etnetera

Linux jako ochránce před (D)DoS útoky

Petr Medonos, Anna Janáčková

O nás

Petr Medonos

- 7 let v ETN
- dohled nad přiváděním projektů k životu, konzultace
- DBA, performance, ofenzivní bezpečnost
- EWA (ewa.etnetera.cz)
- CEH, RHCE, M102, M202

@PetrMedonos

[in] PetrMedonos

Anna Janáčková

- 5 let v ETN
- nahrazení sebe sama strojem (Puppet, ... ;)
- zalohování, automatizace
- M202

@AJanackova

[in]AnnaJanackova

Obsah

- O čem to bude?
  - co je (D)DoS útok
  - tuning network stacku v Linuxu
  - netfilter
  - aplikační úroveň
  - komerční možnosti ochrany
- A o čem naopak ne
  - RTBH
  - IPS
  - škálování - anycast, loadbalancing, GeoIP distribuce

Co je to (D)DoS

Denial of Service (DoS) nebo **distributed Denial of Service (DDoS)** (česky *odmítnutí služby*) je technika útoku na internetové služby nebo stránky, při níž dochází k přehlcení požadavky a pádu nebo minimálně nefunkčnosti a nedostupnosti pro ostatní uživatele. [wikipedia]

Co je to (D)DoS

Typy (D)DoS útoků

- DDoS
  - syn flood
  - udp flood
  - http(s) flood
  - connection flood
  - ...
- DoS
  - slow*, RUDY
  - exploit
  - forkbomby, SQL wildcard, …
  - ...

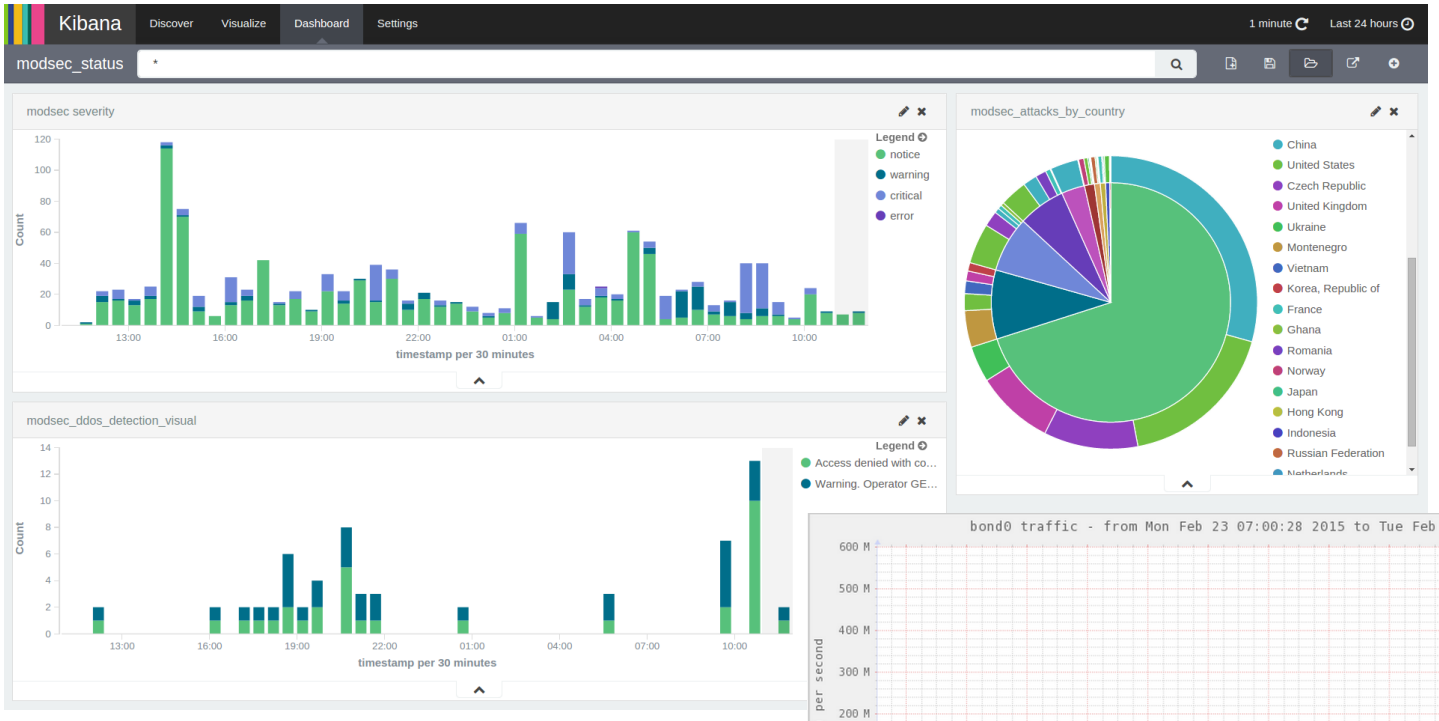
Statistika (D)DoS útoků

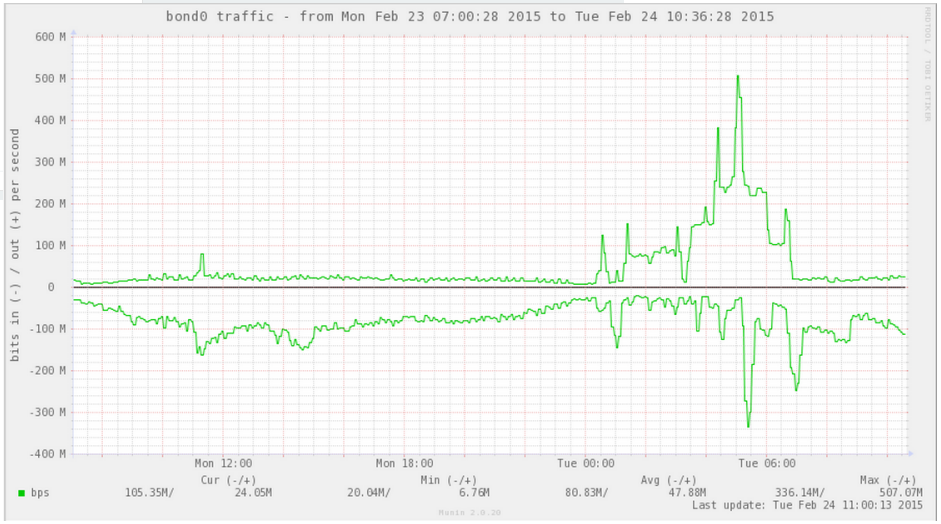
- celosvětově
  - 60% útoků do 1Gbps
  - 75% útoků se zaměřovalo na HTTP
  - 90% trvá méně než hodinu
  - 16% organizací reaguje automaticky
- největší zaznamenaný útok - 400Gbps (Cloudflare)
  - problém v EU
- největší zaznamenaný útok na IPv6 - 6Gbps

Motivace k (D)DoS útoku

- ideály
  - hacktivism
  - …
- peníze
  - vydírání
  - konkurenční boj
  - krytí primárního cíle
  - ...

Detekce (D)DoS útoku





Stack/NIC tuning

- síťové karty a kanály
- přerušení a irqbalance
- Linux stack

Stack tuning

~~net.core.rmem_max = 67108864~~

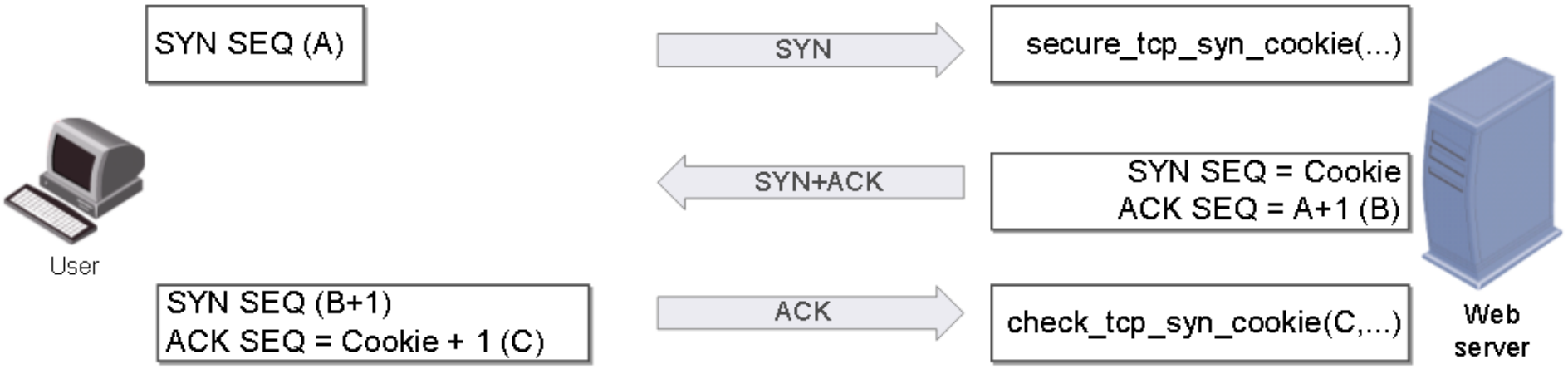
~~net.core.wmem_max = 67108864~~

~~net.ipv4.tcp_rmem = 4096 87380 33554432~~

~~net.ipv4.tcp_wmem = 4096 65536 33554432~~

```
net.ipv4.ip_local_port_range = 4096 65535
net.ipv4.tcp_tw_reuse = 1
net.ipv4.tcp_max_tw_buckets=30000
net.ipv4.tcp_fin_timeout = 10
net.ipv4.tcp_syncookies=1
```

Syncookies


SYN SEQ (A)
SYN
secure_tcp_syn_cookie(...)
SYN+ACK
SYN SEQ = Cookie
ACK SEQ = A+1 (B)
SYN SEQ (B+1)
ACK SEQ = Cookie + 1 (C)
ACK
check_tcp_syn_cookie(C,...)
User
Web server

Syncookies

- nedržíme lokální stav

```
# sysctl -w net.ipv4.tcp_syncookies=[0|1|2]
```

- 1 - aktivní, pokud je překročena velikost backlogu
- 2 - trvale zapnutí

- nedostatky
  - výkon (SHA1)
  - listen lock - 3WHS!
  - ignoruje window size

Conntrack

- ochrana proti ACK a SYN/ACK útokům

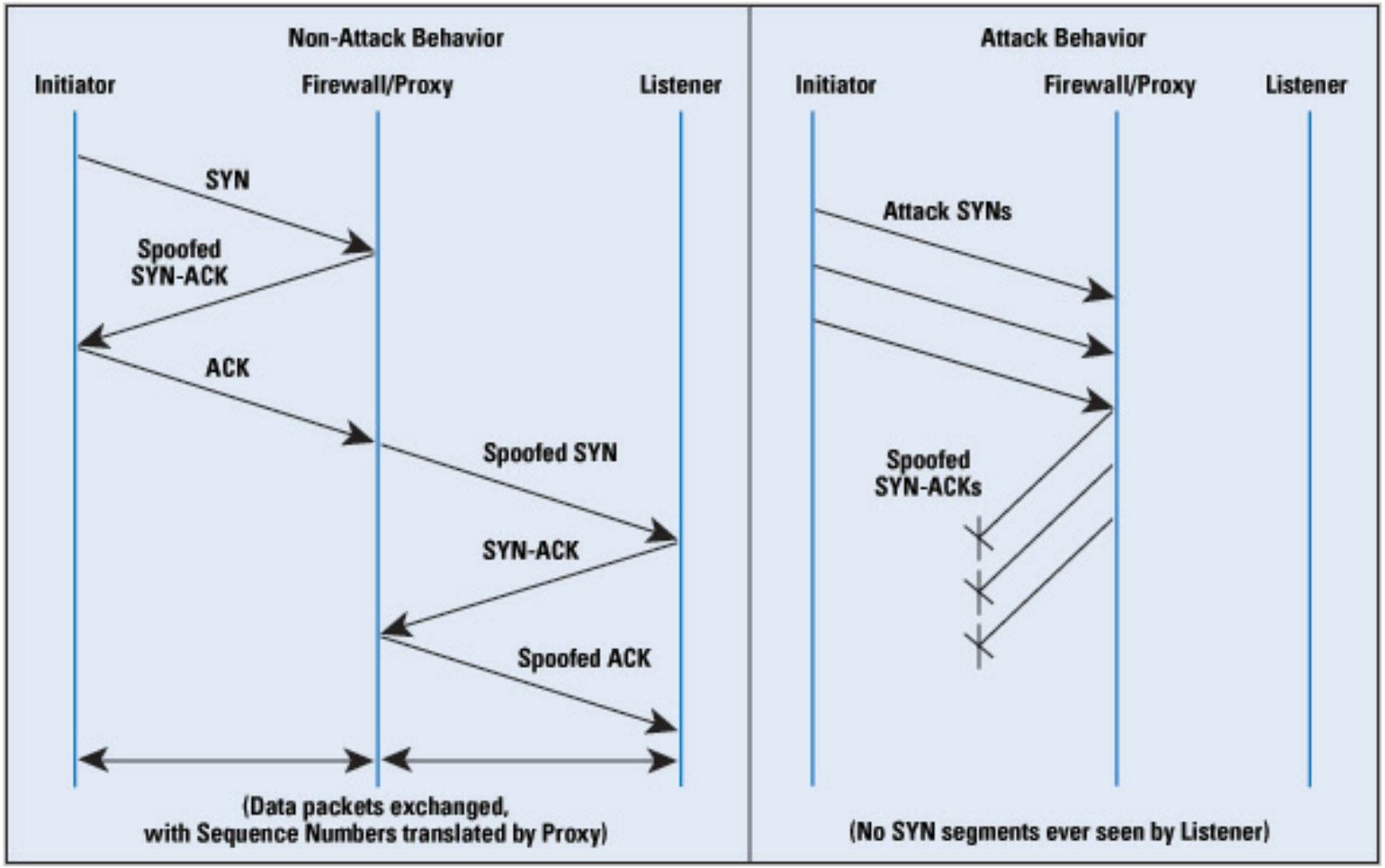
```
# sysctl -w net/netfilter/nf_conntrack_tcp_loose=0
# iptables -m state --state INVALID -j DROP
```

- `net.netfilter.nf_conntrack_max = 2000000`
- `echo 2000000 > /sys/module/nf_conntrack/parameters/hashsize`

Synproxy

- kernel 3.13
- SYN flood
- listen() lock
- conntrack
- pracuje na netrackovaných spojení

Synproxy


Non-Attack Behavior
Initiator
Firewall/Proxy
Listener
SYN
Spoofed SYN-ACK
ACK
Spoofed SYN
SYN-ACK
Spoofed ACK
(Data packets exchanged, with Sequence Numbers translated by Proxy)
Attack Behavior
Initiator
Firewall/Proxy
Listener
Attack SYNs
Spoofed SYN-ACKs
(No SYN segments ever seen by Listener)

Synproxy

```
# iptables -t raw -I PREROUTING -i eth0 -p tcp -m tcp
  -m multiport --syn --dports 80,443 -j CT --notrack

# iptables -A INPUT -i eth0 -p tcp -m tcp -m
  multiport --dports 80,443 -m state --state INVALID,
  UNTRACKED -j SYNPROXY --sack-perm --timestamp --
  wscale 7 --mss 1460

# iptables -A INPUT -m state --state INVALID -j DROP
```

Netfilter

- limitace (např. počet logů za min)

```
iptables -t filter -A INPUT -m hashlimit --hashlimit-
upto 10/min --hashlimit-mode dstip --hashlimit-name
lograte --hashlimit-htable-expire 60000 -j LOG --log-
prefix "Shorewall:INPUT:REJECT:" --log-level 6
```

- hashlimit
  - default 65535

```
--hashlimit-htable-size 2097152
--hashlimit-srcmask 24
```

Nginx zakladní nastavení

```
worker_processes = CPU * 2
worker_connections 1024;

keepalive_timeout 5;
client_body_timeout 10s;
client_header_timeout 10s;
send_timeout 10s;

sendfile on; tcp_nopush on; tcp_nodelay on;
gzip  on;
```

Nginx limitace

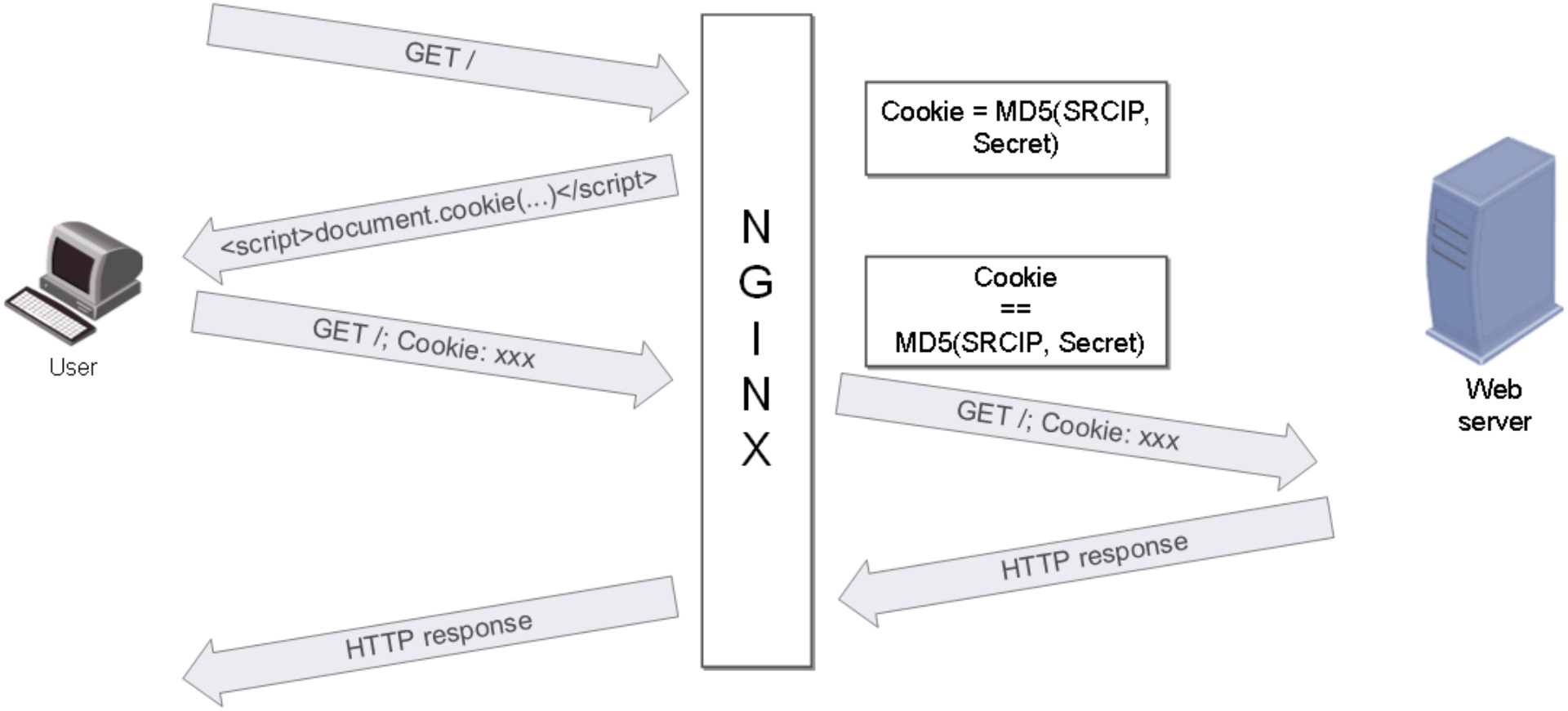
```
limit_conn_zone $binary_remote_addr
zone=conn_limit_per_ip:100m;
limit_req_zone $binary_remote_addr
zone=req_limit_per_ip:100m rate=20r/s;

limit_req_status 499; limit_conn_status 499;

limit_conn conn_limit_per_ip 80;
limit_req zone=req_limit_per_ip burst=20;
```

- SPDY (HTTP/2) - multiplexing!

Nginx test_cookie

- rozlišení bota a člověka
- nepustit požadavek, pokud v odpovědi není validní cookie
- podpora set-cookie i javascriptu
- šifrování cookie pomocí AES
- možnost zavřít spojení po nastavení cookie
- whitelisting

Nginx test_cookie



Nginx mod_security

- dlouho jenom pro Apache
  - podzim 2014 vyšel modul pro Nginx
- ratelimit + blacklist
- tradiční vs. anomaly based přístup
- inbound vs. outbound kontrola
- alternativa Naxsi

Komerční možnosti

- Cloudflare
  - 30 datacenter po celem světě (kapacita 2Tbps)
  - CDN, WAF
  - Anycast DNS/TCP
  - IPv6 gateway
  - free zakladní program
- Checkpoint (D)DoS protector
  - behavior-based engine
  - reakce do 18s
  - síťová i aplikační úroveň
  - řada technik pro minimalizaci false positive stavů
  - do 40Gbps

Zdroje

- https://www.kernel.org/doc/Documentation/sysctl/vm.txt
- http://security-portal.cz/clanky/seznamte-se-%E2%80%93-dos-ddos-%C3%BAtoky
- http://people.netfilter.org/hawk/presentations/devconf2014/iptables-ddos-mitigation_JesperBrouer.pdf
- http://www.root.cz/clanky/konkretni-ukazka-d-dos-utoku-z-pohledu-peeringoveho-uzlu/
- https://t37.net/nginx-optimization-understanding-sendfile-tcp_nodelay-and-tcp_nopush.html
- http://vincent.bernat.im/en/blog/2014-tcp-time-wait-state-linux.html
- https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Performance_Tuning_Guide/main-network.html

QA?