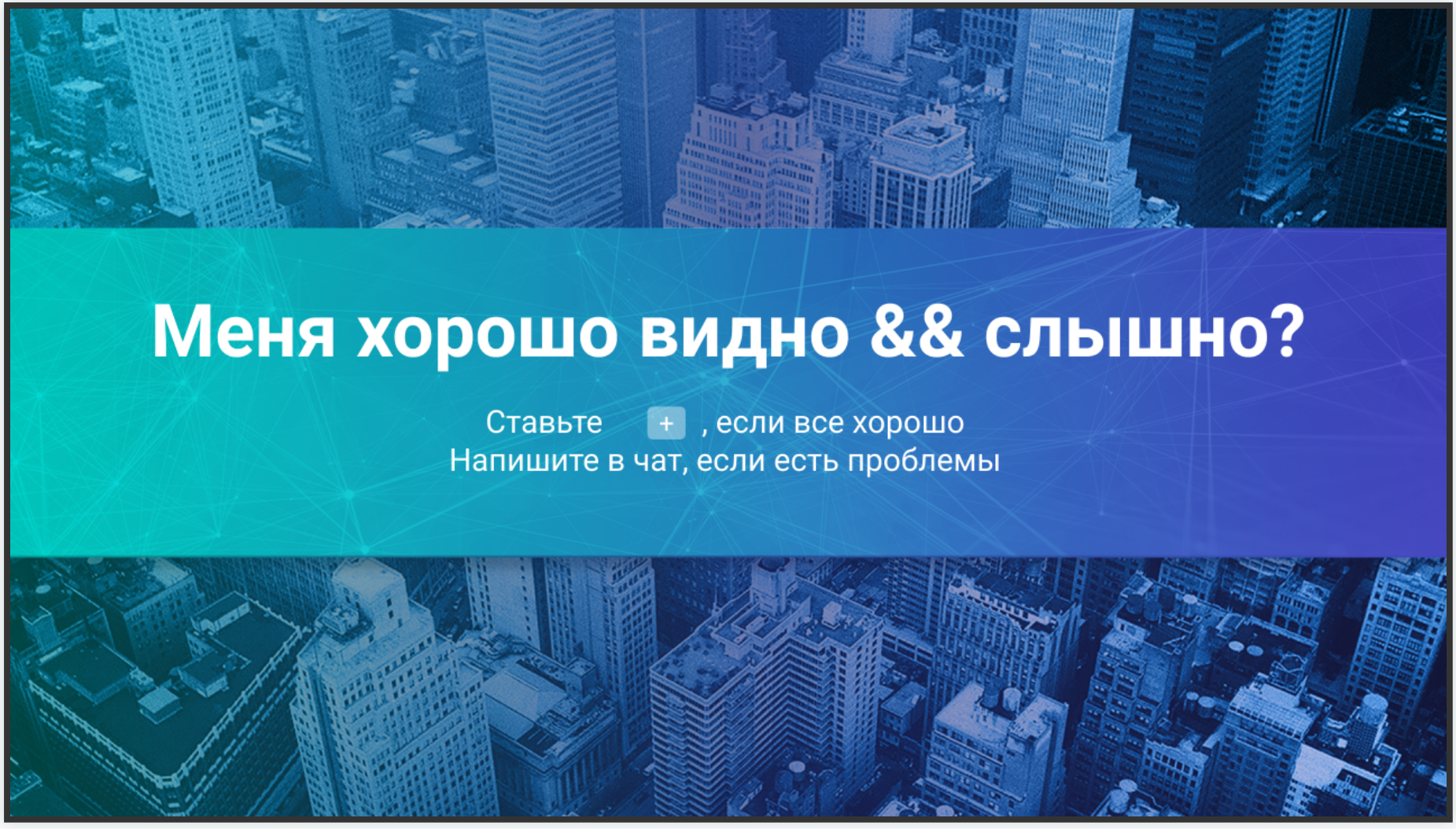

OTUS
ОНЛАЙН-ОБРАЗОВАНИЕ
Онлайн-образование


Меня хорошо видно && слышно?
Ставьте + , если все хорошо
Напишите в чат, если есть проблемы

# НЕ ЗАБЫТЬ ВКЛЮЧИТЬ ЗАПИСЬ!!!

# ФИЛЬТРАЦИЯ ТРАФИКА.

# ОТЛИЧИЯ ОТ IPTABLES

- синтакс
- нет предопределнных таблиц и цепочек
- несколько действий в одном правиле
- отсутствие встроенного счетчика
- не нужно дублировать правила для IPv6
- нативные set и map

# НОВЫЕ ВОЗМОЖНОСТИ

- Опциональные счетчики на правила
  - `nft add rule ip filter input counter`
- Несколько действий в одном правиле
  - 
```
nft add rule ip filter input \
counter log prefix \"packet drop: \" drop
```
- Интерактивный режим
  - `nft -i`
- Режим дебага
  - `nft --debug=all`

# NFTABLES: КОМПОНЕНТЫ

- Family: Семейства протоколов
  - (inet, ip, ipv6, bridge, arp, netdev)
- Ruleset: набор всех таблиц, цепочек и правил

```
nft list ruleset inet
```

- Таблицы: контейнеры для цепочек

```
nft list tables
nftp list tables ip
```

- Цепочки: наборы правил

```
nft list chains
```

# ТАБЛИЦЫ

```
{add | create} table [family] table [ { flags flags } ]
       {delete | list | flush} table [family] table
       list tables
       delete table [family] handle handle
```

# ЦЕПОЧКИ

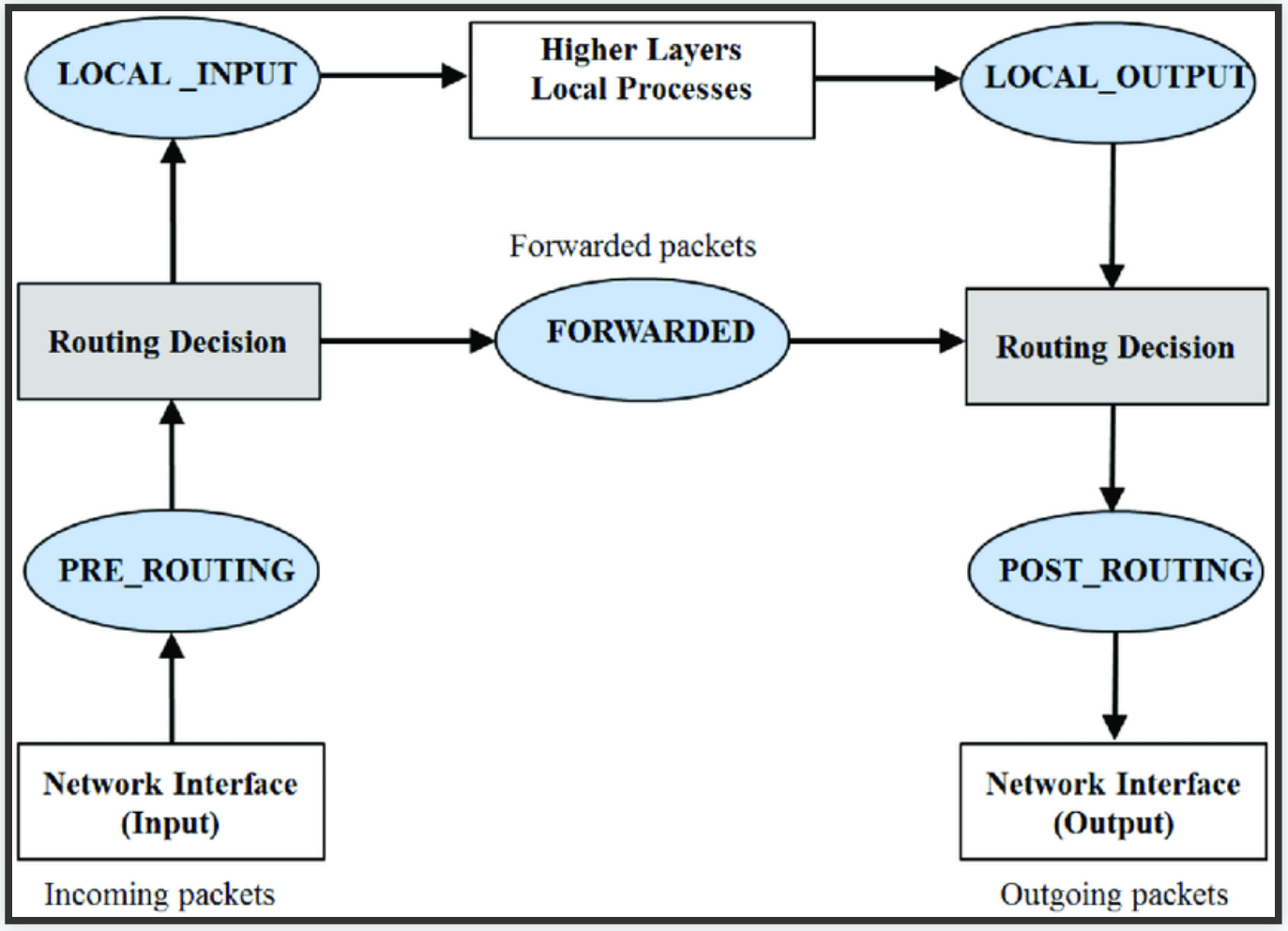
- Base chains
  - точка входа для пакетов из сети
- Regular chains
  - точка входа с другого правила ( jump)

```
{add | create} chain [family] table chain [ { type type hook hook
      {delete | list | flush} chain [family] table chain
      list chains
      delete chain [family] table handle handle
      rename chain [family] table chain newname
```

# ТИПЫ ЦЕПОЧЕК

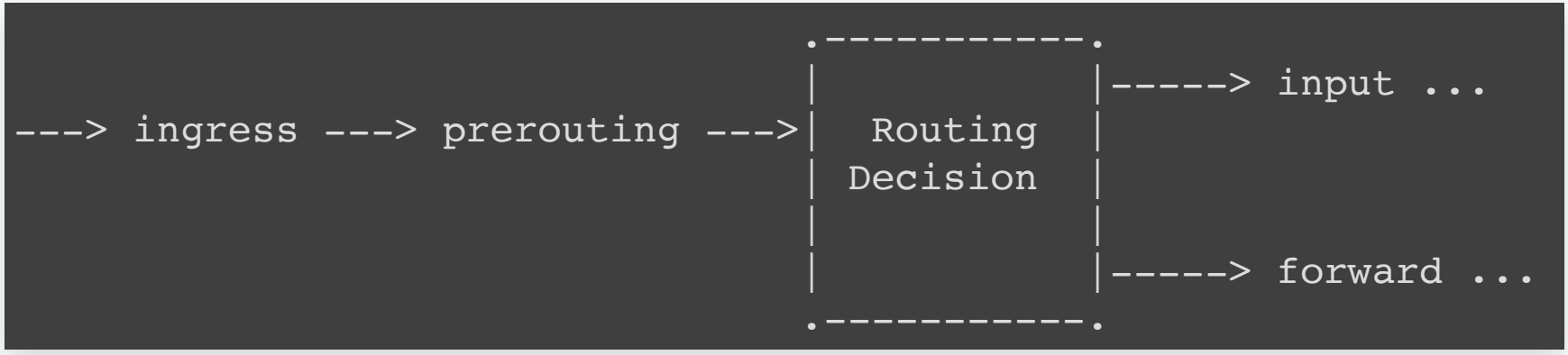
- filter (all)
  - hooks: all (ingress, prerouting, input, forward, output, postrouting)
- nat (ip, ip6)
  - hooks: prerouting, input, output, postrouting
- route (ip, ip6)
  - hooks: output

# HOOKS


LOCAL _INPUT
Higher Layers
Local Processes
LOCAL_OUTPUT
Forwarded packets
Routing Decision
FORWARDED
Routing Decision
PRE_ROUTING
POST_ROUTING
Network Interface
(Input)
Network Interface
(Output)
Incoming packets
Outgoing packets

# INGRESS HOOK

- перед prerouting
- можно фильтровать трафик на L2
- используется для netdev



# ПРИОРИТЕТЫ

- числовое значение со знаком (чем меньше, тем приоритетнее)
- указывает в каком порядке обрабатывать цепочки с одинаковыи hook
- если пакет был принят на одной цепочке и есть следующая цепочка на этом hook, то он будет обработан снова ( до drop )

```
NF_IP_PRI_CONNTRACK_DEFRAG (-400), NF_IP_PRI_RAW (-300),
NF_IP_PRI_SELINUX_FIRST (-225), NF_IP_PRI_CONNTRACK (-200),
NF_IP_PRI_MANGLE (-150), NF_IP_PRI_NAT_DST (-100),
NF_IP_PRI_FILTER (0), NF_IP_PRI_SECURITY (50),
NF_IP_PRI_NAT_SRC (100), NF_IP_PRI_SELINUX_LAST (225),
NF_IP_PRI_CONNTRACK_HELPER (300)
```

# ПОЛИТИКИ ЦЕПОЧЕК

- accept
  - политика по умолчанию
  - пакет продолжает проходить по всему стеку
- drop
  - пакет дропается

# ПРИМЕР С ПРИОРИТЕТАМИ И ПОЛИТИКАМИ

```
table inet filter {
        # this chain is evaluated first due to priority
        chain ssh {
                type filter hook input priority 0; policy accept;
                # ssh packet accepted
                tcp dport http accept
        }

        # this chain is evaluated last due to priority
        chain input {
                type filter hook input priority 1; policy drop;
                # the same ssh packet is dropped here by means of
        }
}
```

# ПРАВИЛА:

```
% nft add rule [<family>] <table> <chain> <matches> <statements>
% nft insert rule [<family>] <table> <chain> [position <position>
% nft replace rule [<family>] <table> <chain> [handle <handle>] <
% nft delete rule [<family>] <table> <chain> [handle <handle>]
```

- position: порядковый номер для вставки
- handle: внутренний идентификатор
- посмотреть handles

```
nft list table filter -n -a
```

# MATCHES: EXAMPLES

- saddr <eth/ip/ipv6 source address>
- daddr <eth/ip/ipv6 destination address>
- version <4/6 protocol version>
- vlan [ id| cfi | pcp | type]
- protocol
- length

# MATCHES: EXAMPLES

- sport
- dport

```
tcp option {eol | noop | maxseg | window | sack-permitted \
| sack | timestamp} tcp_option_field

ct {state | direction | status | mark | expiration | protocol \
| bytes | packets  | zone}

ct {original | reply} {l3proto | protocol | proto-src \
| proto-dst | bytes | packets | avgpkt | zone}

ct {original | reply} {ip | ip6} {saddr | daddr}
```

# META EXPRESSIONS

включение/отключение метаданных, ассоциированных с пакетом

```
meta {length | nfproto | l4proto | protocol | priority}
      [meta] {mark | iif | iifname | iiftype | oif | oifname |
      oiftype | skuid | skgid | nftrace | rtclassid | ibrname |
      obrname | pkttype | cpu | iifgroup |
      oifgroup | cgroup | random | secpath}
```

- length - длина пакета в байтах
- iif - входной интерфейс
- skuid - UID оригинирующего сокета

# TRACING

## meta nftrace set 1

```
nft add rule inet filter input tcp dport 10000 nftrace set 1
nft add rule inet filter input icmp type echo-request nftrace set
nft -nn monitor trace
```

# ПРАВИЛА: ПРИМЕРЫ

```
# Запрет
nft add rule ip filter input tcp dport != 80

# Диапазоны
add rule ip filter input tcp dport 10-1024
add rule ip filter input meta skuid 1000-1100

# Префиксы
add rule ip filter input ip daddr 192.168.10.0/24
add rule ip filter input meta mark 0xffffffff/24

# Состояния
add rule ip filter input ct new,established

# Маркировка пакетов
add rule ip filter input ct mark set 10
```

# ДЕЙСТВИЯ

- accept - принимает пакет и завершает обработку
- drop - дропает пакет и завершает обработку
- queue - поставка пакет в очередь для обработки приложением, завершает обработку
- continue - переход к следующему правилу
- return - выход из текущей цепочки
- jump chain - переход с точкой возврата
- goto chain - переход без точки возврата

# СКРИПТЫ И ЗАГРУЗКА

- /etc/sysconfig/nftables.conf
- `ls /etc/nftables`

```
nft list ruleset > /etc/nftables.rules
nft flush ruleset
nft -f /etc/nftables.rules
```

# COUNTERS

счетчик на каждое правило выставлется опционально

```
#  позиция в синтаксисе имеет значение

nft add rule filter input ip protocol tcp counter
nft add rule filter input ip counter protocol tcp
nft add rule filter input counter ip protocol tcp
```

# LOGGING

```
# включаем лог
nft add rule filter input log

# добавляем префикс лога, в правиле происходит 2 действия: log и
nft add rule filter input tcp dport 22 ct state new log \
prefix "New SSH connection: " accept

# логгируем все пакеты, разрешаем только на порт 22
nft add rule filter input iif lo log tcp dport 22 accept

# отправка лога в очередь на обработку для приложений с libnetfil
nft add rule filter input tcp dport 22 ct state new log \
prefix "New SSH connection: " group 0 accept
```

# JUMPING TO CHAIN

```
table ip filter {
        chain input {
                type filter hook input priority 0; policy accept
                # переход без возврата
                ip saddr 3.3.3.3 ip daddr 2.2.2.2 tcp sport 111
                # переход с возвратом
                ip saddr 1.1.1.1 ip daddr 2.2.2.2 tcp sport 111
                # правило срабатывающее после возврата
                ip saddr 1.1.1.1 ip daddr 2.2.2.2 tcp sport 111
        }

        chain other-chain {
                # this is the 2º matching rule
                counter packets 8 bytes 2020
        }
}
```

# NAT

```
nft add table nat
# snat
nft add chain nat postrouting { type nat hook postrouting priorit
nft add rule nat postrouting ip saddr 192.168.1.0/24 oif eth0 sna
# dnat
nft add chain nat prerouting { type nat hook prerouting priority
nft add rule nat prerouting iif eth0 tcp dport { 80, 443 } dnat 1
# masquerading
nft add rule nat postrouting masquerade
# redirect
nft add rule nat prerouting redirect
nft add rule nat prerouting tcp dport 22 redirect to 2222
```

# DUPLICATING PACKETS

начиная с версии ядра 4.3 есть возможность дублировать пакет

```
# дублировать весь трафик на 172.20.0.2
nft add rule mangle prerouting dup to 172.20.0.2

# дублировать трафик с определенных адресов на конкретные адреса
nft add rule mangle prerouting dup to ip saddr map \
 { 192.168.0.1 : 172.20.0.2, 192.168.0.1 : 172.20.0.3 }
```

# БАЛАНСИРОВКА НАГРУЗКИ

```
# Round Robin
nft add rule nat prerouting dnat to numgen inc mod 2 map { \
                0 : 192.168.10.100, \
                1 : 192.168.20.200 }
```

# SETS

```
# anoynoums
nft add rule ip6 filter input tcp dport {telnet, http, https} acc

# named
nft add set ip filter blackhole { type ipv4_addr\;}
nft add element ip filter blackhole { 192.168.1.4, 192.168.1.5 }

nft add set ip filter ports {type inet_service \; timeout 3h45s \
```

# FLOWTABLES

- выставление лимитов

```
nft add rule ip foo bar ct state new tcp dport
flow table ssh-spammer { \
ip saddr limit rate over 3/second
} log prefix "New SSH connection: " drop
#
nft list flow table ssh-spammer
```

# STATEFUL OBJECTS

```
# counters
nft add counter filter https-traffic
nft add rule filter output counter name tcp dport map { \
          https : "https-traffic", \
          80 : "http-traffic"}

# quotas
nft add quota filter https-quota 25 mbytes
table inet foo {
   quota example { over 100 mbytes used 0 bytes }

   chain dropafterquota {
      type filter hook postrouting priority 0; policy accept;
      udp port 5060 quota name "example" drop
   }
}
```

# MAPS

структуры данных для доступа к значения по ключу

```
# перенаправление трафика в зависимосте от порта
nft add rule ip nat prerouting dnat tcp dport map { 80 : 192.168.

# можно отдельно завести map, добавлять туда элементы, использова
nft add map nat porttoip  { type inet_service: ipv4_addr\; }
nft add element nat porttoip { 80 : 192.168.1.100, 8888 : 192.168
nft add rule ip nat postrouting snat tcp dport map @porttoip
```

# CONCATENATIONS

```
nft add rule ip filter input ip saddr . ip daddr . ip protocol {\
1.1.1.1 . 2.2.2.2 . tcp, \
1.1.1.1 . 3.3.3.3 . udp} counter accept

nft add rule netdev foo bar \
ether saddr . ip saddr . tcp dport { \
c0:fe:00:c0:fe:00 . 192.168.1.123 . 80,
be:ef:00:be:ef:00 . 192.168.1.120 . 22} \ counter accept
```

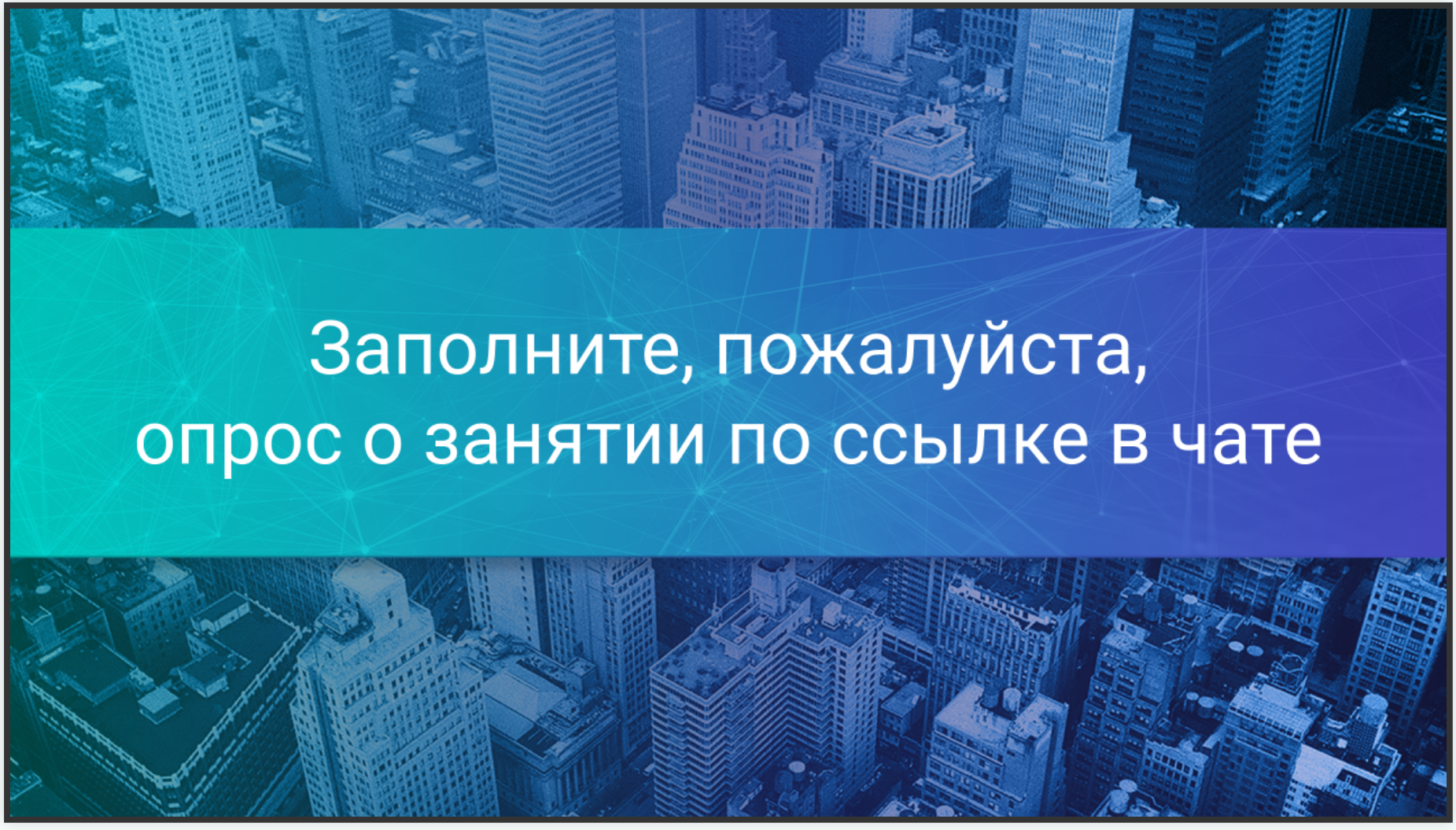
# ТРАНСЛЯЦИЯ ПРАВИЛ

- iptables-translate
- iptables-restore-translate
- ip6tables-translate
- ip6tables-restore-translate

# Рефлексия

Отметьте 3 пункта, которые вам запомнились с вебинара

Что вы будете применять в работе из сегодняшнего вебинара?


Заполните, пожалуйста,
опрос о занятии по ссылке в чате