BGP Flow Specification – V2? draft-hares-idr-flow-spec-combo-01.txt

Susan Hares

February 8, 2016

Agenda

- Overview
- BGP Flow Specification changes – Option 1 (minimal) or Option 2 (ordered),
- Precedence with Other Filters
- BGP Flow Specification Security
- BGP Flow Specification Yang module

Overview

- Review of RFC5575
- Calls for additions to specification
- Why Precedence (ordering) is needed in BGP Flow Specifications (BGP-FS) for currently proposed actions?
-

Flow Spec (RFC5575) Review

RFC 5575 summary

- NLRI
  - For SAFI 133: IPv4 (AFI=1), IPv6 (AFI=2), L2VPN AFI=25)
  - For SAFI 134: IPv4 (AFI=1), IPv6 (AFI=2), L2VPN (AFI=25)
- Validation
  - Originator of flow spec = originator of best-match unicast route for destination embedded in NLRI,
  - No more specific unicast routes, when compared with Flow destination prefix, that have been received from different neighbor AS

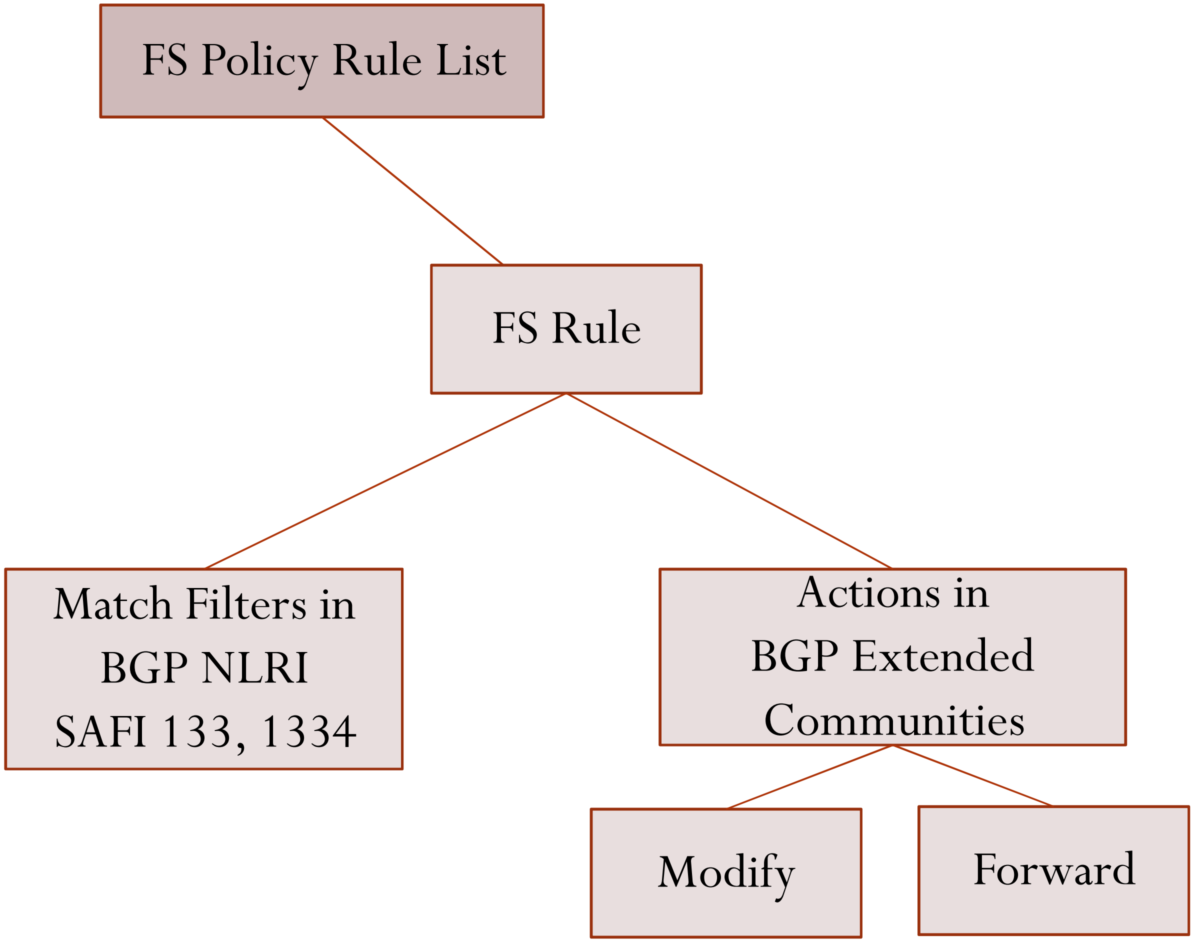
Problems with RFC5575

- Security is – Pre-ROA
- Approved L2VPN doesn't fit

? – No destination check for SDN/VPN

? – without ordering some policies can not be expressed

Flow Specification Policy



FS Policy Rule List

FS Rule

Match Filters in
BGP NLRI
SAFI 133, 1334

Actions in
BGP Extended
Communities

Modify

Forward

BGP Flow Specification is ECA Policy

- **ECA = Event –Condition - Action**
  - Flow-specification event = "packet reception",
  - Condition – match filters in NLRI
  - Action – in Extended communities

- BGP Flow Specifications: last update from BGP peer

Calls for Additions

- 2 IDR WG drafts + 9 proposals - Need rules for combination
  - draft-ietf-idr-flowspec-v6
  - draft-ietf-idr-flowspec-l2vpn
  - draft-eddy-idr-flowspec-packet-rate
  - draft-eddy-idr-flowspec-exp
  - draft-hao-idr-flowspec-nv03
  - draft-hao-flowspec-redirect-tunnel
  - draft-li-idr-flowspec-rpd
  - draft-liang-idr-bgp-flowspec-label
  - draft-liang-idr-flowspec-time
  - draft-litkowski-idr-flowspec-interfaceset
  - draft-vandevelde-idr-flowspec-path-redirect

Why is Precedence needed?

Precedence (order) is needed within BGP Flow Specification

- For filtering – Currently all
  - For ordering policies: use NLRI preference and administrative distance,
  - For ordering filters – by Flow Specification type and precedence - allows L2 and the L3
- For action
  - No order currently, need to add order for option 1 or option

BGP FS Filters types for RFC/WG documents

- RFC 5575 types/v6-draft
  1. Destination prefix
  2. Source prefix
  3. IPv4 protocol / IPv6 Next header
  4. Port (source or destination)
  5. Source port
  6. Destination port
  7. ICMP Type
  8. ICMP Code
  9. TCP Flags
  10. Packet length
  11. Traffic Class
  12. IPv4 Fragment
  13. IPv6 Flow ID

- L2VPN types
  14. Ethernet type
  15. Source MAC
  16. Destination MAC
  17. DSAP in LLC
  18. SSAP in LLC
  19. Control fields in LLC
  20. SNAP
  21. VLAN ID
  22. VLAN COS
  23. Inner VLAN ID
  24. Inner VLAN COS

BGP FS Proposed Filter types

- MF-1: NV03 Delimiter
  - Inner/outer header info
- MF-2: Virtual Network ID (VNID)
- MV-3: Flow ID (NVGRE Flow ID)

- MF-4 : MPLS LSP label or label stack
- MF-5: Interface Grouping
- MF-6: Time matches

Are there others?

- Policy distribution of BGP Flow Specification actions can be handled by Wide-Community actions

BGP FS Filters: Precedence Rules (1)

Precedence logic for BGP Flow Specifications

(RFC5575, draft-idr-bgp-flowspec-l2vpn)

```
flow-rule-cmp (a,b)
{
  comp1 = next_component(a);
  comp2 = next_component(b);
  while (comp1 || comp2) {
   // component_type returns infinity on end of list
   if (component_type(comp1) &lt; component_type(comp2)) {
    return A_HAS_PRECEDENCE;
    }

   if (component_type(comp1) &gt; component_type(comp2)) {
    return B_HAS_PRECEDENCE;
   }
```

BGP FS Filters Precedence Rules (2)

```
// IP values)
 if (component_type(comp1) == IP_DESTINATION || IP_SOURCE) {
  common = MIN(prefix_length(comp1),prefix_length(comp2));
          cmp = prefix_compare (comp1,comp2,common);
          // not equal, lowest value has precedence
          // equal, longest match has precedence;
 } else if (component_type (comp1) == MAC_DESTINATION ||
     MAC_SOURCE) {
                  common = MIN(MAC_address_length(comp1),
                      MAC_address_length(comp2));
                  cmp = MAC_Address_compare(comp1,comp2,common);
                  //not equal, lowest value has precedence
                  //equal, longest match has precedence
         } else {
   common = MIN(component_length(comp1),
                      component_length(comp2));
           cmp = memcmp(data(comp1), data(comp2), common);
                  //not equal, lowest value has precedence
                  //equal, longest string has precedence
  }
 }
}
```

Flow Specification Actions

Approved Actions

(RFC 5575 & RFC 7674)

- Traffic rate in bytes (0x8006)
- Traffic Action (0x8007) with S(sample) T (terminal) flags
- Redirect to IP VPN via Route Target
  - RD 2 octet AS, 4 byte value (0x8008)
  - RD 4 octet IP, 2 byte value (0x8108),
  - RD 4 octet AS, 2 byte value (0x8208)

Proposed Actions

- (FA1) Traffic Rate in packets
- (FA2) Traffic Action with "R" for refer to more policy in BGP Attribute
- (FA3) Redirect to Tunnel
- (FA4) VLAN Action
- (FA5) TPID action
- (FA6) MPLS label action (push, pop, swap)
- (FA7) change validation to ROA or bgpsec-protocol
- (FA8a) interface set
- (FA8b) ACL+BGP FS

Default Precedence for BGP FS actions

- **Filters – AND**
  - 01-13: IP Protocol
  - 14-16: NVO3 matches [MF1-MF3]
  - 17: Segment ID
  - 18-29: MPLS [MF-4 + others]
  - 30-40: L2VPN matches (14-24)
  - 41: Interfaces matches (MF-5)
  - 42: Time matches (MF-6)
  - 43: IPv4 Neighbor
  - 44: IPv6 Neighbor
  - 45: AS Neighbor

Action

1. Alternate NLRI validation (FA-7)
2. Traffic rate in bytes (0x8006)
3. Traffic rate in packets (FA-1)
4. Traffic Action (0x8007)
5. Extended Traffic Action (FA-2)
6. Redirect to IP VPN (0x8008, 0x8108, 0x8208)
7. Redirect to tunnel (FA-3)
8. VLAN action (FM-4)
9. TPID action (FM-5)
10. Label Action (FM-6)
11. Interface Set (FM-8a)
12. Protocol Filter precedence (FM-8b)

Possible Conflicts

	Possible conflicts										
Action	Traffic rate Bytes	Traffic Rate Pkts	Traffic Action	Ext. Traffic Action	Redirect To IP VPN	Redirect to IP Tunnel	VLAN	TPID	Label	Intf Set	BGP valid
Traffic Rate Bytes		X									
Traffic rate Pkts	X										
Traffic action				X							
Ext. Traffic action			X								

Possible Conflicts

	Possible conflicts										
Action	Traffic rate Bytes	Traffic Rate Pkts	Traffic Action	Ext. Traffic Action	Redirect To IP VPN	Redirect to IP Tunnel	VLAN	TPID	Label	Intf Set	BGP valid
Redirect IP VPN						X	X	X	X	X	
Redirect Tunnel					X		X	X	X	X	
VLAN					X	X		X	X	X	
TPID					X	X	X		X	X	
Label					X	X	X	X		X	
Intf. Set					X	X	X	X	X		

BGP-FS Precedence vs. other Protocols

- BGP Flow Spec is filter-Based forwarding
- Precedence between filter-forwarding
  - Routing yang modules
  - ACLs
  - Filter-Based (n-tuple policy)
  - I2RS Filter-Based RIB
  - BGP Flow Specification
- Currently done by local configuration
  - Yang modules require additional specificatoins

Packet/Frame Forwarding Filters

- **Where Forwarding Filters are created**
  - Configuration level: ACLs, PBRs
  - Box/module Ephemeral: I2RS
  - BGP Session Level: BGP Flow Specification

- **Filter-Based Forwarding is Minimalistic ECA Policy**
  - **Event** = packet reception on interfaces
  - **Match Condition** = Match on Filters
  - **Actions** – Modify packet, and Forward (or Drop)

- **Filters should have Yang data modules aligned**
- **Should this impact how BGP Flow filters are passed?**

Precedence between Flow Filters

- Why needed:
  - draft-litkowski-idr-flowspec-interfaceset proposes

    Really two actions
    - Apply policy to group of interfaces
    - Combine ACL + BGP Flow Specification filtering

  - Need Default Precedence + Policy Preference between:
    1) BGP Flow Specification (BGP Session Ephemeral)
    2) I2RS Filter Based RIB (Reboot Ephemeral)
    3) Filter-Based forwarding (aka Policy Routing) – configuration
    4) ACL – configuration

  - Propose Most dynamic (1st) to least dynamic (1-4 above)

BGP-FS Options

- Two options + Use cases
- Changes for each option
- Changes for

Use Case for

- Option 1: Minimal Flow Specification
  - Use Case: Prevent DoS

- Option 2:
  - Use Case: SDN/NFV central controller for paths or segments
  - Why BGP: Peer distribution of some filters from a certain
  - Not: I2RS vs. BGP – but the use of specific filters.

Considering two options

- Option 1: Minimally upgrade BGP-FS for
  - Add optional use of ROA for Security
  - Define default precedence ordering for filters
  - Define default precedence ordering for actions
  - Define precedence between BGP-FS and other packet filters (E.g. I2RS FB-RIB)
  - Define conflict resolution between actions

- Option 2: BGP Flow specification V2
  - Add optional use of ROA for Security
  - Define default precedence ordering for filters within same order
  - Define default precedence ordering for actions within same order
  - Define precedence between BGP-FS and other packet filters (E.g. I2RS FB-RIB)
  - Define conflict resolution between actions

  - BGP-FSv2 NLRI + actions in BGP Wide Communities
    - BGP-FS NLRI supports ordering of filters
    - BGP-FS Wide Community atom and (optional) container type supports ordering of actions

Description of Common actions

- Add optional use of ROA for Security
- Define default precedence ordering for filters
- Define default precedence ordering for actions
- Define precedence between BGP-FS and other packet filters (E.g. I2RS FB-RIB)
- Define conflict resolution between actions

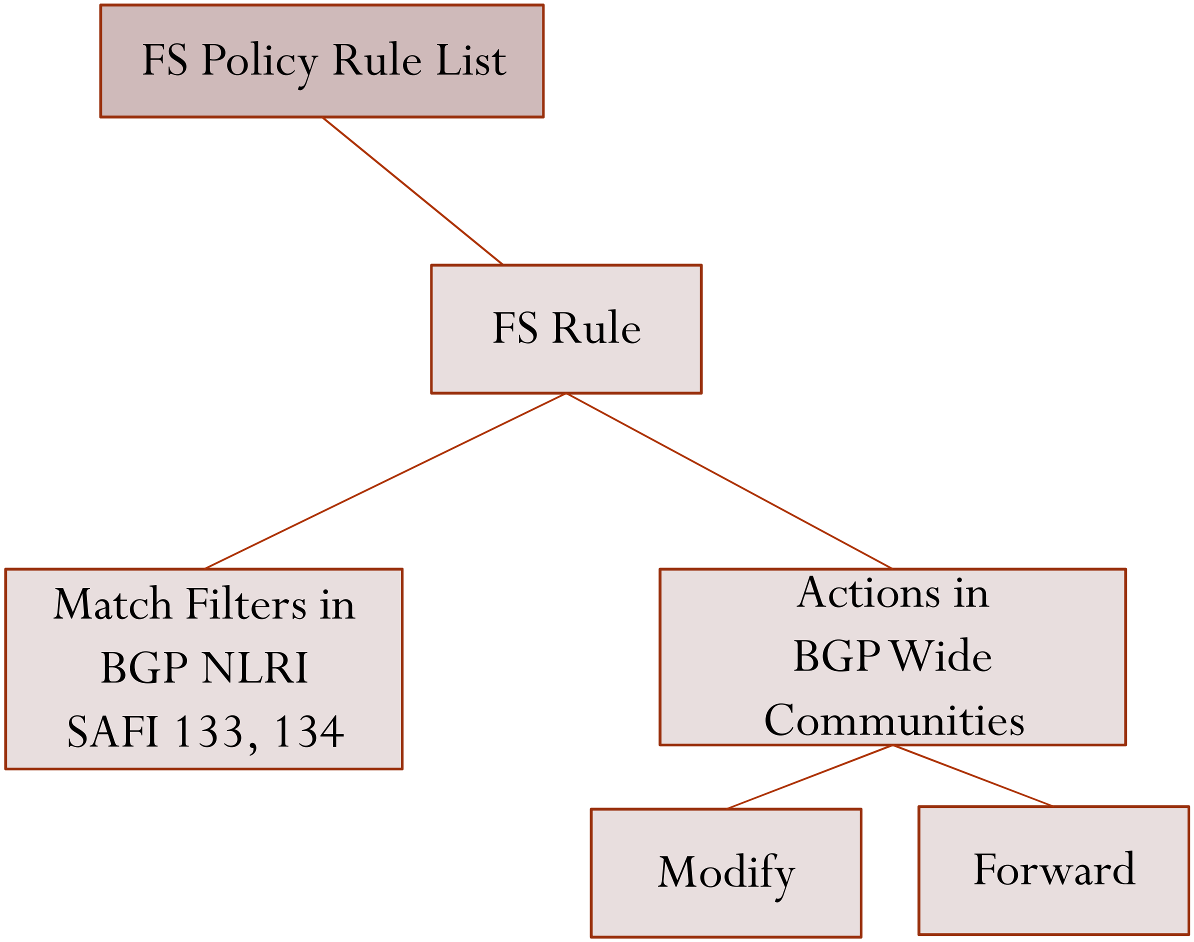
BGP Security Upgrade for BGP FS

- BGP Flow Specification – pre-dates ROA
- Validation using ROA
  - If have ROA: Use to validate transmitter of BGP FlowSpec along with Best-match unicast route for destination (IPv4 or IPv6)
  - If no ROA: Best Match unicast route + no more specific routes

Flow Specification between Protocols

- Key point:
  - Operator-Applied policy = Policy knobs in Vendors to set order and precedence within order
  - Operator-Applied policy must always be allowed
  - Defaults: If no Operator Policy, then default ordering
- Ordering:
  - BGP Flow specification similar
  - I2RS FB-RIB [draft-hares-fb-rib-data-model]
  - Policy Based Routing (config) [draft-hares-rtgwg-fb-rib upcoming]
  - ACL [draft-netmod-acl]
  - Routing configuration [draft-netmod-routing-cfg]
- Precedence for same n-tuple filter based on order
  -

Flow Specification Policy



FS Policy Rule List

FS Rule

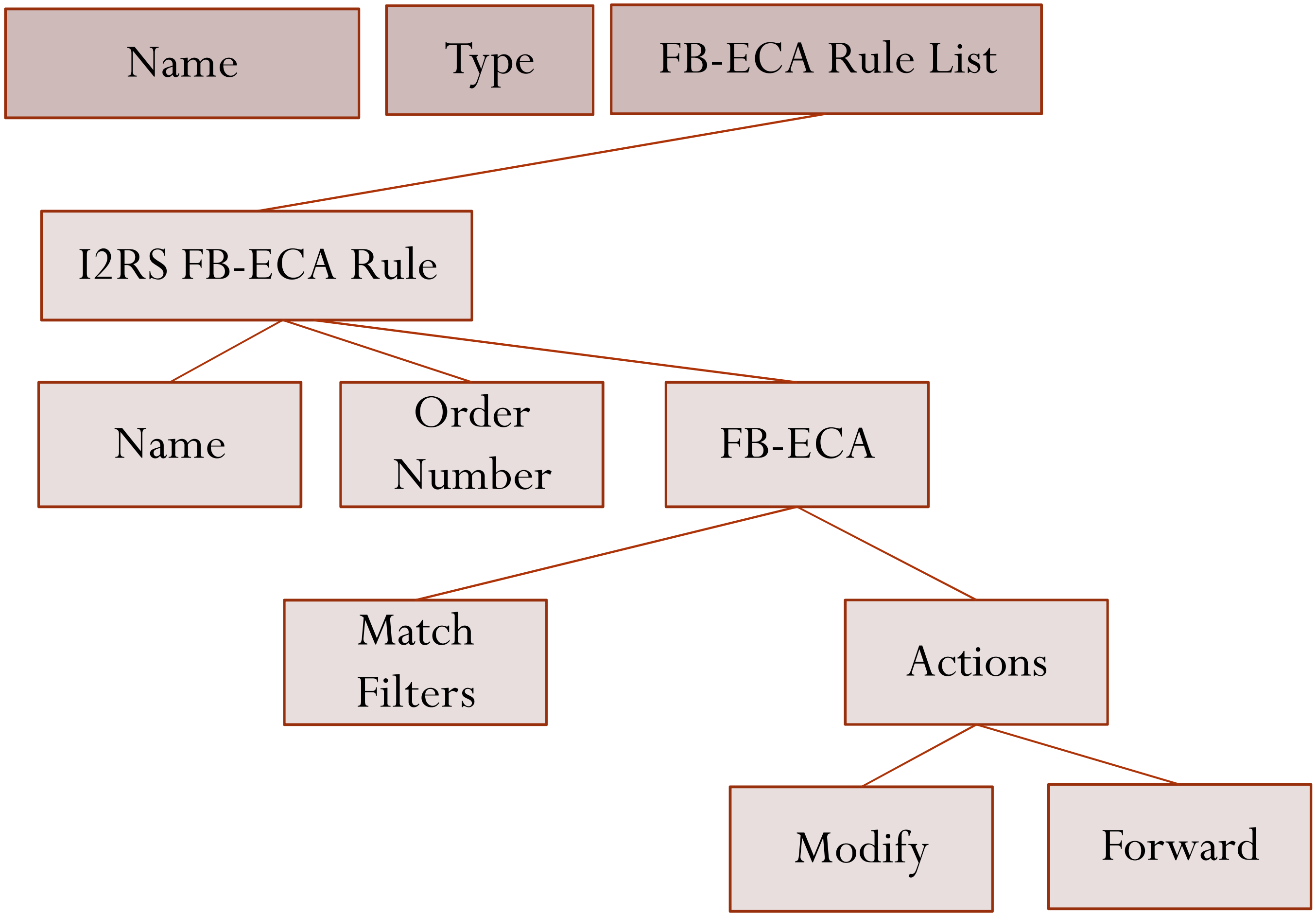
Match Filters in BGP NLRI SAFI 133, 134

Actions in BGP Wide Communities

Modify

Forward

I2RS FB-Rule List



Each Proposal must resolve conflicts

```
action               precedence 1         precedence 2
+----------+       +-----------+
| action 1 |-------|conflict 1 |----|
|          |       +-----------+    |   +----------+
|          |                        |---|conflict 3|
|          |       +-----------+    |   +----------+
|          |-------|conflict 2 |----|
+----------+       +-----------+
```

```
 precedence of conflicts for action 1 {}
  precedence(1) = conflict 1 | conflict 2;
  precedence(2) = conflict 3;
  If precedence (1) found; continue
  if precedence (3) found; exit;
 }
```

BGP-FS Option 2

- NLRI with order
- Actions in BGP Wide Community

New filter match with order

```
+------------------------+
|length (2 octets)       |
+------------------------+
| sub-TLVs (variable)    |
| +====================+ |
| | order (2 octets)   | |
| +--------------------+ |
| | type (2 octets)    | |
| +--------------------+ |
| | length (2 octets)  | |
| +--------------------+ |
| | value (variable)   | |
| |[multiples of       | |
| | 2 octets]          | |
| +====================+ |
+------------------------+
```

Figure 16 - NRLI revision

New Action atom for BGP Wide Communities

```
+--------------------------+
| order (2 octets)         |
+--------------------------+
| Action type (2 octets)   |
+--------------------------+
| Action length (2 octets) |
+--------------------------+
| Action Values (variable) |
| (multiples of 2 octets)  |
+--------------------------+
```

Wide Community Atom
figure 17

BGP-FS Atom added to Wide Community attribute

Wide Communities container (type 1) or

BGP Flow Specification container (type 2) (see below)

BGP-FS Container type 2

```
+-----------------------------+
| Source AS Number  (4 octets)|
+-----------------------------+
| list of atoms (variable)    |
+-----------------------------+
figure 18
```

Summary of January 11th discussion

- Why expand Flow Specification
  - Uses: DoS prevention, SDN/NFV, I2NSF
  - Need ordering for flow Specification
  - True Inter-Domain not as common within Provider with multiple AS-es
- If new mechanism, what about old?
  - Eventually Deprecate old, but allow side-by-side
  - Open Capability separate for New/Old

Summary of February 8 interim

- Clarities questions on draft-hares-idr-flowspec-combo-00

Discussion-1

Should we align all the Yang Modules for Filter-Based RIBs (config (aka policy routing), BGP, I2RS) ?

Yang Modules

Draft-wu-bgp

BGP FS Yang module contains

wu-idr-flowspec-yang-cfg

- Local Configuration of BGP-FS
- Operational state
  - BGP-FS Rules (filters + actions received)
    - Peer received from
    - Selected for installation or not
  - BGP-FS Rules match Statistics

Why Harmonize BGP-FS policies

- Common policy syntax to allow
  - Easy comparison between protocols
  - Easy comparison between received BGP-FS and locally configured BGP-FS

BGP-FS Local Config vs. I2RS FB-RIB

```
 Table 11 - comparison Yang Model Local Configuration
+-------------+----------------------+-----------------------+
| component    | BGP Flow Spec       | I2RS FB-RIB  +        |
|              | Yang                | Packet-ECA Yang       |
+==============+=====================+=======================+
|Policy        |flowspec-policy*     |group* [group-name]    |
| +-name       | [policy-name]       |                       |
| +-vrf        |+-rw vrf-name        | +-rw vrf-name         |
| +-AFI        |+-rw address-family  | +-rw address-famil    |
| +-rules      |+-rw flowspec-rule*  | +-rw group-rule-list  |
|              ||  [rule-name]       | | [rule-name]         |
|  +-rule-name ||+-rw rule-name      | |+-rw rule-name       |
|  +-rule-order||+-rw traffic-filters| |+-rw rule-order      |
|              ||+-rw traffic-actions| +-rw eca-rules        |
|              |                     | | [order-id rule-name]|
|              |                     | | +-rw installer      |
|              |                     | | +-rw eca-matches    |
|              |                     | | +-rw eca-qos-actions|
|              |                     | | +-rw eca-fwd-actions|
+--------------+---------------------+-----------------------+

figure 21 - Comparison of Yang modules (Config state)
```

Bgp Flow Spec vs I2RS Filters

BGP-FS policy received from remote peer

```
+------------+----------------------+-------------------------+
| component  | BGP Flow Spec        | I2RS FB-RIB             |
|            | Yang                 | Packet-ECA Yang         |
+============+======================+=========================+
|opstate     |flowspec-state        |ietf-fb-ribs-oper-status |
| +-rib      |+-ro flowspec-rib     |+-ro fb-rib-oper-status* |
|            |  |                   |  +-ro fb-rib-name       |
|  +-groups  |  |                   |  +-ro group-status      |
|  +-rules   |  +-ro flowspec-entry*|  +-ro rules_opstate     |
|    [index] |    [index]           |  [rule-order, rule-name]|
```

Bgp-FS Statiatiscs vs. I2RS Statistics

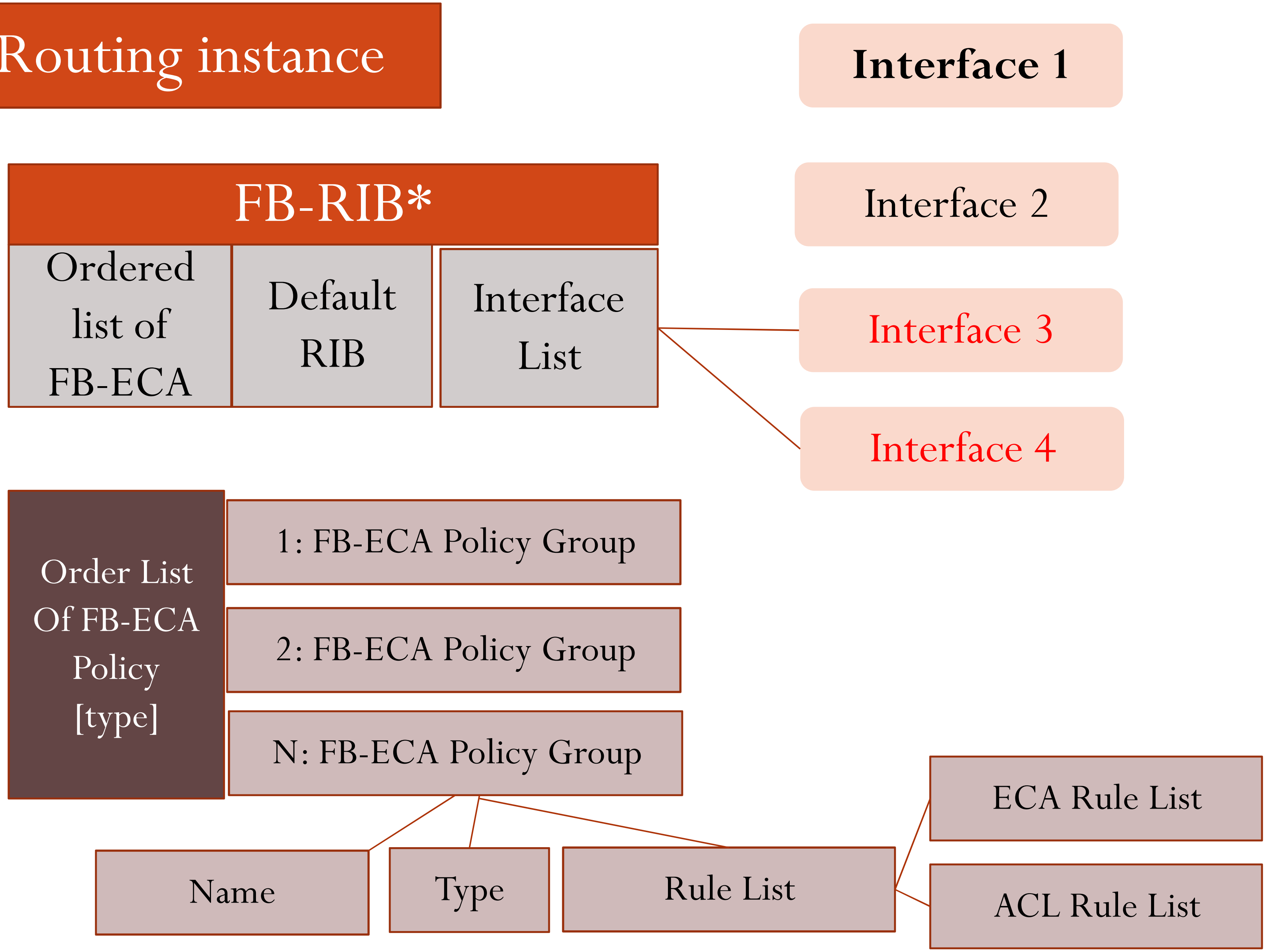
```
+------------+----------------------+-------------------------+
| component  | BGP Flow Spec        | I2RS FB-RIB             |
|            | Yang                 | Packet-ECA Yang         |
+============+======================+=========================+
| +-rules    |+-ro flowspec-stats*  |  +-ro rules_opstats     |
|            |                      |  [rule-order, rule-name]|
|            | +-ro vrf-name        |                         |
|            | +-ro address-family  |                         |
|            | +-ro flowspec-rule-  |                         |
|            | |    stats           |                         |
|            | |                    |                         |
|            | |+-ro traffic-filters|                         |
|            | |+-ro traffic-action |                         |
|            | |+-ro classified-pkts|   | +--ro pkts-match    |
|            | |                    |   | +--ro pkts-modified |
|            | |+-ro drop-pkts      |   | +--ro pkts-dropped  |
|            | |+-ro drop-bytes     |   | +--ro bytes-dropped |
|            |                      |   | +--ro pkts-forwarded|
|            |                      |   | +--ro bytes-forwarded|
+------------+----------------------+-------------------------+
```

Discussion-2

Should we align all the Yang Modules for Filter-Based RIBs (config (aka policy routing), BGP, I2RS) ?

Details on I2RS FB-Rib

Routing instance



Interface 1

Interface 2

Interface 3

Interface 4

FB-RIB*

Ordered list of FB-ECA

Default RIB

Interface List

Order List Of FB-ECA Policy [type]

1: FB-ECA Policy Group

2: FB-ECA Policy Group

N: FB-ECA Policy Group

Name

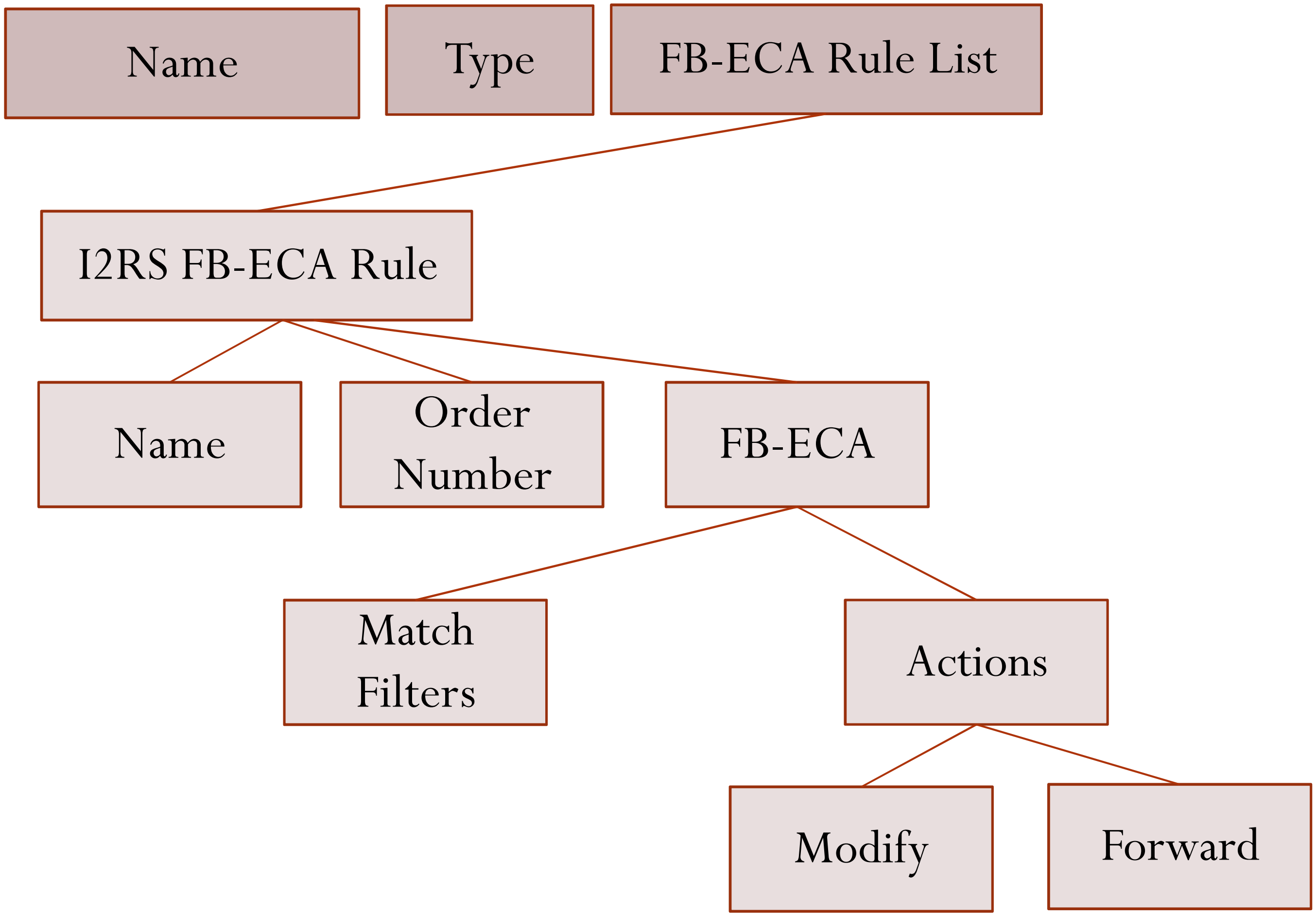
Type

Rule List

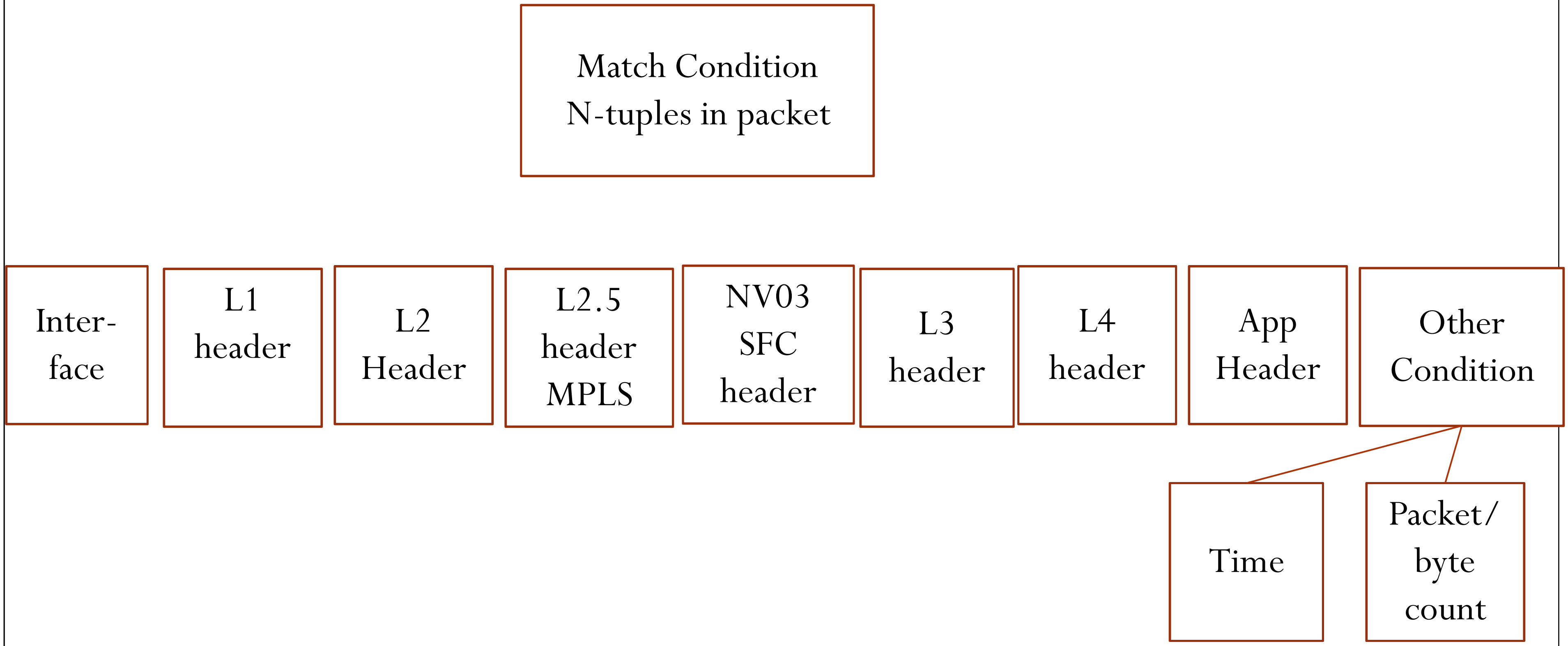
ECA Rule List

ACL Rule List

FB-Rule List


Name
Type
FB-ECA Rule List
I2RS FB-ECA Rule
Name
Order Number
FB-ECA
Match Filters
Actions
Modify
Forward

Filters in I2RS FB-RIB (hares-i2rs-pkt-eca-policy)



Backup slides

Discussion from 1/11/2016

- Should we have a successor to Flow-spec SAFI?
  - Action Criteria: IP Redirect (do-able) with 2 feature; Combination become with Actions is tricky;
    - Choice: combination
    - Precedence: better to specify, but will need to consider actions in combination
    - Redirect actions – interact with each; Modify actions interaction;
    - Traffic filters may
  - Match filters – as AND probably

Discussion Notes from 1/11/2016 (2)

- Flow Specification
  - Combination or separate Flow Spec
  - Rule ordering is reason for Flow-Spec 2,
    - Non-firewall, no SDN –may work
    - Firewall, SDN will not work without the ordering
  - Combination of the two flow-specification
    - If keep 2 SAFIs – two Flow-Specs into the future.
    - Ideal, v2 would have package with it – Date to deprecate V1 – real world doesn't probably won't allow it,
  - Agree with Jeff on backward compatibility
    - [wes] No way to tell which enhancement supported with out pre-knowledge,
    - [Jeff]: We do not have way to discover capabilities
    - [Robert]: We have this problem with the
    - [Jeff]: Redirect IP – possible that flow-specification – action (what does the implementation do with it).
  - Inter-domain flow-specification – not common
    - Service portals rather than inter-AS Flow specification
    - Redirect IP – within a single Provider – within a specific Provider

Discussion Notes from 1/11/2016 (3)

- Centralized mode –
  - Some flow specifications are only centralized controller and not distributed (Lucy Yong)
  - Some have two controllers (DDoS) and another (flow-filters)
    - Need to have precedence of the rules and then fall through (Jeff)
    - SDN (rule), and then the flow-specification rule
    - This requires a flow-specification v2 (jeff) because the existing things do not allow the flow-specification
    - Some the actions may only be appropriate to the list
  - Filter-based RIB
    - Precedence, fall-through – rule chains make sense
    - Take I2RS Filter-Based RIB

Discussion Notes from 1/11/2016 (4)

- Implementation of I2NSF
  - Controller tells the order of the rules
  - Can IDR provide this as well.
  - [Jeff]: More specific hosts, flow specification (longest prefix match will work)
  - [Linda]: Most specific

Discussion Notes from 1/11/2016 (5)

- John Schiel – flow spec rules that have precedence and ordering in flow specification rules.