netfilter/iptables FAQ

Harald Welte, `laforge@gnumonks.org`

Wersja oryginalna: 1.30, 2002/01/14 09:04:47

Oryginał tego dokumentu znajduje się pod adresem: http://netfilter.samba.org/

Tłumaczenie: Łukasz Bromirski, `l.bromirski@mr0vka.eu.org`

Wersja tłumaczenia: 1.1, $Date: 2002/08/22 21:29:41 $

Oryginał tego tłumaczenia znajduje się pod adresem: http://mr0vka.eu.org/tlumaczenia/netfilter-faq.html

---

Dokument ten zawiera Odpowiedzi Na Najczęściej Zadawane Pytania, które do tej pory pojawiły się na liście pocztowej netfilter. Komentarze / uzupełnienia / wyjaśnienia są oczywiście mile widziane, i powinny być kierowane do opiekuna tego FAQ.

---

1. Pytania ogólne

Sekcja ta odpowiada na pytania ogólnie związane z netfilter, które pojawiają się bardzo często.

1.1 Skąd mogę ściągnąć netfilter/iptables?

Netfilter i IPtables są zintegrowane w jądrach linuksa serii 2.4.x. Proszę ściągnij źródła ostatniego kernela z http://www.kernel.org/ lub jednego z jego mirrorów.

Narzędzie działające w przestrzeni użytkownika 'iptables' jest dostępne na stronie domowej netfilter, lub na jednym z mirrorów: http://www.netfilter.org/, http://www.iptables.org/, http://netfilter.samba.org/, http://netfilter.gnumonks.org/ lub http://netfilter.filewatcher.org/.

1.2 Czy jest jakaś wersja netfilter dla Linuksa 2.2?

Nie, aktualnie nie ma. Ale, jeśli ktoś chciałby zacząć to pisać, nie powinno być to zbyt trudne z uwagi na jasny interfejs do stosu sieciowego.

Proszę, poinformujcie nas o jakiejś pracy w tym zakresie.

1.3 Czy jest moduł do śledzenia połączeń i NATu dla ICQ?

Jeśli jesteś przyzwyczajony do maskarady z Linuksa 2.2, używałeś zawsze modułu ip_masq_icq module by uzyskać bezpośrednie połączenie klient-klint dla ICQ.

Nikt nie przepisał tego modułu, ponieważ protokół ICQ jest zbyt brzydki :) Ale, jak sądzę jest to tylko kwestia czasu.

Rusty kiedyś stwierdził, że tylko protokoły z przynajmniej jednym darmowym klientem i jednym darmowym serwerem będą integrowane do głównej dystrybucji netfilter. Jeśli chodzi o ICQ, są tylko darmowe klienty, więc nie pasuje on do tych kryteriów (wolny jak wolność, nie jak wolne piwo, wg. definicji RMS)

1.4 Gdzie podziały się moduły ip_masq_vdolive / ip_masq_quake / ... ?

Niektóre z nich nie są już potrzebne, a niektóre nie zostały jeszcze przeniesione do netfilter. Netfilter zapewnia aktualnie

pełne śledzenie połączeń, nawet dla UDP, i generalnie będzie starał się zmieniać pakiety tak niewiele jak to tylko będzie możliwe, więc czasami programy "po prostu" pracują.

1.5 Co to jest ten patch-o-matic i jak mogę go używać?

Kernele 2.4.x są wersją stabilną, więc nie możemy tak po prostu dołączań swoje aktualne poprawki do głównej dystrybucji. Cały nasz kod jest tworzony i testowany w netfilter patch-o-matic. Jeśli chcesz używać najnowszych funkcji netfilter, będziesz musiał użyć jednego z patchy z patch-o-matic. Możesz znaleźć program patch-o-matic w najnowszej paczce iptables (lub, oczywiście na CVS), którą ściągnąć można ze strony WWW netfilter.

patch-o-matic ma obecnie trzy dostępne opcje:

- make pending-patches (*nałóż zaległe łaty*)
- make most-of-pom (*nałóż większość z opcji*)
- make patch-o-matic

Pierwsza opcja ma na celu dołożenie wszystkich istotnych poprawek (które i tak zostały wysłane koordynatorom kernela) na twój kernel. Druga, `most-of-pom` dodaje dodatkowo wszystkie nowe opcje, które mogą zostać dodane bez konfliktów z już istniejącymi. Trzecia opcja jest dla prawdziwych ekspertów, którzy chcą przejrzeć wszystkie poprawki - ale weźcie pod uwagę, że mogą nie być między sobą zgodne.

patch-o-matic ma miły interfejs. Wprowadź po prostu

```
make most-of-pom (lub pending-patches czy patch-o-matic, zajrzyj wyżej)
```

lub, jeśli źródła twojego kernela nie są w `/usr/src/linux` użyj

```
make KERNEL_DIR={katalog-ze-źródłami} most-of-pom
```

w głównym katalogu paczki iptables. patch-o-matic sprawdza każdy patch pod kątem możliwości zaaplikowania do źródeł kernela. Jeśli uzna że można go użyć, zobaczysz znak zachęty, z poziomu którego możesz uzyskać więcej informacji o patchu, zaaplikwać go, przejść do następnego, ...

Po więcej informacji dotyczących patch-o-matic, zajrzyj na stronę rozszerzeń netfilter, która znajduje się pod adresem http://www.netfilter.org/documentation/index.html#HOWTO.

1.6 Gdzie mogę znaleźć ipnatctl i więcej informacji o nim?

ipnatctl był używany do ustawiania reguł NAT w przestrzeni użytkownika w bardzo wczesnych wersjach netfilter (jeszcze w czasie prac z kernelami serii 2.3.x). Nie jest już potrzebny, więc nie udostępniamy go. Całą jego funkcjonalność przejęło iptables. Obejrzyj NAT HOWTO na stronie domowej netfilter.

2. Problemy podczas procesu kompilacji

2.1 Nie mogę skompilować iptables-1.1.1 z kernelem >= 2.4.0-test4

To znana sprawa. Mechanizm rozpoznający które patche są już zaaplikowane źle działa. Użyj "make build" zamiast "make".

Lepsze rozwiązanie: uaktualnij system do iptables-1.1.2 lub nowszych.

2.2 Nie mogę skompilować iptables 1.1.0 z ostatnimi kernelami (>= 2.3.99-pre8)

Wewnętrzne struktury w iptables zostały zmienione. Uaktualnij iptables do wersji >= 1.1.1

2.3 Niektóre patche z patch-o-matic z iptables-1.2.1a nie działają z kernelem >= 2.4.4

Proszę użyć najnowszej wersji iptables.

2.4 ipt_BALANCE, ip_nat_ftp, ip_nat_irc, ipt_SAME, ipt_NETMAP nie chcą się skompilować

Najprawdopodobniej masz problemy ze skompilowaniem funkcji nazwanej `ip_nat_setup_info`.

Jeśli używasz iptables <= 1.2.2, **MUSISZ** zaaplikować łaty `dropped-table` i `ftp-fixes`.

Jeśli używasz iptables > 1.2.2 lub ostatniej wersji z CVS, proszę **nie** aplikuj patcha 'dropped-table', ponieważ jest niekompatybilny z BALANCE, NETMAP, irc-nat, SAME i talk-nat.

2.5 Używam serii jądra Alana Cox'a 2.4.x-acXX i mam problemy

Podstawowy zespół netfilter bazuje na drzewie Linus'a, więc używasz wersji -ac na własne ryzyko.

3. Problemy w trakcie pracy

3.1 NAT: X dropping untracked packet Y Z aaa.aaa.aaa.aaa -> 224.bbb.bbb.bbb

Wiadomość taką generuje kod NAT, ponieważ pakiety multicastowe trafiają do tabeli NAT, a kod odpowiedzialny za śledzenie połączeń nie potrafi ich obsłużyć. Jeśli nie wiesz co to jest multicast, lub w ogóle tego nie potrzebujesz, użyj:

```
iptables -t mangle -I PREROUTING -j DROP -d 224.0.0.0/8
```

3.2 NAT: X dropping untracked packet Y Z aaa.aaa.aaa.aaa -> bbb.bbb.bbb.bbb

Mój syslog lub moja konsola pokazują taką wiadomość:

```
NAT: X dropping untracked packet Y Z aaa.aaa.aaa.aaa -> bbb.bbb.bbb.bbb
```

Wiadomość taką generuje kod NAT. Odrzuca pakiety, ponieważ by wykonać na nich NAT musi posiadać prawidłową informację o połączeniu. Wiadomość taka jest drukowana dla wszystkich pakietów, których kod śledzący połączenia nie był w stanie zidentyfikować.

Możliwe przyczyny to:

- osiągnięto maksymalny limit wpisów w bazie danych śledzenia połączeń
- nie można było zidentyfikować typu rozgłaszania (multicast, broadcast)
- sypie się kmem_cache_alloc (brak pamięci)
- odpowiedź na niepotwierdzone połączenie
- pakiet multicastowy (sprawdź poprzednie pytanie)
- pakiet icmp zbyt krótki
- pakiet icmp jest sfragmentowany
- zła suma kontrolna pakietu icmp

Jeśli chciałbyś prowadzić dużo bardziej szczegółowe logowanie tych pakietów (np. podejrzewasz że jest to zdalna próba skanowania), użyj następującej reguły:

```
iptables -t mangle -A PREROUTING -j LOG -m state --state INVALID
```

I tak, musisz wstawić tą regułę do tabeli przekształceń (ang. *mangle*), ponieważ pakiety będą odrzucane przez NAT zanim osiągną tabelę filtrowania.

3.3 Nie mogę używać netfilter z kodem mostującym Linuksa

Chcesz stworzyć kompletnie transparentną ścianę ogniową? Wspaniały pomysł!

Począwszy od wersji kernela 2.4.16 musisz dodać tylko łatkę do kernela, dostępną pod adresem http://bridge.sourceforge.net/.

3.4 Moduł IRC nie potrafi obsłużyć DCC RESUME

Cóż, to połowa prawdy. Tylko tabela NAT nie potrafi go obsłużyć. Jeśli używasz filtrowania pakietów bez NAT wszystko powinno działać.

3.5 Jak działa SNAT dla wielu adresów?

Netfilter stara się modyfikować pakiety minimalnie, na tyle ile to możliwe. Więc jeśli mamy maszynę po restarcie, i ktoś za komputerem robiącym SNAT otwiera połączenie do lokalnego portu 1234, netfilter zmienia jedynie adres IP a port zostaje ten sam.

Ale jeśli ktoś inny otworzy połączenie z tym samym portem źródłowym, netfilter musi zmienić i IP i port, jeśli ma tylko jedno IP dla SNAT.

Jeśli są dostępne inne, **ponownie** zmienia tylko adres IP.

3.6 ip_conntrack: maximum limit of XXX entries exceeded

Jeśli zobaczysz powyższy komunikat w syslog, wygląda na to że baza danych śledzenia połączeń nie ma wystarczająco dużo miejsca w twoim otoczeniu. Śledzenie połączeń domyślnie obsługuje tylko pewną liczbę jednoczesnych połączeń. Liczba ta jest zależna od ilości pamięci operacyjnej (przy 64MB: 4096, 128MB: 8192, ...).

Możesz łatwo zwiększyć numer maksymalnie śledzonych połączeń, ale zwróć uwagę że każde połączenie zabiera około 350 bajtów pamięci nie wymienialnej!

By zwiększyć limit to dnp. 8192, napisz:

```
echo "8192" > /proc/sys/net/ipv4/ip_conntrack_max
```

3.7 Jak mogę wylistować wszystkie śledzone / maskaradowane połączenia, tak jak przy użyciu 'ipchains -L -M' w 2.2.x ?

Istnieje plik w systemie plików /proc, nazywa się `/proc/net/ip_conntrack`. Możesz wydrukować go na wyjście używając komendy

```
cat /proc/net/ip_conntrack
```

3.8 Jak wylistować wszystkie dostępne tabele IP?

Wszystkie dostępne tabele IP można wylistować przez

```
cat /proc/net/ip_tables_names
```

3.9 iptables-save / iptables-restore z iptables-1.2 powodują segfault

Znany błąd. Uaktualnij się do najnowszej wersji z CVS lub używaj iptables >= 1.2.1 jak tylko będzie dostępne.

3.10 Komenda iptables -L bardzo długo listuje reguły

Dzieję się tak ponieważ iptables sprawdza DNS przy każdym adresie IP. Ponieważ każda reguła składa się z jednego lub dwóch adresów, w najgorszym przypadku mamy dwa sprawdzenia na regułę.

Problem zaczyna się, gdy używasz prywatnych adresów IP (jak np. 10.x.x.x czy 192.168.x.x), a DNS nie jest w stanie ich rozwinąć i w końcu się poddaje. Suma tych wszystkich time-out'ów może trwać _bardzo_ długo, wszystko zależy od twojego zestawu reguł.

Użyj opcji -n (numerycznie) dla iptables, by zapobiec sprawdzaniu DNSu.

3.11 Jak mogę powstrzymać cel LOG przed logowaniem na konsolę?

Musisz prawidłowo skonfigurować swój syslogd: Cel LOG loguje do facility kern z priorytetem ostrzeżenia (ang. *warning*, 4). Zajrzyj do podręcznika syslogd.conf by dowiedziec się więcej.

Domyślnie, wszystkie wiadomości kernela z priorytetem wyższym niż debug (7) wysyłane są na konsolę. Jeśli podniesiesz go do 4, zamiast 7, zapobiegniesz wysyłaniu informacji z LOG na konsole.

Zwróć jednak uwagę, że powstrzyma to również inne, istotne wiadomości od pojawiania się na konsoli (nie wpływa to na syslog).

3.12 Jak zbudować transparentne proxy przy użyciu squid'a i iptables?

Po pierwsze, potrzebujesz właściwej reguły DNAT lub REDIRECT. Użyj REDIRECT tylko jeśli squid pracuje na tej samej maszynie. Na przykład:

```
iptables -t nat -A PREROUTING -p tcp --dport 80 -j DNAT --to 192.168.22.33:3128
```

Następnie, musisz właściwie skonfigurować squid'a. Możemy tu tylko dać krótkie wskazówki, zajrzyj do dokumentacji squid'a po dalsze informacje.

Plik squid.conf dla Squid'a 2.3 potrzebuje następujących linijek:

```
http_port 3128
httpd_accel_host virtual
httpd_accel_port 80
httpd_accel_with_proxy  on
httpd_accel_uses_host_header on
```

Squid 2.4 potrzebuje **dodatkwej** linijki:

```
httpd_accel_single_host off
```

3.13 Jak mogę używać celu LOG / Jak i LOGować i DROPować?

Cel LOG jest tym co nazywamy "nie-ostatecznym celem", tzn. nie kończy on sprawdzania pakietu. Jeśli używasz celu LOG, pakiet zostanie zalogowany, ale przeglądanie reguł odbywa się dalej.

Więc jak możesz zalogować i odrzucić pakiet jednocześnie? Nic prostszego, tworzysz osobny łańcuch który zawiera dwie reguły:

```
iptables -N logdrop
```

```
iptables -A logdrop -j LOG
iptables -A logdrop -j DROP
```

Teraz za każdym razem gdy chcesz zalogować i odrzucić pakiet, używasz celu "-j logdrop".

3.14 kernel loguje: Out of window data xxx

Możesz użyć patcha tcp-window-tracking patch z patch-o-matic, którego kod śledzi akceptowalne parametry dla przepuszczanych strumieni TCP zgodnie z numerami sekwencyjnymi/potwierdzeń, rozmiarami segmentów, itd. Kiedy wykryje że pakiet jest nie do zaakceptowania (poza oknem), zaznacza go jako INVALID i wypisuje wiadomość jak powyżej.

Nowsze wersje logują pakiet i dokładnie co spowodowało taką akcję:

- ACK is under the lower bound (prawdopodobnie zbyt opóźniony ACK)
- ACK is over the upper bound (dane potwierdzane jeszcze nie dotarły)
- SEQ is under the lower bound (retransmitujemy już potwierdzone dane)
- SEQ is over the upper bound (ponad oknem odbiorczym)

Również w nowszych wersjach logowanie można całkowicie powstrzymać przez sysctl

```
echo 0 > /proc/sys/net/ipv4/netfilter/ip_ct_tcp_log_out_of_window
```

3.15 Dlaczego system śledzenia połączeń śledzi połączenia w stanieUNREPLIED (*nieodpowiedziane*) z dużą wartością wygasania (timeout)?

Więc sprawdzałeś /proc/net/ip_conntrack i znalazłeś bardzo wysokie wpisy połączeń UNREPLIED (które nie są oczywiście połączeniami)?

Odpowiedź jest prosta: wpisy UNREPLIED są tymczasowe, tzn. jak tylko skończy się pula połączeń (osiągnięta zostanie granica /proc/sys/net/ipv4/ip_conntrack_max), kasujemy stare wpisy UNREPLIED. Innymi słowy, staramy się raczej by inne, prawidłowe i całe połączenia mogły dochodzić do skutku.

4. Pytania dotyczące programowania netfilter

4.1 Nie rozumiem jak używać celu QUEUE z przestrzeni użytkownika

Dostępna jest biblioteka nazwana libipq, która służy do obsługi pakietów w przestrzeni użytkownika. Jest też dokumentacja dla niej w postaci stron podręcznikowych. Musisz zbudować i zainstalować komponenty programistyczne iptables:

```
make install-devel
```

a następnie sprawdź libipq(3).

Może cię również zainteresować Perlipq zawierająca powiązania dla Perla do libipq: http://www.intercode.com.au/jmorris/perlipq/. Same powiązania są przykładem jak używać biblioteki.

Inne przykłady kodu obejmują:

- testsuite/tools/intercept.c z CVS netfilter
- ipqmpd (sprawdź http://www.gnumonks.org/projects/)
- nfqtest, część netfilter-tools (sprawdź http://www.gnumonks.org/projects/)
- symulator WAN Jerome Etienne'go (sprawdź http://www.off.net/~jme/)

4.2 Chciałbym dodać trochę kodu, ale nie mam pojęcia co zrobić

Główny zespół netfilter utrzymuje listę TODO (do zrobienia), gdzie wpisuje wszystkie najbardziej pożądane zmiany / nowe możliwości. Możesz ściągnąć tą listę przez anonimowy CVS, instrukcje znajdują się na stronie domowej netfilter. Możesz również udać się pod adres http://cvs.samba.org/cgi-bin/cvsweb/netfilter/TODO/ i obejrzeć plik przy użyciu CVSweb.

4.3 Poprawiłem błąd lub napisałem rozszerzenie. Jak je dodać?

Jeśli chcesz je opublikować, wyślij pocztę pod adres listy netfilter-devel. Informacje o prenumeracie dostępne są pod adresem http://lists.samba.org/mailman/listinfo/netfilter-devel/.

Prawidłowy sposób na wysłanie patcha jest następujący :

- Temat rozpoczyna się od **[PATCH]**
- Kod wklejony w wiadomości, nie w postaci załącznika MIME.
- wiadomość z diff'a z wpisem z cvs-checkin/Changelog.
- wyjście z `diff -u old new', spoza katalogu głównego (np. może być z dodanym -p1 w katalogu z rozpakowaną zawartością.

Jeśli napisałeś nowe rozszerzenie, lub dodałeś coś do starych funkcji, dobrze jest również uzupełnić dokument netfilter-extensions HOWTO i zawrzeć w nim informacje o nowej funkcjonalności i zasadach działania rozszerzenia. Dodatkowo, zainteresuje to z pewnością większą liczbę użytkowników i zapewni więcej komentarzy.