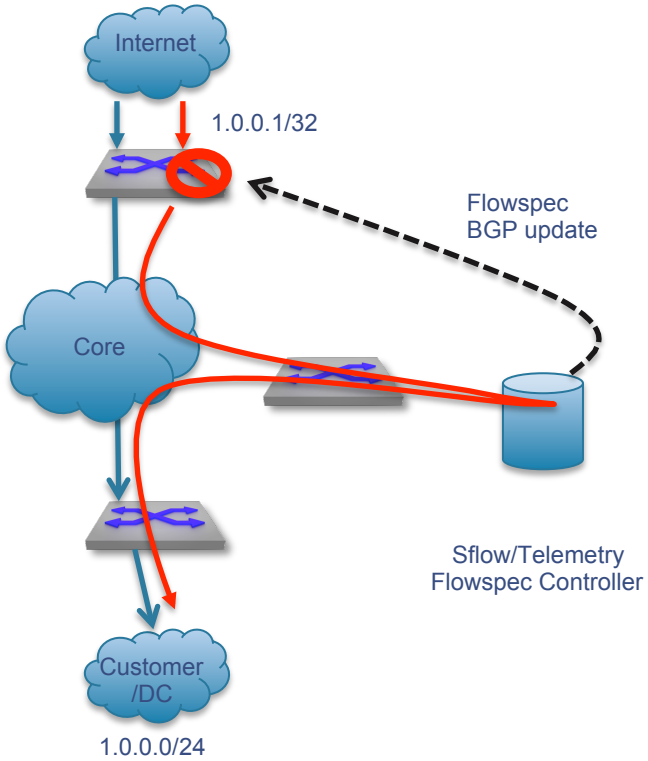
Flowspec

Peter Lundqvist peter@arista.com


Flowspec

Flowspec

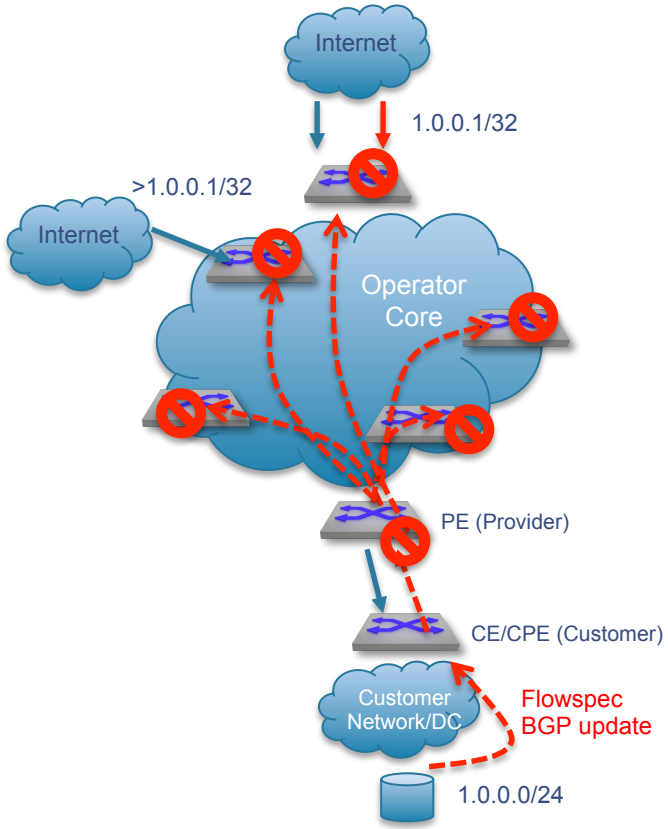
- Been around for some time…
  - However limited attraction until recently…
- Now the indication it's a game changer DDoS/ Mitigation industry
  - Opens up alternatives both regards design and the selection of tools/vendors
- Traffic can now be dropped directly at the BGP edge peering, based on information from mitigation device/controller basis of collected statistic
  - But of course be just shaped alt relay the same way as with onramp&offramp
- However Flowspec can do more…



Flowspec signaled from customer to provider

This is unrealistic scenario !

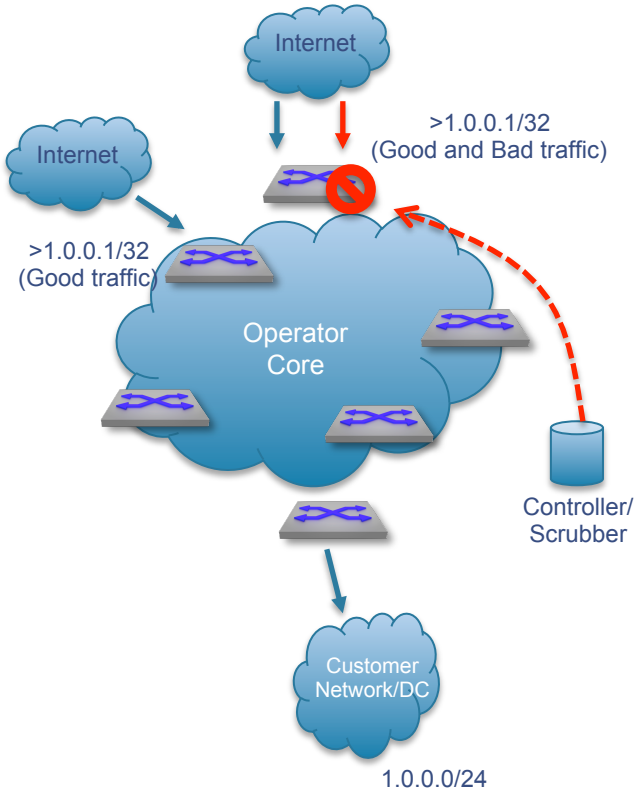
- This scenario comes up from time to time in Whitepapers written by vendors.
- Never really been implemented in real solutions with scale and control
- The problems (to begin with)
  - Trust
  - Scale&control
  - Detection
  - Etc…



Flowspec manage by Operator and controller

This is a realistic scenario

- Flowspec advertisement NOT blindly
- Rules advertised that possible knock out all traffic to a specific host within a specific port range on certain peer/Link, not whole FIB (or network).
- Instead propagated where Statistics (Sflow/ Telemetry) indicate there are DDoS event
- The Gain:
  - Capacity
  - Flowspec installed where needed
  - Mix of actions




Flowspec Theory

Flowspec Theory

- **The main go-to-source regards Flowspec is RFC5575**
  - Then there are RFC and Drafts mostly related to actions
- BGP Flowspec draft was submitted to IETF on August 14th, 2007, and it was later published in 2009 as RFC 5575http://tools.ietf.org/html/rfc5575).
- BGP Flowspec is a DDoS mitigation solution, by installing Flow related ACLs based on N-tuple information.
- The filter can be related to the attacker or the victim or both. The main idea is to walk away from punish whole prefix and thereby cause blackhole for “innocent” sessions, instead punish either individual flows or aggregation of flows.

Flowspec, Multiprotocol extension to BGP

- **Flowspec yet another Multiprotocol extension** to the BGP protocol (SAFI 133)
- Like any other capability its negotiated under the BGP OPEN phase
- Several capabilities more and less default enable when run Flowspec
  - Multiprotocol extension
  - Route-refresh
  - 4 Byte AS
  - ADD-PATH
  - Etc…

```
(…)
Border Gateway Protocol - OPEN Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 49
    Type: OPEN Message (1)
    Version: 4
    My AS: 64515
    Hold Time: 180
    BGP Identifier: 192.168.0.68
    Optional Parameters Length: 20
    Optional Parameter: Capability
            Parameter Type: Capability (2)
            Parameter Length: 6
            Type: Multiprotocol extensions capability (1)
                Length: 4
                AFI: IPv4 (1)
                Reserved: 00
                SAFI: Flow Spec Filter (133)
    Optional Parameter: Capability
            Parameter Type: Capability (2)
            Parameter Length: 2
            Capability: BGP-Extended Message
            Type: BGP-Extended Message (6)
                Length: 0
(…)
```

Flowspec rules

- Flowspec message are MP_REACH_NLRI
- **The Flowspec Rules > FLOW_SPEC_NLRI similar to a "normal IP" route**
- But Rules are ACL to be installed in the TCAM
- Like any ACL, Rules can be simple or in a more detail design
  - Here a basic IP src/dst drop rule
- **Actions are carried > EXT-COMMUNITIES**
  - Here action: rate 0 => drop all in Flowspec terms

```
(…)
        Path Attribute - EXTENDED_COMMUNITIES
             Flags: 0xc0, Optional, Transitive, Complete
             Carried extended communities: (1 community)
                 Flow spec traffic-rate: ASN 0, 0.000 Mbps [Transitive]
                     Type: Transitive Experimental (0x80)
                         1... .... = IANA Authority: Allocated
                         .0.. .... = Transitive across AS: Transitive
                     Subtype (Experimental): Flowspec traffic-rate (0x06)
                     2-Octet AS:
                     Rate shaper: 0
(…)
         Path Attribute - MP_REACH_NLRI
             Flags: 0x80, Optional, Non-transitive, Complete
             Type Code: MP_REACH_NLRI (14)
             Length: 18
             Address family identifier (AFI): IPv4 (1)
             Subsequent address family identifier (SAFI): FlowSpec (133)
             Next hop network address (0 bytes)
             Number of Subnetwork points of attachment (SNPA): 0
             Network layer reachability information (13 bytes)
                 FLOW_SPEC_NLRI (13 bytes)
                     NRLI length: 12
                     Filter type: Destination prefix filter (1)
                     10.10.10.1/32
                     Filter type: Source prefix filter (2)
                     11.11.11.1/32
(…)
```

Flowspec types theory

- **Type 1-12 ====>**
- Excluded 13 “Flow label”

Type	NLRI component provides	Used as
Type 1	prefix	Matches **destination** address in IPv4 packets against this prefix
Type 2	prefix	Matches **source** address in IPv4 packets against this prefix
Type 3	list of (operation, value)	Matches **IP protocol value byte** in IP packets
Type 4	list of (operation, value)	Matches **source or destination TCP/UDP ports**
Type 5	list of (operation, value)	Matches **destination port of a TCP or UDP packet**
Type 6	list of (operation, value)	Matches **source port of a TCP or UDP packet**
Type 7	list of (operation, value)	Matches **ICMP type**
Type 8	list of (operation, value)	Matches **ICMP code**
Type 9	list of (operation, bitmask)	Matches **TCP flags**
Type 10	list of (operation, value)	Matches **IP packet length** (exclude L2 header but include IP header)
Type 11	list of (operation, value)	Matches 6-bit **DSCP field**
Type 12	list of (operation, bitmask)	Matches **fragmentation bits**

Flowspec rules…

- **However with Flowspec, “Rules” can be as wide as you want since flags can be used**
  - Equal (=)
  - Greater than (>)
  - Less than (<)
  - Or Regular-expression combination
- **Example here matches ==>**
  - Destination prefix: 10.10.10.1/32
  - Source prefix: 11.11.11.1/32
  - Next protocol: TCP (6)
  - Destination port: 80-65535
  - Source port: 1025-65535

```
(…)
                FLOW_SPEC_NLRI (23 bytes)
                   NRLI length: 22
                   Filter: Destination prefix filter (10.10.10.1/32)
                       Filter type: Destination prefix filter (1)
                       10.10.10.1/32
                   Filter: Source prefix filter (11.11.11.1/32)
                       Filter type: Source prefix filter (2)
                       11.11.11.1/32
                           Source IP filter prefix length: 32
                           Source IP filter: 11.11.11.1
                   Filter: Protocol / Next Header filter (=6)
                       Filter type: Protocol / Next Header filter (3)
                       Operator flags: 0x81, end-of-list, Value length: 1 byte: 1 <<,equal
                           1... .... = end-of-list: Set
                           .0.. .... = and: Not set
                           ..00 .... = Value length: 1 byte: 1 << (0)
                           .... 0... = Reserved: Not set
                           .... .0.. = less than: Not set
                           .... ..0. = greater than: Not set
                           .... ...1 = equal: Set
                       Decimal value: 6
                   Filter: Destination port filter (>79)
                       Filter type: Destination port filter (5)
                       Operator flags: 0x82, end-of-list, Value length: 1 byte: 1 <<,greater than
                           1... .... = end-of-list: Set
                           .0.. .... = and: Not set
                           ..00 .... = Value length: 1 byte: 1 << (0)
                           .... 0... = Reserved: Not set
                           .... .0.. = less than: Not set
                           .... ..1. = greater than: Set
                           .... ...0 = equal: Not set
                       Decimal value: 79
                   Filter: Source port filter (>1024)
                       Filter type: Source port filter (6)
                       Operator flags: 0x92, end-of-list, Value length: 2 bytes: 1 <<,greaterthan
                           1... .... = end-of-list: Set
                           .0.. .... = and: Not set
                           ..01 .... = Value length: 2 bytes: 1 << (1)
                           .... 0... = Reserved: Not set
                           .... .0.. = less than: Not set
                           .... ..1. = greater than: Set
                           .... ...0 = equal: Not set
                       Decimal value: 1024
```

Flowspec rules…

- **More verbose rule**
- **Here Type 5 (Dst Port) rule section**
- **Example here matches ==>**
  - Destination prefix: 10.10.10.1/32
  - Source prefix: 11.11.11.1/32
  - Next protocol: TCP (6)
  - Destination port: 80,8080-8088
  - Source port: 1024-65535

```
(…)
  Filter: Destination port filter (=80 || >8080 && <8088)
          Filter type: Destination port filter (5)
                       Operator flags: 0x01, Value length: 1 byte: 1 <<, equal
                           0... .... = end-of-list: Not set
                           .0.. .... = and: Not set
                           ..00 .... = Value length: 1 byte: 1 << (0)
                           .... 0... = Reserved: Not set
                           .... .0.. = less than: Not set
                           .... ..0. = greater than: Not set
                           .... ...1 = equal: Set
                       Decimal value: 80
                       Operator flags: 0x12, Value length: 2 bytes: 1 <<, greater than
                           0... .... = end-of-list: Not set
                           .0.. .... = and: Not set
                           ..01 .... = Value length: 2 bytes: 1 << (1)
                           .... 0... = Reserved: Not set
                           .... .0.. = less than: Not set
                           .... ..1. = greater than: Set
                           .... ...0 = equal: Not set
                       Decimal value: 8080
                       Operator flags: 0xd4, end-of-list,Value length: 2bytes: 1 <<less than
                           1... .... = end-of-list: Set
                           .1.. .... = and: Set
                           ..01 .... = Value length: 2 bytes: 1 << (1)
                           .... 0... = Reserved: Not set
                           .... .1.. = less than: Set
                           .... ..0. = greater than: Not set
                           .... ...0 = equal: Not set
                       Decimal value: 8088
                  Filter: Source port filter (>=1024)
                       Filter type: Source port filter (6)
                       Operator flags: 0x93, end-of-list, Value length: 2bytes:1 <<greater than,equal
                           1... .... = end-of-list: Set
                           .0.. .... = and: Not set
                           ..01 .... = Value length: 2 bytes: 1 << (1)
                           .... 0... = Reserved: Not set
                           .... .0.. = less than: Not set
                           .... ..1. = greater than: Set
                           .... ...1 = equal: Set
                        Decimal value: 1024
```


Flowspec CLI example

Flowspec Configuration

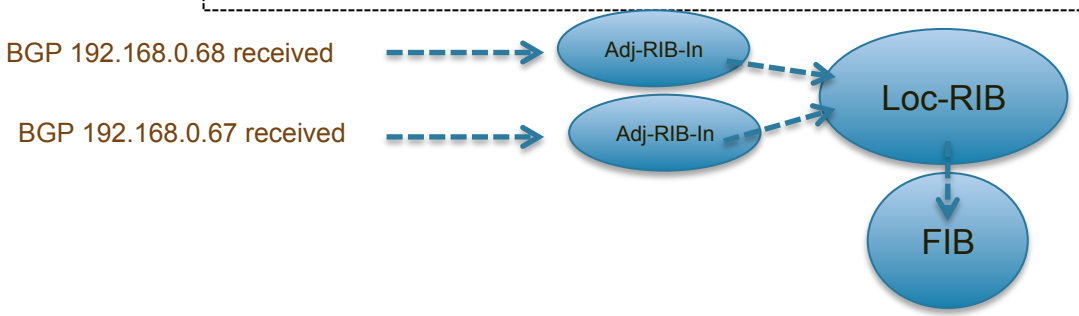
- Flowspec rules installed only on interfaces with "`flow-spec`" keyword enable
- BGP session needs SAFI 133 enable "`address-family flow-spec`"

```
!
interface Ethernet7
   description to attacker trident1
   load-interval 5
   no switchport
   flow-spec ipv4 ipv6
   ip address 6.6.6.2/24
!
(…)
router bgp 1111
   router-id 111.111.111.111
   (…)
   neighbor 64515 peer-group
   neighbor 64515 remote-as 64515
   neighbor 64515 route-map med in
   neighbor 64515 send-community
   neighbor 64515 maximum-routes 0
   (…)
   neighbor 192.168.0.67 peer-group 64515
   neighbor 192.168.0.68 peer-group 64515
   !
   address-family flow-spec ipv4
      neighbor 64515 activate
   !
   address-family ipv4
      no neighbor flow activate
   !
(…)
```

Flowspec RIB/FIB

- Follows the BGP semantics but demands “flow-spec” keyword
- Flowspec have the BGP components
  - **Adj-RIB-In**
  - Loc-RIB
  - FIB

```
qumran-flowspec#sh bgp neighbors 192.168.0.68 flow-spec ipv4 received-routes detail
BGP Flow Specification rules for VRF default

Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for
10.10.10.1/32;11.11.11.1/32;IP:=6;DP:>=80;SP:>=1024;
 Rule identifier: 3890448648
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
   IP Protocol: =6
   Destination Port: >=80
   Source Port: >=1024
 Paths: 1 available
  64515
    from 192.168.0.68 (192.168.0.68)
      Origin IGP, metric -, localpref -, weight 0, valid, external, best
          Actions: Drop
```



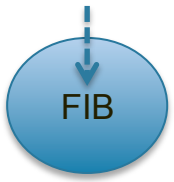
Flowspec RIB/FIB…

- Flowspec have the BGP components
  - Adj-RIB-In
  - **Loc-RIB**
  - FIB


```
qumran-flowspec#sh bgp flow-spec ipv4 detail
BGP Flow Specification rules for VRF default

Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for
10.10.10.1/32;11.11.11.1/32;IP:=6;DP:>=80;SP:>=1024;
 Rule identifier: 3890448648
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
   IP Protocol: =6
   Destination Port: >=80
   Source Port: >=1024
 Paths: 1 available
  64515
    from 192.168.0.68 (192.168.0.68)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
          Actions: Drop

BGP Flow Specification Matching Rule for 120.120.120.1/32;121.121.121.1/32;
 Rule identifier: 3890488512
 Matching Rule:
   Destination Prefix: 120.120.120.1/32
   Source Prefix: 121.121.121.1/32
 Paths: 1 available
  64515
    from 192.168.0.67 (192.168.0.67)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
          Actions: Drop
```

Flowspec EOS RIB/FIB…

- Flowspec have the BGP components
  - Adj-RIB-In
  - Loc-RIB
  - **FIB (TCAM or FSIB)**




```
qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule:
10.10.10.1/32;11.11.11.1/32;IP:=6;DP:>=80;SP:>=1024;
    Rule identifier: 3890448648
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
      Next protocol: 6
      Destination port: 80-65535
      Source port: 1024-65535
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 0 packets

  Flow-spec rule: 120.120.120.1/32;121.121.121.1/32;
    Rule identifier: 3890488512
    Matches:
      Destination prefix: 120.120.120.1/32
      Source prefix: 121.121.121.1/32
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 0 packets
```

Hardware TCAM counters

- On interfaces where Flowspec rules will be installed, static ACL and Flowspec rules do work side by side in most vendors implementations

```
qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: 10.10.10.1/32;11.11.11.1/32;
    Rule identifier: 3882113096
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 31 packets
```

Hardware TCAM profile…

- With example Jericho hardware… TCAM profile changes for Flowspec needed

```
hardware tcam
    profile flowspec
      feature acl port ip
         sequence 45
         key size limit 160
         key field dscp dst-ip ip-frag ip-protocol l4-dst-port l4-ops l4-src-port src-ip tcp-control ttl
         action count drop
         packet ipv4 forwarding bridged
         packet ipv4 forwarding routed
         packet ipv4 forwarding routed multicast
         packet ipv4 mpls ipv4 forwarding mpls decap
         packet ipv4 mpls ipv6 forwarding mpls decap
         packet ipv4 non-vxlan forwarding routed decap
         packet ipv4 vxlan eth ipv4 forwarding routed decap
         packet ipv4 vxlan forwarding bridged decap
      !
      feature flow-spec port ipv4
         key size limit 160
         key field dscp dst-ip ip-frag ip-protocol l4-dst-port l4-ops l4-src-port src-ip tcp-control
         action count redirect
         packet ipv4 forwarding routed
      !
      feature flow-spec port ipv6
         key field dst-ipv6 ipv6-next-header ipv6-traffic-class l4-dst-port l4-ops-3b l4-src-port src-ipv6 tcp-control
         action count redirect
         packet ipv6 forwarding routed
   !
    system profile flowspec
```

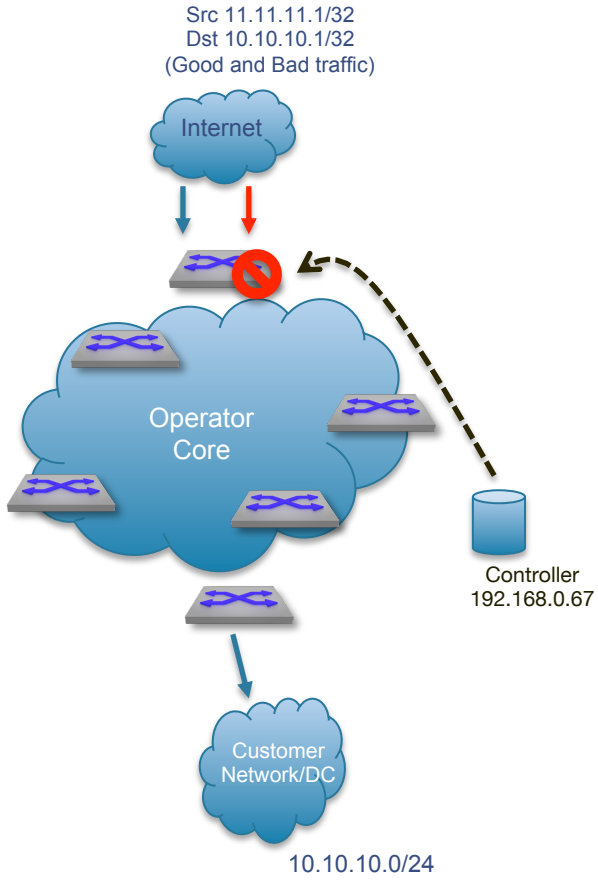

Flowspec actions

Flowspec actions

There are some main actions

- Drop (rate-limit=0)
- Redirect IP nexthop
- Redirect VRF IP nexthop (VRF-Lite)
- Redirect IP nexthop GRE
- Redirect VRF labeled nexthop (L3VPN)
- Etc…

RFC5575 define additional actions (DSCP, Sampling etc…)
Little up to each implementation…



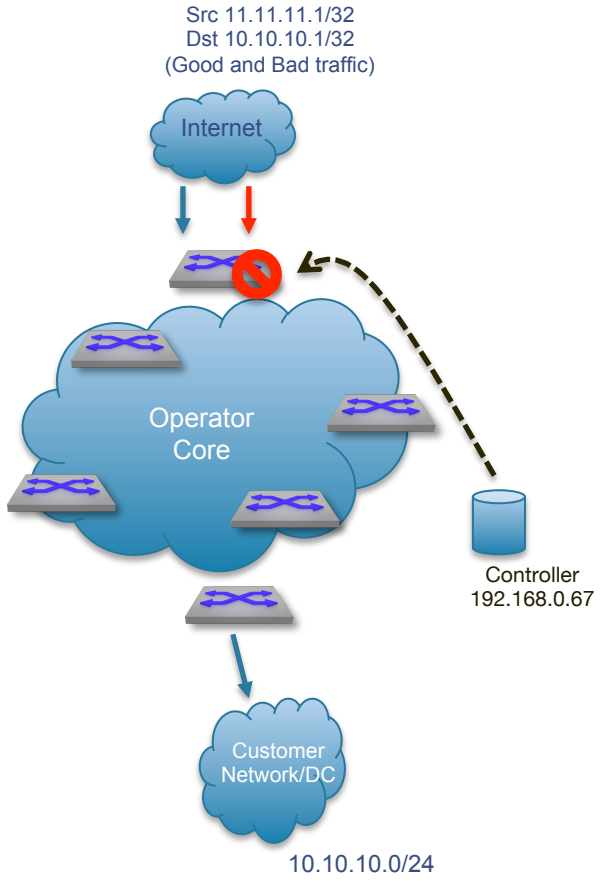
Flowspec drop rule

- **Flowspec action EXTENDED_COMMUNITIES**
- **The Flowspec Rules FLOW_SPEC_NLRI**
- Drop (rate-limit=0 Flowspec term) the most basic of all rules, in brief just drop ANY matching traffic for the installed rule
- Example here matches **====>**
  - Destination prefix: 10.10.10.1/32
  - Source prefix: 11.11.11.1/32
  - Next protocol: ICMP (1)

```
(...)
        Path Attribute - EXTENDED_COMMUNITIES
         Flags: 0xc0, Optional, Transitive, Complete
         Type Code: EXTENDED_COMMUNITIES (16)
         Length: 8
         Carried extended communities: (1 community)
             Flow spec traffic-rate: ASN 0, 0.000 Mbps [Transitive Experimental]
                 Type: Transitive Experimental (0x80)
                     1... .... = IANA Authority: Allocated
                     .0.. .... = Transitive across AS: Transitive
                 Subtype (Experimental): Flow spec traffic-rate (0x06)
                 2-Octet AS: 0

        Path Attribute - MP_REACH_NLRI
         Flags: 0x80, Optional, Non-transitive, Complete
         Type Code: MP_REACH_NLRI (14)
         Length: 21
         Address family identifier (AFI): IPv4 (1)
         Subsequent address family identifier (SAFI): Flow Spec Filter (133)
         Next hop network address (0 bytes)
         Number of Subnetwork points of attachment (SNPA): 0
         Network layer reachability information (16 bytes)
             FLOW_SPEC_NLRI (16 bytes)
                 NRLI length: 15
                 Filter: Destination prefix filter (10.10.10.1/32)
                     Filter type: Destination prefix filter (1)
                         Destination IP filter prefix length: 32
                         Destination IP filter: 10.10.10.1
                 Filter: Source prefix filter (11.11.11.1/32)
                     Filter type: Source prefix filter (2)
                         Source IP filter prefix length: 32
                         Source IP filter: 11.11.11.1
                 Filter: Protocol / Next Header filter (=1)
                     Filter type: Protocol / Next Header filter (3)
                     Operator flags: 0x81, end-of-list, Value length: 1 byte: 1 <<, equal
                         1... .... = end-of-list: Set
                         .0.. .... = and: Not set
                         ..00 .... = Value length: 1 byte: 1 << (0)
                         .... 0... = Reserved: Not set
                         .... .0.. = less than: Not set
                         .... ..0. = greater than: Not set
                         .... ...1 = equal: Set
                     Decimal value: 1
(...)
```

Flowspec action Drop

```
qumran-flowspec(config)#sh bgp flow-spec ipv4 det
BGP Flow Specification rules for VRF default
Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for 10.10.10.1/32;11.11.11.1/32;IP:=1;
 Rule identifier: 3882112672
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
   IP Protocol: =1
 Paths: 1 available
  64515
    from 192.168.0.67 (192.168.0.67)
      Origin IGP, metric -, localpref 100, weight 0, valid, external, best
      Actions: Drop

qumran-flowspec(config)#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: 10.10.10.1/32;11.11.11.1/32;IP:=1;
    Rule identifier: 3882112672
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
      Next protocol: 1
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 32 packets
```



Flowspec redirect to IP rule

- **Flowspec action EXTENDED_COMMUNITIES**
- Redirect to IP nexthop

  Two drafts exist, most vendors support both

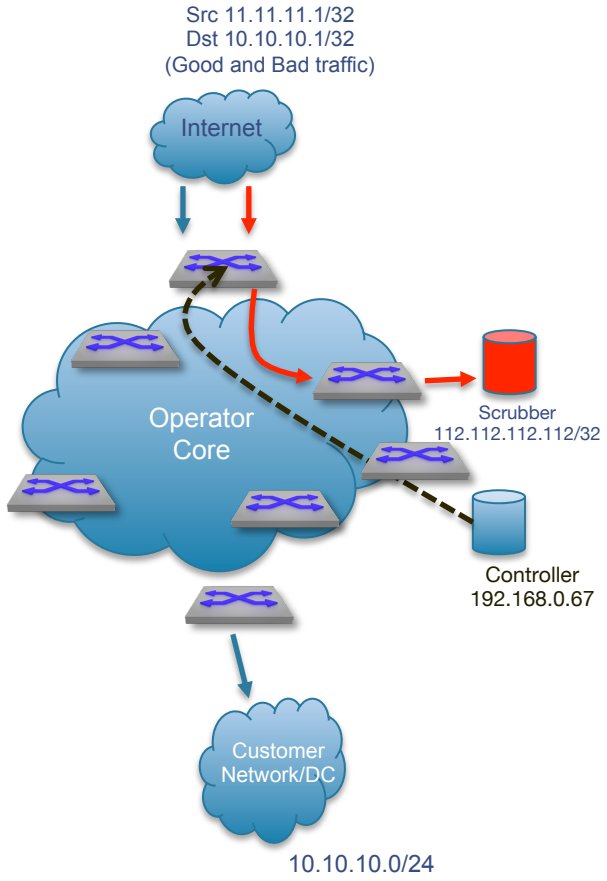
  Redirect to Nexthop (type 0x08)
  - https://tools.ietf.org/html/draft-simpson-idr-flowspec-redirect-02

  Redirect to IP (type 0x8108)
  - https://tools.ietf.org/html/draft-ietf-idr-flowspec-redirect-ip-02

```
(…)
Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
                1... .... = Optional: Set
                .1.. .... = Transitive: Set
                ..0. .... = Partial: Not set
                ...0 .... = Extended-Length: Not set
                .... 0000 = Unused: 0x0
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 8
            Carried extended communities: (1 community)
                Unknown type 0x08 subtype 0x00: 0x0000 0x0000 0x0000
                   Type: Transitive Flow spec redirect/mirror to IP next-hop (0x08)
                        0... .... = IANA Authority: Allocated on Standard Action
                        .0.. .... = Transitive across AS: Transitive
                    Subtype: 0x00
                    Raw Value: 0x0000 0x0000 0x0000
(…)
```

Flowspec action Redirect to IP

```
qumran-flowspec#sh ip ro 112.112.112.112
(…)
 S       112.112.112.112/32 [1/0] via 8.8.8.1, Ethernet5

qumran-flowspec#sh bgp flow-spec ipv4 det
(…)
Rule identifier: 3882061304
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
 Paths: 1 available
  64515
    from 192.168.0.67 (192.168.0.67)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
      Actions: Redirect IP: 112.112.112.112

qumran-flowspec#sh flow-spec ipv4
(...)
Rule identifier: 3882061304
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
    Actions:
      Redirect: VRF default, 112.112.112.112
                Route via next hop 8.8.8.1
    Status:
      Installed: yes
      Counter: 100005 packets
```



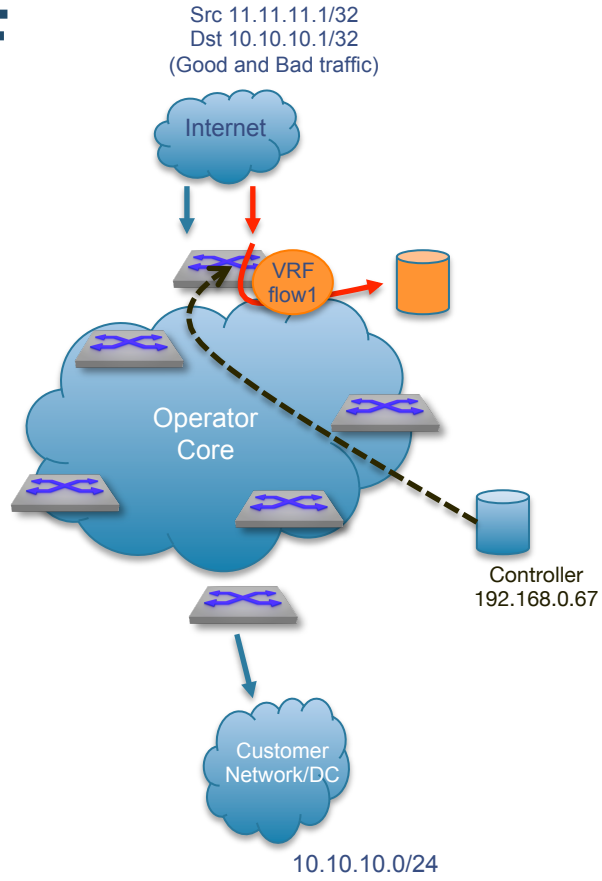
Flowspec redirect VRF rule

- **Flowspec action EXTENDED_COMMUNITIES**
- Redirect to VRF

```
(…)
 Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
                1... .... = Optional: Set
                .1.. .... = Transitive: Set
                ..0. .... = Partial: Not set
                ...0 .... = Extended-Length: Not set
                .... 0000 = Unused: 0x0
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 8
            Carried extended communities: (1 community)
                Flow spec redirect AS 2 bytes: RT 11:11 [Transitive Experimental]
                    Type: Transitive Experimental (0x80)
                        1... .... = IANA Authority: Allocated on First Come
                        .0.. .... = Transitive across AS: Transitive
                    Subtype (Experimental): Flow spec redirect AS 2 bytes (0x08)
                    2-Octet AS: 11
                    4-Octet AN: 11
(…)
```

Flowspec action Redirect to VRF

```
qumran-flowspec(config)#sh bgp flow-spec ipv4 det
BGP Flow Specification rules for VRF default
Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for 10.10.10.1/32;11.11.11.1/32;
 Rule identifier: 3882114952
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
 Paths: 1 available
  64515
    from 192.168.0.67 (192.168.0.67)
      Origin IGP, metric -, localpref 100, weight 0, valid, external, best
      Actions: Redirect VRF: 11:11 (flow1)

qumran-flowspec(config)#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: 10.10.10.1/32;11.11.11.1/32;
    Rule identifier: 3882114952
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
    Actions:
      Redirect: VRF flow1
                Route via next hop 172.16.1.1
    Status:
      Installed: yes
      Counter: 83565 packets
```




Flowspec Operation Gotchas

1) Overlapping Flowspec Rules

- Rules can be overlapping like ACL
- However the main difference between static ACL and Flowspec is that there are no sequence order for Flowspec rules
- Flowspec rules order/precedence instead based on its content
- From RFC5575 Section 5.1

(…)

With traffic filtering rules, more than one rule may match a particular traffic flow. Thus, it is necessary to define the order at which rules get matched and applied to a particular traffic flow. ***This ordering function must be such that it must not depend on the arrival order of the flow*** *specification's rules and must be constant in the network.* ***The relative order of two flow specification rules is determined by comparing their respective components***

(…)

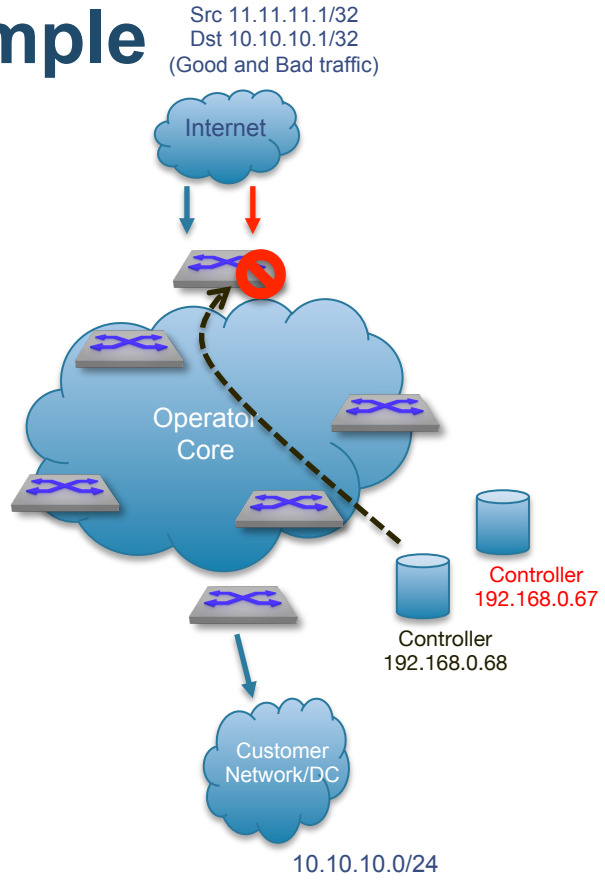
For IP prefix values (IP destination and source prefix) precedence is given to the lowest IP value of the common prefix length; if the common prefix is equal, then the most specific prefix has precedence.

(…)

Overlapping Flowspec Rule example

- Controller #1 (192.168.0.68) advertise rule to the Flowspec Router
- In brief aggregated rule that focus more on a ill-behaved /32 source…

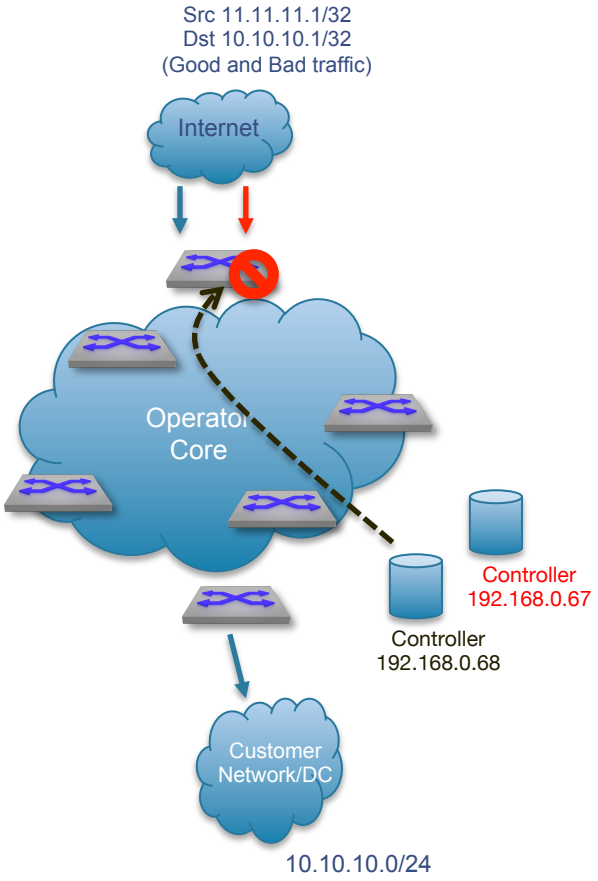
```
(…)
19:01:55 | 1366   | configuration | > neighbor          | '192.168.0.111'
19:01:55 | 1366   | configuration | . local-as          | '64515'
19:01:55 | 1366   | configuration | . peer-as           | '1111'
19:01:55 | 1366   | configuration | . hold-time         | '180'
19:01:55 | 1366   | configuration | > family            |
19:01:55 | 1366   | configuration | . ipv4              | 'flow'
19:01:55 | 1366   | configuration | < family            |
19:01:55 | 1366   | configuration | . router-id         | '192.168.0.68'
19:01:55 | 1366   | configuration | . local-address     | '192.168.0.68'
19:01:55 | 1366   | configuration | > flow              |
19:01:55 | 1366   | configuration | > route             |
19:01:55 | 1366   | configuration | > match             |
19:01:55 | 1366   | configuration | . source            | '11.11.11.1/32'
19:01:55 | 1366   | configuration | < match             |
19:01:55 | 1366   | configuration | > then              |
19:01:55 | 1366   | configuration | . discard           |
(…)
```



Overlapping Flowspec Rules…

- Rule 3882099776 Installed in the Loc-RIB and in the TCAM

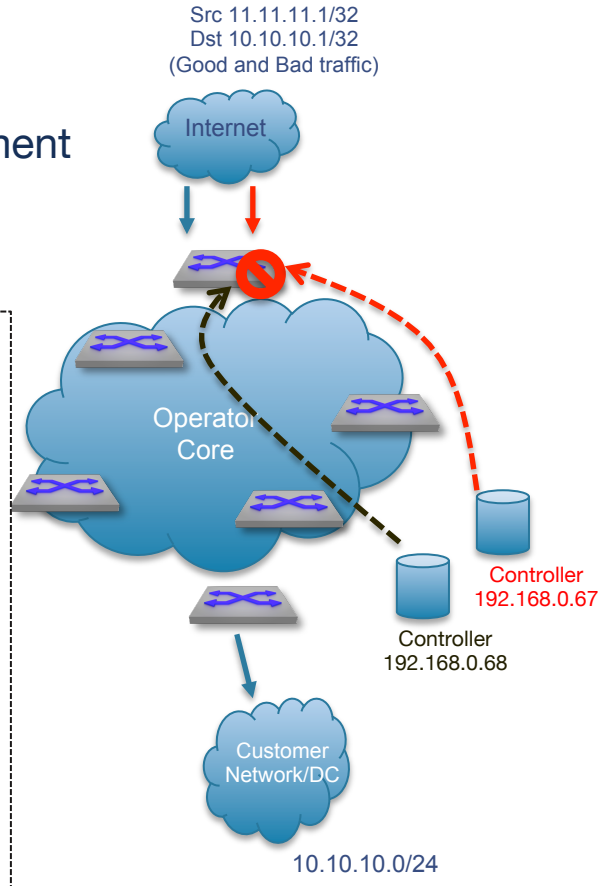
```
qumran-flowspec#sh bgp flow-spec ipv4 det
BGP Flow Specification rules for VRF default
Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for *;11.11.11.1/32;
 Rule identifier: 3882099776
 Matching Rule:
   Destination Prefix: *
   Source Prefix: 11.11.11.1/32
 Paths: 1 available
  64515
    from 192.168.0.68 (192.168.0.68)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
      Actions: Drop

qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: *;11.11.11.1/32;
    Rule identifier: 3882099776
    Matches:
      Source prefix: 11.11.11.1/32
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 14 packets <==
```



Overlapping Flowspec Rules…

- More specific rule from Controller #2 (192.168.0.68) advertisment
- In brief match both source and destination (for the same flow earlier) which means its more specific

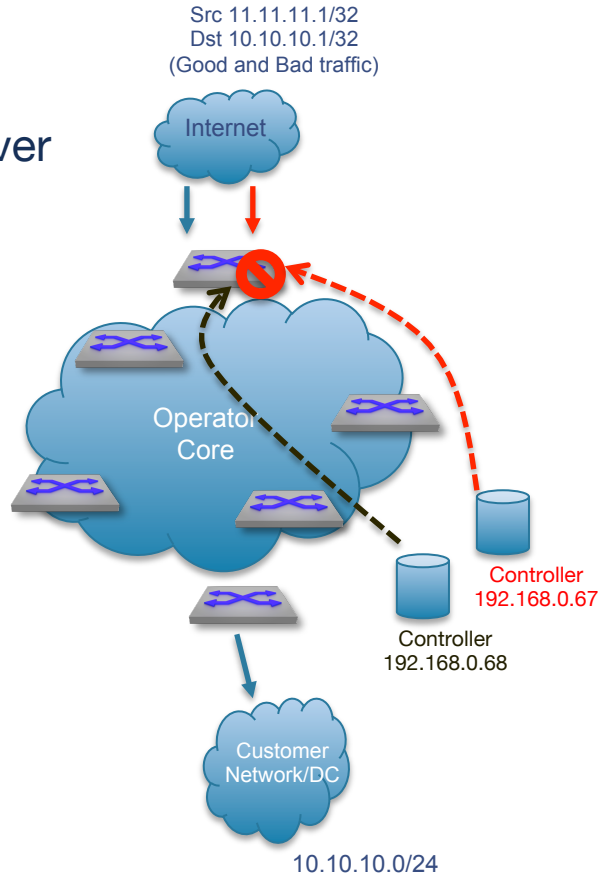
```
(…)
19:03:55 | 1391   | configuration | > neighbor          | '192.168.0.111'
19:03:55 | 1391   | configuration | . local-as          | '64515'
19:03:55 | 1391   | configuration | . peer-as           | '1111'
19:03:55 | 1391   | configuration | . hold-time         | '180'
19:03:55 | 1391   | configuration | > family            |
19:03:55 | 1391   | configuration | . ipv4              | 'flow'
19:03:55 | 1391   | configuration | < family            |
19:03:55 | 1391   | configuration | . router-id         | '192.168.0.67'
19:03:55 | 1391   | configuration | . local-address     | '192.168.0.67'
19:03:55 | 1391   | configuration | > flow              |
19:03:55 | 1391   | configuration | > route             |
19:03:55 | 1391   | configuration | > match             |
19:03:55 | 1391   | configuration | . source            | '11.11.11.1/32'
19:03:55 | 1391   | configuration | . destination       | '10.10.10.1/32'
19:03:55 | 1391   | configuration | < match             |
19:03:55 | 1391   | configuration | > then              |
19:03:55 | 1391   | configuration | . discard           |
(…)
```



Overlapping Flowspec Rules…

- Since its two different rules with the same source (however different destinations), there are no BGP path-selection
- Both installed in the Loc-RIB

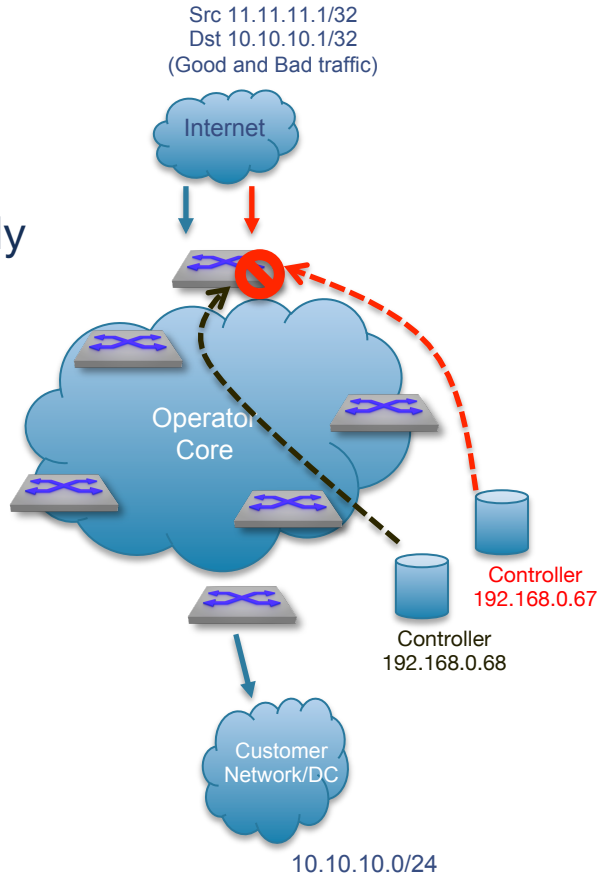
```
qumran-flowspec#sh bgp flow-spec ipv4 det
BGP Flow Specification rules for VRF default
Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for 10.10.10.1/32;11.11.11.1/32;
 Rule identifier: 3882103864
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
 Paths: 1 available
  64515
    from 192.168.0.67 (192.168.0.67)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
      Actions: Drop

BGP Flow Specification Matching Rule for *;11.11.11.1/32;
 Rule identifier: 3882099776
 Matching Rule:
   Destination Prefix: *
   Source Prefix: 11.11.11.1/32
 Paths: 1 available
  64515
    from 192.168.0.68 (192.168.0.68)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
      Actions: Drop
```



Overlapping Flowspec Rules…

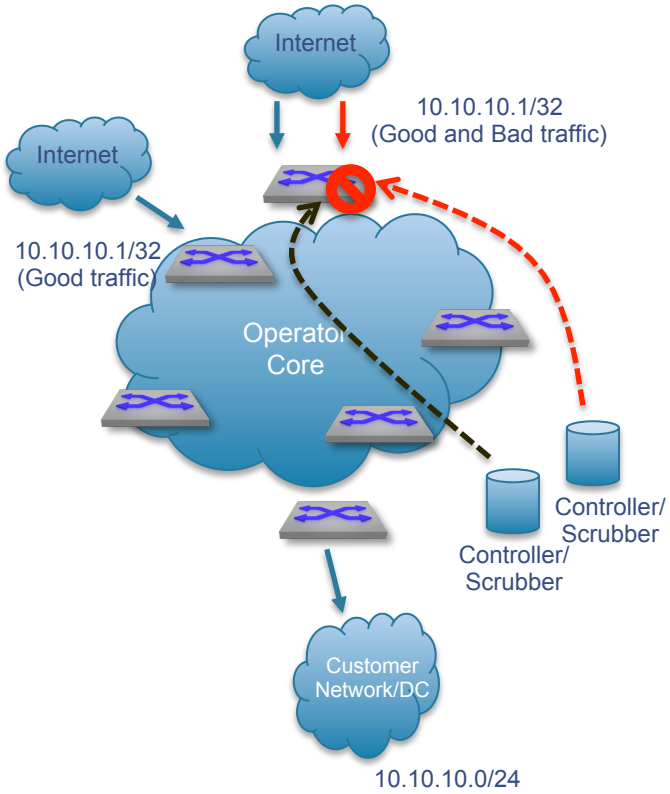
- Note the flow itself have not changed
  - (Src 11.11.11.1/32 and Dst 10.10.10.1/32)
- However the counter now only active and increase only for the more specific flow rule 3882103864

```
qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: 10.10.10.1/32;11.11.11.1/32;
    Rule identifier: 3882103864
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 56 packets <==

  Flow-spec rule: *;11.11.11.1/32;
    Rule identifier: 3882099776
    Matches:
      Source prefix: 11.11.11.1/32
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 0 packets <==
```



2) Flowspec=BGP (thereby Path-selection)

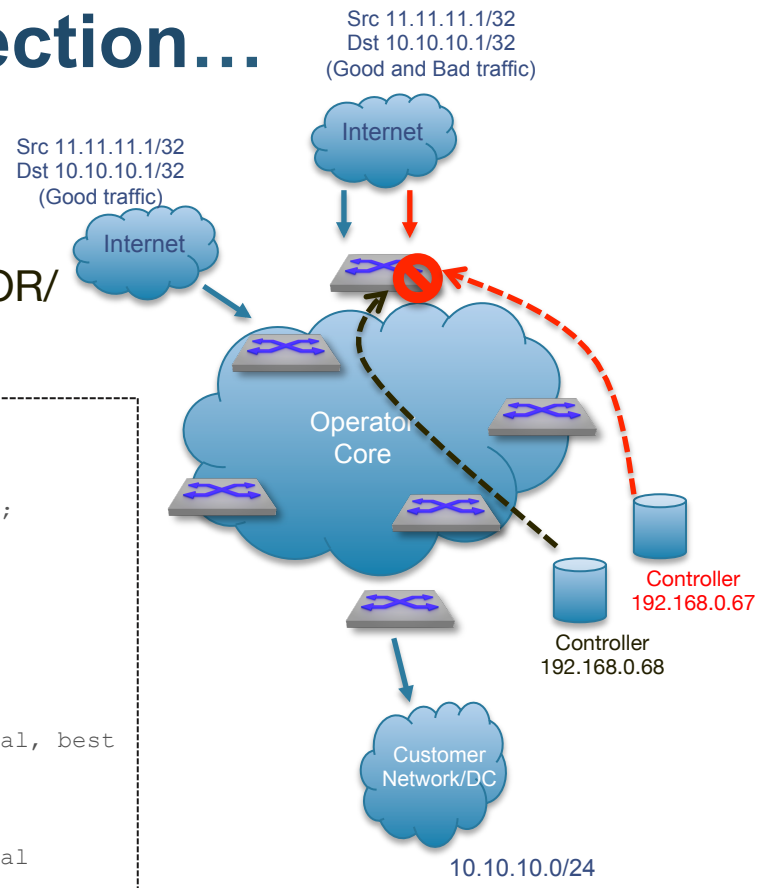
- This is a realistic scenario and needs attention !
  - Controllers can be synced and multiple just for redundancy.
  - Or simple run individually (example have different detection data thereby different mitigation)
- The Rule selection is secondary here, the selection based on the NLRI and BGP attributes
- NOTE: Flowspec NLRI have no NEXT_HOP value to consider
  - Thereby controllers that handle Flowspec Rules relative unaware of Routing logic, unless they also participate in the routing and receive complete RIB (ADD-PATH Clients)



Flowspec=BGP => Path-selection…

- DDoS flow is 11.11.11.1/32>10.10.10.1/32
  - Controller .68 sends Rule with action **Drop**
  - Controller .67 sends Rule with action **Redirect**
- Path-selection => redirect win due to ORIGINATOR/ ROUTER_ID tie-breaker

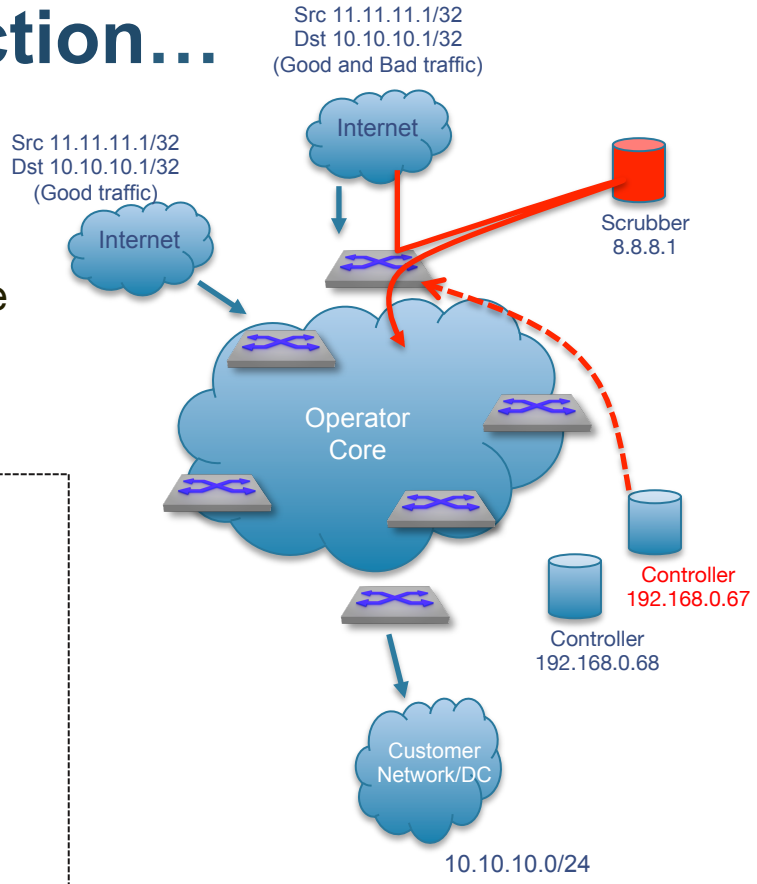
```
qumran-flowspec#sh bgp flow-spec ipv4 detail
BGP Flow Specification rules for VRF default
Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for 10.10.10.1/32;11.11.11.1/32;
 Rule identifier: 3883141696
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
 Paths: 2 available
  64515
    from 192.168.0.67 (192.168.0.67)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
      Actions: Redirect IP: 8.8.8.1
  64515
    from 192.168.0.68 (192.168.0.68)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external
      Not best: Originator/Router ID
      Actions: Drop
```



Flowspec = BGP => Path-selection…

- The actual Flowspec rule installed in the TCAM apply redirect action
- Obviously BGP path-selection needs attention regards controllers and their BGP attribute(s) since in real life probably the scenario is the opposite
  1. First try to redirect
  2. Second (if volume to large to scrub) drop

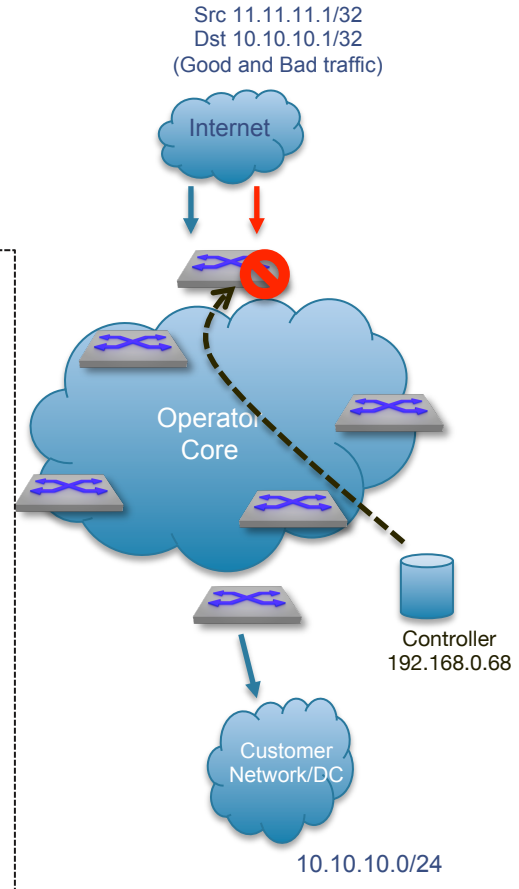
```
qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: 10.10.10.1/32;11.11.11.1/32;
    Rule identifier: 3883141696
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
    Actions:
      Redirect: VRF default, 8.8.8.1
                Route via next hop 8.8.8.1
    Status:
      Installed: yes
      Counter: 0 packets
```



Flowspec Rule error example

- Controller ship rule to the Router
- NOTE: The rule is by purpose a weirdo…

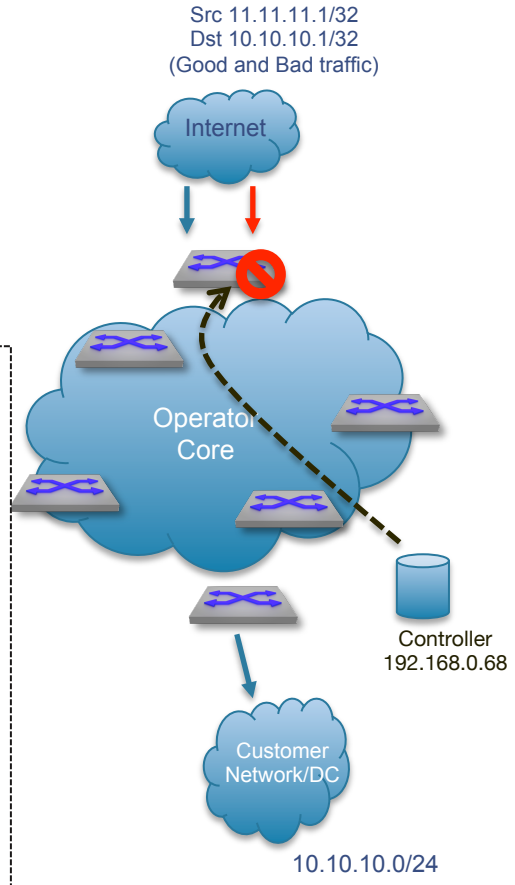
```
20:02:37 | 2354   | configuration | > neighbor          | '192.168.0.111'
20:02:37 | 2354   | configuration | . local-as          | '64515'
20:02:37 | 2354   | configuration | . peer-as           | '1111'
20:02:37 | 2354   | configuration | . hold-time         | '180'
20:02:37 | 2354   | configuration | > family            |
20:02:37 | 2354   | configuration | . ipv4              | 'flow'
20:02:37 | 2354   | configuration | < family            |
20:02:37 | 2354   | configuration | . router-id         | '192.168.0.68'
20:02:37 | 2354   | configuration | . local-address     | '192.168.0.68'
20:02:37 | 2354   | configuration | > flow              |
20:02:37 | 2354   | configuration | > route             |
20:02:37 | 2354   | configuration | > match             |
20:02:37 | 2354   | configuration | . source            | '11.11.11.1/32'
20:02:37 | 2354   | configuration | . destination       | '10.10.10.1/32'
20:02:37 | 2354   | configuration | . destination-port  | '=80&>8080&<8088'
20:02:37 | 2354   | configuration | . source-port       | '>=1024'
20:02:37 | 2354   | configuration | . protocol          | '=tcp'
20:02:37 | 2354   | configuration | < match             |
20:02:37 | 2354   | configuration | > then              |
20:02:37 | 2354   | configuration | . discard           |
20:02:37 | 2354   | configuration | < neighbor          |
```



Flowspec Rules error…

- The BGP RIB accepts the rule
- The reason for acceptance is based on the RFC where BGP part (RIB) only suppose to check BGP and the route Validation function

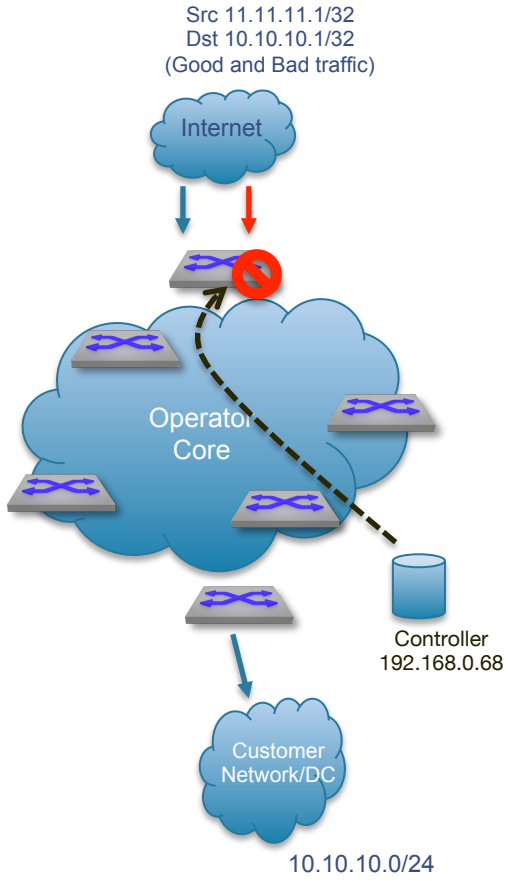
```
qumran-flowspec#sh bgp flow-spec ipv4 det
BGP Flow Specification rules for VRF default
Router identifier 111.111.111.111, local AS number 1111
BGP Flow Specification Matching Rule for
10.10.10.1/32;11.11.11.1/32;IP:=6;DP:=80&>8080&<8088;SP:>=1024;
 Rule identifier: 3882110944
 Matching Rule:
   Destination Prefix: 10.10.10.1/32
   Source Prefix: 11.11.11.1/32
   IP Protocol: =6
   Destination Port: =80 & >8080 & <8088
   Source Port: >=1024
 Paths: 1 available
  64515
    from 192.168.0.68 (192.168.0.68)
      Origin IGP, metric 111, localpref 100, weight 0, valid, external, best
      Actions: Drop
```



Flowspec Rules error…

- However the TCAM FIB do not accept the rule (since its false)
  - Dst port **equal** to 80 and needs to be between 8080-8088 !!!

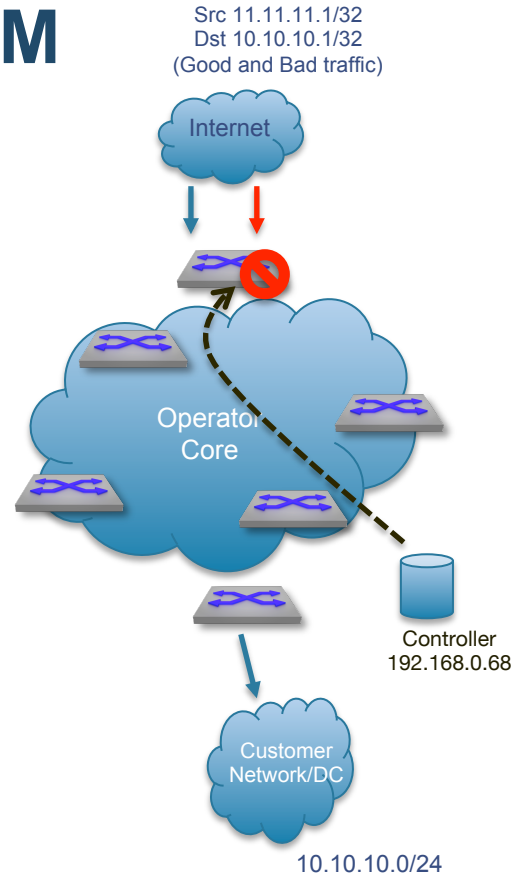
```
qumran-flowspec#sh flow-spec ipv4 infeasible
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule:
10.10.10.1/32;11.11.11.1/32;IP:=6;DP:=80&>8080&<8088;SP:>=1024;
    Rule identifier: 3882110944
    Infeasible due to Destination port
    Status:
      Installed: no(infeasible rule)
```



3) Flowspec Rule scale and TCAM

- Controller ship rule to the Router
- NOTE: The rule simple Src IP with action drop

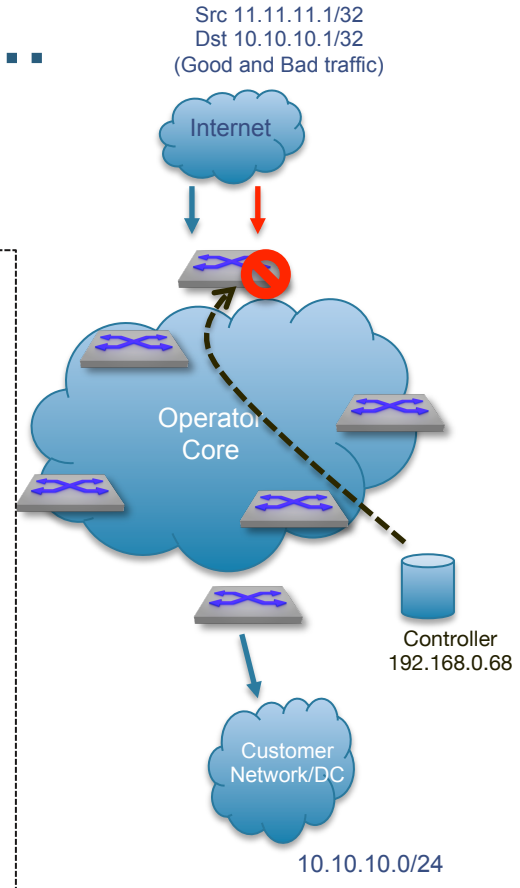
```
(…)
21:29:17 | 2521   | configuration | > neighbor          | '192.168.0.111'
21:29:17 | 2521   | configuration | . local-as          | '64515'
21:29:17 | 2521   | configuration | . peer-as           | '1111'
21:29:17 | 2521   | configuration | . hold-time         | '180'
21:29:17 | 2521   | configuration | > family            |
21:29:17 | 2521   | configuration | . ipv4              | 'flow'
21:29:17 | 2521   | configuration | < family            |
21:29:17 | 2521   | configuration | . router-id         | '192.168.0.68'
21:29:17 | 2521   | configuration | . local-address     | '192.168.0.68'
21:29:17 | 2521   | configuration | > flow              |
21:29:17 | 2521   | configuration | > route             |
21:29:17 | 2521   | configuration | > match             |
21:29:17 | 2521   | configuration | . source            | '11.11.11.1/32'
21:29:17 | 2521   | configuration | < match             |
21:29:17 | 2521   | configuration | > then              |
21:29:17 | 2521   | configuration | . discard           |
(…)
```



Flowspec Rule scale and TCAM…

- TCAM FIB accept the rule and install
- TCAM space consume 1 entry of 24576

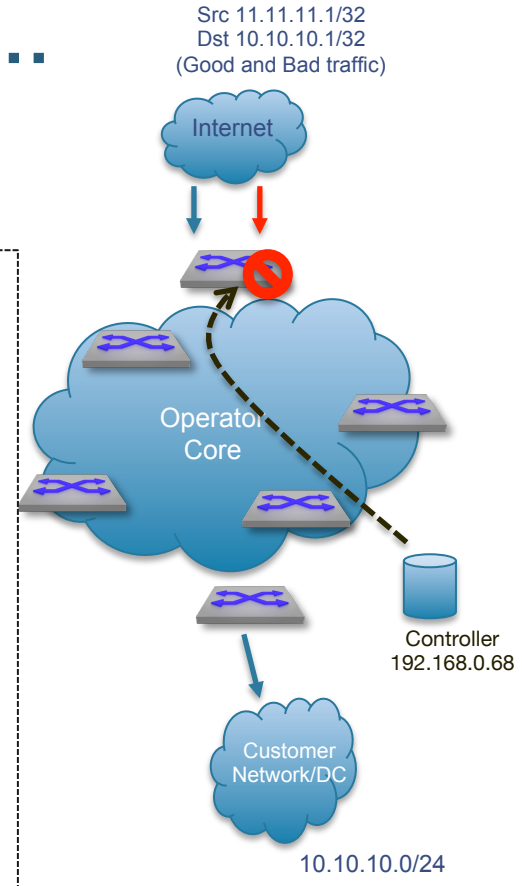
```
qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: *;11.11.11.1/32;
    Rule identifier: 3882112096
    Matches:
      Source prefix: 11.11.11.1/32
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 0 packets

qumran-flowspec#sh hardware capacity | grep Flowspec
TCAM            Flowspec            Jericho0              1
0%       24575          2048          24576             2
```



Flowspec Rule scale and TCAM…

- Controller update the rule to the Router
- NOTE: The rule “verbose” format

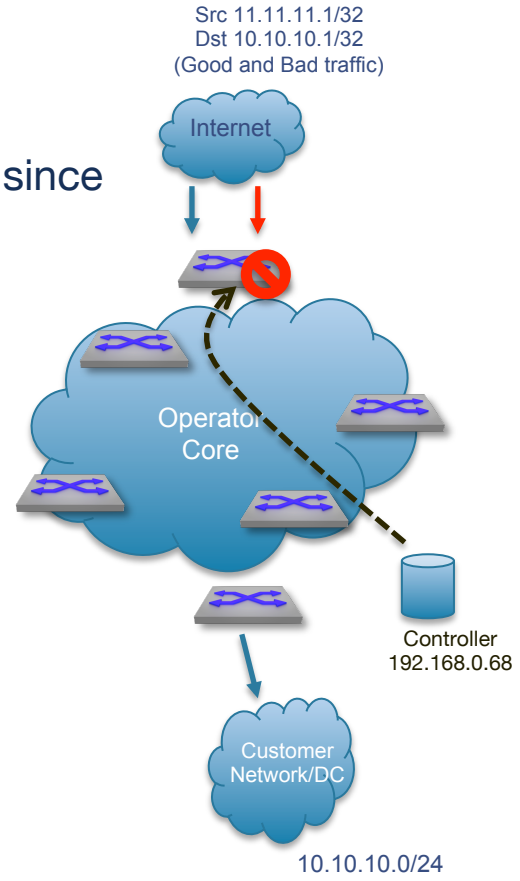
```
(…)
21:34:56 | 2536   | configuration | > neighbor          | '192.168.0.111'
21:34:56 | 2536   | configuration | . local-as          | '64515'
21:34:56 | 2536   | configuration | . peer-as           | '1111'
21:34:56 | 2536   | configuration | . hold-time         | '180'
21:34:56 | 2536   | configuration | > family            |
21:34:56 | 2536   | configuration | . ipv4              | 'flow'
21:34:56 | 2536   | configuration | < family            |
21:34:56 | 2536   | configuration | . router-id         | '192.168.0.68'
21:34:56 | 2536   | configuration | . local-address     | '192.168.0.68'
21:34:56 | 2536   | configuration | > flow              |
21:34:56 | 2536   | configuration | > route             |
21:34:56 | 2536   | configuration | > match             |
21:34:56 | 2536   | configuration | . source            | '11.11.11.1/32'
21:34:56 | 2536   | configuration | . destination       | '10.10.10.1/32'
21:34:56 | 2536   | configuration | . destination-port  | '[' '=80' '>8080&<8088' ']'
21:34:56 | 2536   | configuration | . source-port       | '>=1024'
21:34:56 | 2536   | configuration | . protocol          | '=tcp'
21:34:56 | 2536   | configuration | < match             |
21:34:56 | 2536   | configuration | > then              |
21:34:56 | 2536   | configuration | . discard           |
(…)
```



Flowspec Rules and TCAM…

- TCAM FIB accept the rule and install
- The single rule consume 2 entries of 24576 TCAM space since more complex

```
qumran-flowspec#sh flow-spec ipv4
Flow specification rules for VRF default
Applied on: Ethernet7
  Flow-spec rule: 10.10.10.1/32;11.11.11.1/32;IP:=6;DP:=80|>8080&<8088;SP:>=1024;
    Rule identifier: 3882110920
    Matches:
      Destination prefix: 10.10.10.1/32
      Source prefix: 11.11.11.1/32
      Next protocol: 6
      Destination port: 80
                        8081-8087
      Source port: 1024-65535
    Actions:
      Drop
    Status:
      Installed: yes
      Counter: 0 packets


qumran-flowspec#sh hardware capacity | grep Flowspec
TCAM              Flowspec              Jericho0                2
0%        24574             2048           24576              2
```




Questions