Linuxový firewall - nftables a iptables

Ondřej Caletka

21. října 2020

Linuxový firewall – netfilter

- možnost ovlivnit data během jejich průchodu
- záchytné body tvoří základní řetězy (`iptables`) nebo místa k zavěšení vlastních řetězů (`nftables`)

Záchytné body

`PREROUTING`	provoz těsně po příchodu ze sítě
`INPUT`	příchozí provoz určený lokálnímu procesu
`FORWARD`	příchozí provoz určený k odchodu jinam
`OUTPUT`	odchozí provoz z lokálních procesů
`POSTROUTING`	provoz těsně před odchodem do sítě

RIPE NCC cesnet

Tabulky, řetězy, pravidla

tabulky kontejnery pro řetězy – v `iptables` pevně dané: `filter`, `nat`, `mangle`

řetězy kontejnery pro pravidla – v `iptables` vestavěné a vlastní, v `nftables` jen vlastní, typu `filter`, `nat`, nebo `route`

pravidla procházejí se postupně, v `iptables` má každé počítadlo a nejvýše jeden cíl

iptables(-legacy)

- tradiční rozhraní netfiltru
- rigidní struktura tabulek a řetězců
- samostatná utilita pro každý protokol – `ip`, `ip6`, `arp`, `eb`
- rozšíření pomocí jaderných modulů a userspace knihoven

nftables

- moderní rozhraní netfilteru
- flexibilní pravidla, prázdný výchozí stav
- implementace v jádře používající virtuální stroj
- rozšiřitelnost čistě pomocí userspace
- nástroje `iptables-nft` pro snadný přechod

Kontrola stavu firewallu

```
# iptables --list -vn
# iptables -t nat --list -vn
# iptables-save
# nft -a list ruleset
```

Vyčištění firewallu

```
# iptables --policy INPUT ACCEPT
# iptables --flush
# iptables --delete-chain
# nft flush ruleset
```

RIPE NCC cesnet

Vytváření pravidel

```
# iptables --add INPUT [matchers] [--jump <target>]
```

Cíle (a verdikty) (část)

`ACCEPT`	přijmi
`DROP`	zahoď
`REJECT`	odmítni
`<jméno>`	skoč do příslušného řetězu
`RETURN`	návrat z předchozího řetězu
`LOG`	ulož do logu

RIPE NCC cesnet

Základní matchery v iptables

- `-i` vstupní rozhraní
- `-o` výstupní rozhraní
- `-s` zdrojová IP adresa
- `-d` cílová IP adresa
- `-p` protokol transportní vrstvy
- `-m` rozšiřující modul
- `--sport` zdrojový port (pro tcp, udp)
- `--dport` cílový port (pro tcp, udp)

RIPE NCC cesnet

Connection tracking

- Linux sleduje všechna procházející TCP, UDP a ICMP spojení
- data lze použít pro jednoduchý stavový firewall nebo NAT
- sledovaná spojení lze vyčíst z `/proc/net/nf_conntrack`
- lze vypnout pro určitý provoz a ušetřit prostředky

Možné stavy spojení

`NEW`	paket zahajuje nové spojení
`ESTABLISHED`	paket patří ke známému spojení
`RELATED`	paket se vztahem k existujícímu spojení
`INVALID`	o stavu nejsou informace a nejde o nové spojení
`UNTRACKED`	sledování stavu bylo vypnuto

Jednoduchý firewall s iptables

```
*filter
-A INPUT -i lo -j ACCEPT
-A INPUT -m conntrack --ctstate ESTABLISHED,RELATED -j ACCEPT
-A INPUT -p icmp -4 -m limit --limit 10/s -j ACCEPT
-A INPUT -p ipv6-icmp -6 -m limit --limit 10/s -j ACCEPT
-A INPUT -p tcp --dport 22 -j ACCEPT
-A INPUT -p udp --dport 33434:33499 -j REJECT
-P INPUT DROP
COMMIT
```

Pozn.: Nutno načíst dvakrát, pomocí `iptables-restore` a `ip6tables-restore`

RIPE NCC cesnet

Jednoduchý firewall s nftables

```
table inet firewall {
  chain input {
    type filter hook input priority 0; policy drop;
    iifname "lo" accept
    ct state established,related accept
    meta l4proto icmp meta nfproto ipv4 \
                      limit rate 10/second accept
    meta l4proto ipv6-icmp meta nfproto ipv6 \
                           limit rate 10/second accept
    tcp dport { ssh, https } accept
    udp dport { 33434-33499 } reject
  }
}
```

Priority v nftables

- určují pořadí průchodu řetězu na daném záchytném bodě
- číslo v rozsahu -300 - 100 s případným symbolickým názvem
- u iptables byla priorita určena tabulkou (raw, mangle, filter)

Priority

raw	-300
mangle	-150
dstnat	-100 (pouze prerouting)
filter	0
security	50
srcnat	100 (pouze postrouting)

Ovládání utility nft

```
# nft add table inet myT
# nft -a list table inet myT
# nft add chain inet myT myC { type filter hook input \
                                        priority 0 \; }
# nft chain inet myT myC { policy drop \; }
# nft add rule inet myT myC tcp dport { ssh, https } \
                                              accept
# nft delete rule inet myT myC handle 5
# nft add rule inet myT myC position 5 ct state \
                                invalid counter drop
```

Interoperabilita iptables-nftables

- **nedoporučuje se** míchání tradičních iptables (`x_tables`) a nft (`nf_tables`)
- migrace pouze na nativní nftables je dlouhodobý proces
- přechod usnadňují nástroje z rodiny `iptables-nft`:
  - obsluhují se stejně jako iptables
  - používají nové rozhraní `nf_tables`
  - umožňují volat původní rozšíření napsaná pro iptables

RIPE NCC cesnet

Nativní maškaráda

```
table inet msqT {
  chain msqC {
    type nat hook postrouting priority srcnat; policy accept;
    oif "wlan0" masquerade
  }
}
```

Maškaráda z iptables

```
table ip nat {
  chain POSTROUTING {
    type nat hook postrouting priority srcnat; policy accept;
    oifname "wlan0" counter packets 1 bytes 199 # xt\_MASQUERADE
  }
}
```

Další čtení

- `https://wiki.nftables.org/`
- Root.cz: firewall s nftables

RIPE NCC cesnet

Děkuji za pozornost

Ondřej Caletka
Ondrej.Caletka@cesnet.cz
https://**Ondřej.Caletka.cz**

Prezentace je již nyní k dispozici ke stažení.