====== Selective Packet Discard ======

===== 简介 =====
当路由协议数据包、管理数据包、keepaive等信息进入路由器时需要RP（Route Processor）来处理，或者说目的地址是路由器本身时，也需要由RP来处理。当有针对路由器自身的dos攻击时，如果所有信息都有RP处理，很容易导致路由器瘫痪。此时可通过设置selective packet discard来丢弃一些恶意的数据包，来保证设备的稳定运行。

 * SPD默认是enable的
 * SPD最初只是为pos口设计的，但后来GE口也可以使用spd技术

===== 支持SPD的设备 =====
 * 7200 Series Router
 * 7500 Series Router
 * 12000 Series Router

===== SPD原理 =====
SPD可通过2种方式丢弃数据包：
 * SPD State Check
 * Input Queue Check

==== SPD State Check ====
所有到RP的数据包可分为2类：
 * 如果进入priority queue的，并且priority为7和6的，永远都不会被drop掉
 * 其他数据包被放入general packet queue，并进行spd state check

对于进入general packet queue的数据包，也就是进行spd state check的数据包会进行如下处理:
 * 如果queue的长度小于min-threshold，正常包和畸形包都不会被drop掉
 * 如果queue的长度在min-threshold和max-threshold之间
   * 如果是normal mode，正常包和畸形包会被随机的丢弃
   * 如果是aggresive mode，所有畸形包会被丢弃
 * 如果queue的长度大于max-threshold，那么所有正常包和畸形包都会被drop掉

=== aggressive mode ===

＊如果spd工作在aggressive mode，所有的畸形包会被丢弃，例如invalid checksum、incorrect version、incorrect header length、incorrect packet length等。

＊通过命令ip spd mode aggressive开启aggressive mode

＊12000系列路由器不支持aggressive mode，因为畸形包在会被每个linecard丢弃，而不需要由GRP（gigabit route processor）处理

==== Input Queue Check ====

=== input queue ===

SPD state check是基于RP的，而Input Queue Check是基于interface的。如果不开启spd的话，默认情况下每个interface的queue是75，当queue中的数据大于75时，大于75的部分会被丢弃。这个queue可以通过show interface看到。

```
GigabitEthernet1/2 is up, line protocol is up
 Hardware is GigMac 3 Port GigabitEthernet, address is 0005.5ffd.4082 (bia 0005.5ffd.4082)
 Description: sample
 Internet address is x.x.x.x/30
 MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec, rely 255/255, load 131/255
 Encapsulation ARPA, loopback not set
 Keepalive set (10 sec)
 Full Duplex, 1000Mbps, link type is force-up, media type is LX
 output flow-control is unsupported, input flow-control is unsupported
 ARP type: ARPA, ARP Timeout 04:00:00
 Last input 00:00:00, output 00:00:03, output hang never
 Last clearing of "show interface" counters never
 Queueing strategy: fifo
 Output queue 0/40, 0 drops; input queue 0/75, 25 drops  //[就在这里]
 30 second input rate 613917000 bits/sec, 122041 packets/sec
 30 second output rate 517166000 bits/sec, 123695 packets/sec
    77400124545 packets input, 44369025705444 bytes, 0 no buffer
    Received 5898 broadcasts, 0 runts, 0 giants, 0 throttles
    647964 input errors, 0 CRC, 0 frame, 485923 overrun, 162041 ignored
    0 watchdog, 0 multicast, 0 pause input
    69912443364 packets output, 41951561990047 bytes, 0 underruns
```

Transmitted 1 broadcasts
0 output errors, 0 collisions, 0 interface resets
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 pause output
0 output buffer failures, 0 output buffers swapped out

如果需要修改这个queue的长度，可通过一下命令修改

O-HPM-GSR-1(config-if)#hold-queue ?
<0-4096> Queue length

O-HPM-GSR-1(config-if)#hold-queue 100 ?
in Input queue
out Output queue

O-HPM-GSR-1(config-if)#hold-queue 100 in
O-HPM-GSR-1(config-if)#hold-queue 100 in ?

O-HPM-GSR-1(config-if)#hold-queue 100 in

=== headroom ===

如果开启了SPD，那么priority为7和6的数据包会进入process level input queue（这个queue的名字叫headroom），而其他的数据包仍然会放在interface input queue里。process level input queue的大小默认为100。也就是说当interface总的queue长度175被用满后，priority是7和6的数据包就会被丢弃了。对于GSR来说，这个process level input queue的长度默认是1000，这是由于clear ip bgp时会有很多packet进来，如果还是100的话，很多bgp包会被丢弃，这样就会影响网络收敛的速度。

=== extended headroom ===
由于ospf、isis、ppp、clns这类igp和2层链路间的keepalive的priority和bgp一样，如果在一个很大的bgp网络中，bgp的packet会比igp的多的多，那么他会大量的占据headroom，这就有可能导致igp的中断、或者直接在layer 2链路down掉。因此对于这样的数据包，默认再分配一个值为10的extended headroom，来保证igp和layer2 link的正常工作。

O-HPM-GSR-1#sho ip spd
Current mode: normal.

Queue min/max thresholds: 73/74, Headroom: 1000, Extended Headroom: 10

IP normal queue: 0, priority queue: 0.

SPD special drop mode: none