Comportamento da ACL no PBR no Nexus 7K contendo informações de L3 e L4

Contents

Introduction

Este documento descreve o comportamento do Roteamento Baseado em Políticas (PBR - Policy-Based Routing) em Switches Nexus quando você filtra com base nas informações da Camada 3 (L3) e da Camada 4 (L4).

Informações de Apoio

Se você adicionar uma sequência no PBR para corresponder às informações específicas do L4, à medida que um recurso N7K cria entradas para ACEs (Access Control Entry, entrada de controle de acesso) e uma ACE de fragmento é criada automaticamente que corresponde às informações do L3 especificadas na sequência de correspondência. No caso de pacotes fragmentados, o primeiro pacote conhecido como fragmento inicial contém o cabeçalho L4 e é correspondido corretamente na ACL (Access Control List, lista de controle de acesso). No entanto, os próximos fragmentos conhecidos como não-iniciais não contêm nenhuma informação de L4 e, portanto, se a parte L3 da entrada de ACL corresponder, o fragmento não-inicial será permitido. Portanto, deve-se tomar o máximo cuidado ao filtrar o tráfego com base nas informações de L4, pois os fragmentos não-iniciais podem ser roteados incorretamente na ausência de informações de L4.

Topologia

O Roteador LAN está conectado ao Nexus na interface E2.1, Vlan 700. O requisito é redirecionar o tráfego que corresponde ao protocolo de gerenciamento de rede simples (SNMP - Simple Network Management Protocol), Web etc. para o Otimizer e para todo o tráfego restante diretamente para a interface E2/2 em direção ao firewall. O PBR é configurado no Switch Virtual Interface (SVI) Vlan700 no dispositivo Nexus. A configuração para o mesmo é fornecida aqui. A sequência 70 no mapa de rotas encaminha todo o tráfego restante para o Firewall. Há um novo

requisito de que todo o tráfego com a porta UDP 920x precisa passar pelo Otimizer, para esta Sequência 50 é adicionado no mapa de rota.

Veja aqui como o PBR responde aos pacotes Fragmentados e Não Fragmentados que atingem a sequência 50 e correspondem às informações de L3 e L4.

Esta é a configuração na interface Nexus Vlan700 para redirecionar o tráfego que vem em E2/1:

```
interface Vlan700

 no shutdown

 mtu 9000

 vrf member ABC

 no ip redirects

 ip address 10.11.25.25/28

 ip policy route-map In_to_Out


Nexus# show route-map In_to_Out

route-map In_to_Out, permit, sequence 3

 Match clauses:

   ip address (access-lists): Toolbar

 Set clauses:

   ip next-hop 10.3.22.13

route-map In_to_Out, permit, sequence 5

 Match clauses:

   ip address (access-lists): Internet

 Set clauses:

   ip next-hop 10.11.25.19

route-map In_to_Out, permit, sequence 7

 Match clauses:

   ip address (access-lists): Web

 Set clauses:

   ip next-hop 10.11.25.19

route-map In_to_Out, permit, sequence 10

 Match clauses:

   ip address (access-lists): In_to_Out_Internet
```

```
 Set clauses:
   ip next-hop 10.11.25.23
route-map In_to_Out, permit, sequence 30
 Match clauses:
   ip address (access-lists): In_to_Out_www
 Set clauses:
   ip next-hop 10.11.25.23
route-map In_to_Out, permit, sequence 35
 Match clauses:
   ip address (access-lists): In_to_Out_https
 Set clauses:
   ip next-hop 10.11.25.23
route-map In_to_Out, permit, sequence 40
 Match clauses:
   ip address (access-lists): In_to_Out_8080
 Set clauses:
   ip next-hop 10.11.25.23
route-map In_to_Out, permit, sequence 50
 Match clauses:
   ip address (access-lists): UDP_Traffic
 Set clauses:
   ip next-hop 10.11.25.23 >>>>>>>>>>>>>>>>>>>> Towards Optimizer
route-map In_to_Out, permit, sequence 70
 Match clauses:
   ip address (access-lists): To_Firewall
 Set clauses:
    ip next-hop . 10.22.45.63 >>>>>>>>>>>>>>>> Towards Firewall


Nexus# show ip access-lists UDP_Traffic
IP access list UDP_Traffic
10 permit udp any any eq 9201
20 permit udp any any eq 9202
```

```
30 permit udp any any eq 9203



Nexus# sh ip access-lists To_Firewall

IP access list To_Firewall

       10 permit ip any any
```

Quando o roteamento baseado em política é configurado no SVI, o Nexus cria uma entrada no hardware para o mesmo. Vamos agora analisar a programação de hardware do PBR no módulo 2 do Nexus:

```
Nexus# show system internal access-list vlan 700 input entries detail module 2

Flags: F - Fragment entry  E - Port Expansion

      D - DSCP Expansion  M - ACL Expansion

      T - Cross Feature Merge Expansion

INSTANCE 0x0

---------------

 Tcam 1 resource usage:

 ----------------------

 Label_b = 0x201

  Bank 0

  ------

    IPv4 Class

      Policies: PBR(GGSN_Toolbar)

      Netflow profile: 0

      Netflow deny profile: 0

      Entries:

        [Index] Entry [Stats]

        ---------------------

 [0019:000f:000f] prec 1 permit-routed ip 0.0.0.0/0 224.0.0.0/4   [0]

 [002d:0024:0024] prec 1 redirect(0x5d)-routed tcp 1.1.22.80/28 0.0.0.0/0 eq 80 flow-label 80
[0]

 [002e:0025:0025] prec 1 redirect(0x5d)-routed tcp 1.1.22.80/28 0.0.0.0/0 fragment   [0]

 [002f:0026:0026] prec 1 redirect(0x5d)-routed tcp 1.1.22.80/28 0.0.0.0/0 eq 8080 flow-label
8080  [0]

 [0030:0027:0027] prec 1 redirect(0x5d)-routed tcp 1.1.22.80/28 0.0.0.0/0 fragment   [0]

 [0031:0028:0028] prec 1 redirect(0x5d)-routed tcp 1.1.22.48/28 0.0.0.0/0 eq 80 flow-label 80
```

```
[0]

 [0032:0029:0029] prec 1 redirect(0x5d)-routed tcp 1.1.22.48/28 0.0.0.0/0 fragment   [0]

 [0033:002a:002a] prec 1 redirect(0x5d)-routed tcp 1.1.22.48/28 0.0.0.0/0 eq 8080 flow-label
8080  [0]

 [0034:002b:002b] prec 1 redirect(0x5d)-routed tcp 1.1.22.48/28 0.0.0.0/0 fragment   [0]

 [0035:002c:002c] prec 1 permit-routed ip 1.1.22.24/29 0.0.0.0/0   [0]

 [0036:002d:002d] prec 1 permit-routed ip 1.1.22.32/28 0.0.0.0/0   [0]

 [0037:002e:002e] prec 1 permit-routed ip 1.1.22.64/28 0.0.0.0/0   [0]

 [0038:002f:002f] prec 1 permit-routed ip 1.1.22.80/28 0.0.0.0/0   [0]

 [003d:0033:0033] prec 1 permit-routed ip 1.1.22.96/28 0.0.0.0/0   [0]

 [003e:0034:0034] prec 1 permit-routed tcp 0.0.0.0/0 196.11.146.149/32 eq 25 flow-label 25  [0]

 [0059:004f:004f] prec 1 permit-routed tcp 0.0.0.0/0 196.11.146.149/32 fragment   [0]

 [005a:0050:0050] prec 1 redirect(0x5e)-routed ip 1.1.22.16/29 0.0.0.0/0   [0]

 [005b:0051:0051] prec 1 redirect(0x5e)-routed tcp 0.0.0.0/0 0.0.0.0/0 eq 80 flow-label 80  [0]

 [005c:0052:0052] prec 1 redirect(0x5e)-routed tcp 0.0.0.0/0 0.0.0.0/0 fragment   [0]

 [005d:0053:0053] prec 1 redirect(0x5e)-routed tcp 0.0.0.0/0 0.0.0.0/0 eq 443 flow-label 443
[0]

 [005e:0054:0054] prec 1 redirect(0x5e)-routed tcp 0.0.0.0/0 0.0.0.0/0 fragment   [0]

 [005f:0055:0055] prec 1 redirect(0x5e)-routed tcp 0.0.0.0/0 0.0.0.0/0 eq 8080 flow-label 8080
[0]

 [0060:0056:0056] prec 1 redirect(0x5e)-routed tcp 0.0.0.0/0 0.0.0.0/0 fragment   [0]

********************Sequence 50 is to match the traffic for UDP ports
9201/9202/9203**********************

 [0061:0057:0057] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9201 flow-label 9201
[0]

 [0062:0058:0058] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment   [0]

 [0063:0059:0059] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9202 flow-label 9202
[0]

 [0064:005a:005a] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment   [0]

 [0065:005b:005b] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9203 flow-label 9203
[0]

 [0066:005c:005c] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment   [0]

********************Sequence 70 is to send all other traffic to Firewall**********************

 [0067:005d:005d] prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [23]

 [0068:005e:005e] prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [0]
```

Você vê que, além da entrada da lista de acesso que corresponde ao **udp 0.0.0.0/0 0.0.0.0/0 eq**

9201, há outra entrada que corresponde ao **fragmento udp 0.0.0.0/0 0.0.0.0/0** mas essa entrada não tem nenhuma informação de porta UDP. Essa entrada é equivalente a qualquer outra que corresponda ao pacote UDP, de modo que os pacotes para outras portas UDP também são correspondidos nessa sequência gerada pelo hardware.

Caso de teste 1: Tráfego iniciado do roteador LAN para o firewall

- O pacote que chega ao Nexus não foi fragmentado e, portanto, o tráfego correspondeu como esperado no PBR.
- Ele foi redirecionado corretamente para o Firewall e pode ser visto em depurações executadas no Firewall.

UDP packet -port 500

```
*Mar 26 04:07:48.959: IP: s=1.1.1.1 (GigabitEthernet0/0), d=3.3.3.3, len 28, rcvd 4  -à
```
Traffic entering from Nexus interface

```
*Mar 26 04:07:48.959:     UDP src=500, dst=500
```

TCP packet - port 80

```
*Mar 26 04:07:48.671: IP: s=1.1.1.1 (GigabitEthernet0/1), d=3.3.3.3, len 40, rcvd 4
 -à
```
Traffic entering from Optimizer interface

```
*Mar 26 04:07:48.671:     TCP src=1720, dst=80, seq=0, ack=0, win=0
```

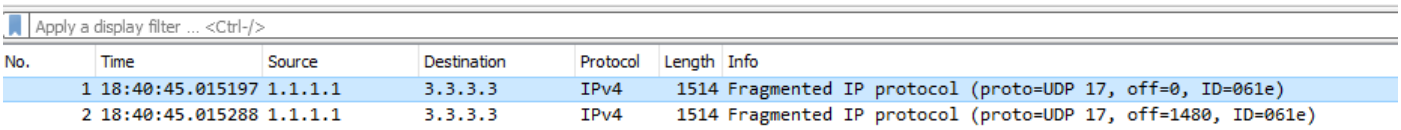
UDP packet -port 9201

```
*Mar 27 09:30:19.879: IP: s=1.1.1.1 (GigabitEthernet0/1), d=3.3.3.3, len 28, input
feature  à
```
Traffic entering from Optimizer interface

```
*Mar 27 09:30:19.879:     UDP src=6000, dst=9201, MCI Check(80), rtype 0, forus FALSE,
sendself FALSE, mtu 0, fwdchk FALSE
```

Teste do caso 2: Tráfego iniciado via Ficheiro Sniffer do Roteador LAN para Firewall com UDP 500

Tráfego com dois fragmentos no arquivo de farejador gerado aqui:



1. Fragmentos iniciais com mapa de rota:

- O primeiro fragmento com **Deslocamento = 0** é conhecido como fragmento inicial e contém o cabeçalho UDP no pacote.

- Como o tráfego é para UDP 500, ele é correspondido na sequência 70 para permitir **ip any any.**

```
        prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [23]
```

- Portanto, o primeiro pacote que tem as informações das Camadas 3 e 4 é roteado corretamente.

2. Pacotes de Fragmentos Não Iniciais com Mapa de Rota:

- O segundo fragmento com **Offset ≠ 0** é conhecido como fragmento não inicial e não contém nenhum cabeçalho UDP. É um pacote puramente IP com o tipo de protocolo UDP (17).
- Como não há informações da Camada 4, elas correspondem na sequência 70 : **permit-routed ip 0.0.0.0/0 0.0.0.0/0.**
- No entanto, na sequência 50, há uma lista de acesso que corresponde ao tráfego para a porta UDP 920x. O hardware cria automaticamente uma entrada para permitir os fragmentos UDP que correspondem às informações especificadas da Camada 3.
- Portanto, cada pacote fragmentado para qualquer informação de Camada 3 com protocolo UDP que é correspondido na sequência 50.

```
            prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9201 flow-label 9201  [0]

            prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment [23]
>>>>>>>>>>>>>>>>>>>>
```

- Dessa forma, há um fragmento que é roteado corretamente e outro roteado por sequência errada.
- O segundo fragmento é modificado para fazer **offset = 0**, e é correspondido na Sequência 70 conforme esperado.
- Esse é um comportamento esperado sempre que os fragmentos da Camada 4 são recebidos.
- A intenção de criar uma entrada extra para permitir fragmentos é permitir os fragmentos não iniciais recebidos sem informações da camada 4.
- No caso, o tráfego era para UDP 9201 e não havia entrada para permitir fragmentos. Em seguida, o segundo fragmento teria correspondido na Sequência 70 para permitir **ip any any** e, portanto, seria roteado incorretamente.

```
Nexus# sh route-map In_to_Out pbr-statistics

route-map In_to_Out, permit, sequence 3

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 5

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 7

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 10
```

```
 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 30

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 35

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 40

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 50 ----------> 2nd Fragment for UDP 500 is matched here

 Policy routing matches: 4397 packets

route-map In_to_Out, permit, sequence 70------------> 1st Fragment for UDP 500 is matched here

 Policy routing matches: 4397 packets
```

- Outra sequência 45 é criada para permitir o tráfego para o UDP 500 e observar que ambos os fragmentos são correspondidos na sequência 45.
- O fragmento inicial correspondeu devido às informações do cabeçalho UDP e não-inicial correspondeu na linha de fragmentos para a sequência 45.

```
Nexus# sh route-map In_to_Out pbr-statistics

route-map In_to_Out, permit, sequence 3

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 5

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 7

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 10

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 30

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 35

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 40

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 45----------> Both fragments matched here
```

```
 Policy routing matches: 213 packets

route-map In_to_Out, permit, sequence 50

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 70

 Policy routing matches: 0 packets


Default routing: 0 packets
```

Lista de acesso para a sequência 45:

```
Nexus# sh ip access-lists udptraffic

IP access list udptraffic

permit udp any any eq isakmp
```

3. Agora vamos ver como a palavra-chave fragments se comporta com ACL e mapa de rota

- A sequência 5 é aplicada para permitir qualquer porta UDP aleatória 56 na porta ACL.

```
Nexus# sh ip access-lists TEST_UDP


IP access list TEST_UDP

statistics per-entry

 5 permit udp any any eq 56 [match=0]

 10 permit udp any any eq isakmp [match=0]

 20 permit ip any any [match=0]
```

- Iniciado um fluxo de tráfego com pacote não inicial fragmentado e observado que ele corresponde na sequência 5. Embora o pacote seja para UDP 500, ele corresponde na sequência 5 para permitir UDP 56.

```
Nexus# sh ip access-lists TEST_UDP


IP access list TEST_UDP

statistics per-entry

 5 permit udp any any eq 56 [match=56]

 10 permit udp any any eq isakmp [match=0]

 20 permit ip any any [match=0]
```

- Os fragmentos são negados na ACL da porta e observa-se que nenhum pacote é

correspondido na ACL para não-inicial, já que o pacote é realmente correspondido na entrada **udp any any any fragments** criada automaticamente pela plataforma.

```
NEXUS# sh ip access-lists TEST_UDP



IP access list TEST_UDP

       statistics per-entry

        fragments deny-all

       5 permit udp any any eq 56 [match=0]

        10 permit udp any any eq isakmp [match=0]

        20 permit ip any any [match=0]



 [0014:000a:000a] prec 3 permit udp 0.0.0.0/0 0.0.0.0/0 eq 56 flow-label 56  [0]-> Here we are
now not seeing any entry to allow UDP fragments

 [0015:000b:000b] prec 3 permit udp 0.0.0.0/0 0.0.0.0/0 eq 500 flow-label 500  [0]

 [0016:000c:000c] prec 3 permit ip 0.0.0.0/0 0.0.0.0/0   [0]

 [0017:000d:000d] prec 3 deny ip 0.0.0.0/0 0.0.0.0/0 fragment   [100]>> Getting matched in
fragments deny statement

 [001e:0014:0014] prec 3 deny ip 0.0.0.0/0 0.0.0.0/0   [0]
```

- Negados os fragmentos na ACL problemática no PBR, no entanto, essa solução não funcionou e os pacotes ainda são vistos como correspondendo na sequência 50 e 70. Isso se deve ao comportamento de programação da lista de acesso e do mapa de rota.

```
          NEXUS# sh ip access-lists UDP_Traffic

       IP access list UDP_Traffic

       statistics per-entry

        fragments deny-all

       10 permit udp any any eq 9201

        20 permit udp any any eq 9202

        30 permit udp any any eq 9203



[0061:0057:0057] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9201 flow-label 9201
[0]

[0062:0058:0058] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment    [8027]

 [0063:0059:0059] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9202 flow-label 9202
[0]
```

```
 [0064:005a:005a] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment    [0]

 [0065:005b:005b] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9203 flow-label 9203
[0]

 [0066:005c:005c] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment    [0]

  [0067:005d:005d] prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [8027]

 [0068:005e:005e] prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [0]
```

- Saídas quando fragmentos negam é aplicado na ACL da porta e na ACL PBR:

```
[0061:0057:0057] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9201 flow-label 9201
[0]

[0062:0058:0058] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment    [8027] ---
> Once the fragments are denied in port CAL, we observed non-initial packets to be getting
dropped (See the mismatch in number of packets between UDP and IP counter)

 [0063:0059:0059] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9202 flow-label 9202
[0]

 [0064:005a:005a] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment    [0]

 [0065:005b:005b] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 eq 9203 flow-label 9203
[0]

 [0066:005c:005c] prec 1 redirect(0x5e)-routed udp 0.0.0.0/0 0.0.0.0/0 fragment    [0]

  [0067:005d:005d] prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [8214]

 [0068:005e:005e] prec 1 permit-routed ip 0.0.0.0/0 0.0.0.0/0   [0]



VDC-1 Ethernet2/1 :

 ====================

INSTANCE 0x0

---------------

 Tcam 0 resource usage:

 ----------------------

 Label_a = 0x200

  Bank 0

  ------

    IPv4 Class

      Policies: PACL(TEST_UDP)

       Netflow profile: 0
```

```
      Netflow deny profile: 0

      Entries:

         [Index] Entry [Stats]

        ---------------------

 [0014:000a:000a] prec 3 permit udp 0.0.0.0/0 0.0.0.0/0 eq 56 flow-label 56  [8027]

 [0015:000b:000b] prec 3 permit udp 0.0.0.0/0 0.0.0.0/0 eq 500 flow-label 500  [8214]

 [0016:000c:000c] prec 3 permit ip 0.0.0.0/0 0.0.0.0/0   [0]

 [0017:000d:000d] prec 3 deny ip 0.0.0.0/0 0.0.0.0/0 fragment   [100]

 [001e:0014:0014] prec 3 deny ip 0.0.0.0/0 0.0.0.0/0   [0]
```

Há várias maneiras possíveis de resolver esse problema ou limitação de pacotes fragmentados com informações de L4:

- O mapa de rotas pode ser ajustado para permitir informações L3 específicas para portas UDP específicas.

Na configuração atual, se as informações de origem e destino L3 forem mencionadas, o pacote não inicial será roteado com base nessas informações específicas. No entanto, isso só é útil quando não há outra sequência antes que ela corresponda às mesmas informações de L3.

```
Nexus# show ip access-lists UDP_Traffic

IP access list UDP_Traffic

10 permit udp host 1.1.1.1 host 3.3.3.3 eq 9201
20 permit udp any any eq 9202

30 permit udp any any eq 9203
```

- O caminho da origem para o destino pode ser verificado para verificar o MTU de modo que o pacote não seja fragmentado.
- A solução alternativa de aplicar outra sequência permite que o UDP acima da sequência problemática funcione, no entanto, o comportamento é o mesmo como explicado anteriormente quando a sequência 45 foi aplicada

```
Nexus# sh route-map In_to_Out pbr-statistics

route-map In_to_Out, permit, sequence 3

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 5

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 7

 Policy routing matches: 0 packets
```

```
route-map In_to_Out, permit, sequence 10

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 30

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 35

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 40

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 45----------> Both fragments matched here

 Policy routing matches: 213 packets

route-map In_to_Out, permit, sequence 50

 Policy routing matches: 0 packets

route-map In_to_Out, permit, sequence 70

 Policy routing matches: 0 packets
```

Lista de acesso para a sequência 45:

```
Nexus# sh ip access-lists udptraffic
```

Tráfego udpda lista de acesso IP:

```
permit udp any any eq isakmp
```

Erro de documento: Bug CSCve05428 N7K Doc || ACL no PBR que contém informações de L3 e L4.