NTK

50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Brace for impact :

Vysokofrekvenční dotazy a obranné mechanismy v NTK

Jan Kolátor

SuAleph 2025 -

28.04.2025

Typy vysokofrekvenčních dotazů za poslední rok

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Legitimní, ale neoptimalizované crawlery

- **Příklad:** AmazonBot, Googlebot, případně jiní indexační roboti.
- **Charakteristika:** Přistupují často a na velké množství URL v krátkém čase.
- **Problém:** I když nejde o škodlivé útoky, bez správné správy (robots.txt, rate limiting) mohou způsobit přetížení katalogu.

Chybně nakonfigurované interní skripty / služby

- **Příklad:** Interní monitoring procesů, testovací skripty, zálohovací nástroje.
- **Charakteristika:** Často pochází z dobrých úmyslů (např. kontrola dostupnosti), ale v důsledku chyby nebo neomezeného dotazování způsobí DDoS-like efekt.
- **Problém:** Obtížněji detekovatelné — pocházejí z důvěryhodného rozsahu IP adres huř se vysvětlují.

Záměrné škodlivé botnet útoky (např. z Číny)

- **Charakteristika:** Automatizované skripty z různých IP adres, často měnící User-Agenty, žádné dodržování pravidel (ignorace robots.txt).
- **Cíl:** Nejen NTK – útoky jsou plošné, typicky s cílem zjistit slabiny (např. otevřené porty, nezabezpečené API, pokusy o login brute force apod.).
- **Problém:** Vyžaduje aktivní obranu – WAF, rate-limiting, geo-blokace, behaviorální detekce.

Počet přístupů na VuFind server z ssl_access_log

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Počet dotazů	IP adresa
18149	216.244.66.229
18849	130.0.6723.69
20112	195.113.241.46
20918	66.249.79.8
21565	1.9.0.20
22119	1.9.1.20
26453	1.9.7.20
26491	216.244.66.202
26889	114.0.1823.43
32268	114.0.0.0
34629	1.9.2.20
35902	1.9.6.20
37592	1.9.5.20
40937	130.0.6723.116
81178	130.0.0.0
120190	10.0.12.132
147509	116.0.1938.76

bing-bot

Nejčastější boti navštěvující NTK

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Názvy:

Googlebot, Googlebot-Image, AdsBot-Google, Bingbot, AdIdxBot, BingPreview, Amazonbot, AlexaBot, Amazon, CloudFront facebookexternalhit Facebot InstagramBot Applebot AppleNewsBot YandexBot YandexImages YandexVideo DuckDuckBot Baiduspider Baiduspider-image Baiduspider-news SeznamBot Twitterbot

Příklady User-Agent stringů:

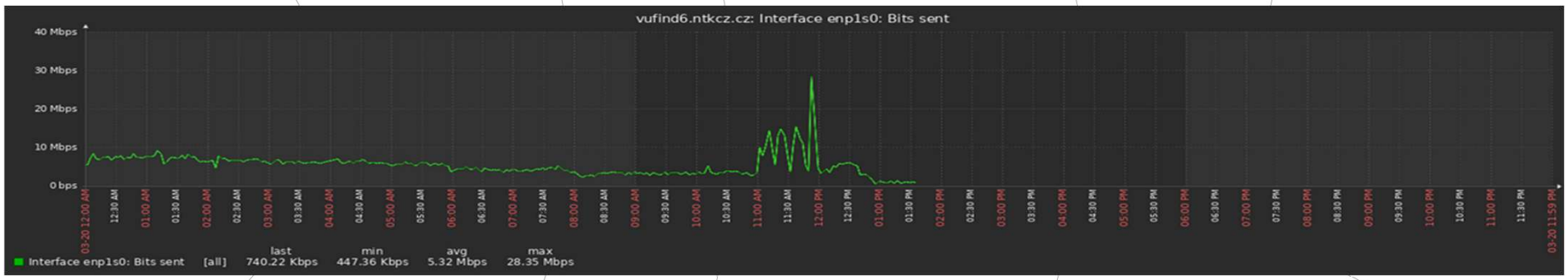
Mozilla/5.0 (compatible; Amazonbot/1.0; +https://www.amazon.com/bot.html)

facebookexternalhit/1.1 (+http://www.facebook.com/externalhit_uatext.php)

Příklad chybného volání X-serveru

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Zabbix:



Alpeh:

```
aleph@aleph24(a24_1):~/a24_1/log>grep 'ip address:' www_server_4991.log | awk '{print $3}' | sort | uniq -c | sort
 125622 10.0.12.127
  58349 127.0.0.1
    764 10.0.8.120
    322 10.0.12.241
    265 10.0.1.43
```

Úzké hrdlo Alephu - logování

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Standardní dotaz nalezen v :

2025-03-24 13:02:36 21 [000] [vrb] IN 20250324 130236

ip address: 127.0.0.1 427

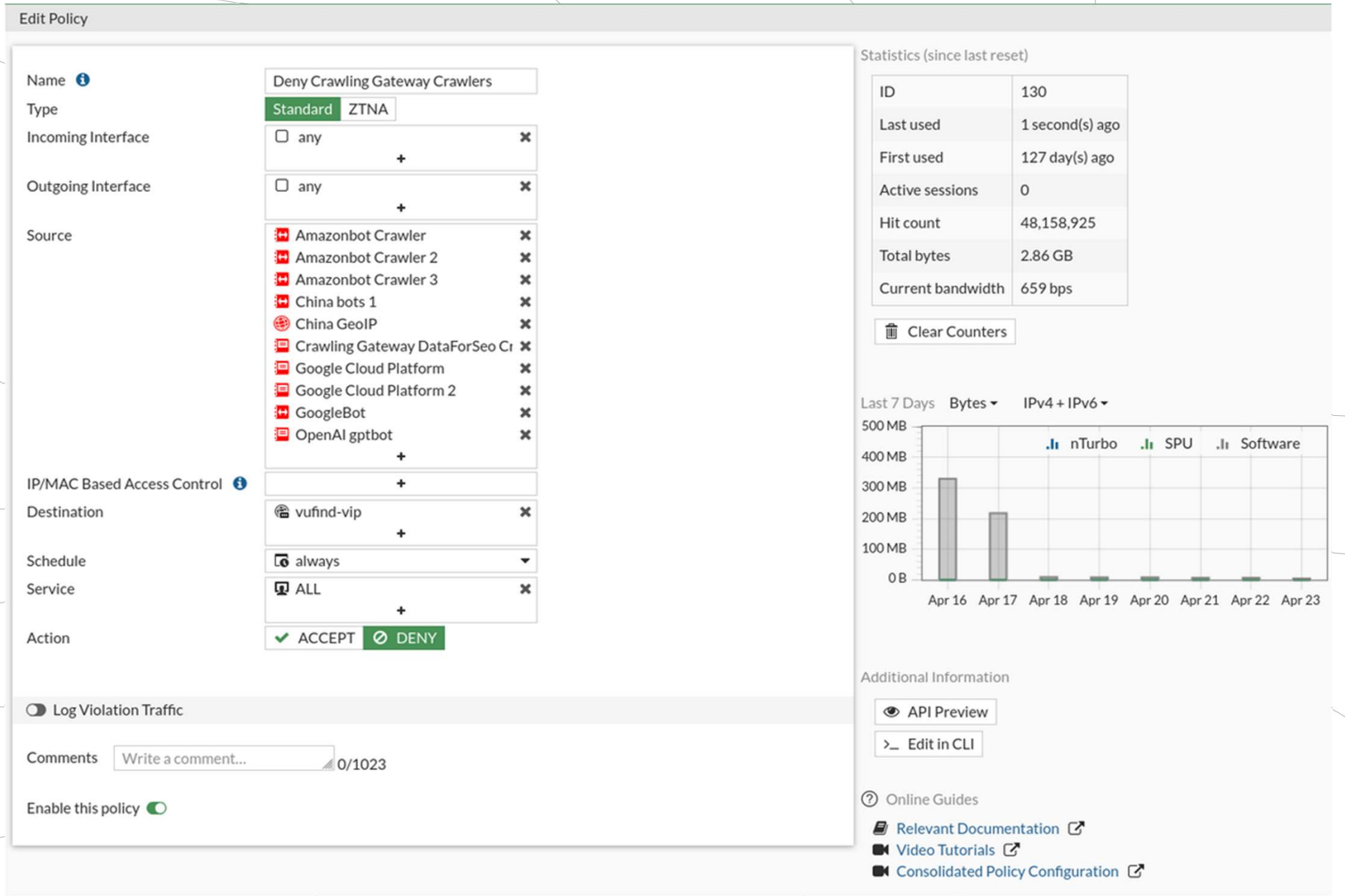
request: "/P?op=get-item-list&id=STK01000933114&institution=&loaned=&year=&volume=&sublibrary=&patron=&view=full&con_lng=cze&no_items=&start_pos=&sublibs=&collection=&holid=&rest_uri=https://aleph24.ntkcz.cz:8443/rest-dlf/record/STK01000933114/items"

DLF SERVICE: GET-ITEM-LIST

Dotaz na Aleph vir01 – z05 (začátek problému)

Datum	Hodina	Počet řádků
2025-03-24	03	873
2025-03-24	04	435
2025-03-24	05	80157
2025-03-24	06	120576

Záměrně škodlivé botnet útoky



Name	Source	Destination	Schedule	Service	Action	N...	S...	Log	Bytes	Type
any → any 121										
Deny Crawling Gateway Crawlers	Crawling Gateway DataForSeo Crawlers Amazonbot Crawler Amazonbot Crawler 2 Amazonbot Crawler 3 GoogleBot China bots 1 China GeoIP Google Cloud Platform Google Cloud Platform 2 OpenAI gptbot	vufind-vip	always	ALL	DENY			Disabled	2.86 GB	Standard

Address	GoogleBot
Type	IP Range
IP Range	66.249.64.172 - 66.249.64.172
Interface	any
References	1

Edit

Address	China GeoIP
Type	Geography
Geography	China
Interface	any
References	1

Edit

Nejčastější typy útoků

Flood útok (také známý jako **DoS - Denial of Service** útok)

TCP SYN flood: Útočník posílá požadavky na navázání TCP spojení, ale neprovádí odpověď na potvrzení, což způsobí vyčerpání serverových zdrojů.

UDP flood: Útočník posílá velké množství UDP paketů na náhodné porty cílového systému, což zvyšuje zátěž a způsobuje vyčerpání systémových zdrojů.

Nasazování nových řešení – příklady možných

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Anubis

https://github.com/TecharoHQ/**anubis**
Před připojením na web donutí tvůj stroj něco spočítat (Proof of work)

Fail2Ban

Open-source nástroj pro prevenci průniků, který monitoruje logy a blokuje IP adresy s podezřelou aktivitou.

Závěr

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Připravte se na to že vaše systémy nebudou dostupné.

Připravte se na to že můžete být hackuniti.

Pěkný den