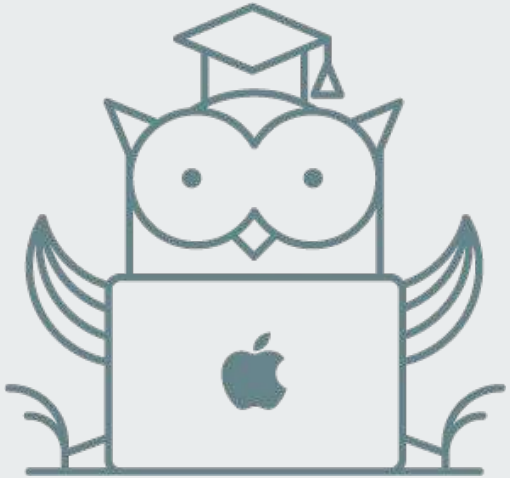
Курс «Администратор Linux»

Фильтрация пакетов

Занятие # 12

Дмитрий Молчанов
Григорий Ожегов



Фильтрация пакетов

Фильтрация трафика в linux осуществляется с помощью компоненты ядра linux - netfilter.

Компоненты/утилиты управления

- Компоненты
  - xtables
    - ip_tables
    - ip6_tables
    - arp_tables
  - nf_conntrack
  - ebtables
- Утилиты управления
  - iptables/ip6tables/ip6?tables-(save|restore)
  - ipset
  - ebtables
  - arptables
  - firewalld

ip_tables/ip6_tables

Предназначены для фильтрации L3-4 трафика.

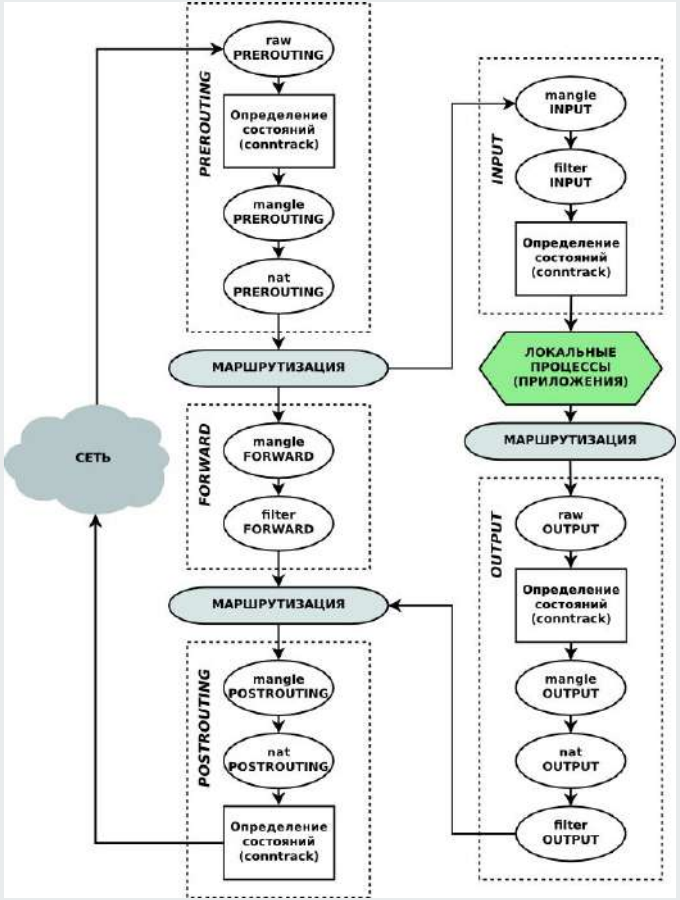
Принцип работы

Пакетный фильтр это набор правил. В iptables правила организованы по таблицам и цепочкам.
Каждая таблица состоит из цепочек.
Каждая цепочка - упорядоченный набор правил, которые просматриваются последовательно начиная с первого.
Каждое правило состоит из:

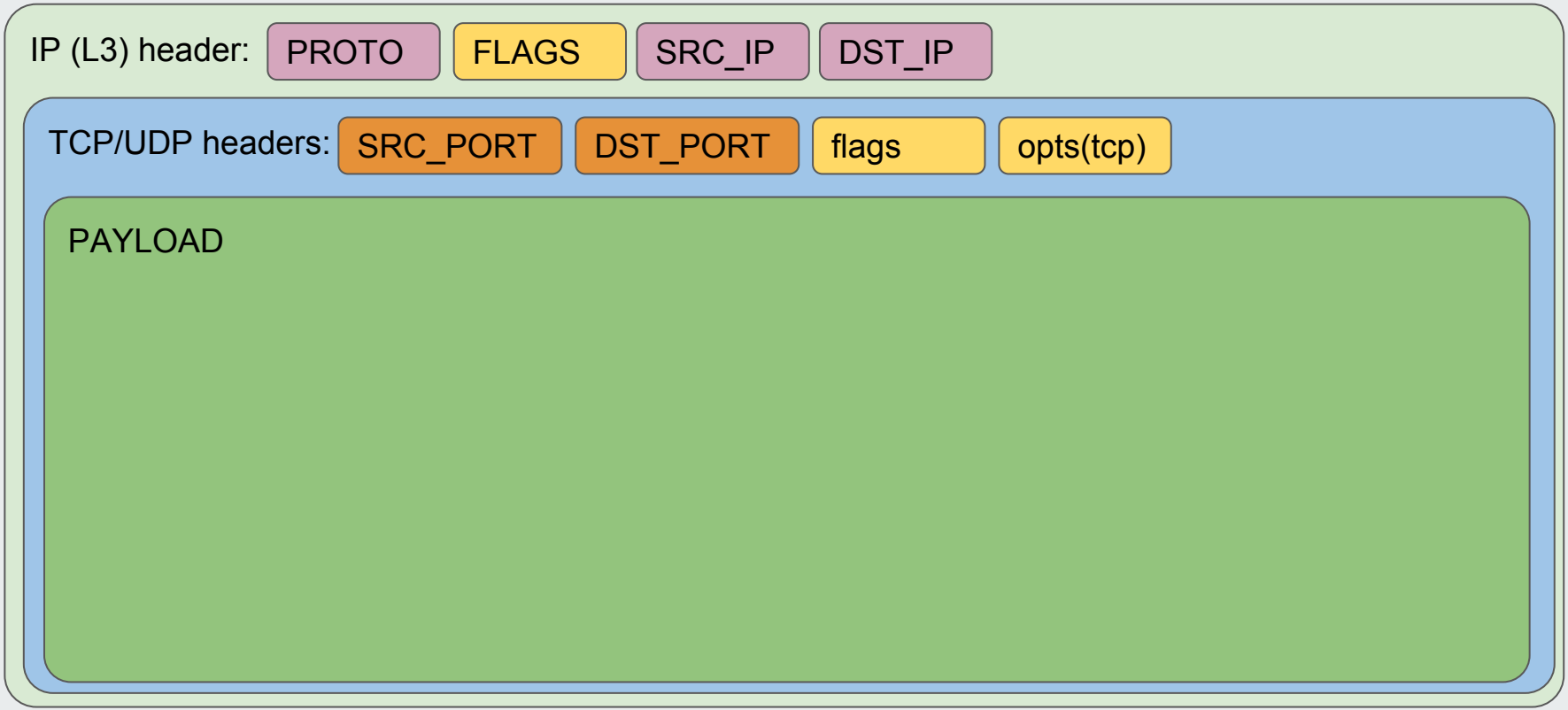
- Критериев срабатывания
- Действия

Каждое правило имеет счетчики срабатываний.

Схема работы



IP-пакет


IP (L3) header:
PROTO
FLAGS
SRC_IP
DST_IP
TCP/UDP headers:
SRC_PORT
DST_PORT
flags
opts(tcp)
PAYLOAD

Критерии срабатывания(общие, L3)

-i/--in-interface	Входящий интерфейс
-o/--out-interface	Исходящий инерфейс
-s/--source	Адрес источника
-d/--destination	Адрес назначения
-p/--protocol	IP-Протокол (tcp,udp,icmp…)
-f/--fragment	Является ли фрагментом (2+ в серии)

Критерии срабатывания (L4)

TCP/UDP

--sport port[:port]	порт или диапазон портов источника
--dport port[:port]	порт или диапазон портов назначения

TCP

--tcp-flags mask flags (SYN,ACK,RST,FIN SYN)	флаги TCP
--syn	взведен SYN

ICMP

--icmp-type	тип icmp-пакета

Действия

ACCEPT	T	Принять пакет
DROP	T	Отбросить пакет
REJECT	T	Отбросить пакет и сообщить источнику icmp-сообщением
RETURN	T	вернуться в вышестоящую цепочку или применить правило по-умолчанию
LOG	N	
chain_name	N	Перейти в цепочку *chain_name*
DNAT	T	Destination NAT
SNAT	T	Source NAT
MASQUERADE	T	Source NAT для динамически-конфигурируемых интерфейсов
SET		Добавление/удаления адреса в ipset

Дополнительные модули срабатывания

- conntrack/state - критерии срабатывания основанные на состоянии соединения.
- multiport - критерий срабатывания позволяющий указывать список портов, а не диапазон.
- iprange - критерий срабатывания, который позволяет указать ip-range, вместо cidr-префикса.
- mark/connmark - критерий срабатывания основаный на маркировке пакета/соединения.
- set - критерий срабатывания основаный на ipset
- u32 - гибкий критерий срабатывания который позволяет работать напрямую с заголовками пакетов и отдельными битами.

conntrack

Подсистема отслеживания состояния соединений. Базово соединения имеют состояния:

- NEW - новое соединение. Отбираются пакеты устанавливающие соединения.
- ESTABLISHED - Установленное соединение. Отбираются пакеты не !syn/syn+ack, !rst/fin, которые относятся к уже отслеживаемым соединениям
- RELATED - Относящиеся к другому, уже установленному соединению (passive ftp, icmp-messages)
- INVALID - пакеты принадлежность которых к отслеживаемым соединениям установить не удалось.

Подсистема очень удобна при невысокой нагрузке, при высокой же требует настройки, иначе может приводить к потере пакетов или связности в целом из-за переполнения таблицы conntrack или слишком большого её размера. в HL рекомендуется отключать, т.к. просмотр таблиц добавляет время к обработке пакета повышая latency приложения.

ipset

Использование iptables для фильтрации и защиты от злоумышленников вкупе с автоматизированными средствами управлением пакетным фильтром ведет к неизбежному росту количества правил, который, в свою очередь, ведет к задержках в обработке пакетов, что приводит к дополнительным задержкам в сети и, как следствие меньшей скорости работы приложений.
В определенных ситуациях (геотаргетирование, борьба с DDoS) возникает необходимость добавлять большое количество правил, которые можно сгруппировать по какому-либо признаку. Например это адреса с которых были запроссы отвечающие определенному паттерну, коих могут быть тысячи или сети принадлежащие одной стране.

Пусть, допустим, сверка с одним правилом, занимает ~1us, правил у нас 2000, в лучшем случае проверка iptables завершится через 1us (первое правило), в худшем - через 2ms, а это уже существенно, когда речь идет об одном пакете в рамках соединения.

ipset

Для того, чтобы можно было сократить количество правил фильтрации, объединяя правила по определенным признакам, можно использовать инструмент ipset, который позволяет строить списки адресов, вместо списков правил, что существенно упрощает обработку больших списков адресов.

Таблицы

- raw - изначальная обработка, до conntrack
- mangle - модификация заголовков пакетов и маркировка пакетов
- nat - трансляция адресов
- filter - фильтры
- security - работа с SELinux

Цепочки

- PREROUTING - до принятия решения о марщрутизации
- POSTROUTING - после принятия решения о маршрутизации
- OUTPUT - пакеты сгенерированные локальными приложениями
- INPUT - пакеты предназначающиеся локальной системе
- FORWARD - пакеты проходящие через систему

raw

Предназначена для базовой обработки пакетов до conntrack, в частности для управления conntrack в отношении некоторых пакетов.
Цепочки:

- PREROUTING
- OUTPUT

Действия:

- NOTRACK - отключить conntrack для пакетов попадающих в правило
- CT - настроить работу с модулем conntrack, включает NOTRACK через --notrack
- DROP - отбросить пакет.

mangle

Предназначена для маркировки и классификации пакетов, модификации заголовков (tos, mss, ttl)

Цепочки:

- PREROUTING
- INPUT
- FORWARD
- OUTPUT
- POSTROUTING

Действия:

- TTL - установить ttl
- MARK/CONNMARK - установить метку (fwmark) пакета/соединения
- CLASSIFY - классифицировать пакет для обработки в шейпере
- TCPMSS - установить TCP Maximum Segment Size (если не работает PMTU disco)

Предназначена для манипуляций с адресами источника/назначения.
Цепочки:

- PREROUTING
- INPUT
- OUTPUT
- POSTROUTING

Действия:

- SNAT/MASQUERADE - Source NAT
- DNAT - Destination NAT
- REDIRECT - Подмена dst_ip:dst_port на свои собственные (частный случай DNAT)

filter

Основная таблица, где происходит фильтрация пакетов.

Цепочки:

- INPUT
- FORWARD
- OUTPUT

Действия:

- ACCEPT
- REJECT
- DROP

Таблица предназначена для работы совместно с selinux.
Цепочки:

- INPUT
- FORWARD
- OUTPUT

Действия:

- SECMARK/CONNSECMARK - установить SELinux context для пакета/соединения

Полезные ссылки

- https://ru.wikibooks.org/wiki/Iptables
- http://linuxgazette.ru/rus/articles/iptables-tutorial.html

Спасибо за внимание

Дмитрий Молчанов
Григорий Ожегов