
TOMORROW
starts here.
CISCO
Cisco live!

SP Security using BGP FlowSpec

BRKSPG-2618

Nicolas Fevrier, SP Routing Technical Marketing Engineer

#clmel

Cisco live!

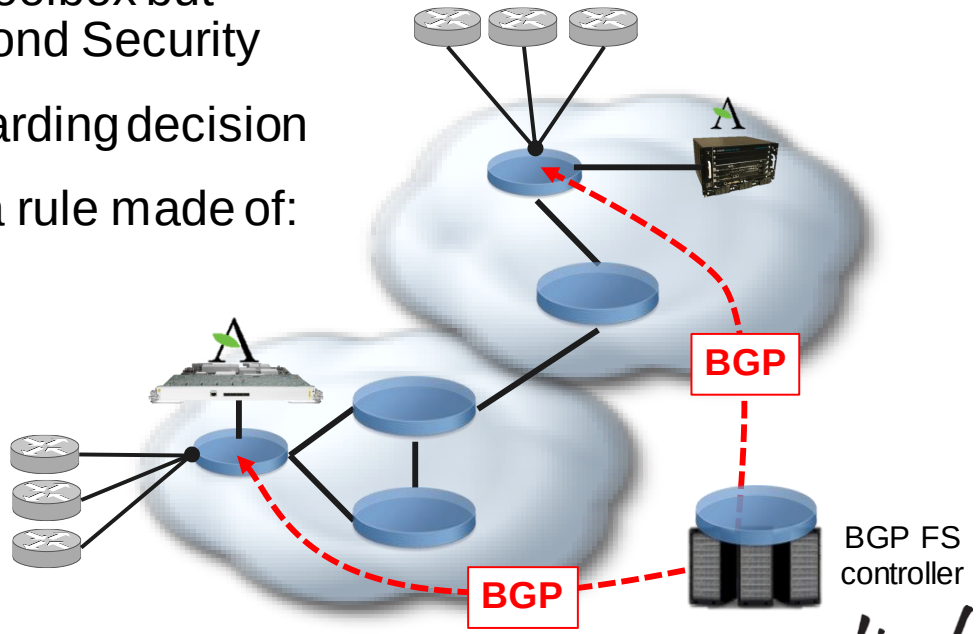
Agenda

- Introduction
- BGP FS Protocol Description
- Use-cases for DDoS Mitigation
- Other Use-cases
- Configuration, Troubleshooting and Monitoring
- Conclusion

Introduction

Introduction to BGP FlowSpec

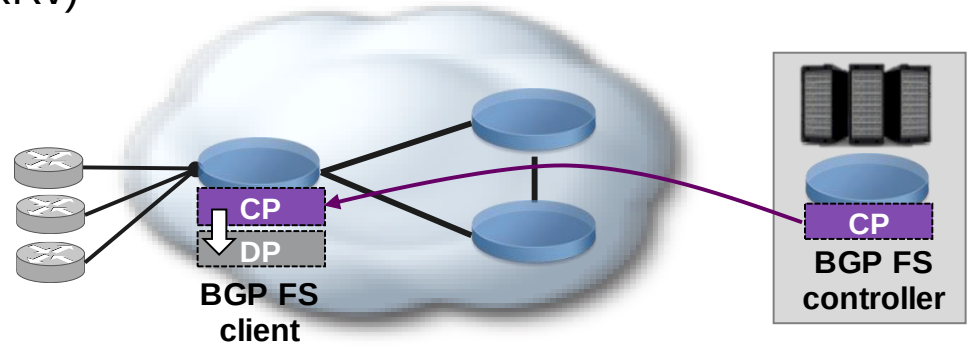
- Separation of controlling and forwarding plane. Sounds familiar ?
- A powerful tool in the SP Security toolbox but Use-cases are expending way beyond Security
- A remote controller programs forwarding decision
- BGP is used to program remotely a rule made of:
  - A traffic description
  - An action to apply on this traffic
- Three elements:
  - Controller
  - Client
  - Optional Route-reflector



BGP FlowSpec Components

Controller

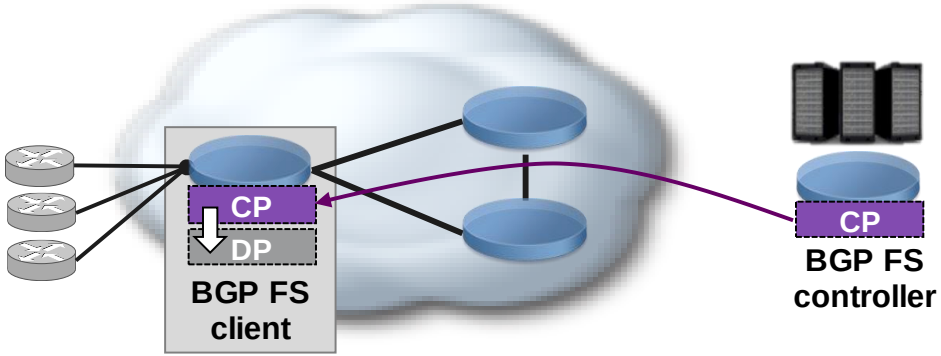
- Injects rules remotely in the clients
- Needs to implement at the minimum the Control Path (CP)
- Examples of BGP FS Controllers:
  - router (ASR9000, CRS, NCS6000, XR12000, …)
  - server (ExaBGP, Arbor Peakflow SP Collector Platform, …)
  - virtual router (XRv)



BGP FlowSpec Components

Client

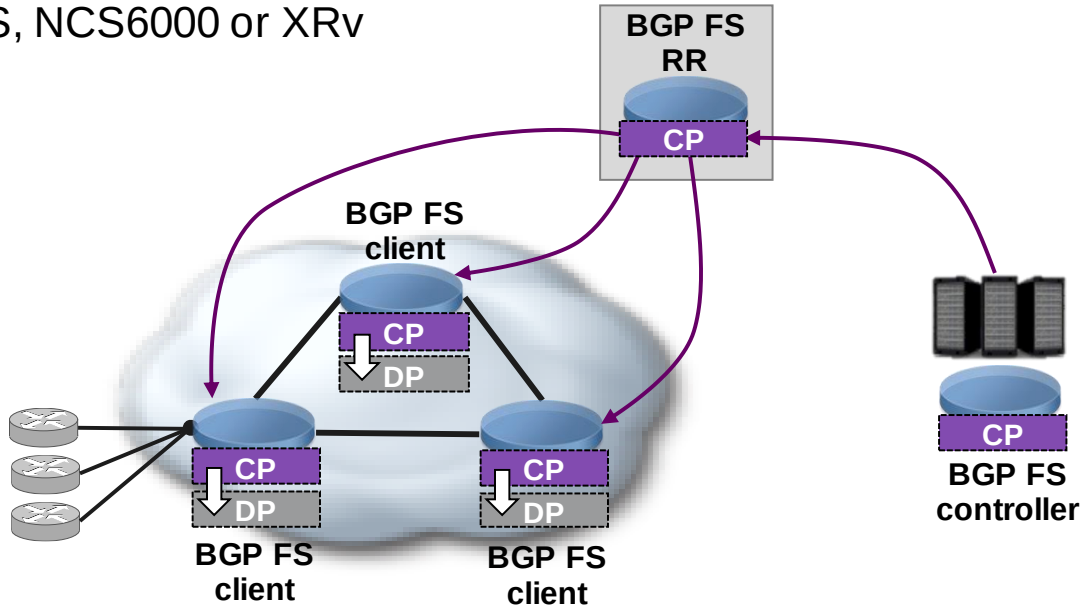
- Receives rules from Controller(s) and programs the match/actions in hardware
- Needs to implement both Control Plane (CP) and Data Plane (DP)
- Examples of BGP FS Clients:
  - router (ASR9000, CRS, NCS6000 …)



BGP FlowSpec Components

Route-Reflector

- Receives rules from Controller(s) and distributes them to Clients
- Examples of BGP FS Router-Reflector:
  - ASR9000, CRS, NCS6000 or XRv



BGP FlowSpec Protocol Description

Cisco live!

RFC5575

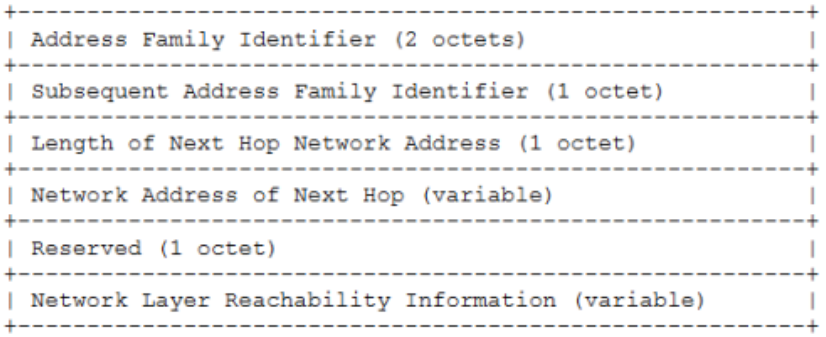
Dissemination of Flow Specification Rules

- Why using BGP?
  - Simple to extend by adding a new NLRI
    - MP_REACH_NLRI / MP_UNREACH_NLRI
  - Already used for every other kind of technology
    - IPv4
    - IPv6
    - VPN
    - Multicast
    - Labels
    - Etc…
  - Point to multipoint with Route-Reflectors
  - Inter-domain support
  - Networking engineers and architects understand perfectly BGP

RFC5575

Dissemination of Flow Specification Rules: Traffic Matching

- New NLRI defined (AFI=1, SAFI=133) to describe the traffic of interest
  1. Destination IP Address (1 component)
  2. Source IP Address (1 component)
  3. IP Protocol (+1 component)
  4. Port (+1 component)
  5. Destination port (+1 component)
  6. Source Port (+1 component)
  7. ICMP Type
  8. ICMP Code
  9. TCP Flags
  10. Packet length
  11. DSCP
  12. Fragment



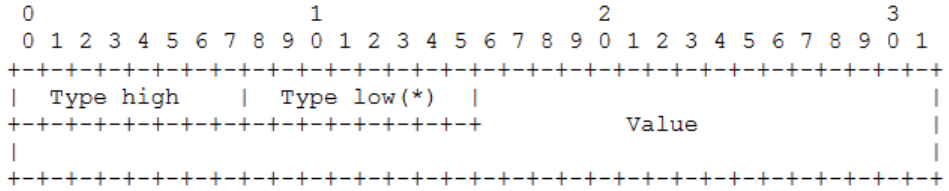
The MP_REACH_NLRI – RFC 4760

From RFC 5575: “Flow specification components must follow strict type ordering. A given component type may or may not be present in the specification, but if present, it MUST precede any component of higher numeric type value.”

RFC5575

Dissemination of Flow Specification Rules: Actions

- Traffic Action is defined in extended-communities (RFC4360)



Type	Description	Encoding
0x8006	Traffic-rate	2 bytes ASN; 4 bytes as float
0x8007	Traffic-action	Bitmask
0x8008	Redirect	6 bytes RT (Route Target)
0x8009	Traffic-marking	DSCP Value

IETF Drafts

Extensions for RFC5575

On top of the RFC implementation, our XR routers supports:

- IPv6 extensions: draft-ietf-idr-flow-spec-v6-03
- Redirect IP extension: draft-simpson-idr-flowspec-redirect-02
- IBGP extension: draft-ietf-idr-bgp-flowspec-oid-01
- Persistence Support: draft-uttaro-idr-bgp-persistence-02 (in IOS XR5.2.2)
- HA/NSR Support

Cisco IOS XR Routers BGP FS Implementation

Platform Hardware	Control Plane Support	Data Plane Support
ASR9k – Typhoon LC	5.2.0	5.2.0
ASR9k – Thor LC	5.2.0	5.2.2
ASR9001	5.2.0	5.2.2
ASR9k – Tomahawk	Target 5.3.x	Target 5.3.x
CRS – Taiko LC	5.2.0	5.2.0
CRS – Topaz LC	5.2.0	Target 5.3.1
XRVR	5.2.0	N.A.
C12K	5.2.0	Not planned
NCS6000	Target 5.2.3/5.2.4	Target 5.2.3 (EFT) /5.2.4

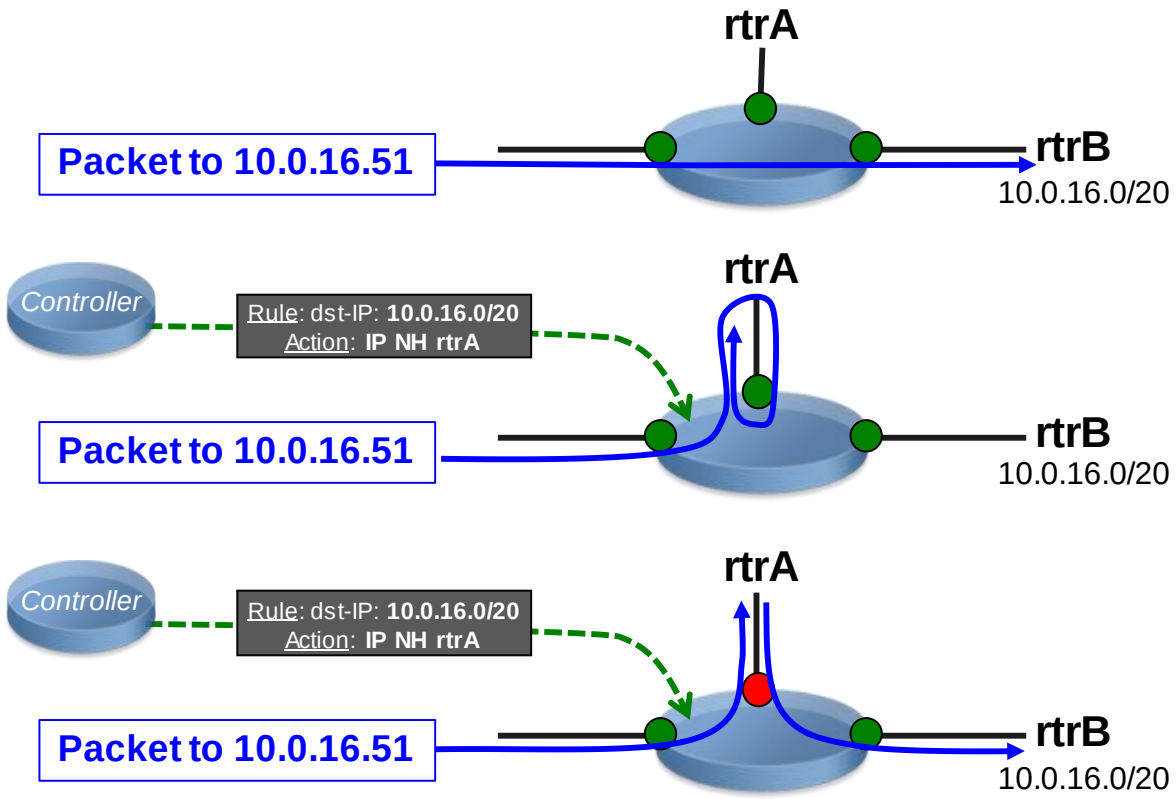
Cisco IOS XR Routers BGP FS Implementation

NLRI type	Match fields	Value input method	XR PI	ASR9000	CRS	NCS6000
Type 1	IPv4 Destination address	Prefix length	✓	✓	✓	✓
Type 2	IPv4 Source address	Prefix length	✓	✓	✓	✓
Type 3	IPv4 protocol	Multi value range	✓	✓	✓	✓
Type 4	IPv4 source or destination port	Multi Value range	✓	✓	✓	✓
Type 5	IPv4 destination port	Multi Value range	✓	✓	✓	✓
Type 6	IPv4 Source port	Multi Value range	✓	✓	✓	✓
Type 7	IPv4 ICMP type	Multi value range	✓	✓	✓	✓
Type 8	IPv4 ICMP code	Multi value range	✓	✓	✓	✓
Type 9	IPv4 TCP flags (2 bytes include reserved bits)	Bit mask	✓	Only Lower byte Reserved and NS bit not supported	Only Lower byte Reserved and NS bit not supported	Only Lower byte Reserved and NS bit not supported
Type 10	IPv4 Packet length	Multi value range	✓	✓	✓	✓
Type 11	IPv4 DSCP	Multi value range	✓	✓	✓	✓
Type 12	IPv4 fragmentation bits	Bit mask	✓	Only indication of fragment	✓	✓

Cisco IOS XR Routers BGP FS Implementation

NLRI type	Match fields	Value input method	XR PI	ASR9000	CRS	NCS6000
Type 1	IPv6 Destination address	Prefix length	✓	✓	✓	✓
Type 2	IPv6 Source address	Prefix length	✓	✓	✓	✓
Type 3	IPv6 Next Header	Multi value range	✓	✓	✓	✓
Type 4	IPv6 source or destination port	Multi Value range	✓	✓	✓	✓
Type 5	IPv6 destination port	Multi Value range	✓	✓	✓	✓
Type 6	IPv6 Source port	Multi Value range	✓	✓	✓	✓
Type 7	IPv6 ICMP type	Multi value range	✓	✓	✓	✓
Type 8	IPv6 ICMP code	Multi value range	✓	✓	✓	✓
Type 9	IPv6 TCP flags (2 bytes include reserved bits)	Bit mask	✓	Only Lower byte Reserved and NS bit not supported	Only Lower byte Reserved and NS bit not supported	Only Lower byte Reserved and NS bit not supported
Type 10	IPv6 Packet length	Multi value range	✓	✓	✓	✓
Type 11	IPv6 Traffic Class	Multi value range	✓	✓	✓	✓
Type 12	Reserved	N/A	N/A	N/A	N/A	N/A
Type 13	IPv6 Flow Based (20 bytes)	Multi value range	✘	✘	✘	✘

IOS XR Implementation Improvements



BGP FlowSpec Enabled

BGP FlowSpec Disabled

BGP FS is applied to the whole router but can be activated or deactivated on particular interfaces via CLI configuration. Particularly useful in Distributed DDoS mitigation architecture.

IOS XR Implementation

Application on Interface

- In current implementation, rules are applied in ingress physical or logical interfaces (Link-bundles and dot1q) but not on tunnels
- Up to 3000 simple rules per line card using the TCAM. When the rules are complex using multi-value ranges for BGP tuples, it will consume more TCAM cells and will reduce overall scale
- Scale of other TCAM based features like ACL, QOS in the linecard will decrease the space available for BGP flowspec

IOS XR Implementation

Application on Interface

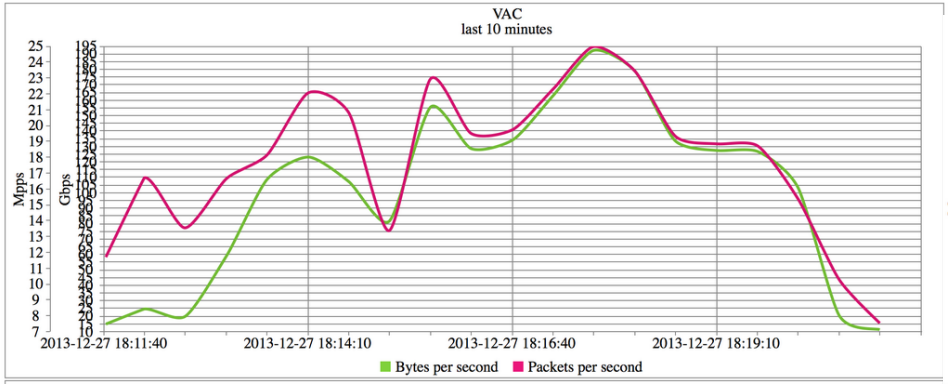
- Uses the PBR infrastructure with similar performance penalty than other PBR features like ABF. Performance cost will vary depending upon the action
  - DSCP marking will be least expensive
  - redirect action pointing to recursive TE tunnel path being most expensive
- Can coexist with other features like QoS or ACL (and sharing TCAM space) but not with other PBR features applied on the same interface
- Interface can be in the Global Routing Table or on a VRF (L3VPN or VRF-Lite)

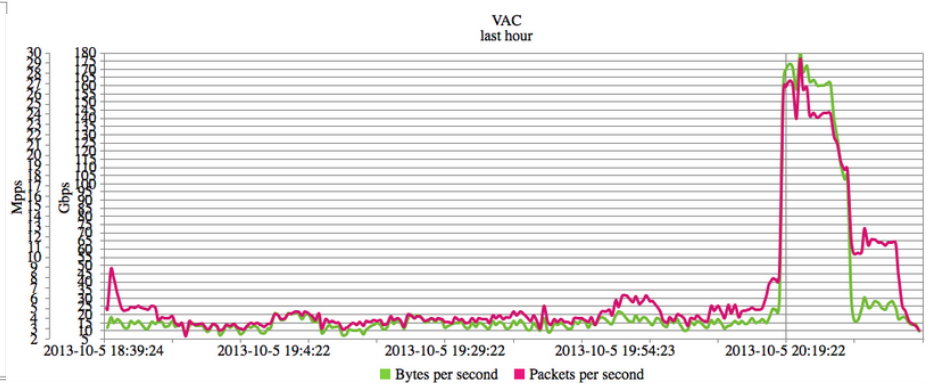
Use-cases: DDoS Mitigation

Cisco live!

DDoS Attacks

- No longer necessary to explain the risk
  - Distributed Denial of Service (DDoS) is a lucrative activity for attackers
  - ISP, Hosting Services, Enterprises: it can jeopardise your business. Everyone is at risk.
  - Just scratching the surface, attack complexity is increasing
- DDoS Mitigation is about business continuity

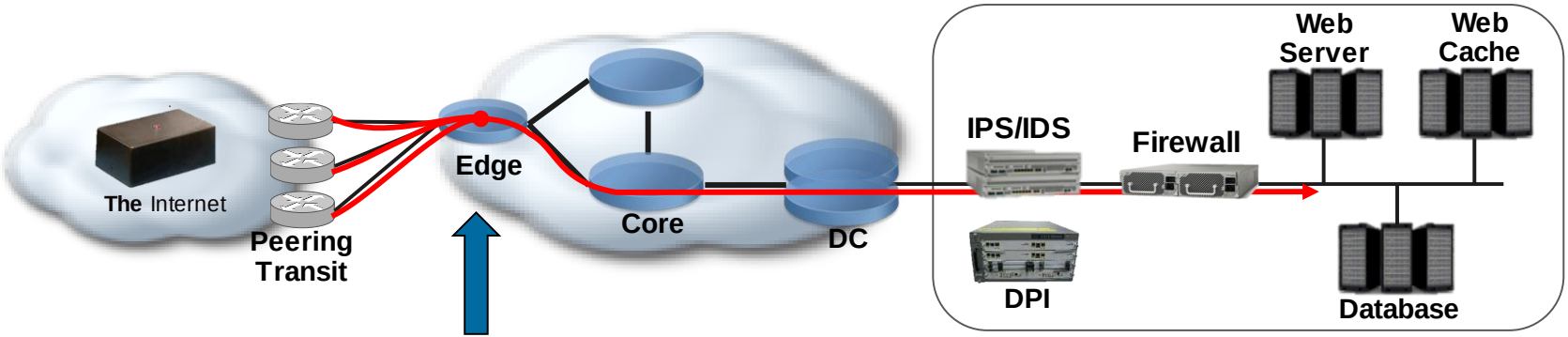




https://twitter.com/olesovhcom/status/416667262146195456/photo/1

DDoS Attacks

- Denial of Service attacks are of different natures:
  - Application-layer attacks
    - Detected and handled by Firewalls, IDS or at the Server level
  - Volumetric attacks (including Protocols attacks)
    - Can NOT be mitigated in data centre or server farm (too late)
    - Should be handled in the backbone or at the border



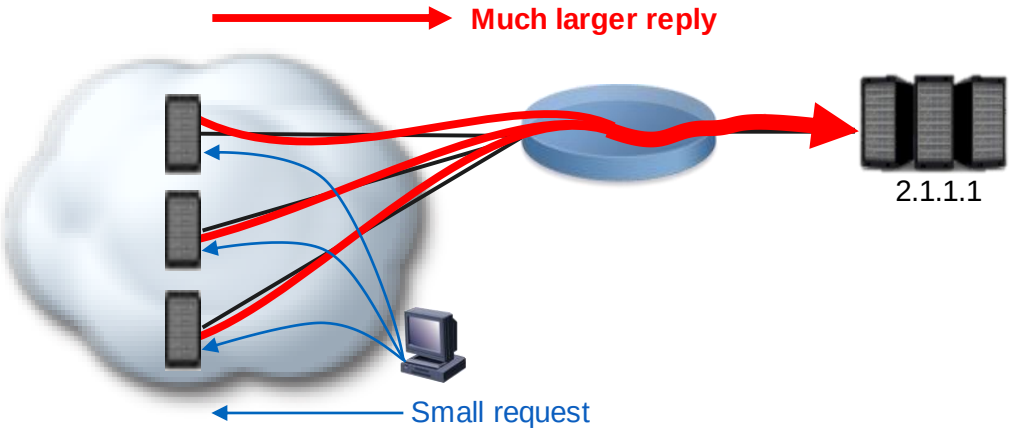
DDoS Attacks Mitigation

- BGP FS was initially designed with DDoS Mitigation use-case in mind
- Distributed attack received from all transit and peering points
- We use a mitigation system in a VSM card or an appliance connected to our IOS-XR router
- We differentiate arbitrarily three DDoS attack families:
  - Stateless Amplification
  - Stateless L3 / others
  - Stateful / up to L7 on application resources

DDoS Mitigation with BGP FS

Rate-limiting / Filtering Stateless Attacks: Amp Attacks

- Stateless attacks are not using a full handshake and are based on spoofed source addresses
- First example: Amplification attacks using vulnerable protocols on high bandwidth servers

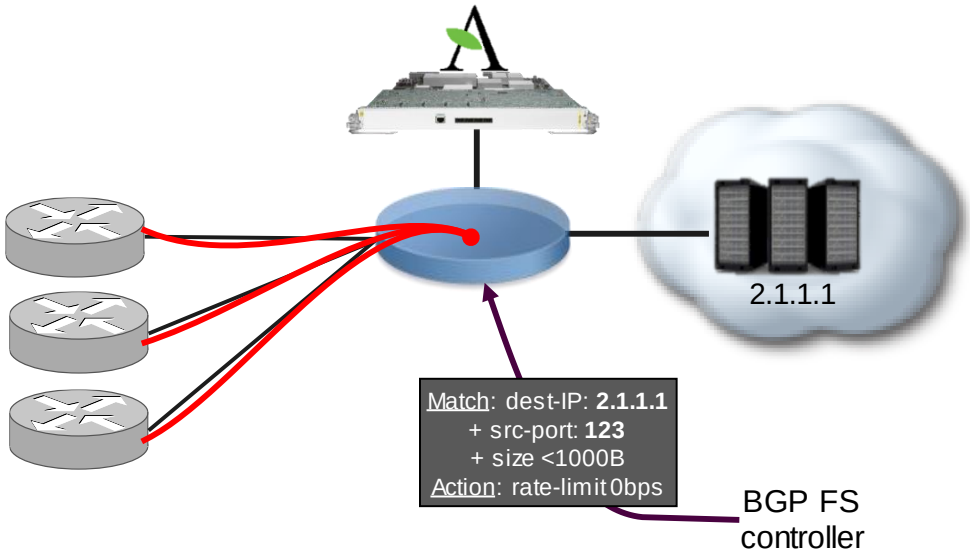


– DNS
– NTP
– CharGen
– SNMP
– …

DDoS Mitigation with BGP FS

Rate-limiting / Filtering Stateless Attacks: Amp Attacks

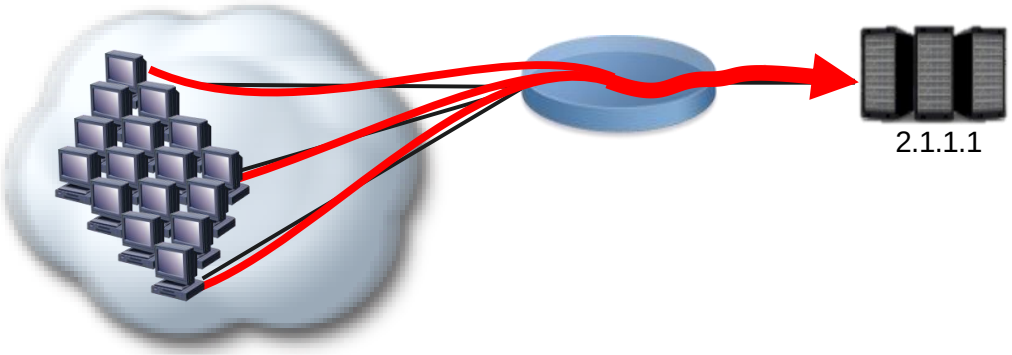
- Amplification attacks, example NTP
  - Don't need to be sent to a "smart" scrubbing system to be mitigated
  - Identified by precisely matching the traffic pattern and filtered at the edge router level



DDoS Mitigation with BGP FS

Rate-limiting / Filtering Stateless Attacks: L3/L4 Protocol Attacks

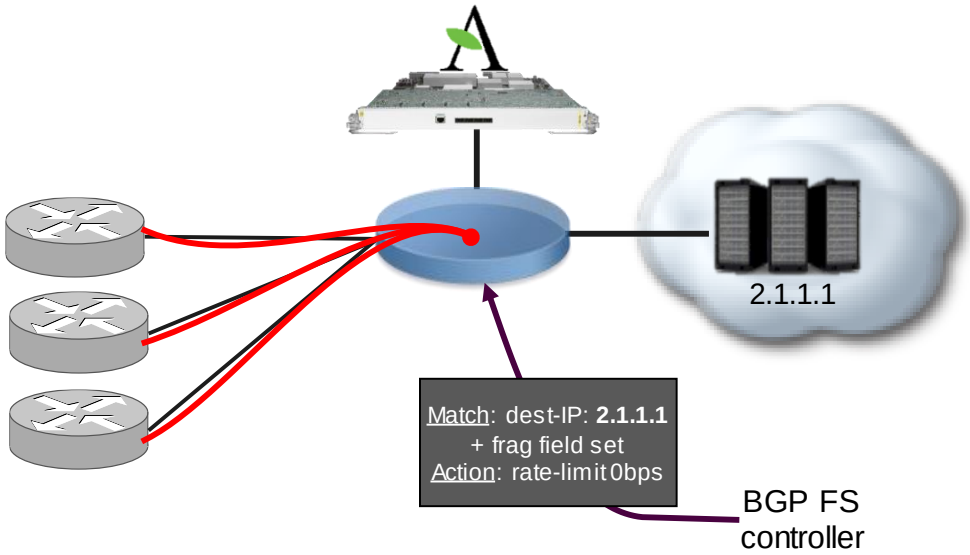
- Generic family covering
  - TCP SYN
  - UDP Frag
  - ICMP Flood
- Source address could be forged or not, the botnet members are corrupted hosts



DDoS Mitigation with BGP FS

Rate-limiting / Filtering Stateless Attacks: L3/L4 Protocol Attacks

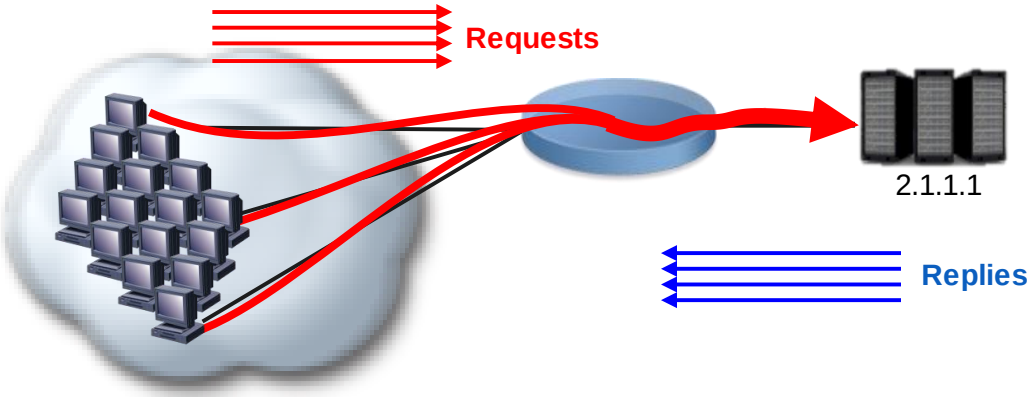
- L3/L4 attacks, like the Amp attacks can be filter at the edge router via BGP FS
- Example with a fragmentation attack:



DDoS Mitigation with BGP FS

Addressing More Sophisticated Attacks: L7

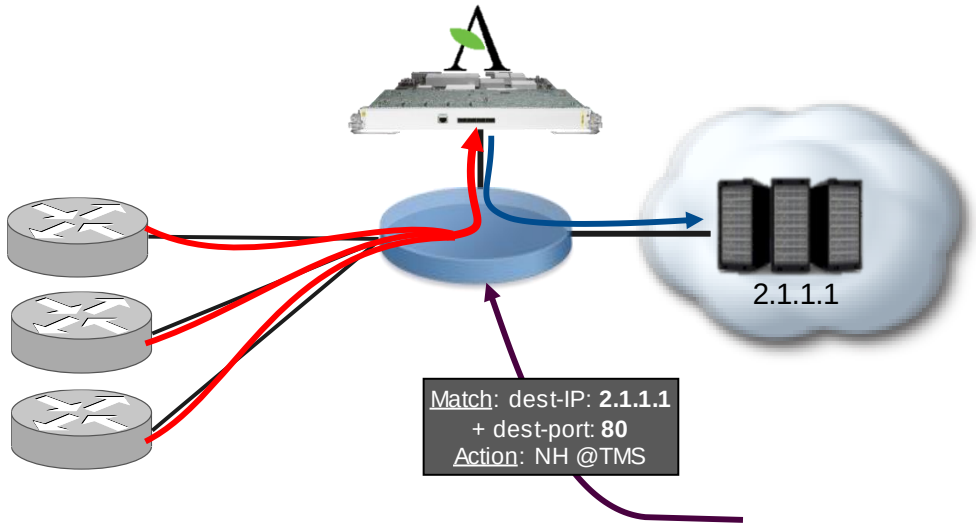
- More advanced attacks using Botnets or even real users (LOIC) needs to be addressed differently by a specific scrubbing device. Examples:
  - HTTP: bots mimicking the behaviour of a real web browser
  - SSL
  - SIP
  - …



DDoS Mitigation with BGP FS

Addressing More Sophisticated Attacks: L7

- BGP FlowSpec will be used to program a different action here
  - Diversion to a next-hop address
  - Diversion to a different VRF



DDoS Mitigation with BGP FS

Benefits

- Single point of control to program rules in many clients
- Allows a very precise description/matching of the attack traffic
- Can be used for both mitigation and diversion of the attack traffic, without impact the course of the rest of the traffic targeted to the victim
- Filtering stateless attacks on the edge route permits mitigation of millions of PPS of dirty traffic while liberating precious CPU cycle on the scrubbing device for more advanced mitigation needs
- The Cisco ASR9000 supports Arbor Peakflow SP TMS software on the VSM service card

DDoS Mitigation on ASR9000

Cisco / Arbor Partnership

- Peakflow SP TMS is an Arbor product, could be embedded in different hardware
  - Arbor Chassis or Appliance, connected to a L3 device

  - ASR9000 Service Card: **VSM**

DDoS Mitigation on ASR9000

Virtualised Service Module

- Supported with
  - RSP440 onwards (not RSP2)
  - All 9000 chassis except **9001**
- Multi-purpose service card
  - CGN
  - IPsec
  - Mobile GW
  - DPI
  - ASAv
  - DDoS Mitigation
- Service chaining
- KVM virtualised environment

DDoS Mitigation Demo

Cisco live!

Improving Existing DDoS Mitigation Models

Cisco live!

DDoS Mitigation Models

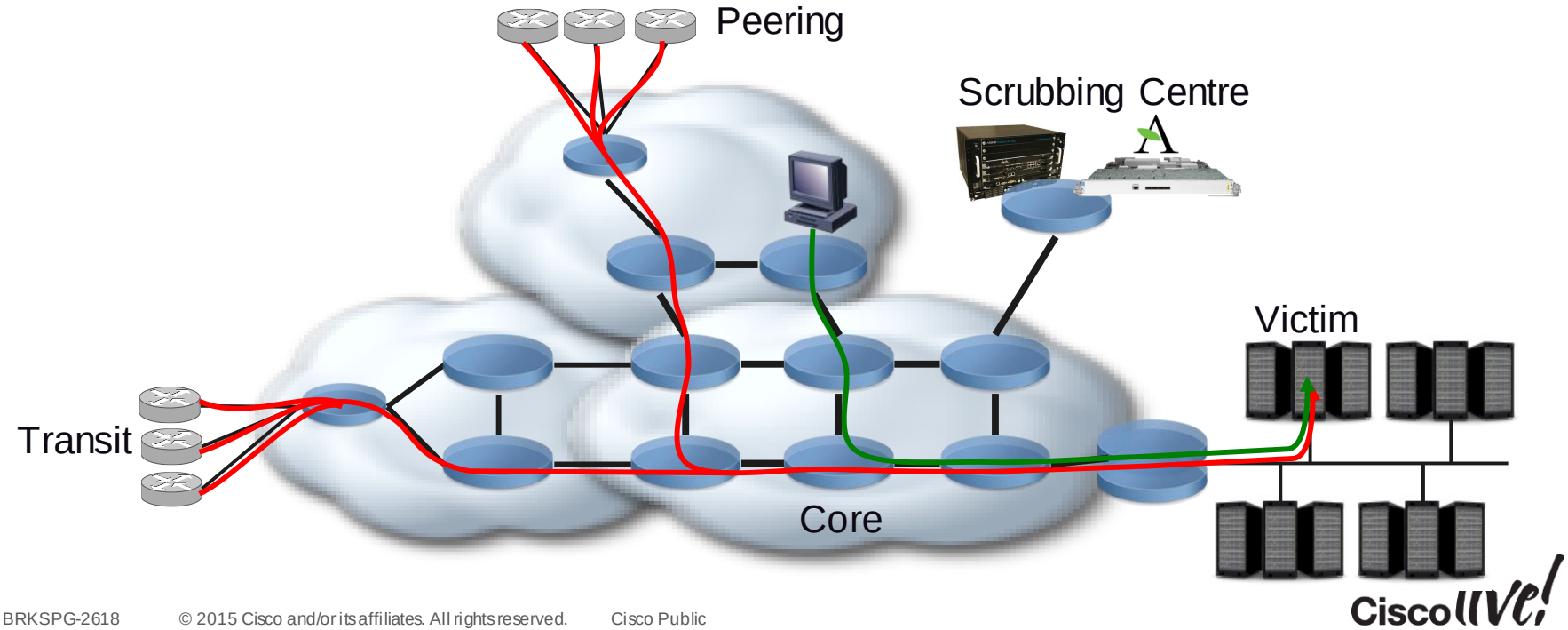
Network Design

- Several approaches exist in the design of a DDoS mitigation solution
- No real “best practices” in this field, it mainly depends on
  - The topology
  - The protocols and services: IP only, MPLS transport, L2/L3VPN
- They all consist in:
  - Diverting the traffic targeted to the victim to push it into scrubbing devices
  - Performing an analysis of the packets to discriminate legit packets from attack packets
  - Re-injecting the legit traffic into the network
- Following examples are real-case used in very large production networks

DDoS Mitigation Models

Centralised

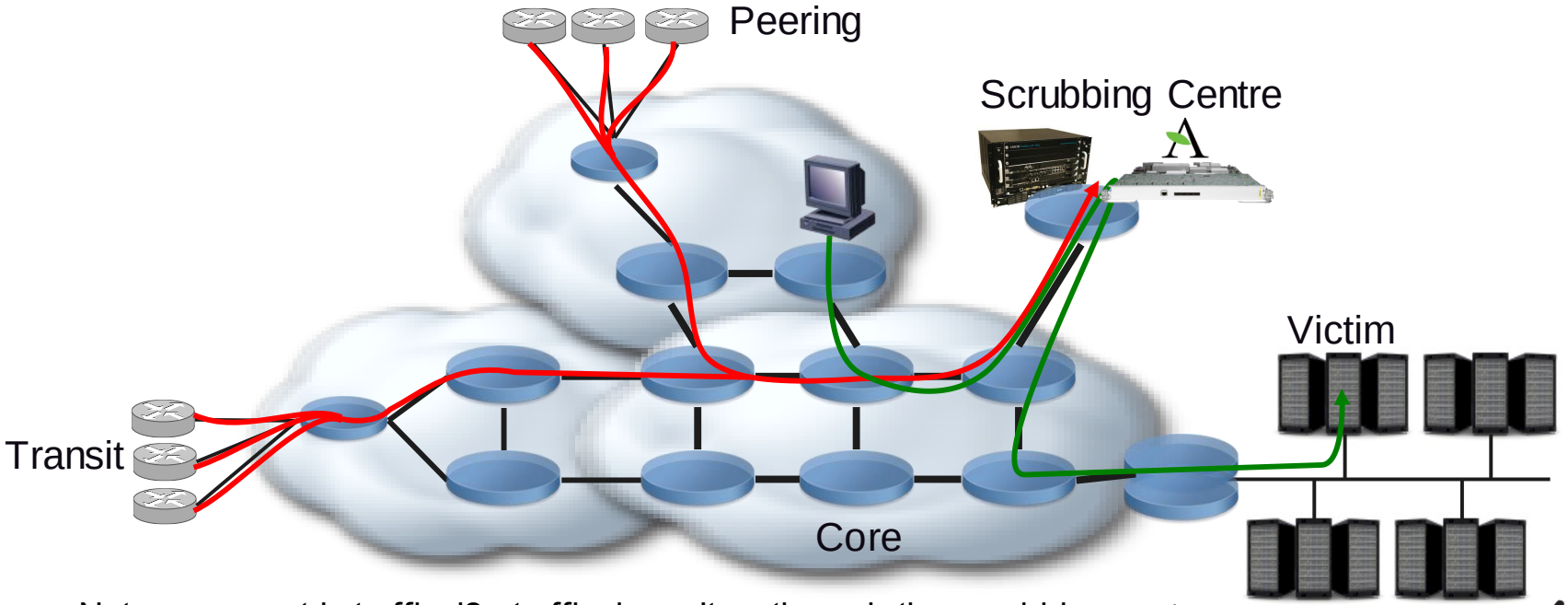
- A central point in the network is dedicated for hosting scrubbing devices



DDoS Mitigation Models

Centralised

- Traffic target to the victim is diverted to this place for analysis

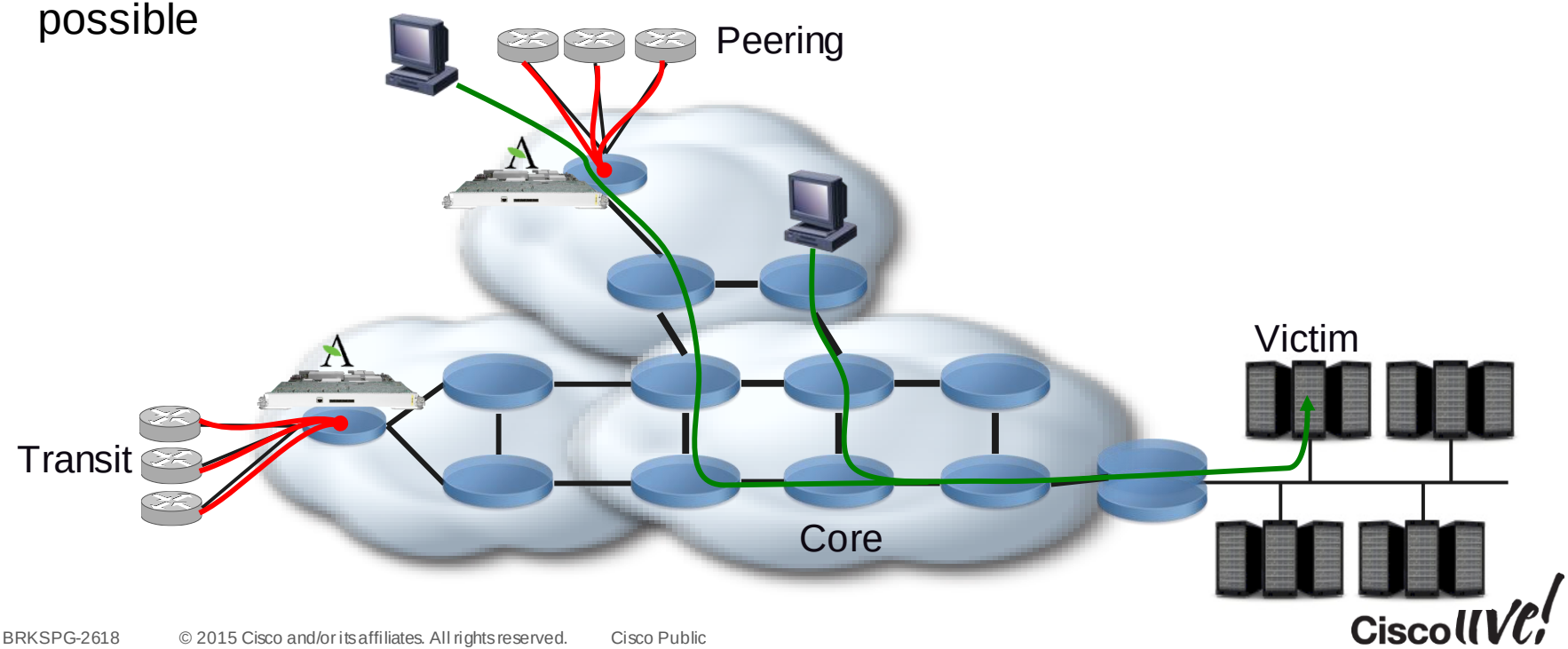


Note: asymmetric traffic, i2o traffic doesn't go through the scrubbing centre

DDoS Mitigation Models

Distributed

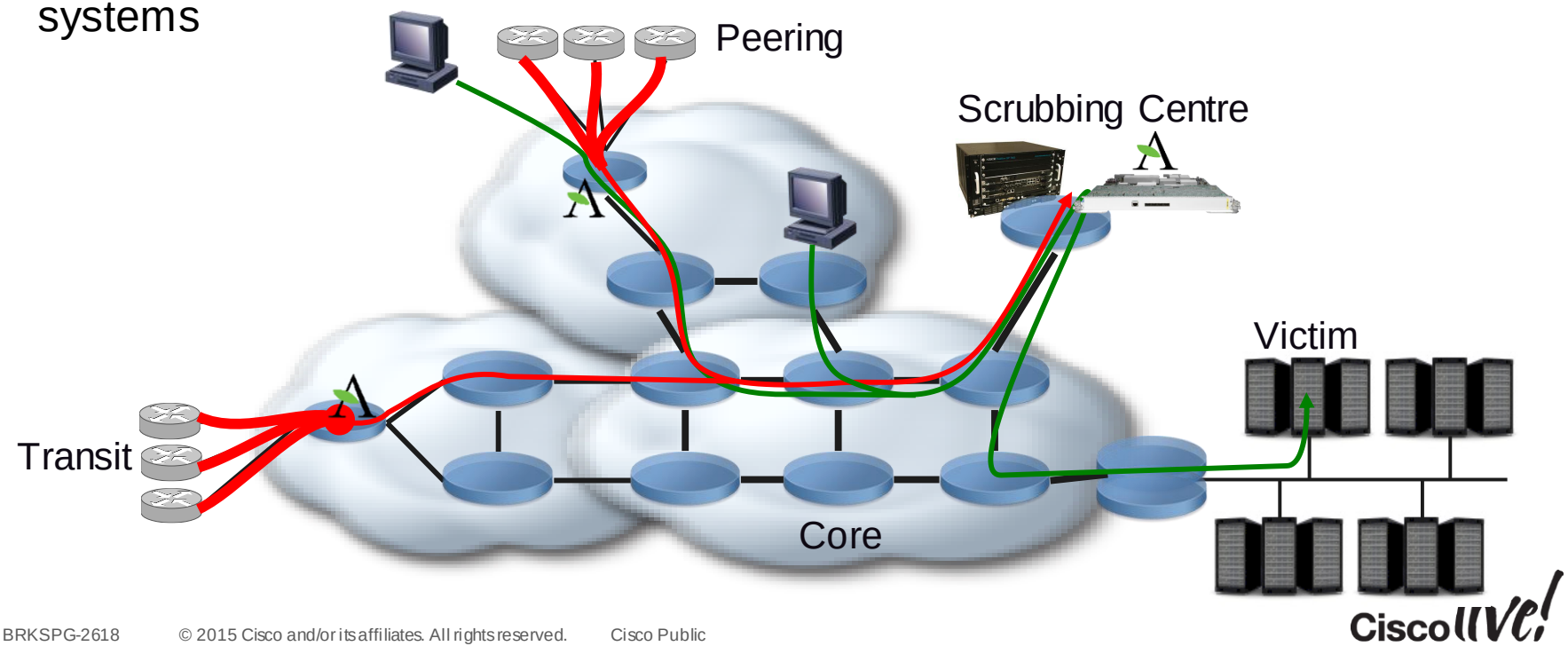
- We install scrubbers at the edge of the backbone to tackle the attack as early as possible



DDoS Mitigation Models

Mixed

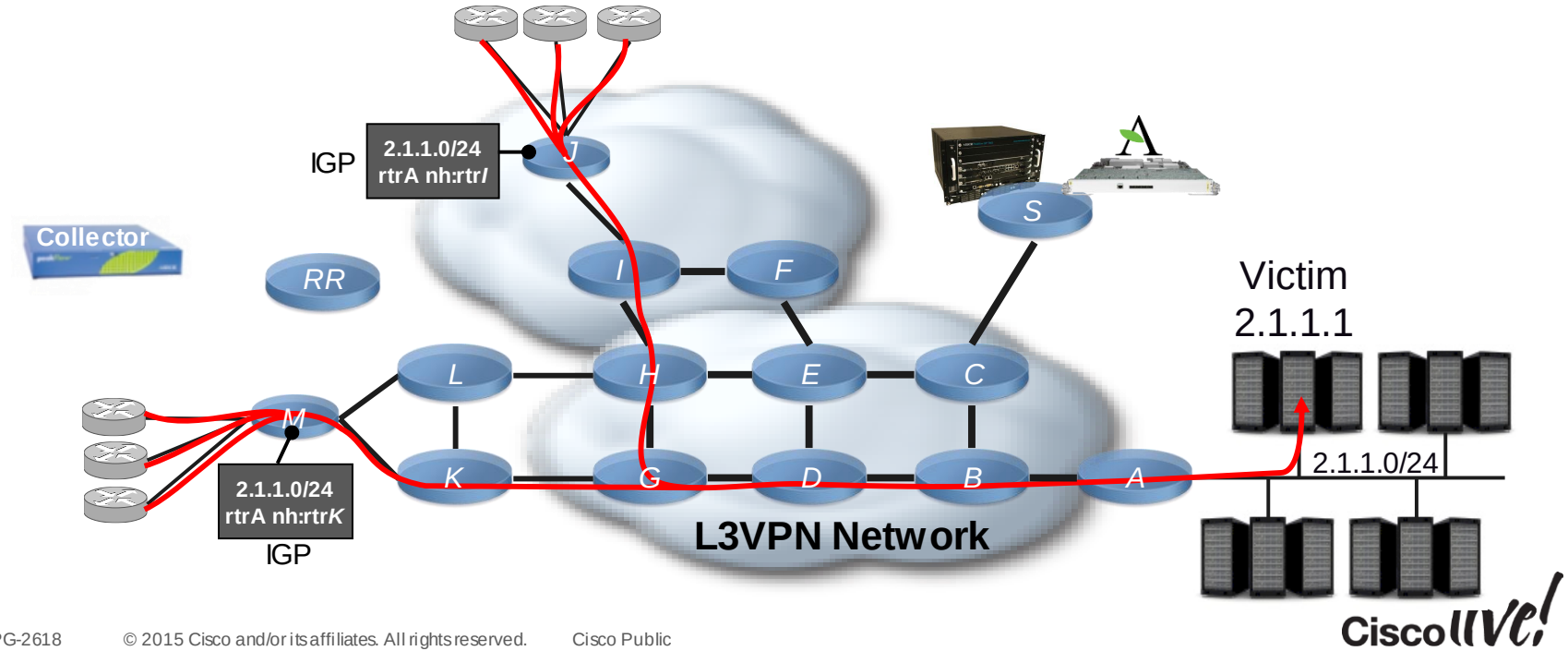
- Specific attacks can be handled in the central point or to off-load the edge systems



L3VPN Network w/ Scrubbing Centre

Currently deployed

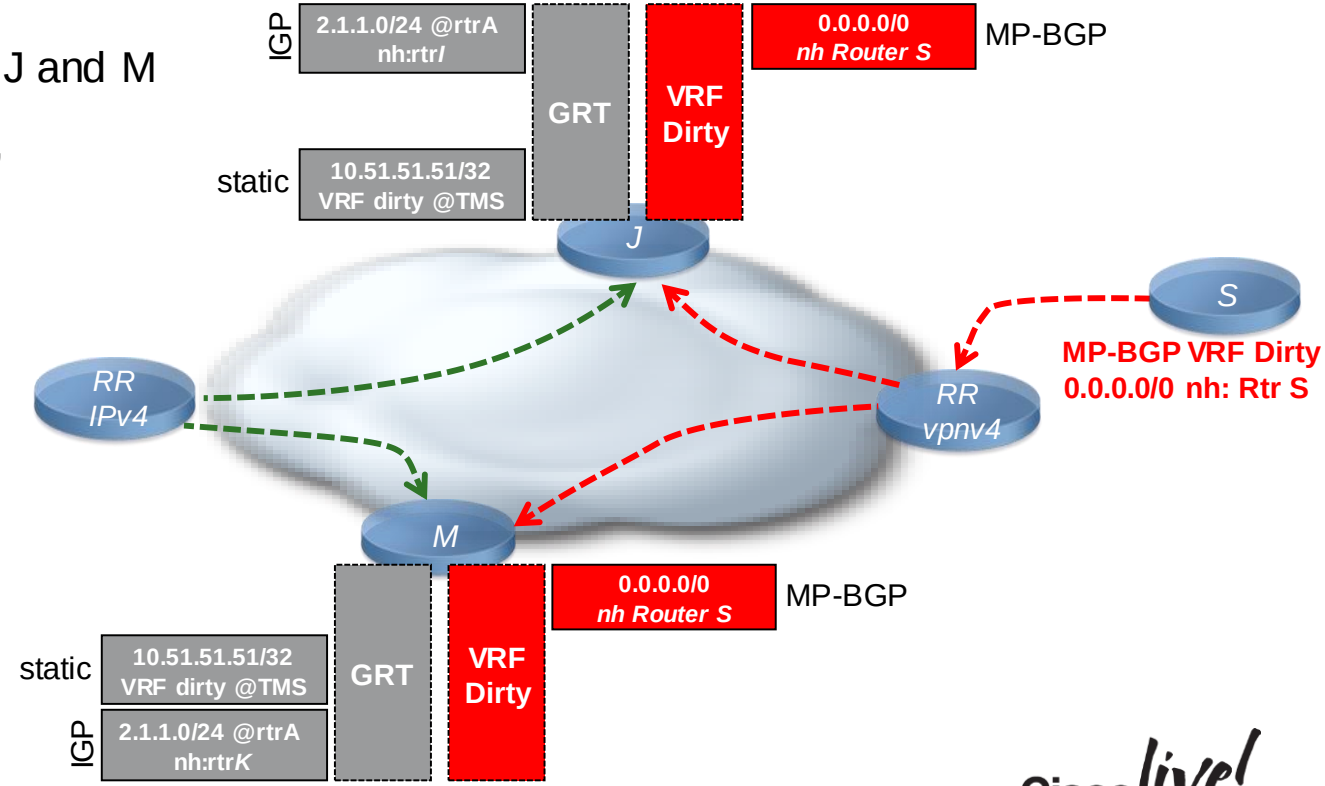
- 2.1.1.1 is under attack. Traffic is transported in the GRT or a VRF Internet



L3VPN Network w/ Scrubbing Centre

Currently deployed

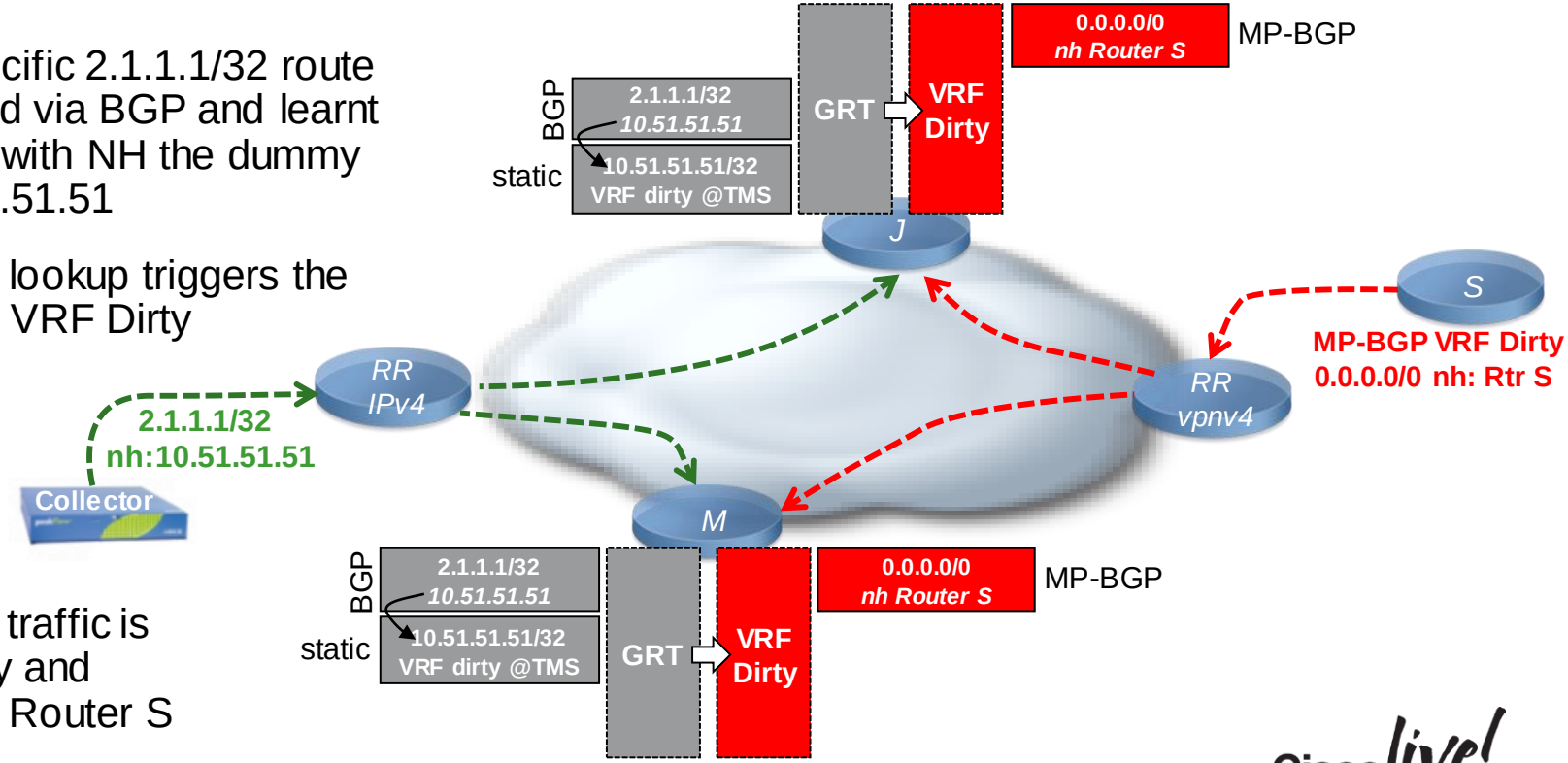
- VRF Dirty is configured on J and M
- MP-BGP is configured too, default route is advertised from @TMS in VRF Dirty
- On edge routers J and M, we configure static entries for a dummy host route (10.51.51.51/32) with a NH in VRF Dirty. If matched, traffic will leak into this VRF Dirty
- Now, traffic to 2.1.1.1 uses the IGP route 2.1.1.0/24



L3VPN Network w/ Scrubbing Centre

Currently deployed

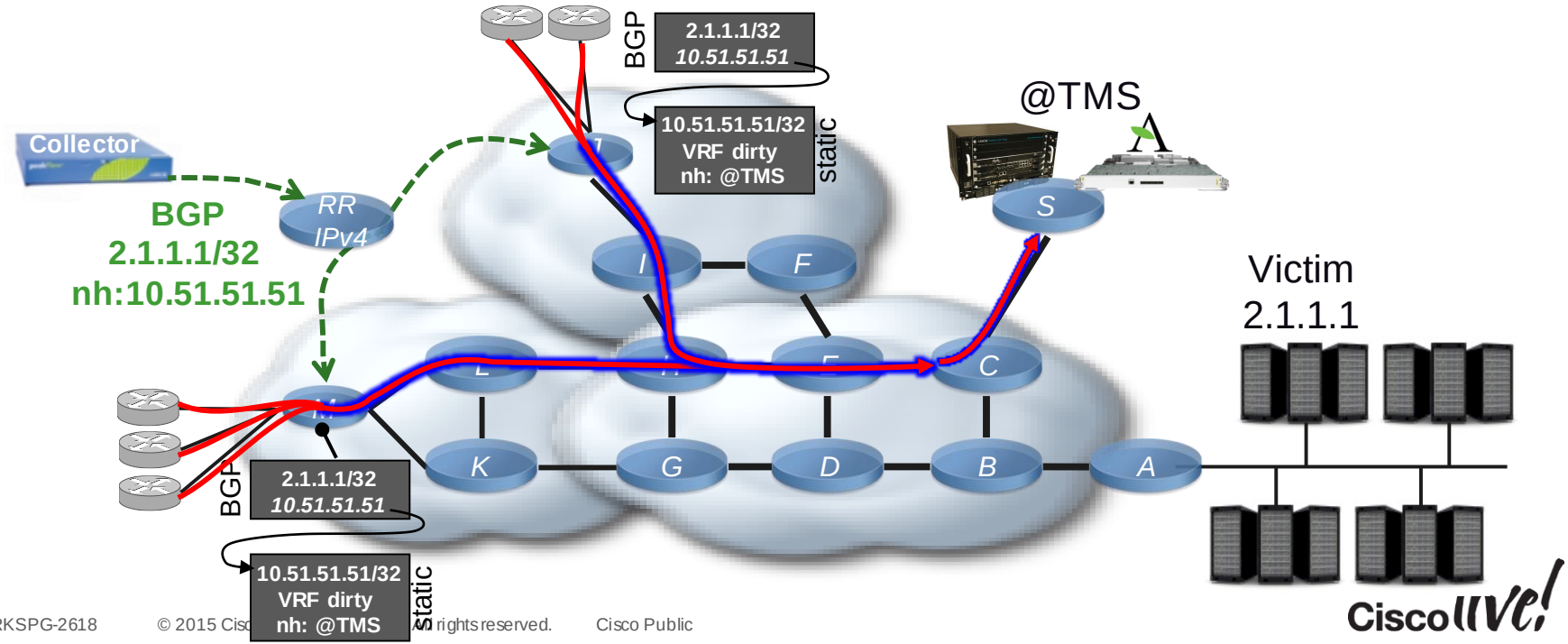
- A more specific 2.1.1.1/32 route is advertised via BGP and learnt in the GRT with NH the dummy route 10.51.51.51
- A recursive lookup triggers the leaking into VRF Dirty
- Now attack traffic is in VRF Dirty and attracted to Router S



L3VPN Network w/ Scrubbing Centre

Currently deployed

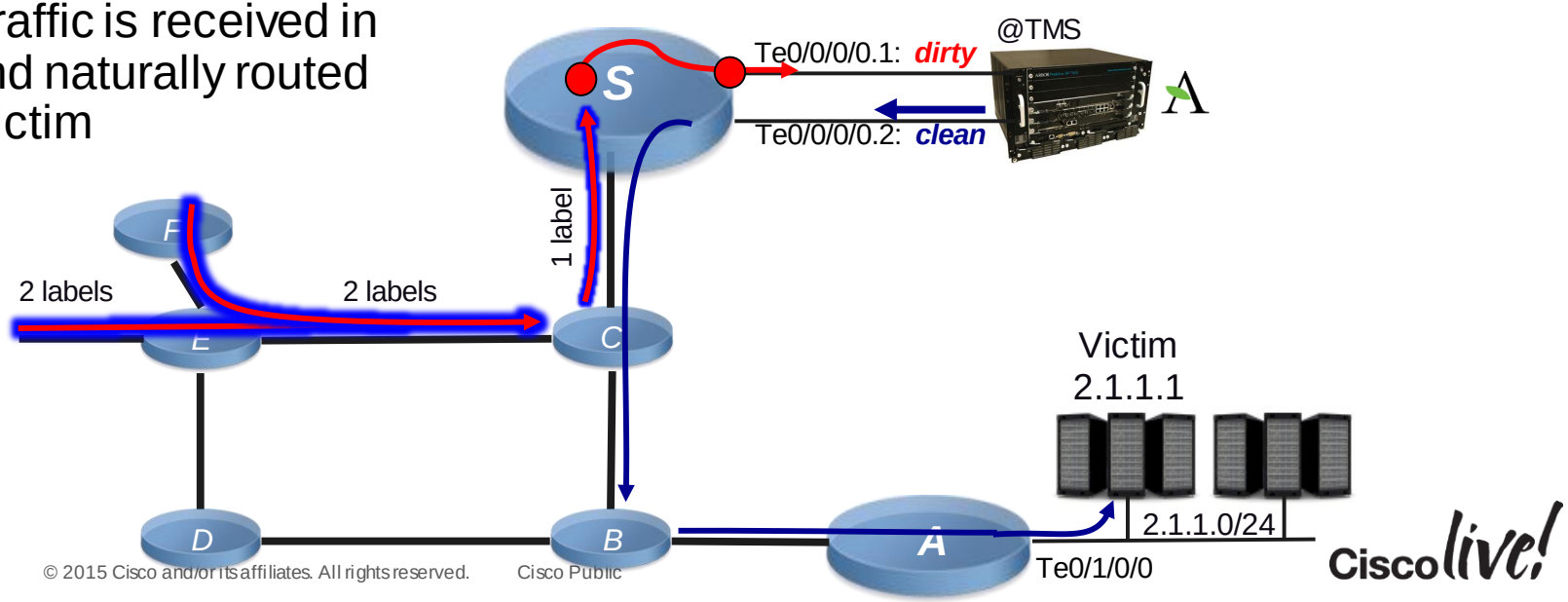
- CP advertises a BGP route for 2.1.1.1/32 with next-hop the dummy 10.51.51.51



L3VPN Network w/ Scrubbing Centre

Currently deployed

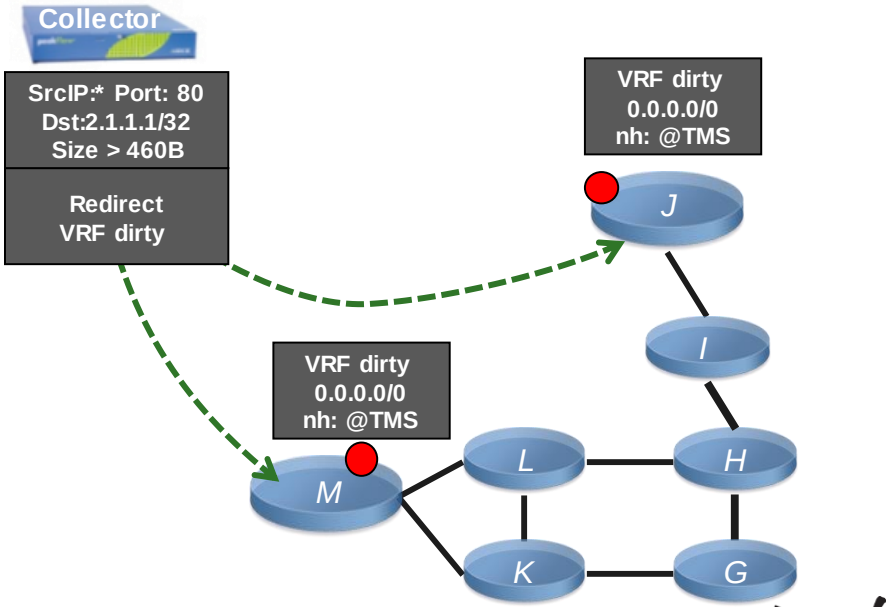
- Traffic with a VRF label Dirty is dragged to router S
- Router S is pushing unlabeled traffic to the TMS via an interface in VRF Dirty
- Clean traffic is received in GRT and naturally routed to the victim



L3VPN Network w/ Scrubbing Centre

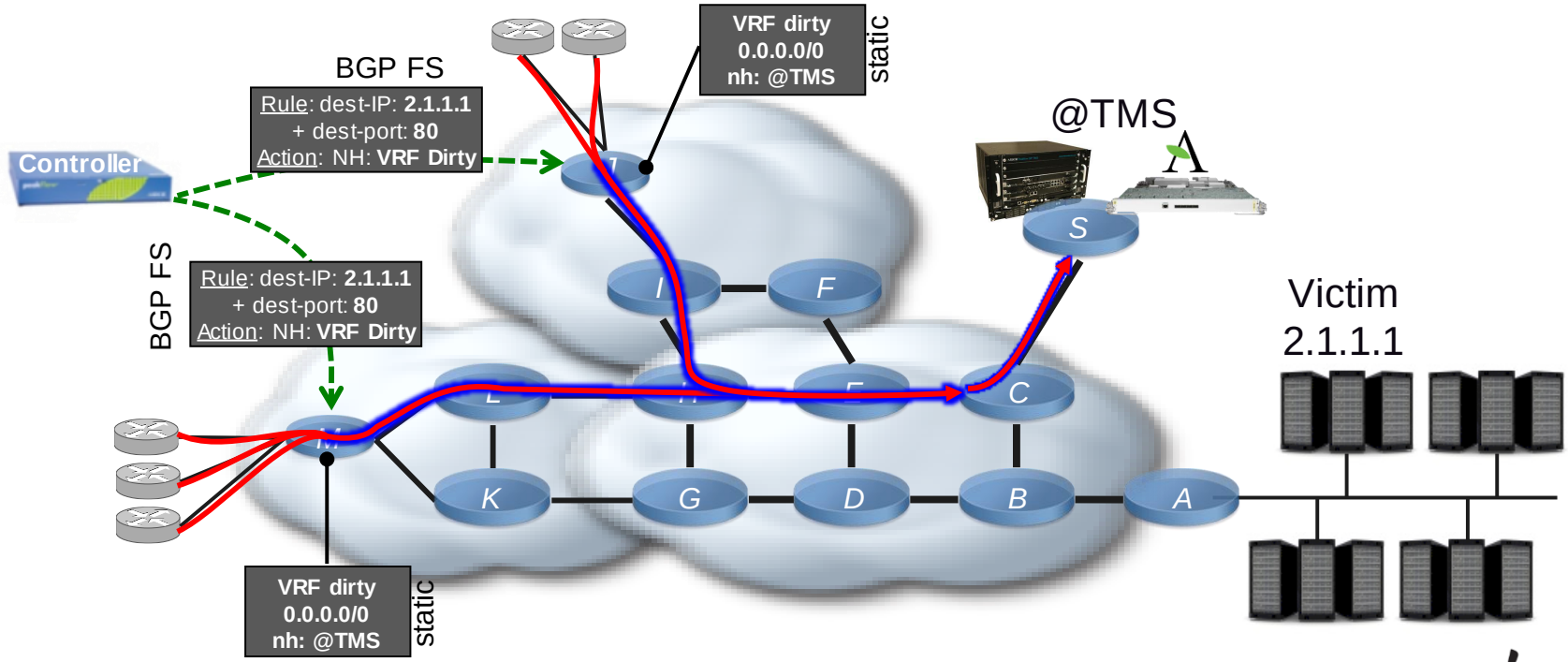
Improved with BGP FlowSpec

- BGP FlowSpec inject rules to redirect attack traffic into VRF dirty
- No more dummy route needed
- Only a default route in dirty VRF is needed to reach the scrubber
- More granular "matching" parameters: only the packets with specific protocol/port/packet-size/etc are diverted in Dirty VRF



L3VPN Network w/ Scrubbing Centre

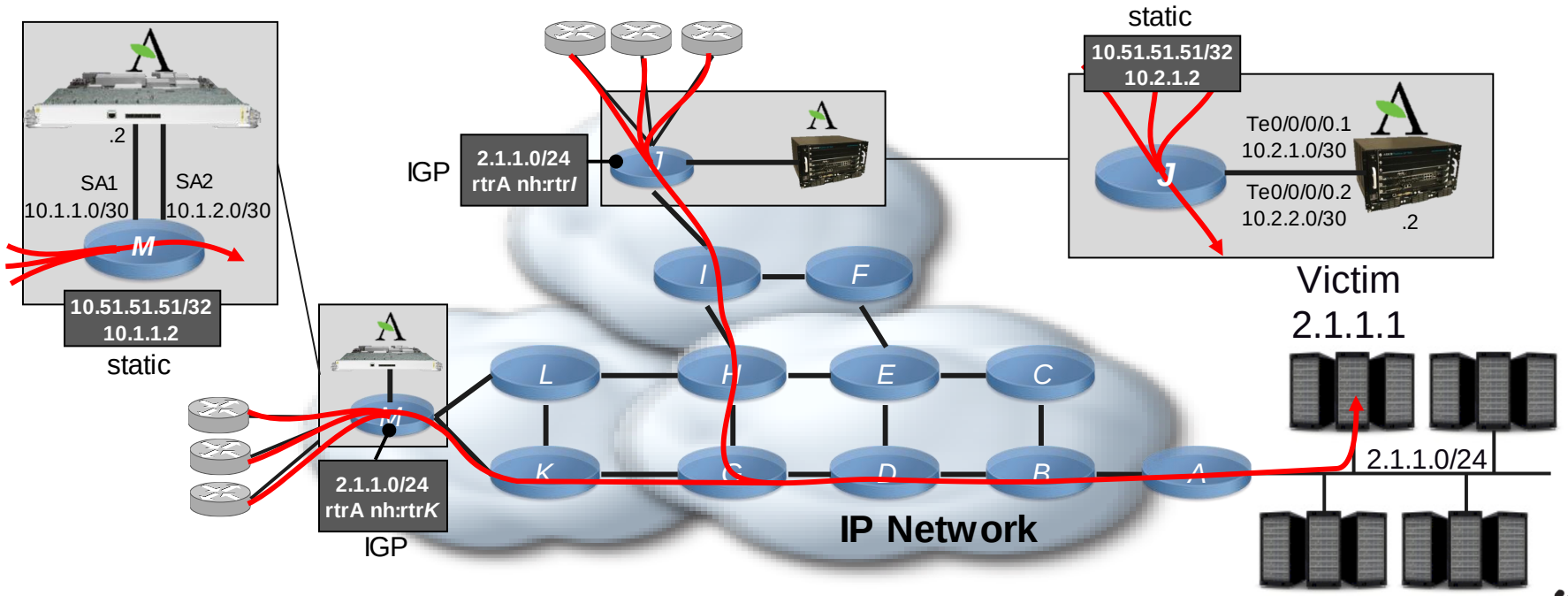
Improved with BGP FlowSpec



IP-only Network w/ Distributed TMS

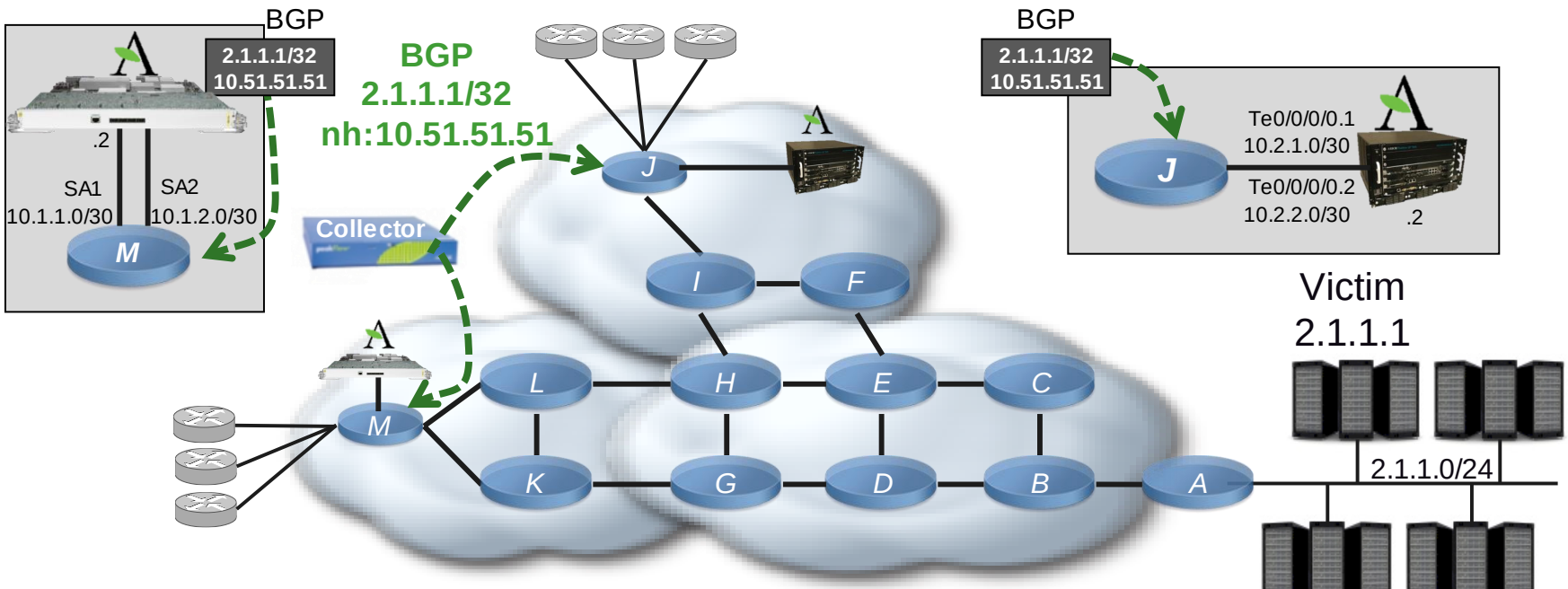
Currently deployed

- A static route for 10.51.51.51 is defined on routers M and J pointing to local TMS



IP-only Network w/ Distributed TMS

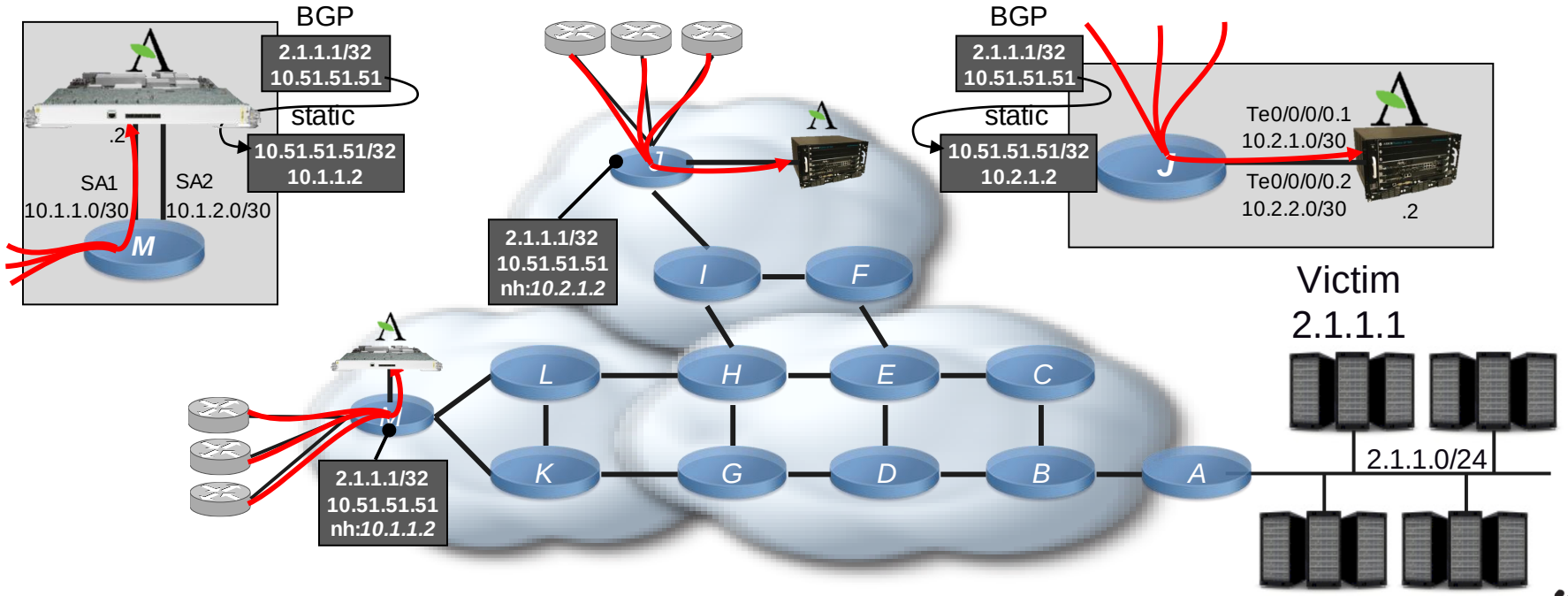
Currently deployed



10.51.51.51 is a dummy route, advertised to trigger the redirection

IP-only Network w/ Distributed TMS

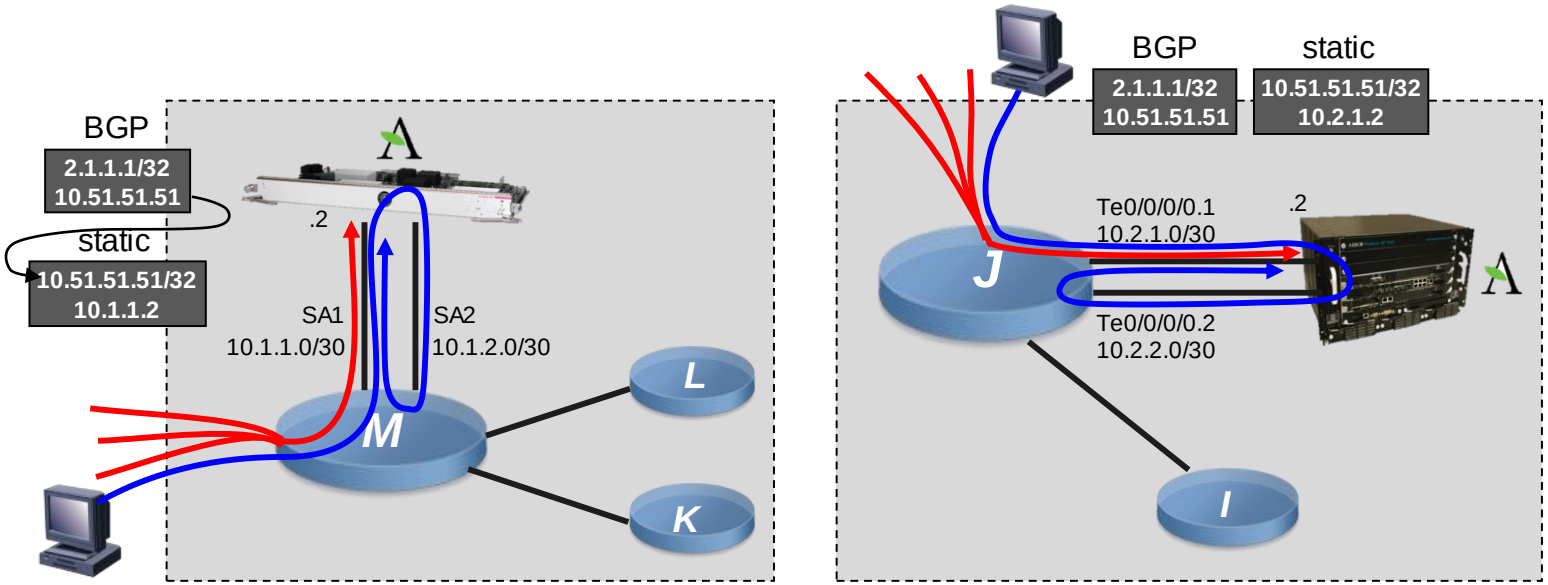
Currently deployed



IP-only Network w/ Distributed TMS

Currently deployed

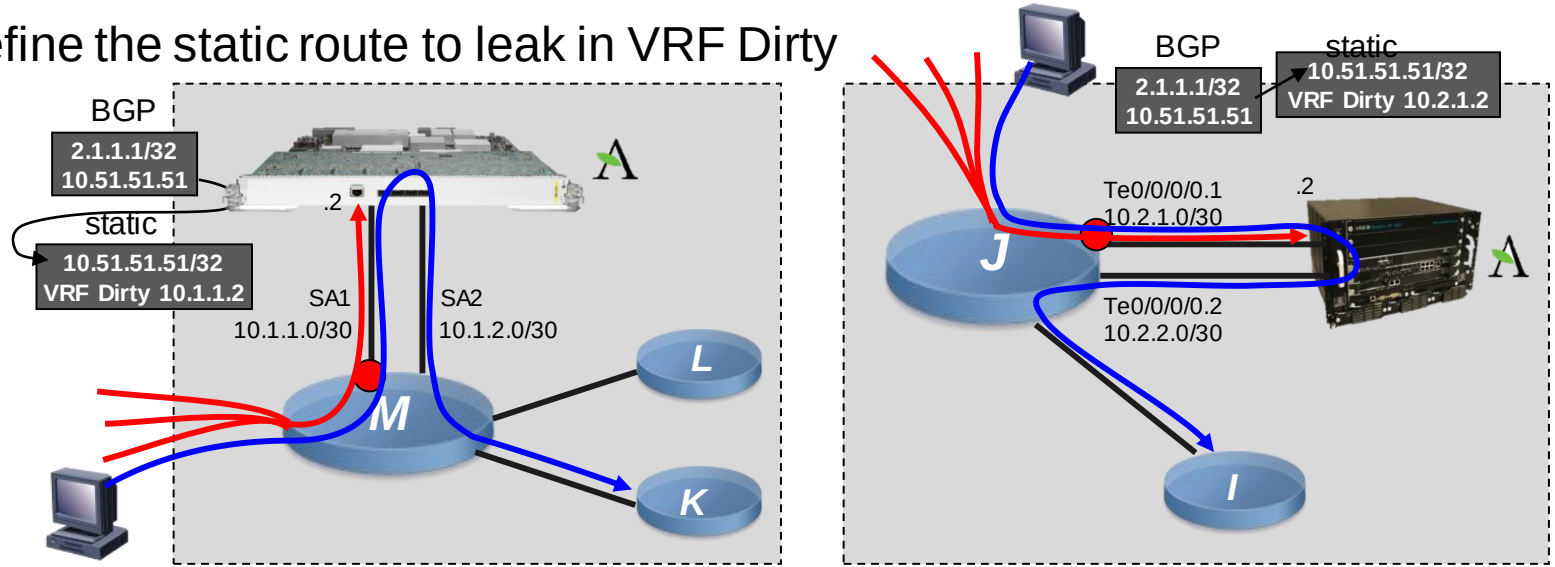
- With the specific route received we now have to deal with a routing loop for the legit traffic going out of the TMS device. We need solutions to prevent it



IP-only Network w/ Distributed TMS

First Solution to Avoid the Routing Loop

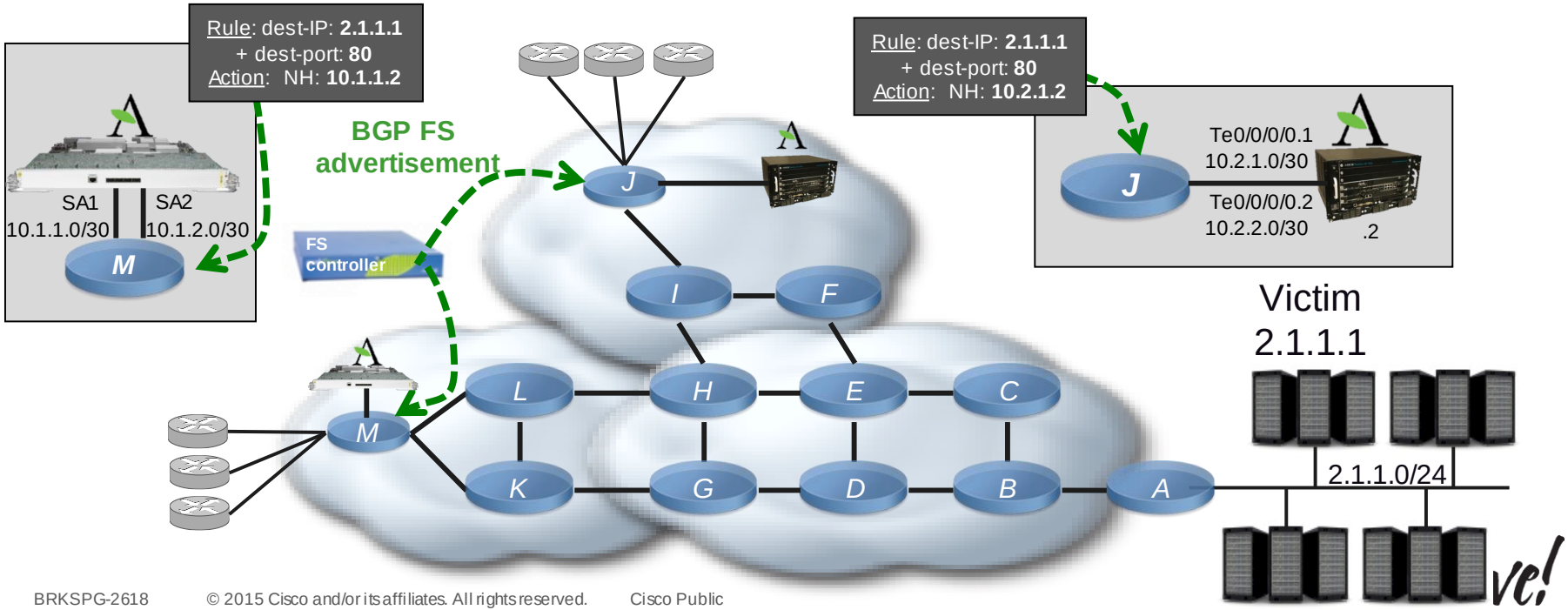
- Define an VRF-Lite Dirty and assigned the ingress TMS interfaces to it:
  - SA1 in the VSM/TMS case
  - Te0/0/0/0.1 in the Appliance case
- Define the static route to leak in VRF Dirty



IP-only Network w/ Distributed TMS

BGP FlowSpec Improvement: Granularity

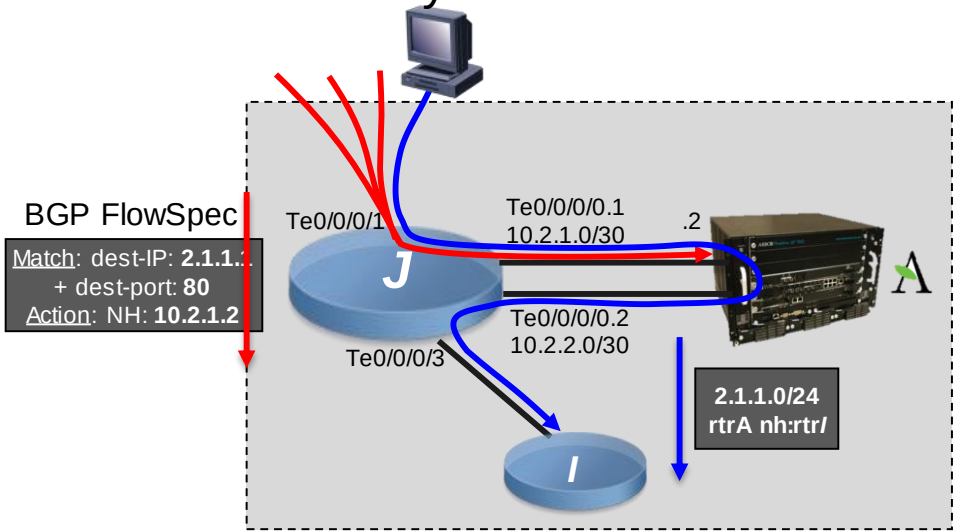
- BGP FS defines precisely the flow to divert to the local TMS

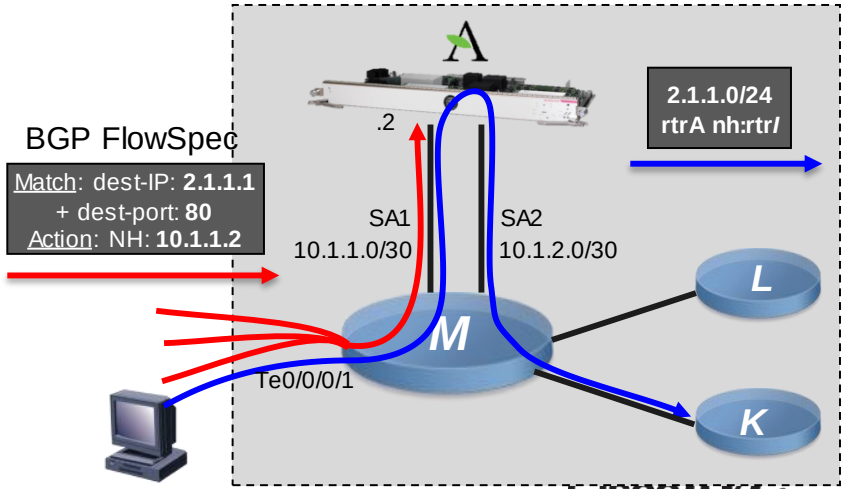


IP-only Network w/ Distributed TMS

BGP FlowSpec Improvement: No VRF-Lite needed

- BGP FlowSpec is activated on Te0/0/0/1, dirty traffic targeted to 2.1.1.1:80 is forwarded to the TMS address 10.2.1.2
- BGP Flowspec is deactivated on port te0/0/0/0.2, clean traffic from TMS is routed naturally via IGP route 2.1.1.0/24 to router I



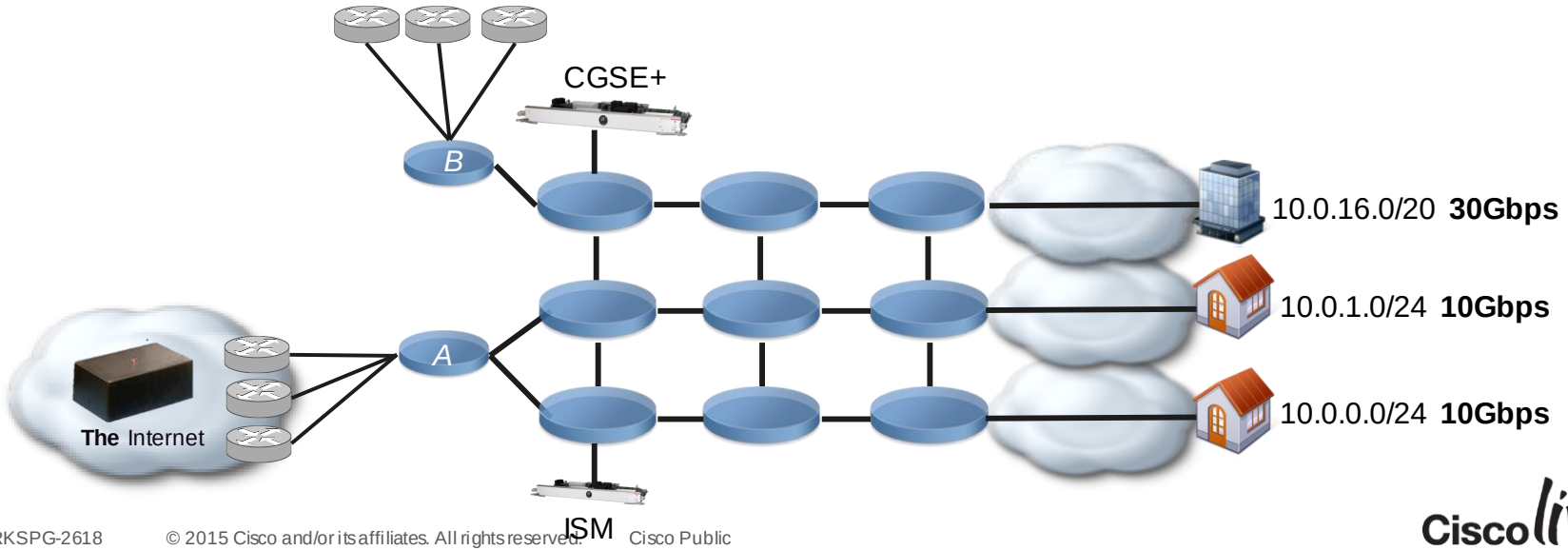


Other Use-Cases

Other BGP FS Use-Cases

Unequal Load-Balancing

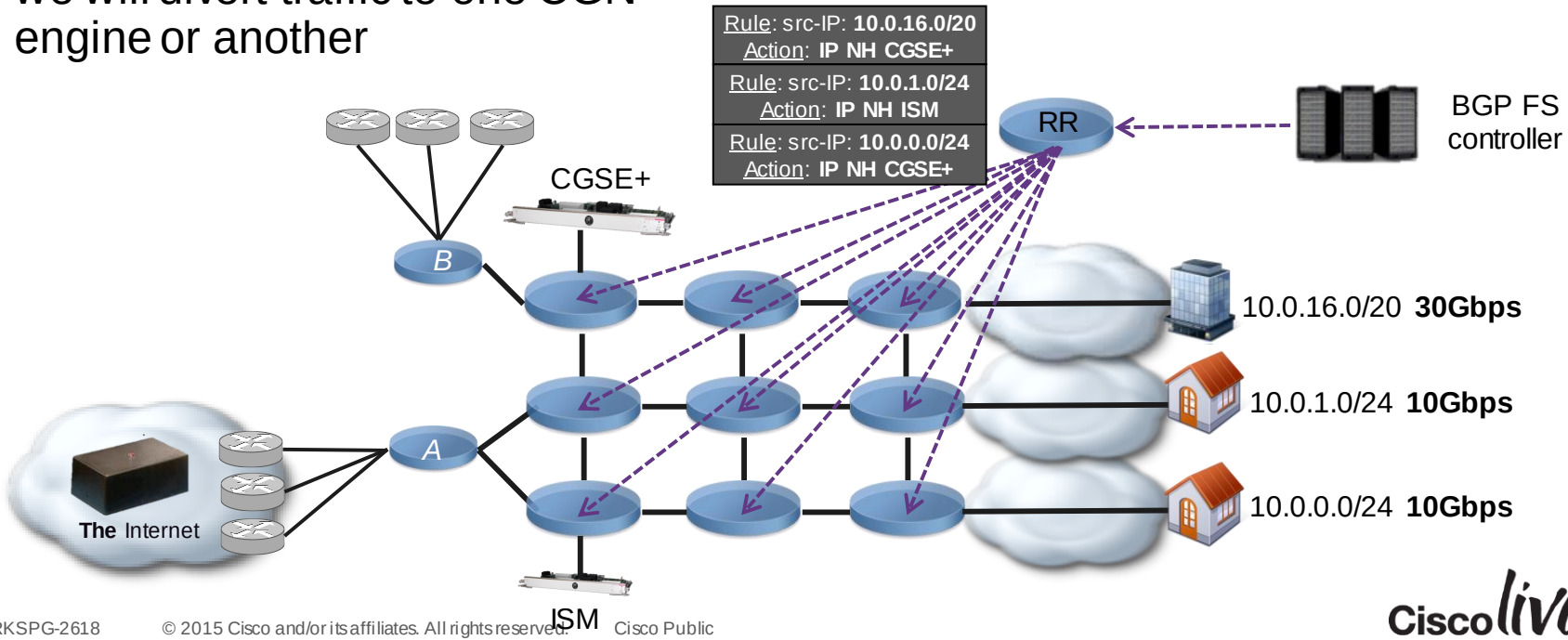
- Different peering / transit points
- Different NATing points with different performances / capabilities



Other BGP FS Use-Cases

Unequal Load-Balancing

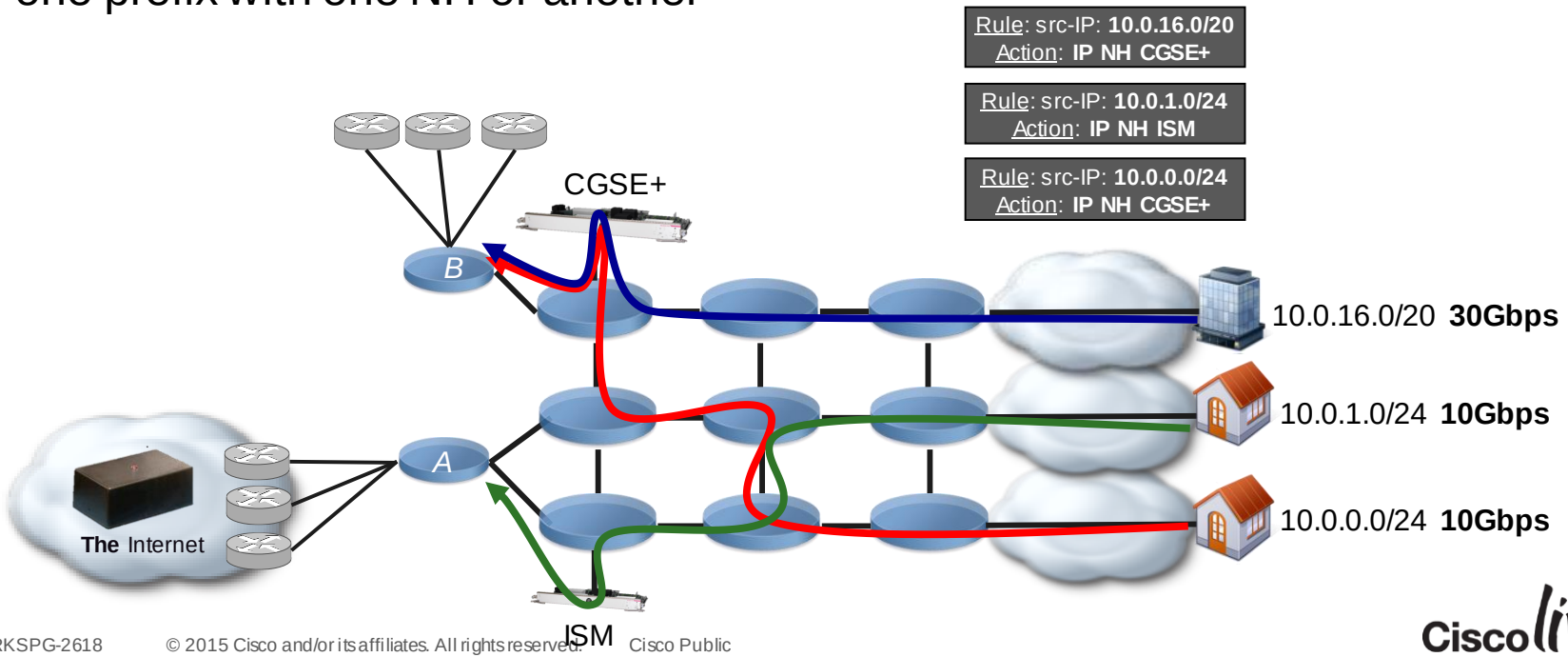
- Based on the source ranges, we will divert traffic to one CGN engine or another



Other BGP FS Use-Cases

Unequal Load-Balancing

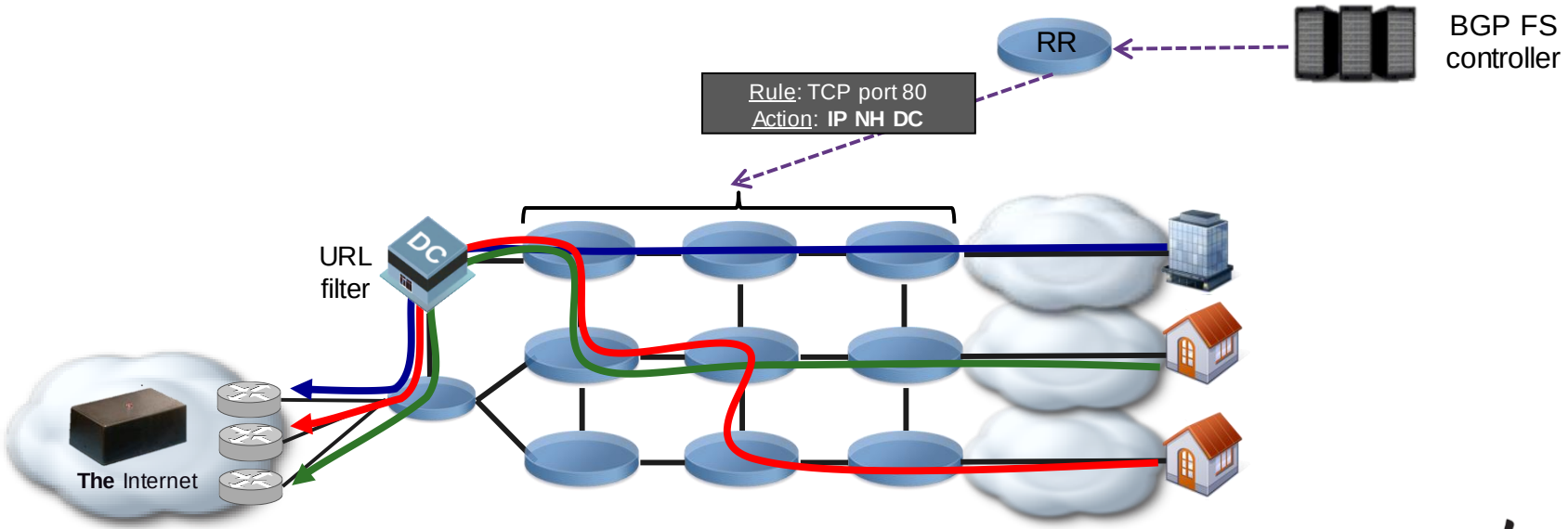
- This approach allows fine tuning of the traffic in the NAT engines, advertising one prefix with one NH or another



Other BGP FS Use-Cases

URL Filtering

- FlowSpec offers the granularity to divert only the HTTP traffic, the rest will be routed naturally



Configuration, Troubleshooting and Monitoring

Cisco live!

Configuring BGP FlowSpec on IOS XR Routers

Signalisation: Use of a new Address-Family flowspec

Controller

```
router bgp 1
 bgp router-id 6.6.6.6
 address-family ipv4 flowspec
 !
 neighbor-group ibgp-flowspec
  remote-as 1
  update-source loopbook0
  address-family ipv4 flowspec
  !
 !
 neighbor 25.2.1.3
  use neighbor-group ibgp-flowspec
 !
 neighbor 25.2.1.4
  use neighbor-group ibgp-flowspec
 !
!
flowspec
 address-family ipv4
  service-policy type pbr FS
```

Advertise policy FS

Client

```
router bgp 1
 bgp router-id 3.3.3.3
 address-family ipv4 flowspec
 !
 neighbor-group ibgp-flowspec
  remote-as 1
  update-source loopback0
  address-family ipv4 flowspec
  !
 neighbor 25.2.1.11
  use neighbor-group ibgp-flowspec
 !
!
flowspec
 local-install interface-all
!
```

Install all rules on all interfaces

Configuring BGP FlowSpec on IOS XR Routers

Verifying the Session Establishment (on Client)

```
RP/0/RP0/CPU0:Client#sh bgp ipv4 flowspec summary

BGP router identifier 3.3.3.3, local AS number 1
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0x0   RD version: 7072
BGP main routing table version 7072
BGP NSR Initial initsync version 0 (Reached)
BGP NSR/ISSU Sync-Group versions 7072/0
BGP scan interval 60 secs

BGP is operating in STANDALONE mode.

Process       RcvTblVer   bRIB/RIB   LabelVer  ImportVer  SendTblVer  StandbyVer
Speaker            7072       7072       7072       7072        7072        7072

Neighbor        Spk    AS MsgRcvd MsgSent   TblVer  InQ OutQ  Up/Down  St/PfxRcd
25.2.1.11         0     1  106269  105679     7072    0    0     1w1d        1001

RP/0/RP0/CPU0:Client#
```

Configuring BGP FlowSpec on IOS XR Routers

Configuring Rules on the Controller

- In many aspects, the rules configuration on the controller is similar to the MQC (Modular QoS Configuration)
- Rules are defined in Cisco Common Classification Policy Language (C3PL) format:
  - Traffic Matching is defined in class-map
  - Action is defined in a policy-map and refers a class-map
  - This policy-map is advertised by the “service-policy type pbr”

Configuring BGP FlowSpec on IOS XR Routers

Configuring Rules on the Controller

```
class-map type traffic match-all match-UDP53
 match destination-port 53
 match protocol udp
 end-class-map
!
class-map type traffic match-all match-src-ipv4-addr
 match destination-address ipv4 25.1.104.0 255.255.255.0
 end-class-map
!
```

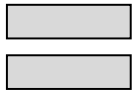
```
policy-map type pbr FS
 class type traffic match-src-ipv4-addr
  police rate 100000 bps
  !
 !
 class type traffic match-UDP53
  redirect next 192.42.52.125
  !
 !
 class type traffic class-default
 !
 end-policy-map
```

```
flowspec
 address-family ipv4
  service-policy type pbr FS
```

Configuring BGP FlowSpec on IOS XR Routers

Configuring Rules on the Controller

```
class-map type traffic match-all MATCH-UDP123
 match destination-port 123
 match protocol udp
 end-class-map
!
class-map type traffic match-all MATCH-SRCv4
 match destination-address ipv4 2.1.1.0/24
 end-class-map
!
policy-map type pbr FS1
 class type traffic MATCH-SRCv4
  police rate 100000 bps
  !
 end-policy-map
!
policy-map type pbr FS2
 class type traffic MATCH-UDP123
  redirect nexthop 192.168.2.5
  !
end-policy-map

flowspec
 address-family ipv4
  service-policy type pbr FS1
  service-policy type pbr FS2
```



```
class-map type traffic match-all MATCH-UDP123
 match destination-port 123
 match protocol udp
 end-class-map
!
class-map type traffic match-all MATCH-SRCv4
 match destination-address ipv4 2.1.1.0/24
 end-class-map
!
policy-map type pbr FS
 class type traffic MATCH-SRCv4
  police rate 100000 bps
  !
 class type traffic MATCH-UDP123
  redirect nexthop 192.168.2.5
  !
end-policy-map

flowspec
 address-family ipv4
  service-policy type pbr FS
```

BGP FS Matching Fields and Actions

NLRI type	Match fields	Match fields
Type 1	IPv4 Destination address	IPv6 Destination address
Type 2	IPv4 Source address	IPv6 Source address
Type 3	IPv4 protocol	IPv6 Next Header
Type 4	IPv4 source or destination port	IPv6 source or destination port
Type 5	IPv4 destination port	IPv6 destination port
Type 6	IPv4 Source port	IPv6 Source port
Type 7	IPv4 ICMP type	IPv6 ICMP type
Type 8	IPv4 ICMP code	IPv6 ICMP code
Type 9	IPv4 TCP flags (2 bytes include reserved bits)	IPv6 TCP flags (2 bytes include reserved bits)
Type 10	IPv4 Packet length	IPv6 Packet length
Type 11	IPv4 DSCP	IPv6 Traffic Class
Type 12	IPv4 fragmentation bits	Reserved
Type 13	N/A	IPv6 Flow Based (20 bytes)

Type	Action
0x8006	Traffic-rate
0x8007	Traffic-action
0x8008	Redirect
0x8009	Traffic-marking

Configuring BGP Flowspec

Configuring a Type 1 Match "Destination Address"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-RULE
RP/0/0/CPU0:Ctrl(config-cmap)#match destination-address ipv4 81.253.193.0/24
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail

AFI: IPv4
 Flow            :Dest:81.253.193.0/24
  Actions        :Traffic-rate: 100000 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                         0/0
   Transmitted          :                         0/0
   Dropped              :                         0/0
RP/0/RP0/CPU0:Client#sh flowspec ipv4 nlri

AFI: IPv4
 NLRI (Hex dump) :       0x011851fdc1
  Actions        :Traffic-rate: 100000 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type	Prefix length	Prefix
1 byte	1 byte	Variable
1	/24	81.253.193
0 x01	**0x18**	**0x 51 fd c1**

0x011851fdc1

Configuring BGP Flowspec

Configuring a Type 1 Match "Destination Address"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-RULE
RP/0/0/CPU0:Ctrl(config-cmap)#match destination-address ipv4 81.253.193.0/24
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#show contr pse tcam summary location 0/0/CPU0

<SNIP>

TCAM Device Information for Ingress PSE, CAM bank 1:
Device size: 20M (256K array entries of 80-bits), 261122 available
Current mode of operation: Turbo
<SNIP>
Feature specific information:
<SNIP>
        Flowspec IPv4 (id 32):
                Owner client id: 20.   Limit 245760 cells
                Total 1 regions using 4 CAM cells
<SNIP>
```

Configuring BGP Flowspec

Configuring a Type 2 Match "Source Address"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-RULE
RP/0/0/CPU0:Ctrl(config-cmap)#match source-address ipv4 2.2.0.0/16
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail

AFI: IPv4
 Flow             :Source:2.2.0.0/16
  Actions         :Traffic-rate: 100000 bps  (bgp.1)
  Statistics                              (packets/bytes)
   Matched                :                       0/0
   Transmitted            :                       0/0
   Dropped                :                       0/0
RP/0/RP0/CPU0:Boca#sh flowspec ipv4 nlri

AFI: IPv4
 NLRI (Hex dump) :        0x02100202
  Actions         :Traffic-rate: 100000 bps  (bgp.1)
RP/0/RP0/CPU0:Boca#
```

Type	Prefix length	Prefix
1 byte	1 byte	Variable
2	/16	2.2
0x 02	**0x 10**	**0x 02 02**

0x02100202

Configuring BGP Flowspec

Configuring a Type 3 Match “IPv4 Protocol Type” / “IPv6 Next Header”

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-RULE
RP/0/0/CPU0:Ctrl(config-cmap)#match protocol udp tcp
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail

AFI: IPv4
 Flow            :Proto:=0|=17|=6
  Actions        :Traffic-rate: 100000 bps  (bgp.1)
  Statistics                        (packets/bytes)
   Matched              :                  0/0
   Transmitted          :                  0/0
   Dropped              :                  0/0
RP/0/RP0/CPU0:Client#sh flowspec ipv4 nlri

AFI: IPv4
 NLRI (Hex dump) :        0x03010001118106
  Actions        :Traffic-rate: 100000 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Option Byte						
End	And	Len	0	Lt “<“	Gt “>”	Eq “=“
1b	1b	2b	1b	1b	1b	1b

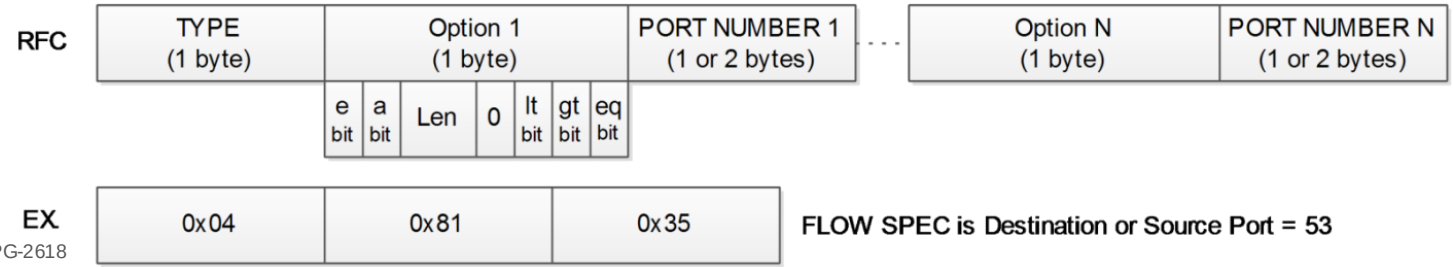
Type	Option1	IP proto1	Option2	IP proto2	Option3	IP proto3
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1	0b00000001	0x00	0b00000001	17 = 0x11	0b10000001	0x06
0x 03	**01**	**00**	**01**	**11**	**81**	**06**

Configuring BGP Flowspec

Configuring a Type 4 Match "Source or Destination Ports"

- We can receive Type4 messages on client but can not generate it on the controller due to C3PL limitation

```
RP/0/0/CPU0:Ctrl(config)#show config failed
<SNIP>
class-map type traffic match-any MATCH-TYPE-4
 match source-port 123
 match destination-port 123
 end-class-map
!
!!% Policy manager does not support this feature: Match all is the only mode supported
for match type "source-port" in class-map type "traffic"
End
```



Configuring BGP Flowspec

Configuring a Type 5 Match "Destination Port"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-TYPE5
RP/0/0/CPU0:Ctrl(config-cmap)#match destination-port 80 443 8080
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#show flowspec afi-all detail
AFI: IPv4
 Flow           :DPort:=80|=443|=8080
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                        (packets/bytes)
   Matched              :                  0/0
   Transmitted          :                  0/0
   Dropped              :                  0/0
RP/0/RP0/CPU0:Client#show flowspec ipv4 nlri
AFI: IPv4
 NLRI (Hex dump) :       0x0501501101bb911f90
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type (1B)	Option x (1B)	Dest Port (1B or 2B)
5	equal/length=0 Not last	d80 = x50
0 x05	**0x01**	**0x50**
-	equal/length=1 Not last	d443 = x1BB
-	**0x11**	**0x01BB**
-	equal/length=1 last	d8080 = x1F90
-	**0x91**	**0x1F90**

0x0501501101bb911f90

Option Byte							
	End	And	Len	0	Lt "<"	Gt ">"	Eq "="
01	0	0	00	0	0	0	1
11	0	0	01	0	0	0	1
91	1	0	01	0	0	0	1

Configuring BGP Flowspec

Configuring a Type 6 Match “Source Port”

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-TYPE6
RP/0/0/CPU0:Ctrl(config-cmap)#match source-port 80-100
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail
AFI: IPv4
 Flow           :SPort:>=80&<=100
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                        (packets/bytes)
   Matched              :                 0/0
   Transmitted          :                 0/0
   Dropped              :                 0/0
RP/0/RP0/CPU0:Client#sh flowspec ipv4 nlri
AFI: IPv4
 NLRI (Hex dump) :       0x060350c564
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type (1B)	Option 1 (1B)	Dest Port
6	0000 0011 greater+equal/le=0/not last	80
0 x06	0x03	0x50
-	1100 0101 lower+equal/le=0/last	100
-	0xc5	0x64

0x060350c564

Option Byte							
	End	And	Len	0	Lt “<“	Gt “>”	Eq “=“
03	0	0	00	0	0	1	1
c5	1	1	00	0	1	0	1

Configuring BGP Flowspec

Configuring a Type 7+8 Match “ICMP Type” + “ICMP Code”

```
RP/0/0/CPU0:Ctrl(config-cmap)# match ipv4 icmp-type 3
RP/0/0/CPU0:Ctrl(config-cmap)# match ipv4 icmp-code 13
RP/0/0/CPU0:Ctrl(config-cmap)#commit
```

```
RP/0/RSP0/CPU0:Client#show flowspec afi-all detail
AFI: IPv4
 Flow           :ICMPType:=3,ICMPCode:=13
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                       0/0
   Dropped              :                       0/0
RP/0/RSP0/CPU0:Client#show flowspec ipv4 nlri
AFI: IPv4
 NLRI (Hex dump) :       0x07810308810d
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RSP0/CPU0:Client#
```

Type (1B)	Option 1 (1B)	ICMP
7	1000 0001	03
0 x07	**0x81**	**0x03**
8	100 0001	13
0 x08	**0x81**	**0x0d**

0x07810308810d

Option Byte							
	End	And	Len	0	Lt “<“	Gt “>”	Eq “=“
81	1	0	00	0	0	0	1

Configuring BGP Flowspec

Configuring a Type 9 Match "TCP Flag Component"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-TYPE9
RP/0/0/CPU0:Ctrl(config-cmap)#match tcp-flag 2
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail
AFI: IPv4
 Flow            :TCPFlags:=0x02
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                 8/496
Dropped               :                  0/0
RP/0/RP0/CPU0:Client#sh flowspec ipv4 nlri
AFI: IPv4
 NLRI (Hex dump) :       0x098102
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type (1B)	Option 1 (1B)	Flag
9	1000 0001	x02
0 x0	**0x81**	**0x02**

0x098102

Option Byte							
	e bit	a bit	Len	0	0	Not bit	m bit
81	1	0	00	0	0	0	1

- Ex: http://rapid.web.unc.edu/resources/tcp-flag-key/
  - **0x02: SYN**
  - 0x12: SYN-ACK
  - 0x10: ACK

Configuring BGP Flowspec

Configuring a Type 10 Match "Packet Length"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-TYPE10
RP/0/0/CPU0:Ctrl(config-cmap)#match packet length 100
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#show flowspec afi-all detail
AFI: IPv4
 Flow            :Length:=100
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                       0/0
   Transmitted          :                       0/0
   Dropped              :                       0/0
RP/0/RP0/CPU0:Client#show flowspec ipv4 nlri
AFI: IPv4
 NLRI (Hex dump) :        0x0a8164
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type (1B)	Option 1 (1B)	Pkt Length
10	1000 0001	100
0 x0a	**0x81**	**0x64**

0x0a8164

Option Byte							
	End	And	Len	0	Lt "<"	Gt ">"	Eq "="
81	1	0	00	0	0	0	1

Configuring BGP Flowspec

Configuring a Type 11 Match "IPv4/IPv6 DSCP"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-TYPE11
RP/0/0/CPU0:Ctrl(config-cmap)#match dscp ef
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#show flowspec afi-all detail
AFI: IPv4
 Flow            :DSCP:=46
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                       0/0
   Transmitted          :                       0/0
   Dropped              :                       0/0
RP/0/RP0/CPU0:Client#show flowspec afi-all nlri
AFI: IPv4
 NLRI (Hex dump) :       0x0b812e
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type (1B)	Option 1 (1B)	DSCP
11	1000 0001	ef
0 x0b	0x81	0x2e

0x0a812b

Option Byte							
	End	And	Len	0	Lt "<"	Gt ">"	Eq "="
81	1	0	00	0	0	0	1

Configuring BGP Flowspec

Configuring a Type 12 Match "IPv4 Fragment"

```
RP/0/0/CPU0:Ctrl(config)#class-map type traffic match-all MATCHING-TYPE12
RP/0/0/CPU0:Ctrl(config-cmap)#match fragment-type is-fragment last-fragment
RP/0/0/CPU0:Ctrl(config-cmap)#
```

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail
AFI: IPv4
 Flow            :Frag:=LF:IsF
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched                :                    0/0
   Transmitted            :                    0/0
   Dropped                :
0/0RP/0/RP0/CPU0:Client#sh flowspec ipv4 nlri
AFI: IPv4
 NLRI (Hex dump) :        0x0c810a
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Type (1B)	Option 1 (1B)	Pkt Length
11	1000 0001	LF + IsF
0 x0b	0x81	0x0a

0x0a810a

Option Byte							
	End	And	Len	0	Lt "<"	Gt ">"	Eq "="
81	1	0	00	0	0	0	1

Bitmask							
	0	0	0	lf	ff	isf	df
0a	0	0	0	1	0	1	0

Configuring BGP Flowspec

Mixing Several Matching Statements

```
class-map type traffic match-all MATCHING-RULE1
 match source-port 10 20 30-40 50-52 60-70
 match protocol udp
 match dscp ef
 match packet length 10-100 102-200 202-400 402-1500
 match destination-port 80
 match destination-address ipv4 11.200.4.0 255.255.255.0
 end-class-map
```

```
RP/0/RSP0/CPU0:Client#sh flowspec afi-all detail
AFI: IPv4
 Flow
:Dest:11.200.4.0/24,Proto:=17,DPort:=80,SPort:=10|=20|>=30&<=40|>=50&<=52|>=60&<=70,Length:>=10&<
=100|>=102&<=200|>=202&<=400|>=402&<=1500,DSCP:=46
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched             :                        0/0
   Dropped             :                        0/0
RP/0/RSP0/CPU0:Client#sh flowspec afi-all nlri
AFI: IPv4
 NLRI (Hex dump) :
0x01180bc804038111058150060l0a0114031e45280332453403 3cc5460a030a4564036645c803ca550190130192d505d
c0b812e
  Actions       :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RSP0/CPU0:Client#
```

Configuring BGP Flowspec

Configuring an Action: Police

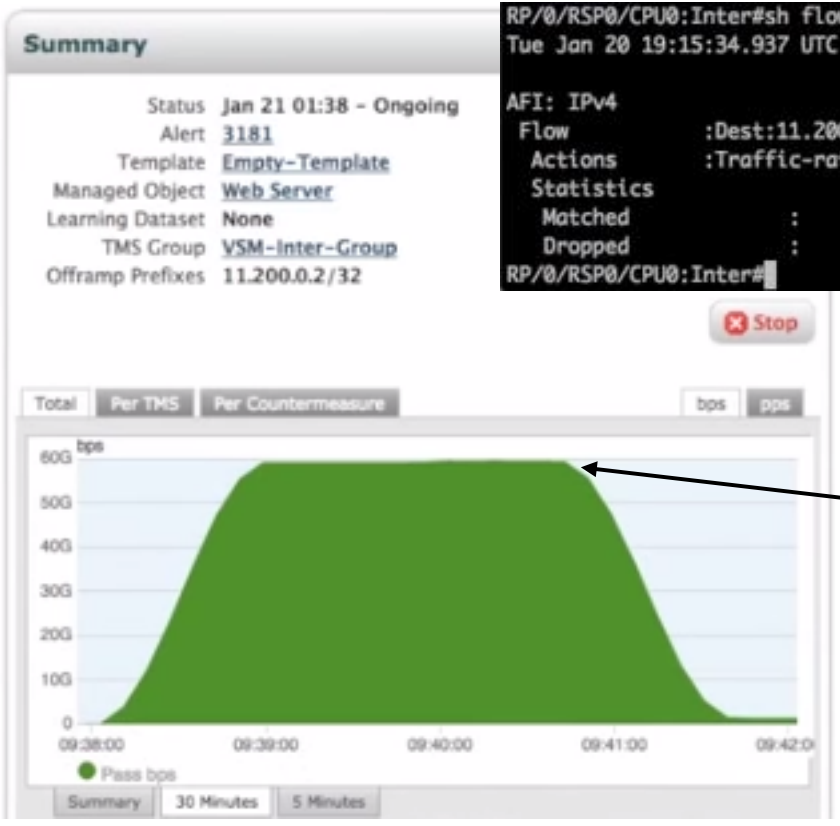
```
RP/0/0/CPU0:Ctrl(config)#policy-map type pbr FS
RP/0/0/CPU0:Ctrl(config-pmap)# class type traffic MATCHING-RULE1
RP/0/0/CPU0:Ctrl(config-pmap-c)#police ?
  rate  Committed Information Rate
RP/0/0/CPU0:Ctrl(config-pmap-c)#police rate ?
  <1-4294967295>  Committed Information Rate
RP/0/0/CPU0:Ctrl(config-pmap-c)#police rate 1000 ?
  bps       Bits per second (default)
  cellsps   Cells per second
  gbps      Gigabits per second
  kbps      Kilobits per second
  mbps      Megabits per second
  <cr>
RP/0/0/CPU0:Ctrl(config-pmap-c)#police rate 1000
RP/0/0/CPU0:Ctrl(config-pmap-c)#
```

	TYPE (2 bytes)	ASN (only the last 2 bytes) (2 bytes)	Rate (bytes/s) (4 bytes)
RFC			
EX.	0x8006	0x1234	0x4a3ebc20

→ Hex 4a3ebc20 = 31,125,000 Bytes/sec
= 25 Mbps

Configuring BGP Flowspec

Configuring an Action: Police




```
RP/0/RSP0/CPU0:Inter#sh flowspec ipv4 detail
Tue Jan 20 19:15:34.937 UTC

AFI: IPv4
 Flow           :Dest:11.200.0.0/24,Proto:=17,DPort:=123,SPort:=123,Length:>=900&<=1500
  Actions       :Traffic-rate: 0 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :             92207278/115259100220
   Dropped              :             92207278/115259100220
RP/0/RSP0/CPU0:Inter#
```

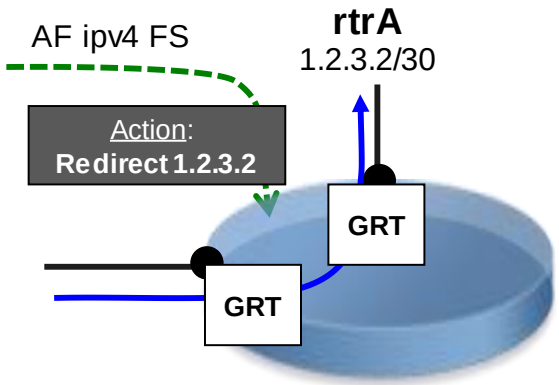
Application FS Rule

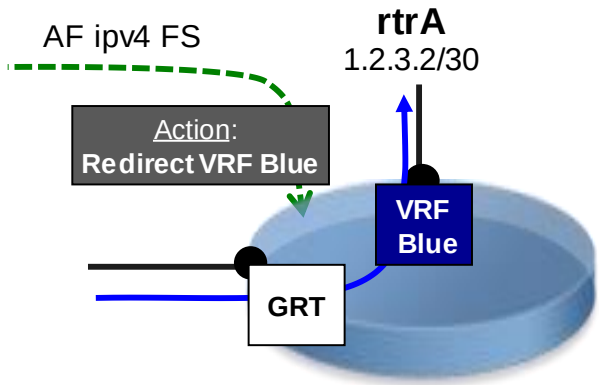
Drop = Police at 0bps

Configuration Flowspec

Action: Redirection

- If the ingress interface is in the Global Routing Table, the flowspec rule should be advertised via an “address-family IPv4 flowspec”
- Redirection to an NH address implies the egress interface is in the GRT too
- Redirection to a different VRF can not specify the destination address, a second lookup in this target VRF will happen to the destination address of the packet

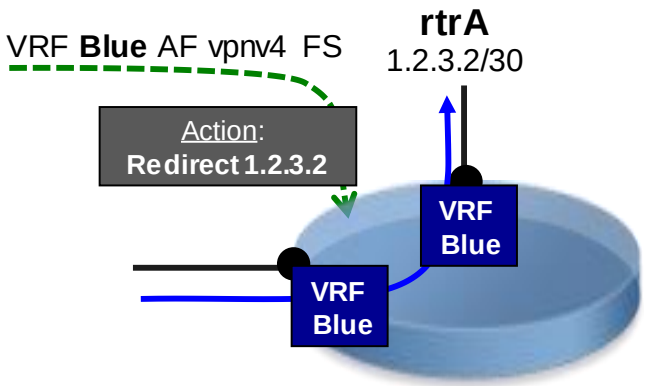


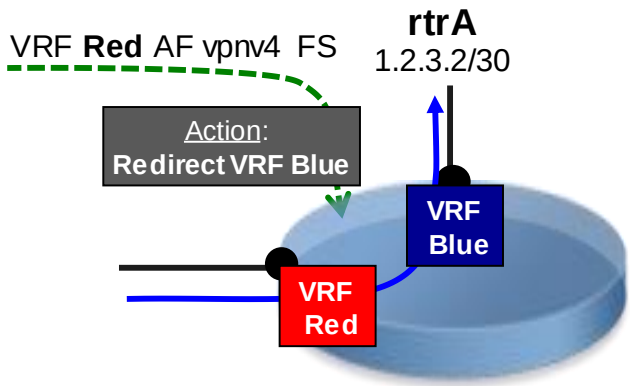


Configuring Flowspec

Action: Redirection

- If the ingress interface is in a VRF, the flowspec rule should be advertised via an "address-family vpnv4 flowspec" under the VRF statement in BGP
- In the same VRF Blue, we can apply an redirect action to an IP address
- Or we can apply a redirect to a different VRF Red where a new lookup will happen





Configuring BGP Flowspec

Configuring an Action: Redirect in VRF / IP address

Controller Configuration

```
policy-map type pbr TEST
 class type traffic MATCHING-RULE1
  redirect nexthop 25.3.9.3
 !
 class type traffic class-default
 !
 end-policy-map
!
traffic MATCHING-RULE1
class-map type traffic match-all MATCHING-
RULE1
 match protocol udp
 match packet length 500-1550
 match destination-address ipv4 25.1.102.1
255.255.255.255
 end-class-map
!
```

Client View

```
RP/0/RSP0/CPU0:Client#show bgp ipv4 flowspec
<SNIP>
Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale, N Nexthop-
discard
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop            Metric LocPrf Weight
Path
*>iDest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550/128
                      25.3.9.3                 100      0 i

Processed 1 prefixes, 1 paths

RP/0/RSP0/CPU0:Client#show flowspec afi-all detail

AFI: IPv4
 Flow           :Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550
  Actions       :Nexthop: 25.3.9.3  (bgp.1)
  Statistics                        (packets/bytes)
   Matched              :                 0/0
   Dropped              :                 0/0
RP/0/RSP0/CPU0:Client#
```

Configuring BGP Flowspec

Gotchas with Redirect Action

- A rule is advertised from controller only if the configured NH is reachable
- Not necessary reachable on the client side but mandatory on the controller side

```
RP/0/0/CPU0:Ctrl#sh route 25.1.102.1

% Network not in table

RP/0/0/CPU0:Ctrl#
```

```
RP/0/RSP0/CPU0:Client#sh bgp ipv4 flowspec

RP/0/RSP0/CPU0:Client#sh bgp ipv4 flowspec sum
Neighbor        Spk    AS MsgRcvd MsgSent   TblVer  InQ OutQ  Up/Down  St/PfxRcd
25.2.1.11         0     1   16488   16457      596    0    0 00:32:57           0

RP/0/RSP0/CPU0:Client#
```

```
RP/0/0/CPU0:Ctrl#sh run router static
router static
 address-family ipv4 unicast
  25.3.9.3/32 GigabitEthernet0/0/0/0
!
!
```

```
RP/0/RSP0/CPU0:Client#show bgp ipv4 flowspec

Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale, N Nexthop-discard
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop            Metric LocPrf Weight Path
*>iDest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550/128
                      25.3.9.3                 100      0 i

Processed 1 prefixes, 1 paths
RP/0/RSP0/CPU0:Client#
```

Configuring BGP Flowspec

Gotchas with Redirect Action

- If the NH is not reachable in the Client, the rule will be ignored

```
RP/0/RSP0/CPU0:Client#sh route 11.22.33.44

% Network not in table

RP/0/RSP0/CPU0:Client#
```

```
RP/0/0/CPU0:Ctrl#sh run policy-map type pbr TEST
policy-map type pbr TEST
 class type traffic MATCHING-RULE1
  redirect nexthop 11.22.33.44
 !
 class type traffic class-default
 !
 end-policy-map
!
RP/0/0/CPU0:XRv-service#sh run router static
router static
 address-family ipv4 unicast
  11.22.33.44/32 GigabitEthernet0/0/0/0
 !
!
RP/0/0/CPU0:Ctrl#
```

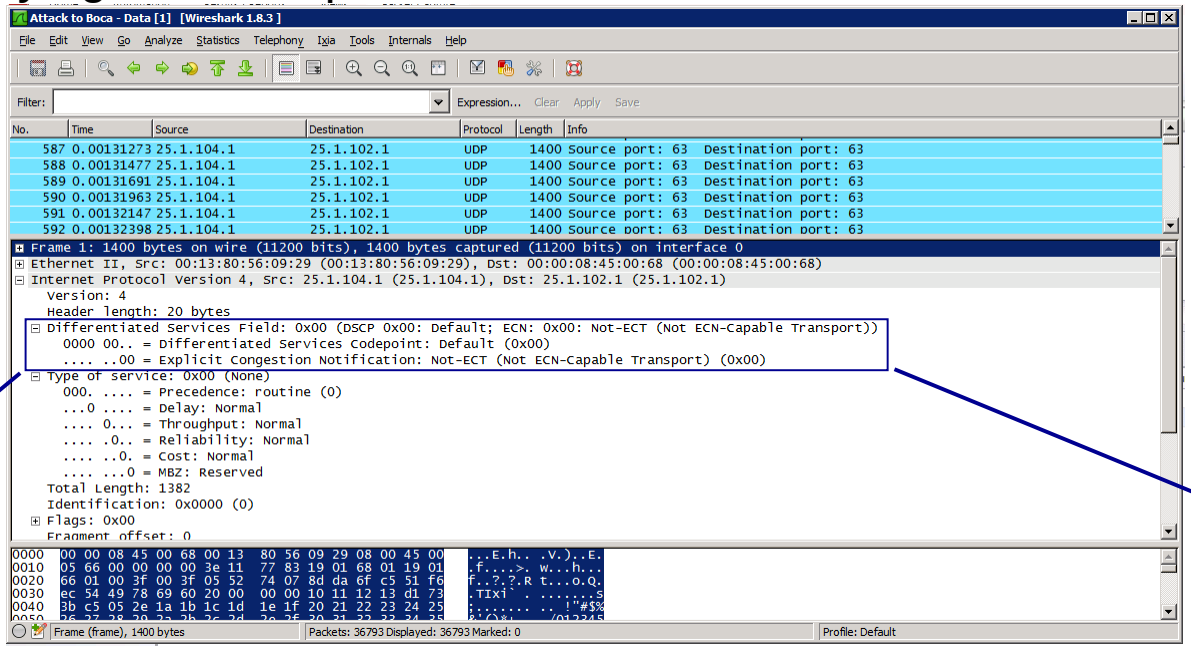
```
RP/0/RSP0/CPU0:Client#show bgp ipv4 flowspec
Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550/128
detail
BGP routing table entry for
Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550/128
<SNIP>
Last Modified: Feb  8 12:55:45.095 for 00:01:19
Paths: (1 available, no best path)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Flags: 0x4000000000020005, import: 0x20
  Not advertised to any peer
  Local
    11.22.33.44 (inaccessible) from 25.2.1.11 (6.6.6.6)
      Origin IGP, localpref 100, valid, internal
      Received Path ID 0, Local Path ID 0, version 0
      Extended community: FLOWSPEC Redirect-IP:0

RP/0/RSP0/CPU0:Client#show flowspec afi-all detail

RP/0/RSP0/CPU0:Client#
```

→ No blackhole

Configuring BGP Flowspec

Action: Set DSCP

Before applying the rules, packets are received with DSCP = 0x00



Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00: Not-ECT (Not ECN-Capable Transport))
0000 00.. = Differentiated Services Codepoint: Default (0x00)
.... ..00 = Explicit Congestion Notification: Not-ECT (Not ECN-Capable Transport) (0x00)

Configuring BGP Flowspec

Action: Set DSCP

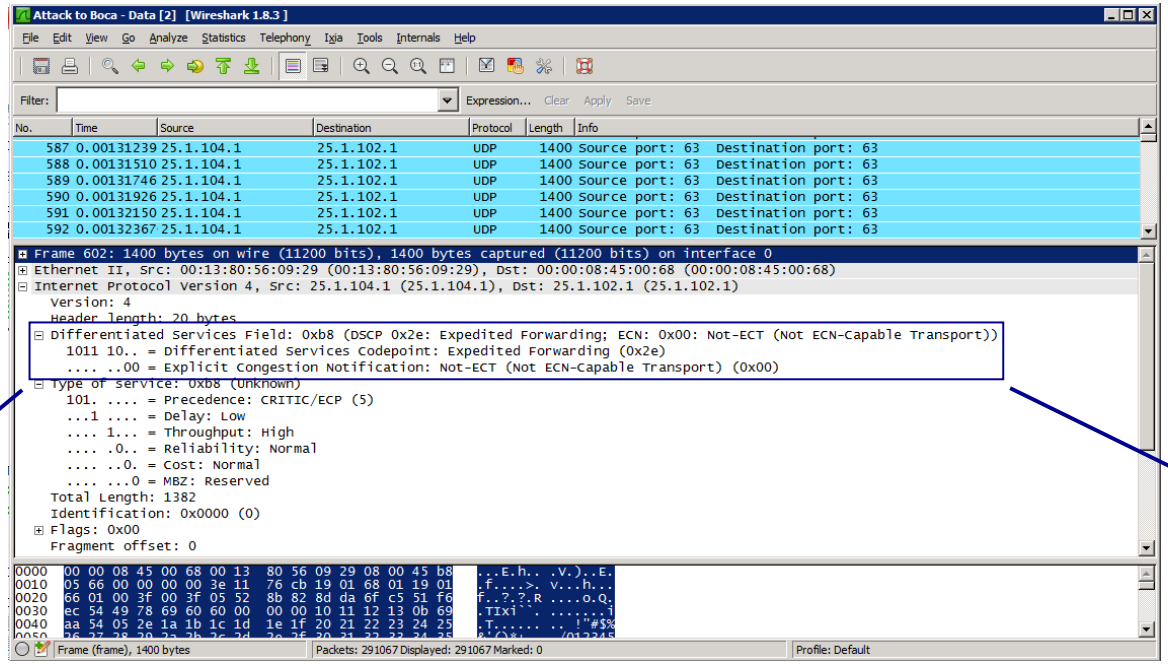
```
RP/0/0/CPU0:Ctrl#sh run policy-map type pbr TEST
policy-map type pbr TEST
 class type traffic MATCHING-RULE1
  set dscp ef
 !
 class type traffic class-default
 !
 end-policy-map
!

RP/0/0/CPU0:Ctrl#
```

```
RP/0/RSP0/CPU0:Client#show flowspec afi-all detail

AFI: IPv4
 Flow            :Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550
  Actions        :DSCP: ef  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :           594839090/832774726000
   Dropped              :                    0/0
RP/0/RSP0/CPU0:Client#
```

Configuring BGP Flowspec

Action: Set DSCP

After applying the rules, packets are received with DSCP = 0x2e (ef)



Differentiated Services Field: 0xb8 (DSCP 0x2e: Expedited Forwarding; ECN: 0x00: Not-ECT (Not ECN-Capable Transport))
1011 10.. = Differentiated Services Codepoint: Expedited Forwarding (0x2e)
.... ..00 = Explicit Congestion Notification: Not-ECT (Not ECN-Capable Transport) (0x00)

Configuring BGP Flowspec

Mixing Multiple Actions

- We can mix several Actions:
  - Rate-limit + Redirect VRF/IP
  - Rate-limit + DSCP Marking
  - Redirect VRF/IP + DSCP Marking
  - Rate-limit + Redirect VRF/IP + DSCP Marking

- It's not possible to mix:
  - Redirect VRF + Redirect NH IP
  - Redirect NH IP@A + Redirect NH IP@B

```
RP/0/RP0/CPU0:Client#sh flowspec ipv4 detail
AFI: IPv4
 Flow            :Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550
  Actions        :Traffic-rate: 100000 bps DSCP: ef Nexthop: 25.3.9.3   (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :             75899782/106259694800
   Dropped              :             75686514/105961119600
RP/0/RP0/CPU0:Client#
```

Configuring BGP Flowspec

Order of Matching Types

- Not dependent on the arrival order of the flow specification's rules
- The algorithm starts by comparing the left-most components of the rules.
- If the types differ, the rule with lowest numeric type value has higher precedence (and thus will match before) than the rule that doesn't contain that component type.

NLRI type	Match fields
Type 1	IPv4 Destination address
Type 2	IPv4 Source address
Type 3	IPv4 protocol
Type 4	IPv4 source or destination port
Type 5	IPv4 destination port
Type 6	IPv4 Source port
Type 7	IPv4 ICMP type
Type 8	IPv4 ICMP code
Type 9	IPv4 TCP flags (2 bytes include reserved bits)
Type 10	IPv4 Packet length
Type 11	IPv4 DSCP
Type 12	IPv4 fragmentation bits

Configuring BGP Flowspec

Order of Matching Types

- If the component types are the same, then a type-specific comparison is performed.
- For IP prefix values (IP destination and source prefix) precedence is given to the lowest IP value of the common prefix length; if the common prefix is equal, then the most specific prefix has precedence.
- For all other component types, unless otherwise specified, the comparison is performed by comparing the component data as a binary string using the memcmp() function as defined by the ISO C standard.
- For strings of different lengths, the common prefix is compared. If equal, the longest string is considered to have higher precedence than the shorter one.

Configuring BGP Flowspec

```
class-map type traffic match-all MATCHING-RULE1
 match protocol udp
 match packet length 500-1550
 match destination-address ipv4 25.1.102.1 255.255.255.255
 end-class-map
!
class-map type traffic match-all MATCHING-RULE2
 match protocol udp
 match packet length 500-1550
 match destination-address ipv4 25.1.102.0 255.255.255.0
 end-class-map
!
policy-map type pbr TEST1
 class type traffic MATCHING-RULE1
  redirect nexthop 25.4.9.3
 class type traffic class-default
 !
 end-policy-map
!
policy-map type pbr TEST2
 class type traffic MATCHING-RULE2
  redirect nexthop 25.3.9.3
 class type traffic class-default
 !
 end-policy-map
flowspec
 address-family ipv4
  service-policy type pbr TEST1
  service-policy type pbr TEST2
 !
```

Controller

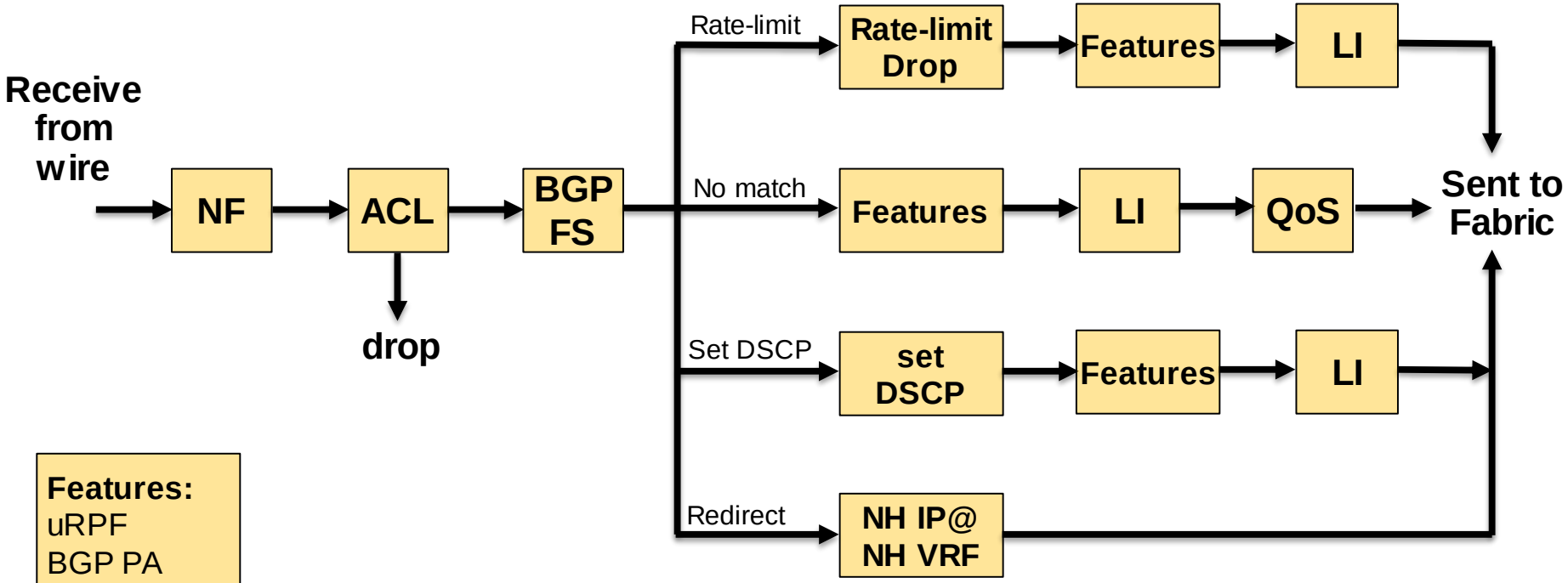
```
RP/0/RSP0/CPU0:Client#show flowspec afi-all detail

AFI: IPv4
 Flow
:Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550
  Actions        :Nexthop: 25.4.9.3   (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :             304006799/425609518600
   Dropped              :                     0/0
 Flow
:Dest:25.1.102.0/24,Proto:=17,Length:>=500&<=1550
  Actions        :Nexthop: 25.3.9.3   (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                     0/0
   Dropped              :                     0/0

RP/0/RSP0/CPU0:Client#
```

Client

25.1.102.1/32 more specific than 25.1.102.0/24

Configuring BGP Flowspec

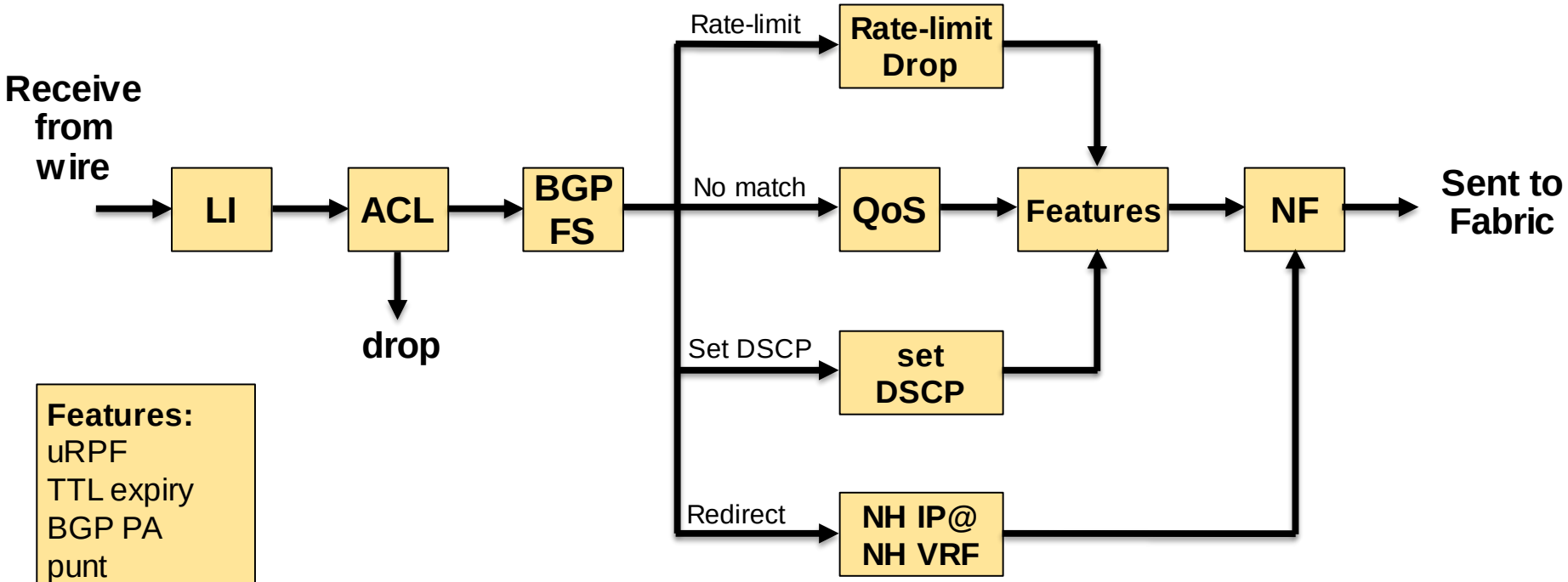
Order of Operation in CRS ASICs



Refer to back up slides for detailed tested use-cases

Configuring BGP Flowspec

Order of Operation in ASR9000 ASICs



Features:
uRPF
TTL expiry
BGP PA
punt
…

Refer to back up slides for detailed tested use-cases

Conclusion

BGP FlowSpec in SP Security

- Very powerful addition to your countermeasure tools
- Interoperable, Standard-based solution to remotely program actions on precisely identified flows
- Particularly useful in DDoS mitigation architectures
  - Filtering the stateless attacks on the Edge router, it offloads the scrubbing devices
  - Allow redirection of only the attack traffic into the scrubbing device
- Works perfectly with the ASR9000/VSM running Arbor Peakflow SP software
- XRv can be used as a controller
  - Free to test with a CCO account


Q & A
Cisco live!

Complete Your Online Session Evaluation

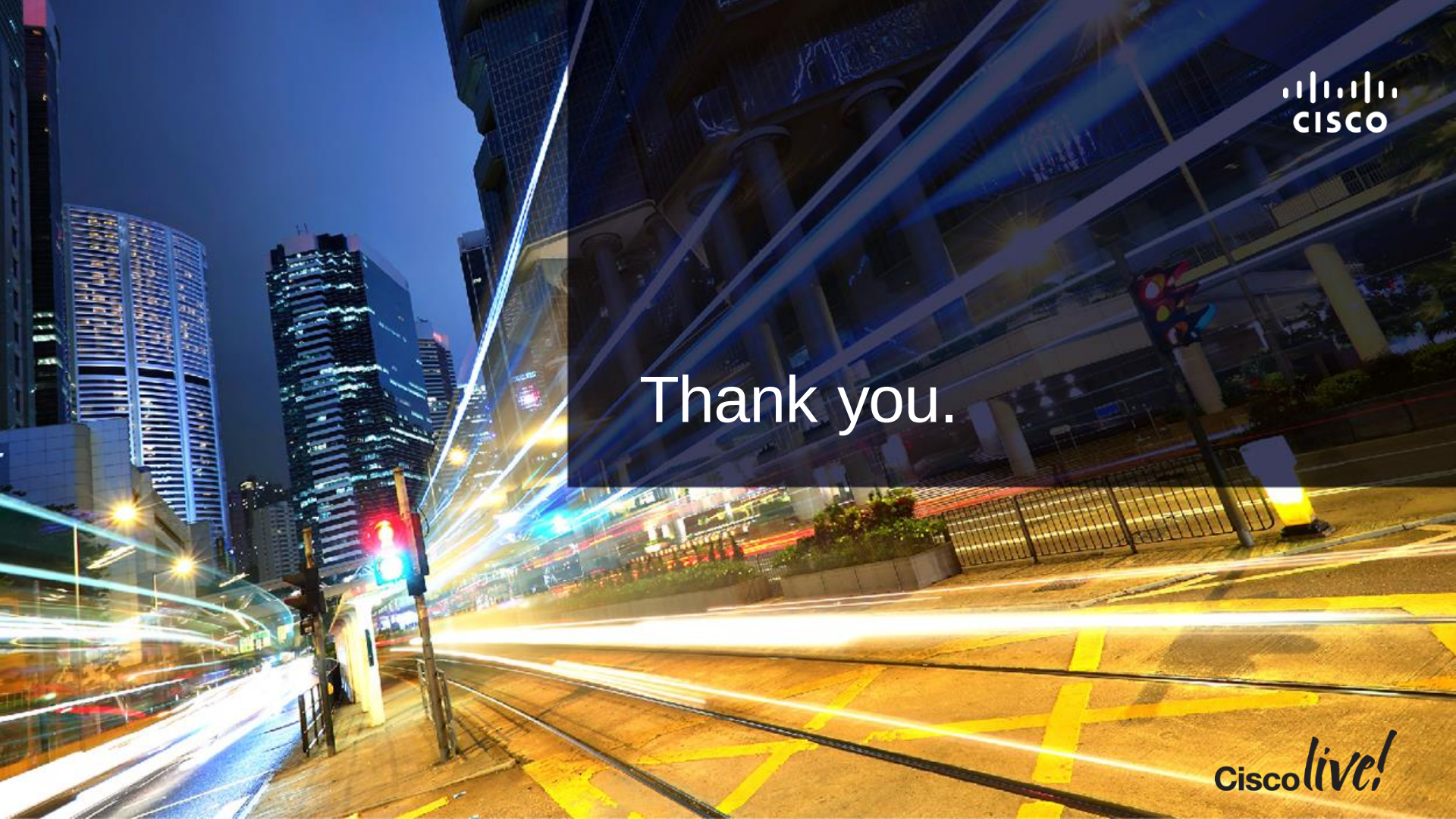
Give us your feedback and receive a Cisco Live 2015 T-Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site http://showcase.genie-connect.com/clmelbourne2015
- Visit any Cisco Live Internet Station located throughout the venue

T-Shirts can be collected in the World of Solutions on Friday 20 March 12:00pm - 2:00pm

Learn online with Cisco Live!
Visit us online after the conference for full access to session videos and presentations. www.CiscoLiveAPAC.com


Thank you.
CISCO
Cisco live!


CISCO


MONITORING

Multiple Features on the Interface

Let's try several scenarios to illustrate the order of operation.

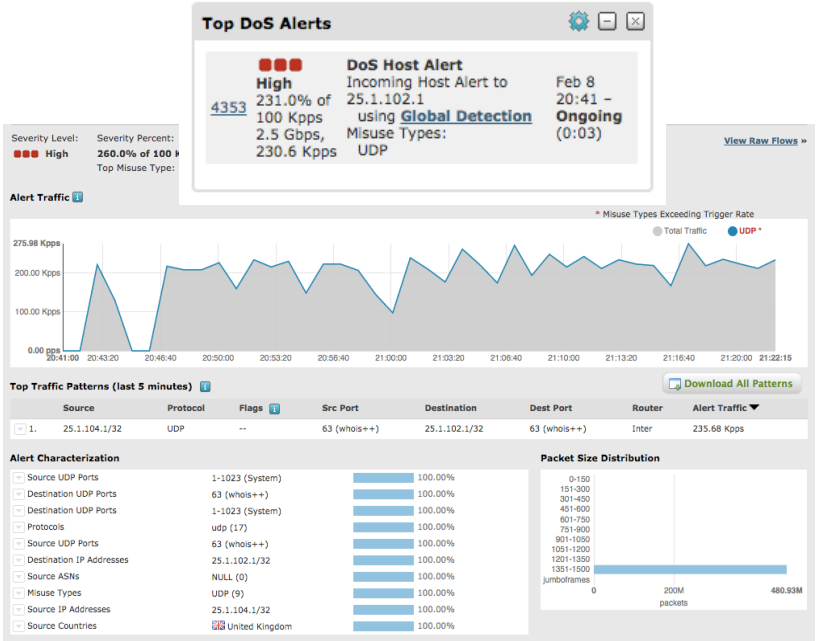
- ABF configured on interface vs BGP FS rule (Drop or Redirect)
- ACL configured on interface vs BGP FS rule
- Netflow configured on interface vs BGP FS Drop rule
- QoS configured on interface vs BGP FS rule

Netflow Sampling vs BGP FlowSpec

Even if a BGP FlowSpec rule drops the packets, they are sampled and handled by the linecard CPU.

```
RP/0/RSP0/CPU0:Client#sh run int hundredGigE 0/0/0/0
interface HundredGigE0/0/0/0
 description *** to Boca ***
 cdp
 ipv4 address 25.1.9.4 255.255.255.0
 load-interval 30
 flow ipv4 monitor MON-MAP-IP sampler SAM-MAP ingress
!

RP/0/RSP0/CPU0:Client#sh flowspec ipv4 detail
AFI: IPv4
 Flow            :Proto:=17,Length:>=500&<=1550
  Actions        :Traffic-rate: 0 bps   (bgp.1)
  Statistics                              (packets/bytes)
   Matched                :                146077011/182594343700
   Dropped                :                146077011/182594343700
RP/0/RSP0/CPU0:Client#
```

Attack still detected



Netflow Sampling vs BGP FlowSpec

Before applying the BGP FlowSpec rules, we check the NF cache:

```
RP/0/RSP0/CPU0:Client#sh flow monitor MON-MAP-IP cache location 0/0/CPU0
Cache summary for Flow Monitor MON-MAP-IP:
Cache size:                             1000000
Current entries:                         164916
Flows added:                            2043769
<SNIP>
Flows exported                          1878853

IPV4SrcAddr      IPV4DstAddr      L4SrcPort  L4DestPort BGPDstOrigAS BGPSrcOrigAS BGPNextHopV4     IPV4DstPrfxLen
IPV4SrcPrfxLen  IPV4Prot IPV4TOS  InputInterface  OutputInterface L4TCPFlags   ForwardStatus        FirstSwitched
LastSwitched    ByteCount    PacketCount  Dir SamplerID  InputVRFID                          OutputVRFID
100.102.8.178    11.200.0.2       123        123        0           0           0.0.0.0          24
0               udp      0       Hu0/0/0/0       Te0/2/0/1       0           Fwd                  12 15:47:40:093
12 15:47:40:093 1402         1           Ing 1          default                          default
100.2.42.67      11.200.0.2       123        123        0           0           0.0.0.0          24
0               udp      0       Hu0/0/0/0       Te0/2/0/1       0           Fwd                  12 15:47:51:618
12 15:47:51:618 1182         1           Ing 1          default                          default
100.77.86.28     11.200.0.2       123        123        0           0           0.0.0.0          24
0               udp      0       Hu0/0/0/0       Te0/2/0/1       0           Fwd                  12 15:48:31:530
12 15:48:31:530 1082         1           Ing 1          default                          default

RP/0/RSP0/CPU0:Client#
```

Netflow Sampling vs BGP FlowSpec

After applying the BGP FlowSpec rules, we check the NF cache:

```
RP/0/RSP0/CPU0:Client#sh flow monitor MON-MAP-IP cache location 0/0/CPU0
Cache summary for Flow Monitor MON-MAP-IP:
Cache size:                             1000000
Current entries:                          12706
Flows added:                            1467559
<SNIP>
Flows exported                          1454853

IPV4SrcAddr      IPV4DstAddr      L4SrcPort  L4DestPort BGPDstOrigAS BGPSrcOrigAS BGPNextHopV4     IPV4DstPrfxLen
IPV4SrcPrfxLen  IPV4Prot IPV4TOS  InputInterface  OutputInterface L4TCPFlags   ForwardStatus        FirstSwitched
LastSwitched    ByteCount    PacketCount  Dir SamplerID  InputVRFID                          OutputVRFID
100.37.17.132    11.200.0.2       123        123        0            0           0.0.0.0          24
0               udp      0       Hu0/0/0/0       0              0           DropACLDeny          12 15:45:00:310
12 15:45:00:310 1362         1           Ing 1          default                          0
100.47.47.62     11.200.0.2       123        123        0            0           0.0.0.0          24
0               udp      0       Hu0/0/0/0       0              0           DropACLDeny          12 15:45:01:850
12 15:45:01:850 1122         1           Ing 1          default                          0
100.11.100.55    11.200.0.2       123        123        0            0           0.0.0.0          24
0               udp      0       Hu0/0/0/0       0              0           DropACLDeny          12 15:45:00:947
12 15:45:00:947 1462         1           Ing 1          default                          0

RP/0/RSP0/CPU0:Client#
```

ACL vs BGP FlowSpec

It's important that ACL is applied before the BGP FlowSpec action.

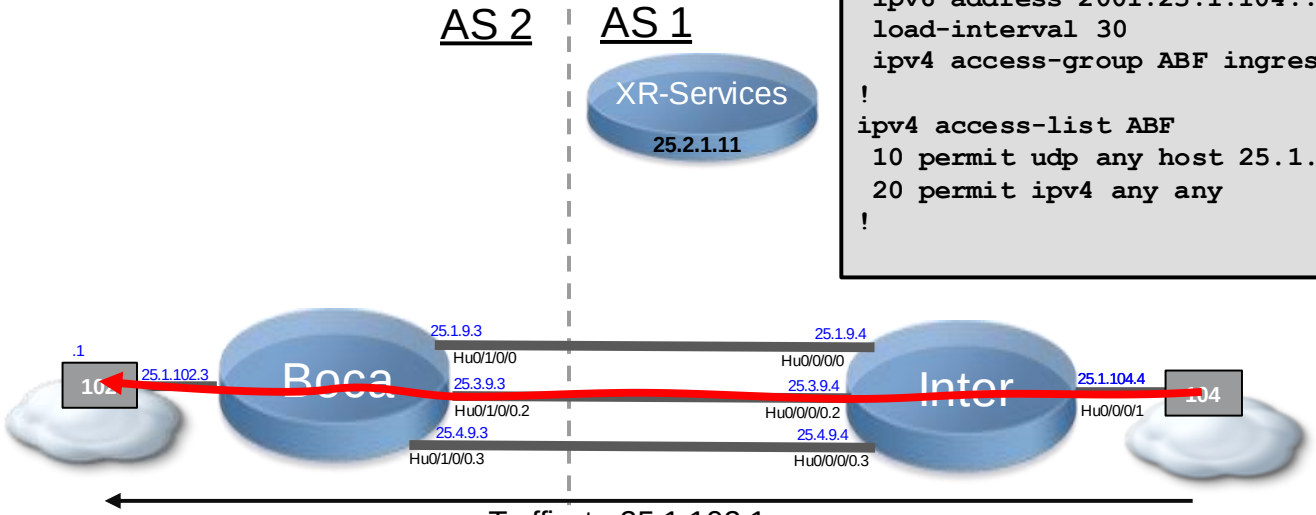
```
RP/0/RSP0/CPU0:Client#sh int hundredGigE 0/0/0/1 accounting rates
HundredGigE0/0/0/1
                              Ingress                        Egress
  Protocol              Bits/sec         Pkts/sec     Bits/sec         Pkts/sec
  IPV4_UNICAST       5065311000           458150         1000                2

RP/0/RSP0/CPU0:Client#sh flowspec ipv4 detail
AFI: IPv4
 Flow           :Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550
  Actions       :Nexthop: 25.3.9.3  (bgp.1)
  Statistics                        (packets/bytes)
   Matched             :                  0/0
   Dropped             :                  0/0

RP/0/RSP0/CPU0:Client#sh access-lists ipv4 INFRA-ACL hardware ingress location 0/0/CPU0

ipv4 access-list INFRA-ACL
 10 deny udp any host 25.1.102.1 counter INFRA-ACL-COUNT (230292976 hw matches)
 20 permit ipv4 any any
RP/0/RSP0/CPU0:Client#
```

ACL-Based Fwd (PBR) vs BGP FlowSpec

Which one will take precedence?

Before applying the BGP FS rule, on the Client side:

```
interface HundredGigE0/0/0/1
 ipv4 address 25.1.104.4 255.255.255.0
 ipv6 address 2001:25:1:104::4/64
 load-interval 30
 ipv4 access-group ABF ingress
!
ipv4 access-list ABF
 10 permit udp any host 25.1.102.1 nexthop1 ipv4 25.3.9.3
 20 permit ipv4 any any
!
```

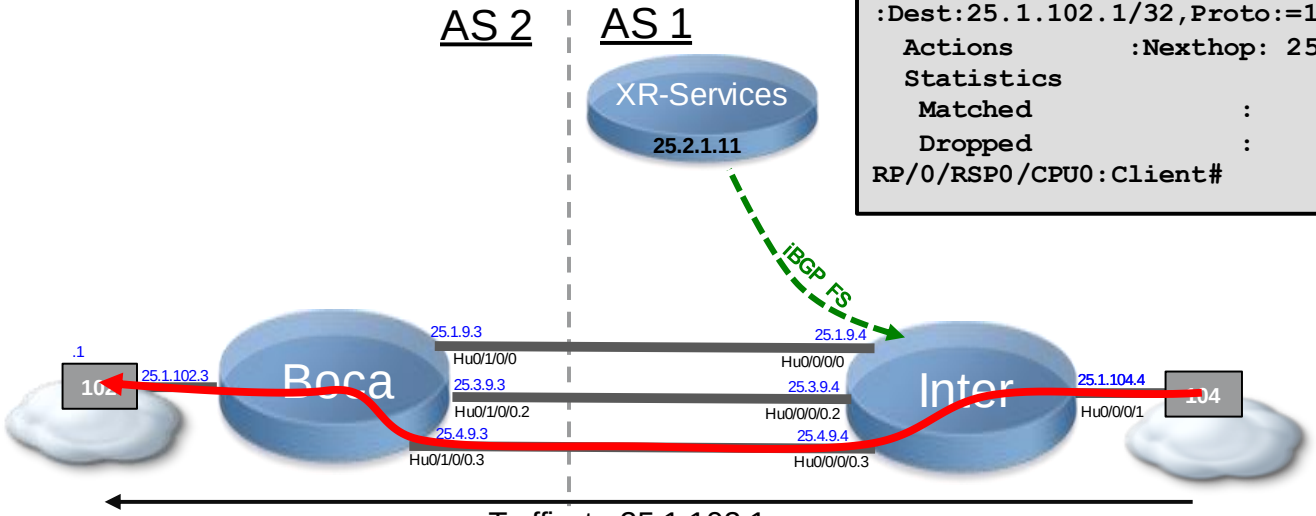


ACL-Based Fwd (PBR) vs BGP FlowSpec

BGP FlowSpec action takes precedence over ABF/PBR

After applying the rule, traffic follows the BGP FlowSpec Redirect action.

```
RP/0/RSP0/CPU0:Client#sh flowspec ipv4 detail

AFI: IPv4
 Flow
:Dest:25.1.102.1/32,Proto:=17,Length:>=500&<=1550
  Actions        :Nexthop: 25.4.9.3   (bgp.1)
  Statistics                          (packets/bytes)
   Matched                :            2217686/3104760400
   Dropped                :                  0/0
RP/0/RSP0/CPU0:Client#
```




MONITORING

Show Commands to Check BGP FlowSpec Operation

- First, we verify the BGP session for the address-family FlowSpec

```
RP/0/RP0/CPU0:Client#show bgp ipv4 flowspec

BGP router identifier 3.3.3.3, local AS number 2
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0x0   RD version: 16
BGP main routing table version 16
BGP NSR Initial initsync version 0 (Reached)
BGP NSR/ISSU Sync-Group versions 16/0
BGP scan interval 60 secs

Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale, N Nexthop-discard
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop            Metric LocPrf Weight Path
*> SPort:=80/24       0.0.0.0                                  0 1 i

Processed 1 prefixes, 1 paths
RP/0/RP0/CPU0:Client#
```

Show Commands

- Then, we can get more details for this particular rule

```
RP/0/RP0/CPU0:Client#show bgp ipv4 flowspec SPort:=80/24 detail

BGP routing table entry for SPort:=80/24
NLRI in Hex: 068150/24
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker                 16          16
    Flags: 0x04001001+0x00000000;
Last Modified: Feb  5 04:00:37.373 for 00:03:29
Paths: (1 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Flags: 0x4000000001060001, import: 0x20
  Not advertised to any peer
  1
    0.0.0.0 from 25.2.1.11 (6.6.6.6)
      Origin IGP, localpref 100, valid, external, best, group-best
      Received Path ID 0, Local Path ID 1, version 16
      Extended community: FLOWSPEC Traffic-rate:1,39269
RP/0/RP0/CPU0:Client#
```

Show Commands

- Globally, we verify which interfaces are enable for FlowSpec

```
RP/0/RP0/CPU0:Client#show policy-map transient targets type pbr

1) Policymap: __bgpfs_default_IPv4    Type: pbr
     Targets (applied as main policy):
       HundredGigE0/1/0/0 input
       HundredGigE0/0/0/0 input
       ServiceInfra7 input
       TenGigE0/2/0/5 input
       TenGigE0/2/0/8 input
       TenGigE0/2/0/4 input
     Total targets: 6

RP/0/RP0/CPU0:Client#
```

Show Commands

- We verify also how are reconstructed these policies

```
RP/0/RP0/CPU0:Client#show policy-map transient type pbr pmap-name
__bgpfs_default_IPv4

policy-map type pbr __bgpfs_default_IPv4
 handle:0x36000002
 table description: L3 IPv4 and IPv6
 class handle:0x7600000a  sequence 1024
   match source-port 80
  police rate 314152 bps
   conform-action transmit
   exceed-action drop
  !
 !
 class handle:0xf6000002  sequence 4294967295 (class-default)
 !
 end-policy-map
!
RP/0/RP0/CPU0:Client#
```

Show Commands

- Globally, we verify which interfaces are enable for FlowSpec

```
RP/0/RP0/CPU0:Client#show flowspec afi-all detail

AFI: IPv4
 Flow            :SPort:=80
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                     0/0
   Transmitted          :                     0/0
   Dropped              :                     0/0
RP/0/RP0/CPU0:Client#

RP/0/RP0/CPU0:Client#show flowspec ipv4 nlri

AFI: IPv4
 NLRI (Hex dump) :        0x068150
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
RP/0/RP0/CPU0:Client#
```

Show Commands

```
RP/0/RP0/CPU0:Client#show flowspec ipv4 internal
AFI: IPv4
 Flow            :SPort:=80
  Actions        :Traffic-rate: 314152 bps  (bgp.1)
   Client Version: 0
   Unsupported:    FALSE
   RT:
    VRF Name Cfg:   0x00
    RT Cfg:         0x00
    RT Registered:  0x00
    RT Resolved:    0x00
   Class handles:
    Handle [0]:         300000007600000a
   Class Handle Version:     1
   Sequence:                 1024
   Synced:                   TRUE
   Match Unsupported:        None
   Ref Count:                1
   Last Error:               0:No error
   Last Batch:               9
  Statistics                          (packets/bytes)
   Matched             :                    0/0
   Transmitted         :                    0/0
   Dropped             :                    0/0
RP/0/RP0/CPU0:Client#
```

Show Commands

- On a CRS client, we check the TCAM usage on the linecard

```
RP/0/RP0/CPU0:CRS-3#show contr pse tcam summary location 0/0/CPU0

<SNIP>

TCAM Device Information for Ingress PSE, CAM bank 1:
Device size: 20M (256K array entries of 80-bits), 261122 available
Current mode of operation: Turbo
<SNIP>
Feature specific information:
<SNIP>
        Flowspec IPv4 (id 32):
                Owner client id: 20.   Limit 245760 cells
                Total 1 regions using 4 CAM cells
<SNIP>
```

Show Commands

- On a ASR9000 client, we can also check the TCAM entries in some extend

```
RP/0/RSP0/CPU0:ASR9000#sh prm server tcam summary all PBR np0 location 0/0/CPU0

                Node: 0/0/CPU0:
----------------------------------------------------------------
TCAM summary for NP0:

  TCAM Logical Table: TCAM_LT_L2 (1)
    Partition ID: 0, priority: 2, valid entries: 1, free entries: 2047
    Partition ID: 1, priority: 2, valid entries: 0, free entries: 2048
    Partition ID: 2, priority: 1, valid entries: 0, free entries: 2048
    Partition ID: 3, priority: 1, valid entries: 0, free entries: 8192
    Partition ID: 4, priority: 0, valid entries: 1, free entries: 83967
  TCAM Logical Table: TCAM_LT_ODS2 (2), free entries: 89723, resvd 128
    ACL Common Region: 448 entries allocated. 448 entries free
    Application ID: NP_APP_ID_PBR (5)
      Total: 0 vmr_ids, 0 active entries, 0 allocated entries.
  TCAM Logical Table: TCAM_LT_ODS8 (3), free entries: 15204, resvd 127
    ACL Common Region: 448 entries allocated. 448 entries free
    Application ID: NP_APP_ID_PBR (5)
      Total: 1 vmr_ids, 2 active entries, 2 allocated entries.

RP/0/RSP0/CPU0:ASR9000#
```

Show Commands

- To help TAC progress faster to identify a problem

```
On the Controller:
- show run class-map
- show class-map

On the Client:
- debug flowspec all
- show flowspec trace manager event error
- show flowspec trace client event error
- show flowspec client internal
- show logging | inc FLOW
- show flowspec vrf all afi-all summary internal
- show flowspec vrf all afi-all internal
- show tech flowspec
```

Show Commands

- To measure the traffic matched, no SNMP but CLI and Netconf/XML.

```
RP/0/RP0/CPU0:Client#show flowspec ipv4 detail

AFI: IPv4
 Flow           :Dest:25.1.104.0/24
  Actions       :Traffic-rate: 100000 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :          21946725652/13958117514672
   Transmitted          :              236878/150654408
   Dropped              :          21946488774/13957966860264
 Flow           :Proto:=17,DPort:=53
  Actions       :Traffic-rate: 1234000000 bps  (bgp.1)
  Statistics                          (packets/bytes)
   Matched              :                  0/0
   Transmitted          :                  0/0
   Dropped              :                  0/0
RP/0/RP0/CPU0:Client#
```

Counters for each rule are available per VRF / address-family, not per interface.

Show Commands

- On the Client, Netconf/XML

```
<rpc message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
 <get>
  <filter>
   <Operational>
<FlowSpec></FlowSpec>
</Operational>
</filter>
</get>
</rpc>]]>]]>
```

```
<<<SNIP>>>
            <FlowTable>
              <Flow>
                <Naming>
                  <FlowNotation>
                    Dest:25.1.104.0/24
                  </FlowNotation>
                </Naming>
                <FlowStatistics>
                  <Classified>
                    <Packets>
                      21946725652
                    </Packets>
                    <Bytes>
                      13958117514672
                    </Bytes>
                  </Classified>
                  <Dropped>
                    <Packets>
                      21946488774
                    </Packets>
                    <Bytes>
                      13957966860264
                    </Bytes>
                  </Dropped>
                </FlowStatistics>
              </Flow>
<<</SNIP>>>
```


CISCO