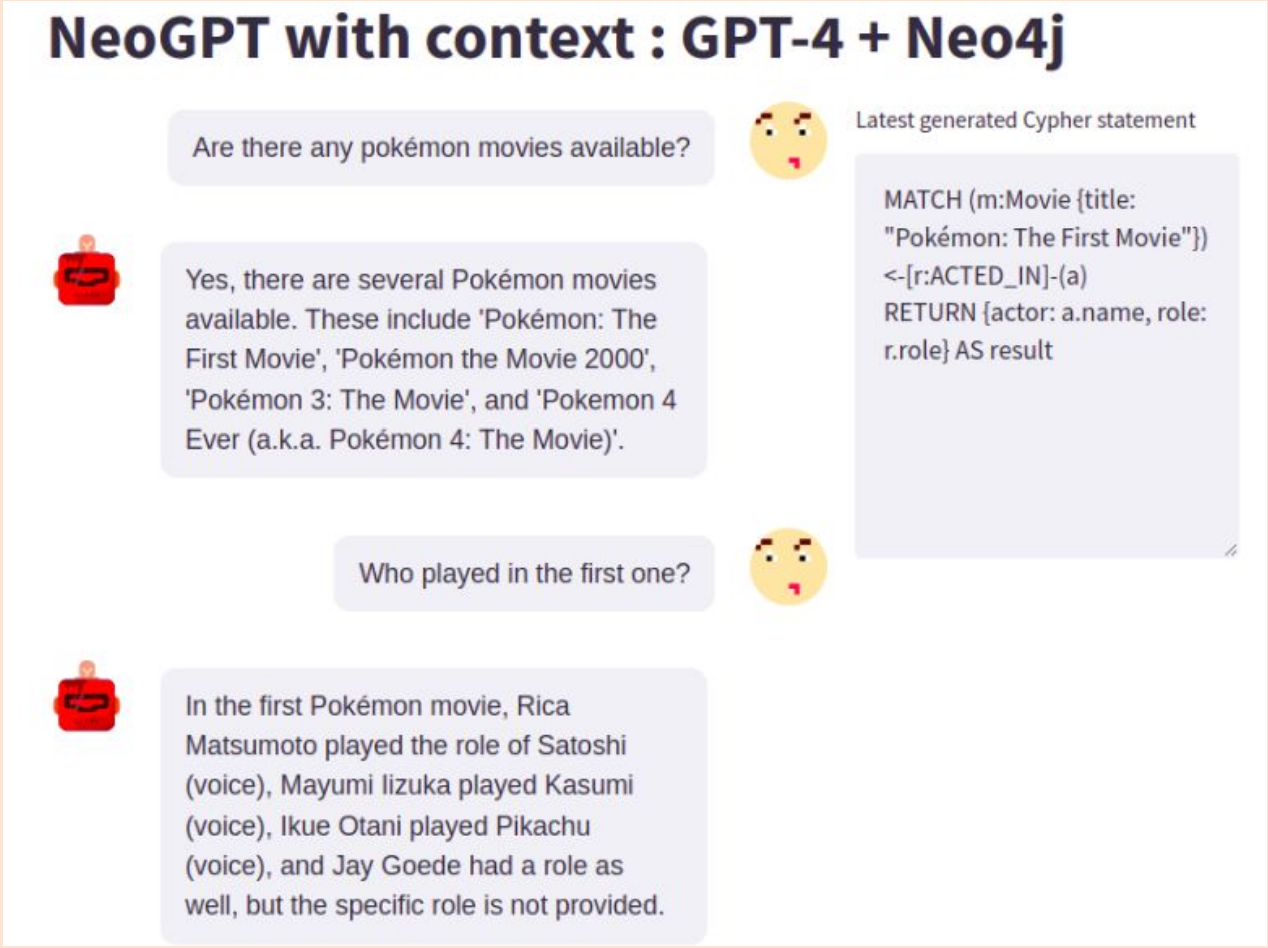
Figure 2: Overview of strategy to use LLMs for penetration testing.

ners such as Nexus [30] and OpenVAS [31]. Consequently, we explicitly instruct the LLMs to refrain from using these tools. However, we follow the LLMs' recommendations for utilizing other tools designed to validate specific vulnerability types (e.g., sqlmap [32] for SQL injections). Occasionally, versioning discrepancies may lead the LLMs to provide incorrect instructions for tool usage. In such instances, our penetration testing experts evaluate whether the instructions would have been valid for a previous version of the tool. They then make any necessary adjustments to ensure the tool's correct operation.

SURF

NeoGPT

- Currently WIP
- Ask questions to DB
- Let DB generate custom Reports by using components



Game changers in VM: LLM's & Exploitation

LLM's

- FP prevention
- A little bit of intelligence is enough
- Chatbot that explains your data
- Data enrichment
- In future: partial pen testing capabilities

Exploitation

If we actually use found holes? Then we could

- Test for privilege escalation
- Do lateral movement
- Find candy

We can partially already create attack chains!

- Agent
- Outside scanner
- Inside scanner

SURF

Results and Risk

Results

01
Improvement on our own VM tool (Outpost24)

02
Improvement on Tenable Nessus

03
LFI dutch government

04
SSTI → RCE in system of dutch government

05
RCE at NASA

06
Finding full student databases exposed for 30.000+ students at school in NL

07
Success in bug bounty

SURF

Risk presented in the industry

CVSS
What is the impact of a vulnerability?

Often bloated
Lots of reasons for vendors to find ‘high risks’.

Unfair indication
Is a vulnerability really exploitable?

How long is this already exploitable?

Is it at the main website surf.nl) or on a less important website conferentie2017.surf.nl

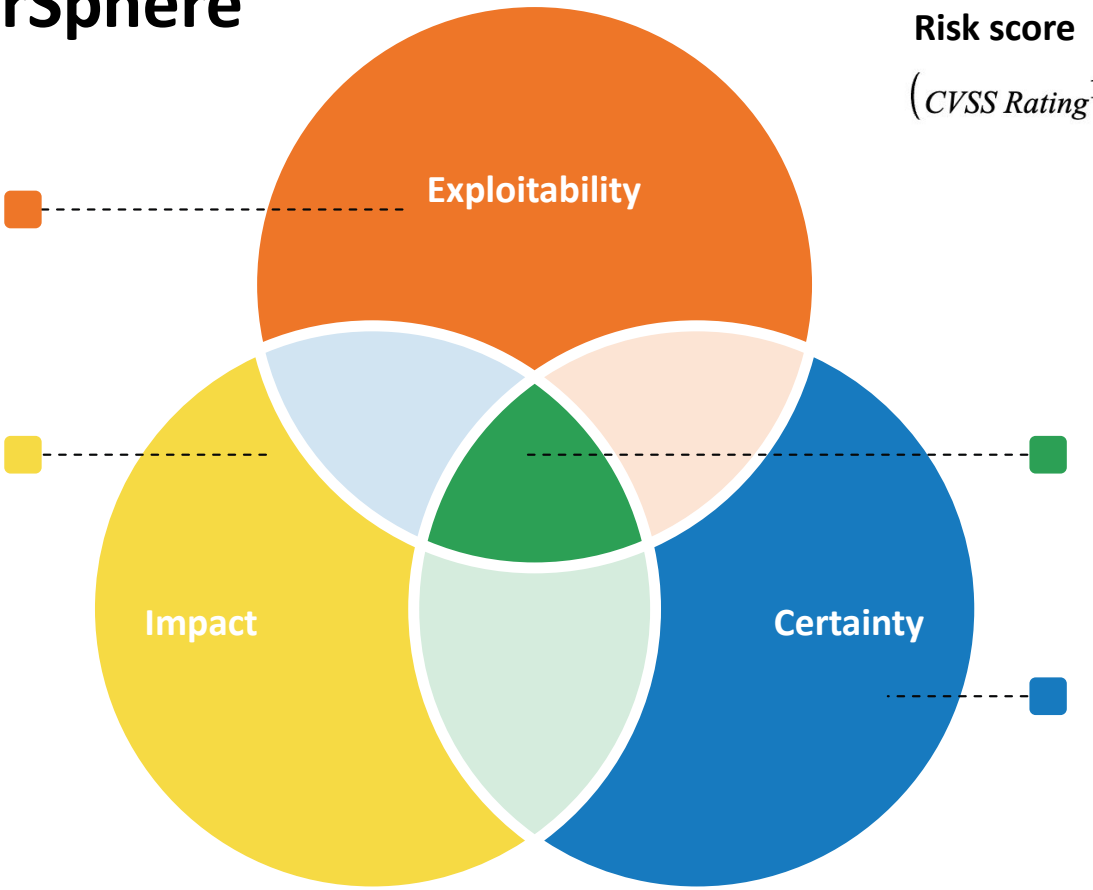
Risk in CyberSphere

Risk score

$$\left(CVSS\ Rating^{1.2} \cdot EPSS \cdot Confidence^{1.7}\right) \cdot 7$$

EPSS
How easy is a vulnerability exploitable?

CVSS
How impactful is a vulnerability?



Risk
Combination of multiple factors

Certainty
How certain are we about a vulnerability?

SURF

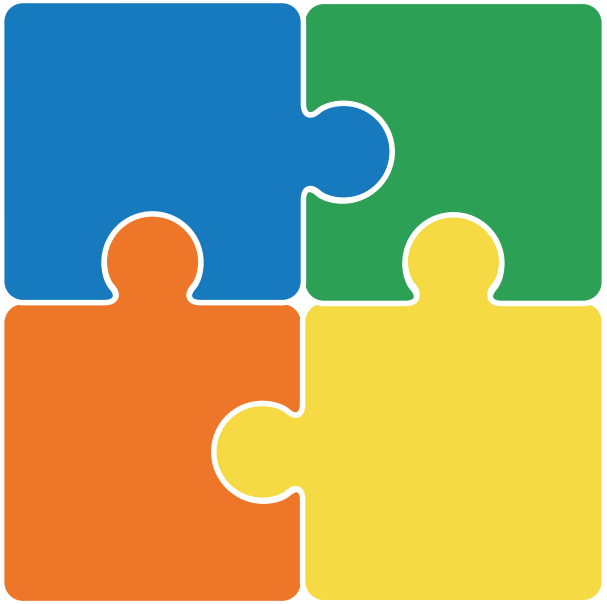
Risk in CyberSphere

Fluid

Risk scores are fluid, they can change over time, when an exploit is posted on Github for example.

Worsen

Internet facing vulnerabilities become worse when they are exposed to the internet for longer.



Host based

Traffic, service and keyword based identification of more and less important hosts.

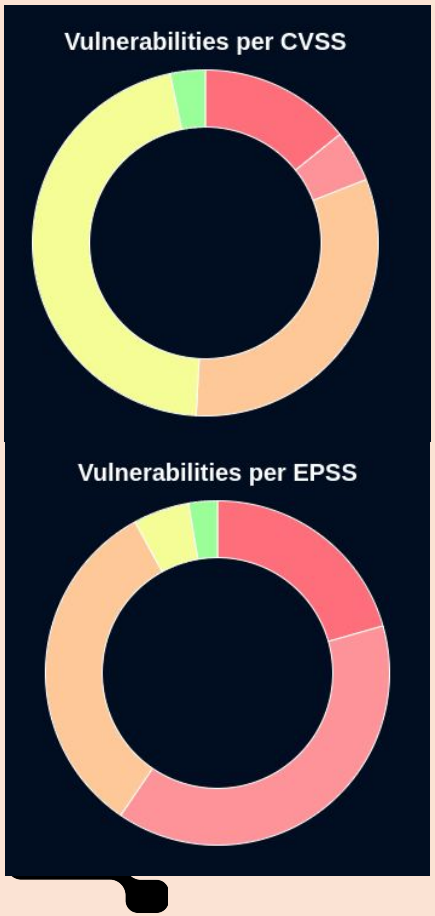
Risk over time

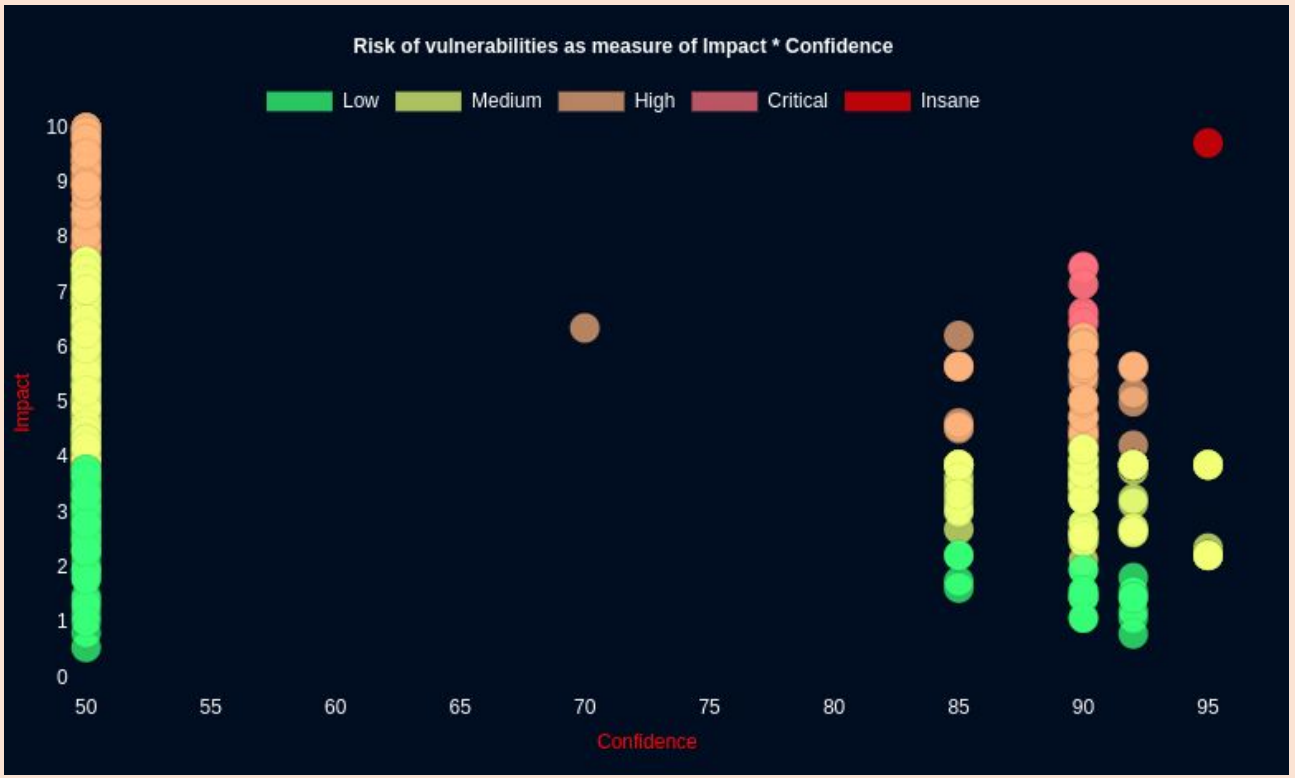
Periodic scanning shows the development of fluid risks over time, so you can track how a problem arises.

SURF

Reports

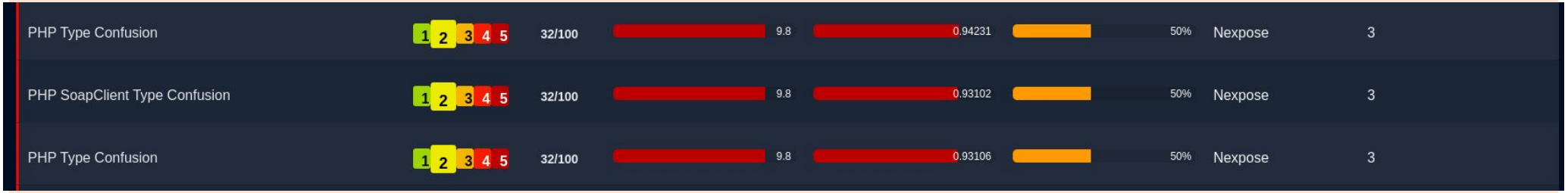
The risk spectrum



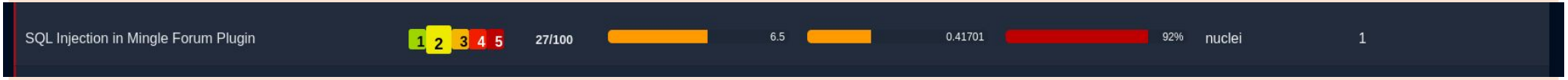


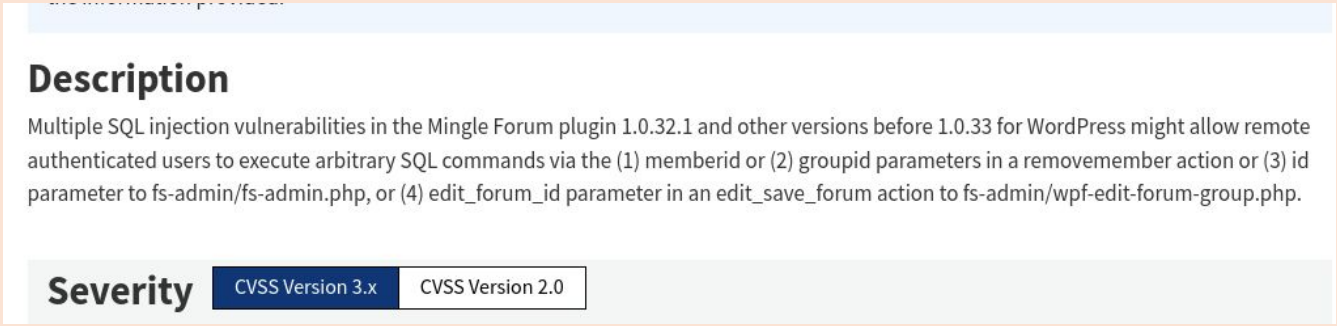
Some examples

Version based



Non exploitable





Description

Multiple SQL injection vulnerabilities in the Mingle Forum plugin 1.0.32.1 and other versions before 1.0.33 for WordPress might allow remote authenticated users to execute arbitrary SQL commands via the (1) memberid or (2) groupid parameters in a removemember action or (3) id parameter to fs-admin/fs-admin.php, or (4) edit_forum_id parameter in an edit_save_forum action to fs-admin/wpf-edit-forum-group.php.

Severity

CVSS Version 3.x | CVSS Version 2.0

SURF

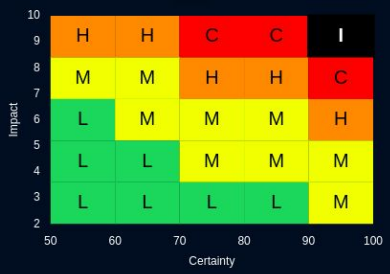
Report

Apache Struts Command Execution

Apache Struts 2.0.0 through 2.3.15 allows remote attackers to execute arbitrary OGNL expressions via a parameter with a crafted (1) action:, (2) redirect:, or (3) redirectAction: prefix.

Risk

We define Risk as Likelihood * Impact. Where Impact = Severity * Exploitability.



Risk Score	1 2 3 4 5	100/100
Certainty		95%
Severity		9.3
Exploitability		0.9993600000

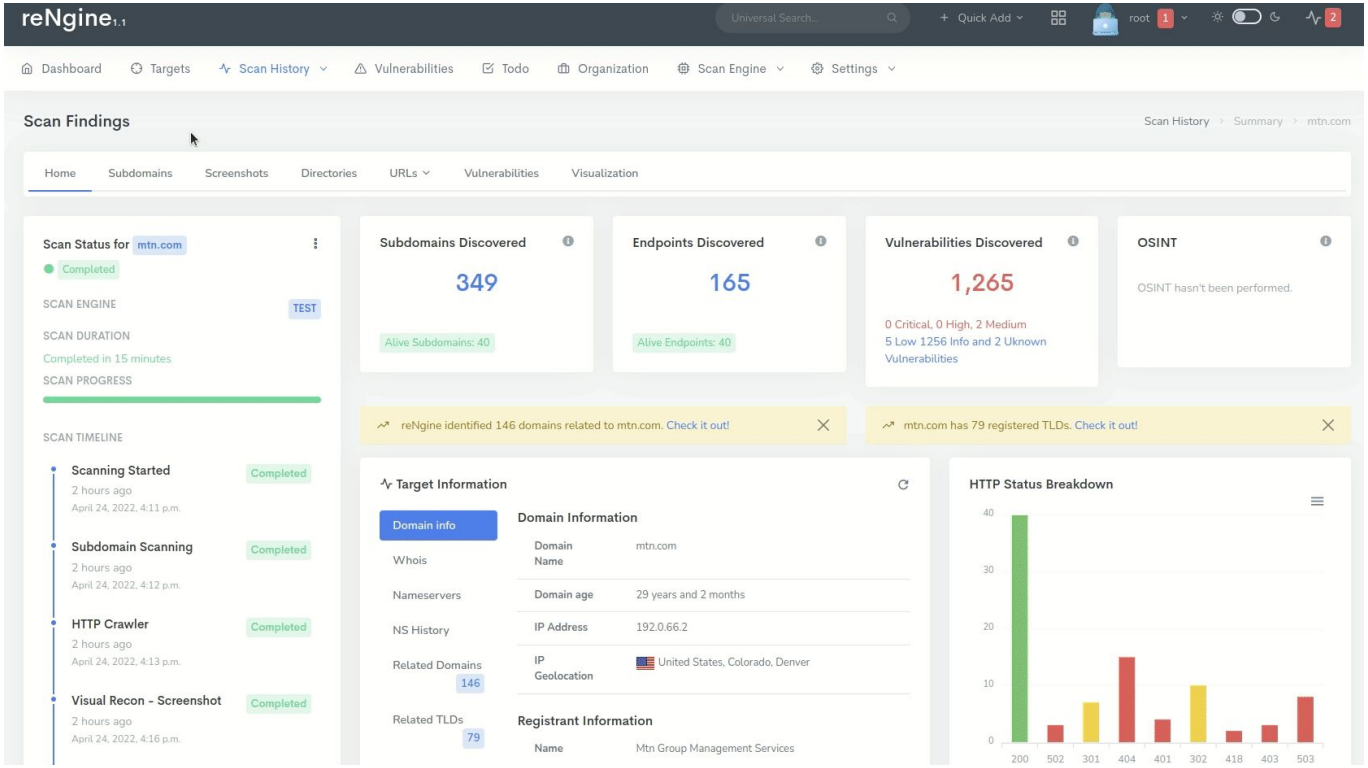
Report (2)

Proof of Concept	HT cor this Loc cor plu red 201
CVE	CVE-2013-2251
Asset info	CentOS Linux
Service info	HTTPD 2.2.15
CWE	None
Impact	Availability: COMPLETE, Confidentiality: COMPLETE, Integrity: COMPLETE
Access	Authentication: NONE, Complexity: MEDIUM, Vector: NETWORK
CVSS Vector	AV:N/AC:M/Au:N/C:C/I:C/A:C
Nexpose risk score (max 1000)	1,000
Published	2013-07-20T03:37:00
metasploit	exploit/multi/http/struts_default_action_mapper

Sensitive

Near future

Frontend



SURF

Further additions



docker

RED TEAM

';--have i been pwned?

open source

SURF

Questions?

- Any questions?
- Want to be scanned?
- Want to compare?

Feel free to contact us!
joost.grunwald@surf.nl
patrick.kuin@surf.nl

