IP Rule of Management Network and Service Network

Introduction

IP rule is a policy routing function of Linux systems. Compared with the classic routing algorithms used on the internet that make routing decisions based only on the destination address of packets, IP rule is more flexible which can support more filter attributes for route forwarding. IP rule can select routes by executing some set of policy routing rules and could set priorities of the rules.

Usage of IP Rule

```
Usage: ip rule [ list | add | del ] SELECTOR ACTION
SELECTOR := [ from PREFIX ] [ to PREFIX ] [ tos TOS ][ dev STRING ] [ pref NUMBER ]
ACTION := [ table TABLE_ID ] [ nat ADDRESS ][ prohibit | reject | unreachable ]
          [ flowid CLASSID ]
TABLE_ID := [ local | main | default | new | NUMBER ]
```

IP rule supports configuring **SELECTOR** of the following attributes for choosing a forwarding path:

From - source address

To - destination address (here we can choose the rules, also used to search the routing entry)

Tos - TOS (type of sevice) field in IP header

Dev - physical interface

Fwmark - firewall parameters

IP rule supports configuring the **ACTION** on how to process the packets if the rule selector matches:

Table - the routing table identifier to lookup if the rule selector matches

Nat - translate the source address of the IP packet into some other value

Prohibit - drop the packets and generate a 'Communication is administratively prohibited' error
Reject - drop the packets

Unreachable - drop the packets and generate a 'Network is unreachable' error

Policy Routing Rules

Linux supports up to 255 routing tables, each routing table has its own table name and table ID. IP rule action defines tables to lookup if the rule selector matches. IP rule also defines the priority parameter which indicates the priority of this rule. Higher number means lower priority, and rules get processed in order of increasing number. Each rule should have an explicitly set unique priority value.

When executing **ip rule** command on Linux shell, we can find all the IP rules of the current system.

```
admin@Xorplus$ip rule
1000:  from all lookup [l3mdev-table]
1500:  from all lookup local
2000:  from 10.10.51.142 lookup main
2001:  from all to 10.10.51.142/24 lookup main
2010:  from all lookup 252
32766: from all lookup main
32767:  from all lookup default
```

By default, the kernel has three rules setting:

- Priority: 1500, Selector: match anything, Action: lookup routing table **local** (ID 255). The local table is a special routing table containing high priority control routes for local and broadcast addresses.
- Priority: 32766, Selector: match anything, Action: lookup routing table **main** (ID 254). The main table is the normal routing table containing all non-policy routes and all the management network routes.
- Priority: 32767, Selector: match anything, Action: lookup routing table **default** (ID 253). The default table is empty. It is reserved for some post-processing if no previous default rules selected the packet.

On the basis of the default rules, PICOS adds three new rules before the rule with priority 32766.

- Priority: 1000, Selector: match anything, Action: lookup routing table **l3mdev-table**. The **l3mdev-table** is a VRF associated routing table.
- Priority: 2000, Selector: match packets from all source to destination address of **eth0_subnet**, Action: lookup routing table main (ID 254). The **eth0_subnet** represents the subnet address of eth0 interface, for example, if the IP address of eth0 interface is 10.10.51.195, then eth0_subnet will be 10.10.51.195/24.
- Priority: 2001, Selector: match from source address of packets **eth0_address**, Action: lookup routing table **main** (ID 254). The **eth0_address** represents the IP address of eth0 interface, for example, 10.10.51.195.
- Priority: 2010, Selector: match anything, Action: lookup routing table **252** (ID 252, both table name and table ID are 252). The **252** table contains all the IPv4 service network routes.

NOTE:

- If data packets match the routes in both 252 table and main table, the routing table entries in 252 table are used preferentially for route forwarding as the priority of 252 table is higher than the main table.
- 252 table only supports IPv4 routing entries, IPv6 routing entries are still in the main table.

Example

Here is an example explaining how IP rule works on management network routes and service network routes.

1. Configure IP addresses for service port and eth0 management port.

   #Configure the IP address for service port.

```
admin@Xorplus# set vlans vlan-id 3
admin@Xorplus# set interface gigabit-ethernet te-1/1/2 family ethernet-switching native-vlan-id 3
admin@Xorplus# set l3-interface vlan-interface vlan-3 address 192.168.2.1 prefix-length 24
admin@Xorplus# set vlans vlan-id 3 l3-interface vlan-3
admin@Xorplus# commit
Commit OK.
Save done.
```

   #Assign an IP address to the eth0 management port by default method of DHCP. Use **ifconfig eth0** command to find the IP address of eth0.

```
admin@Xorplus$ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:18:23:30:dd:52
          inet addr:10.10.51.142  Bcast:10.10.51.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:6866 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1622 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:889873 (869.0 KiB)  TX bytes:200121 (195.4 KiB)
```

2. Configure the next hop of 10.10.20.0/24 as the IP address of the service network segment.

```
admin@Xorplus# set protocols static route 10.10.20.0/24 next-hop 192.168.2.5
admin@Xorplus# commit
Commit OK.
Save done.
```

Check the routing table. The above routing entry is only in 252 table and not in the main table because the next hop is the IP address of the service network segment.

```
admin@Xorplus# run show route ipv4
IPv4 Routing table: 3 routes
10.10.20.0/24      [static(1)/1]
                   > to 192.168.2.5 via vlan-3/vlan-3
192.168.2.1/32     [local(0)/0]
                   > via vlan-3/vlan-3
192.168.2.0/24     [connected(0)/0]
                   > via vlan-3/vlan-3
Codes: K - kernel route, C - connected, S - static, R - RIP,
       O - OSPF, I - IS-IS, B - BGP, E - EIGRP, N - NHRP,
       T - Table, v - VNC, V - VNC-Direct, A - Babel, D - SHARP,
       F - PBR, f - OpenFabric,
       > - selected route, * - FIB route, q - queued route, r - rejected route

C>* 10.10.51.0/24 is directly connected, eth0, 01:33:30
S>* 10.10.20.0/24 [1/0] via 192.168.2.5, vlan3, weight 1, 01:05:18
C>* 192.168.2.0/24 is directly connected, vlan3, 01:05:18
admin@Xorplus# quit
admin@Xorplus> quit

admin@Xorplus$ip route list table 252
10.10.20.0/24 via 192.168.2.5 dev vlan.3  proto xorp  metric 1
192.168.2.0/24 via 192.168.2.1 dev vlan.3  proto xorp

root@Xorplus$ip route list table main
default via 10.10.51.1 dev eth0
10.10.51.0/24 dev eth0  proto kernel  scope link  src 10.10.51.142
192.168.2.0/24 dev vlan.3  proto kernel  scope link  src 192.168.2.1
```

3. Configure next hop of default route as IP address of the management network gateway.

NOTE:

The management port does not support the configuration of network segment routing, you can only configure the default route.

```
admin@Xorplus# set protocols static route 0.0.0.0/0 next-hop 10.10.51.1
admin@Xorplus# commit
Commit OK.
Save done.
```

Check the routing table. The above routing entry is only in main table and not in 252 table because the next hop is the IP address of the management network segment.

```
admin@Xorplus$ip route list table main
default via 10.10.51.1 dev eth0
10.10.51.0/24 dev eth0  proto kernel  scope link  src 10.10.51.142
192.168.2.0/24 dev vlan.3  proto kernel  scope link  src 192.168.2.1

admin@Xorplus$ip route list table 252
10.10.20.0/24 via 192.168.2.5 dev vlan.3  proto xorp  metric 1
192.168.2.0/24 via 192.168.2.1 dev vlan.3  proto xorp
```

4. Configure the next hop of default route as the IP address of the service network segment.

```
admin@Xorplus# set protocols static route 0.0.0.0/0 next-hop 192.168.2.88
admin@Xorplus# commit
Commit OK.
Save done.
admin@Xorplus# quit
admin@Xorplus> quit
```

Check the routing table. The above routing entry is only in 252 table and not in main table because the next hop is the IP address of the service network segment.

```
admin@Xorplus$ip route list table 252
default via 192.168.2.88 dev vlan.3  proto xorp  metric 1
10.10.20.0/24 via 192.168.2.5 dev vlan.3  proto xorp  metric 1
192.168.2.0/24 via 192.168.2.1 dev vlan.3  proto xorp

admin@Xorplus$ip route list table main
default via 10.10.51.1 dev eth0
10.10.51.0/24 dev eth0  proto kernel  scope link  src 10.10.51.142
192.168.2.0/24 dev vlan.3  proto kernel  scope link  src 192.168.2.1
```

There are default routing entries in both 252 table and main table, the default routing entry in the main table is automatically generated by the system when assigning the IP address by DHCP. When the packet matches no routing entry in the routing table, it will then match the default routing entry. In this case, the default routing entry in 252 table is used preferentially for route forwarding as the priority of 252 table is higher than the main table.

5. If the source IP address carried in a packet is empty and the packet matches no routing entry in the routing table, the default route in the 252 table and the service port is used for packet forwarding.

#For example, ping 10.10.50.22 without source IP.

```
admin@Xorplus$ping 10.10.50.22
PING 10.10.50.22 (10.10.50.22) 56(84) bytes of data.
From 192.168.2.1 icmp_seq=1 Destination Host Unreachable
From 192.168.2.1 icmp_seq=2 Destination Host Unreachable
^C
--- 10.10.50.22 ping statistics ---
9 packets transmitted, 0 received, +8 errors, 100% packet loss, time 8003ms
pipe 4
```

When the source address carried in a packet is the IP address of the eth0 management interface, the packet will match the IP rule: "*2000: from 10.10.51.142 lookup main*". For example, ping 10.10.50.22 with source IP 10.10.51.142.

```
admin@Xorplus$ping -I 10.10.51.142 10.10.50.22
PING 10.10.50.22 (10.10.50.22) from 10.10.51.142 : 56(84) bytes of data.
64 bytes from 10.10.50.22: icmp_req=1 ttl=63 time=0.183 ms
64 bytes from 10.10.50.22: icmp_req=2 ttl=63 time=0.153 ms
^C
--- 10.10.50.22 ping statistics ---
11 packets transmitted, 11 received, 0% packet loss, time 9999ms
rtt min/avg/max/mdev = 0.139/0.151/0.183/0.014 ms
```