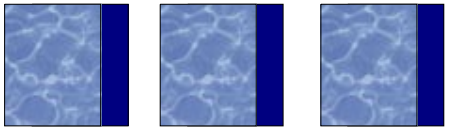
Sieb im Netz
Das Netfilter Framework

11.03.2001 – 3. Chemnitzer Linux–Tag

Agenda

- Thnx an Harald Welte <laforge@gnumonks.org>

- Netfilter basics / concepts
- IP−Paketfilterung mit iptables und Netfilter
- NAT
- Packet mangling
- Advanced netfilter concepts



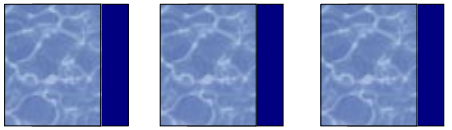
Was ist das Netfilter Framework?

- mehr als ein Firewall Subsystem
- einheitliches Framework (unabhängig vom Protokoll)
- Hooks im Netzwerk−Stack
- Kernel−Module können sich an Hooks registrieren
- Asynchrone Verarbeitung von Paketen in UserSpace
- IP Tables für alle Module nutzbar

- Traditionelle Packet−Filterung / NAT (Masquerading) / ...
  wurde auf Basis des Frameworks implementiert

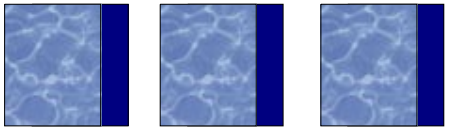
Wieso brauchen wir das neue Netfilter Framework?

- keine Möglichkeit Pakete in den UserSpace zu leiten
- „transparent proxying“ extrem schwierig
- Filterregeln abhängig von Interface−Adressen
- Masquerading und Paketfilterung bisher nicht getrennt implementiert
- Code ist zu komplex
- weder modular noch erweiterbar

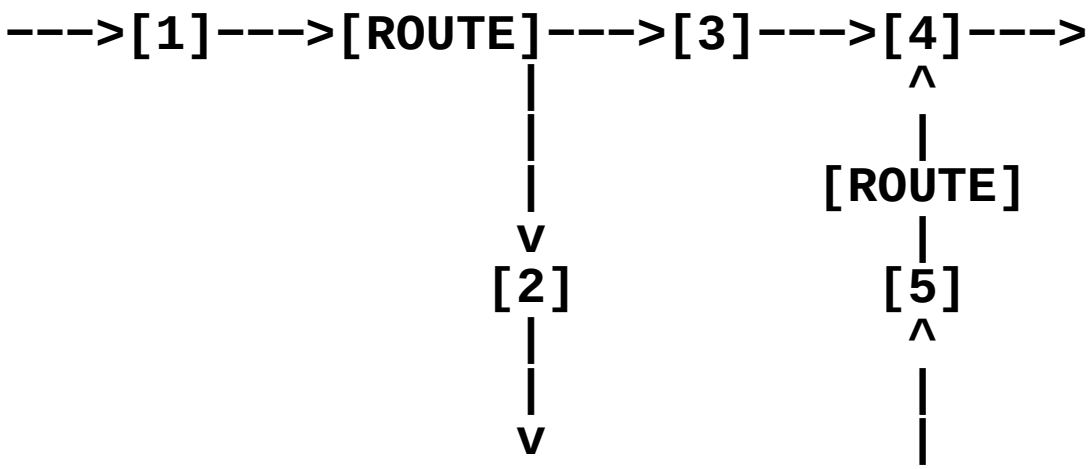


Die Autoren

- Paul „Rusty“ Russel
  - Co−Autor von iptables in Linux 2.2
  - 1 Jahr durch Watchguard bezahlt; jetzt Linuxcare
- James Morris
  - UserSpace−Queuing & REJECT Target
- Marc Boucher
  - NAT, Paketfilterung (iptables), Mangle Table
- Harald Welte
  - IRC conntrac + NAT helper, UserSpace Logging, Ipv6
- Non−core team contributors (siehe Scoreboard:)



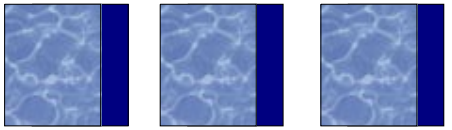
Netfilter basics
Architektur (Ipv4)



1=NF_IP_PRE_ROUTING
2=NF_IP_LOCAL_IN
3=NF_IP_FORWARD
4=NF_IP_POST_ROUTING
5=NF_IP_LOCAL_OUT

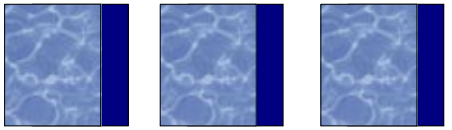
Netfilter basics
Arbeitsweise

- jedes Modul kann Callback−Funktionen an den Hooks anmelden
- mögliche Rückgabewerte:
  - NF_ACCEPT Weiterleiten (normal)
  - NF_DROP Paket verwerfen!
  - NF_STOLEN TakeOver durch Modul; „vergiss es“
  - NF_QUEUE Weiterleiten an UserSpace
  - NF_REPEAT Hook wiederholen



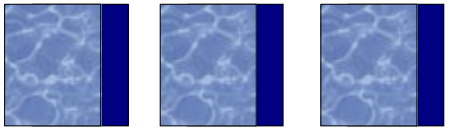
Netfilter basics

- Kernel stellt normale IP tables zur Verfügung
- Module können eigene IP tables erstellen
- Erweiterte Paket−Verarbeitung im 2.4’er
  - Tabelle zur Paketfilterung ’filter’
  - NAT−Tabelle ’nat’
  - ’mangle’ Tabelle



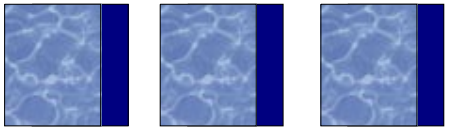
IP–Paketfilterung

- Implementiert über drei Hooks
  - NF_IP_LOCAL_IN
    Pakete, die den lokalen Rechner erreichen sollen
  - NF_IP_FORWARD
    Pakete, die durch den Rechner weitergeleitet werden
  - NF_IP_LOCAL_OUT
    Pakete vom lokalen Rechner selbst
- an diese Hooks werden Chains (Input, Output und Forward) der Tabelle ’filter’ registriert
- Achtung: jedes Paket passiert nur einen dieser Chains!



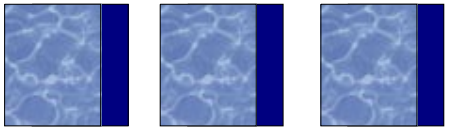
Chains und Tables Mgmt.

- jede Regel in einer Chain besteht aus
  - match welches Paket trifft auf die Regel
  - target was ist dann zu tun?
- matches und targets können Kernel−Build−In oder via Module realisiert sein
- UserSpace−Tool ’iptables’ sehr flexibel
  - alle unterschiedlichen Arten von IP tables
  - unterstützt plugin/shlib Interface für target/match spezifische Optionen



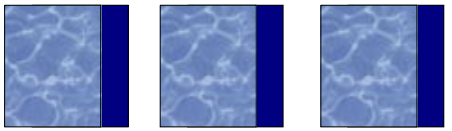
Einfache iptables cmds

- ein kompletter Befehl besteht aus
  - mit welcher Tabelle wird gearbeitet
  - welche Chain der Tabelle soll genutzt werden
  - eine Operation (insert, add, delete, modify)
  - match und target
- also
  iptables −t table −Operation chain −j target match(es)
- zum Beispiel
  iptables −t filter −A INPUT −j ACCEPT −p tcp −−dport smtp



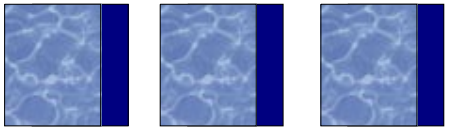
Targets

- Build−In im Kernel
  - ACCEPT — Paket akzeptieren
  - DROP — Paket verwerfen ohne Meldung
  - QUEUE — Enqueue in UserSpace
  - RETURN — Rückkehr zur vorherigen Chain
  - foobar — zu nutzerdefinierte Chain
- als Modul implementiert
  - REJECT — Drop mit Info an Absender − ICMP
  - MIRROR — Src−IP + Dst−IP tauschen und senden :)
  - LOG / ULOG — Logging via Syslog / UserSpace



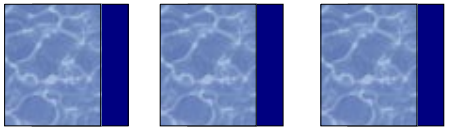
Matches

- Basic
  - −p Protokol (tcp/udp/icmp/...)
  - −s/−d Source / Destination Address (ip/mask)
  - −i/−o Incoming / Outgoing Interface
- Extensions
  - −−sport / −−dport Source / Destination Port
  - −−mac−source MAC
  - −−mark nfmark
  - −−limit Rate−Limit (Pakets / Timeframe)
  - −−tos, −−ttl



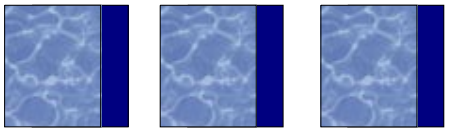
NAT

- bisher nur Masquerading; jetzt alle Möglichkeiten
  - SNAT Source NAT
    Ersetzen Src−IP in NF_IP_POST_ROUTING
  - DNAT Destination NAT
    Ersetzen von Destination−IP in NF_IP_PRE_ROUTING
- MASQUERADE ist spezielle Form von SNAT
- REDIRECT ist spezielle Form von DNAT



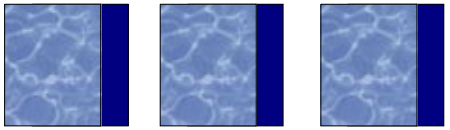
SNAT Beispiel

- iptables −t nat −A POSTROUTING −j SNAT −−to−source 1.2.3.4 −s 10.0.0.0/8
- Masquerading ist fast das Gleiche nur wird IP−Adresse vom Interface genutzt (wegen dynamischer Änderung bei DHCP/PPP/...)
  - iptables −t nat −A POSTROUTING −j MASQUERADE −o ppp0



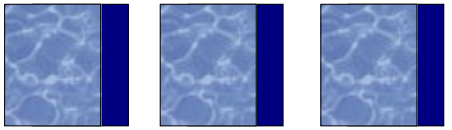
DNAT Beispiel

- iptables −t nat −A PREROUTING −j DNAT −−to−destination 1.2.3.4:8080 −p tcp −−dport 80 −i eth1
- REDIRECT setzt Destination−IP auf Adresse des Incoming−Interface
  - iptables −t nat −A PREROUTING −j REDIRECT −−to−port 3128 −i eth1 −p tcp −−dport 80



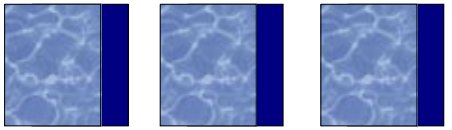
Mangling

- Verändern von Teilen des Pakets
- alle Matches möglich, um Paket auszuwählen
- z.Z. unterstützte Targets
  - TOS Verändern der TOS bits
  - TTL Setzen/Verringern/Erhöhen der TTL
  - MARK Modifizieren von nfmark
- iptables −t mangle −A PREROUTING −j MARK −−set−mark 10 −p tcp −−dport 80
- Nutzen der fwmask durch iproute2−Tools zum Class Based Queueing (CBQ) möglich!



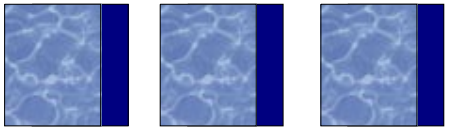
Connection tracking :))

- Unabhängig von NAT implementiert
- Stateful filtering möglich :))
- Hooks in NF_IP_PRE_ROUTING zum Tracking
- in NF_IP_POST_ROUTING / NF_IP_LOCAL_IN zum Entfernen gefilterter Verbindungen
- Protokol−Modules TCP/UDP/ICMP
- Application−Helpers (z.Z. FTP, IRC−DCC)
- Conntrack unterscheidet vier Kategorien
  - NEW, ESTABLISHED, RELATED, INVALID



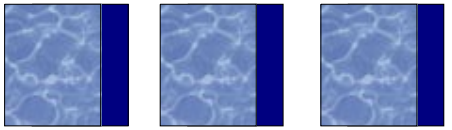
Erweiterungen

- Userspace Logging
  - Ersatz für syslogd−Logging via multicast−netlink−Sockets
  - easy−to−use Library (libipulog) und Daemon verfügbar
- Queuing
  - Pakete nach UserSpace via unilink−netlink−Sockets
  - libipq und experimental queue multiplex daemon (ipqmpd)



Current Development und Zukunft

- in Entwicklung aber ’proved very stable’
- Weiterentwicklungen
  - Möglichkeit von Conntrack/NAT−Helpers in UserSpace
  - TCP sequence number tracking :))
  - Conntrack für Multicast
  - weitere Matches (maxconn, ...)
  - weitere Conntrack / NAT−Module (RPC, SNMP, SMB)
  - Verbesserung von Ipv6−Unterstützung



Links und Rechts

- Harald’s Seiten
  - http://www.gnumonks.org
- Netfilter Homepage
  - http://netfilter.samba.org
    - http://netfilter.kernelnotes.org
    - http://netfilter.filewatcher.org
- Weitere Docs und Netfilter−Erweiterungen (ulogd, ipqmpd, ...)
  - http://www.gnumonks.org/projects
- http://avalon.tuts.nu/~mwei/archiv/ und Archiv TU−Chemnitz

mwei@linuxvortrag:~ > logout

http://www.tu-chemnitz.de/~als/images/penguins/