BGP Flow Specification
Multi Vendor and Inter AS Interoperability

Loibl, Christoph (next layer communications)
christoph.loibl@nextlayer.at, c@tix.at

Bacher, Martin (T-Mobile Austria)
martin.bacher@t-mobile.at

January 2, 2017

Abstract

BGP flow specification (RFC 5575) defines a protocol to rapidly deploy access control lists and forwarding policies (flow-specification filters and actions) amongst all participating routers via a BGP address family.

This paper shows the current limitations of some of the major BGP flow specification implementations focusing on inter AS deployments in a multi vendor environment.

Based on bugs observed during initial configuration of an interoperability multi vendor lab it demonstrates, that some of the bugs can potentially lead to a remotely triggered complete fail of a provider network caused by terminating BGP sessions. It shows that current implementations show different behaviours and that all of the implementations are - at least in one manner - violating the current Internet standard. It proposes changes to the flow specification standard (RFC 5575) to improve interoperability and to guard against observed misunderstandings and different behaviours.

Because of missing features and the bugs in current BGP flow specification implementations it recommends to carefully consider a inter AS flow specification deployment.

Contents

1 Introduction

This paper is the result of a research that was carried out in order to evaluate the current technical restrictions of inter-AS and multi vendor deployments of flow-specification (RFC5575 [4]) and potentially produce known working configurations that address some of the identified security issues. It also addresses some of the stability issues seen in current products.

RFC5575 [4] (Dissemination of Flow Specification Rules) defines a protocol to rapidly deploy access control lists and forwarding policies (flow-specification filters and actions) amongst all participating routers via a BGP address family. Since BGP messages are not restricted to a single AS (iBGP), neighbouring ASs may also choose to exchange flow-specification filters (eBGP).

Most of today's flow-specification deployments are iBGP-only and are used for DDoS mitigation within a single AS [1]. Such deployments are relatively easy to manage since all flow-specification announcements are issued by a single entity (the carrier's network management/security center) and do not propagate over network borders. Even though there are use cases for exchanging flow-specification in an inter-AS manner, carriers seem to hesitate introducing such deployments.

There are several technical reasons why such deployment are rarely seen in the Internet[1]:

- **Scalability[2]:**

  Internet routers are designed to keep a big destination based forwarding table (FIB) in their hardware, but when it comes to access control lists and forwarding policies, the underlaying hardware is much more limited and may not scale very well when a large number of flow-specifications learnt from the entire Internet needs to be programmed.

- **Security:**

  While RFC5575 defines a simple, yet effective way to validate the flow-specification rules learnt by a remote router this bare minimum of verification is insufficient in an inter AS setting, where service providers want to introduce additional policies or restrictions to the flow-specification announcements learnt from their customers/peers. Such policy frameworks that have been implemented by the majority of the vendors for the major address families [3] are simply missing for flow-specifications.

[1] Not necessarily a complete list of technical issues

[2] Scalability of flow-specification is not addressed in this paper.

[3] ie. IPv4, IPv6, VPN-V4, VPN-V6, ...

For example most of the current implementations may allow remote peers to redirect Internet traffic into any arbitrary MPLS-VPN on the remote router by simply setting the traffic-redirect community that cannot be filtered by the implementations. It may also allow remote peers to map Internet traffic in any forwarding-class (DiffServ class) usually protected by filters on the border routers.

- **Stability:**

  The stability of eBGP sessions is crucial to a network's external availability. The most basic validation of flow-specification rules usually requires to run the IPv4 and flow-specification address-families to a particular external neighbour within a single BGP session (MP-BGP). Everything that goes wrong within flow-specification announcements may potentially lead to a BGP notification to be sent out followed by a BGP session to be termination[4]. This leads to an outage not only for flow-specification but also for IPv4 routing over a particular external link and routing flaps.

While there may be plenty of other reasons (including organisational) why inter-AS flow-specification deployments are rarely seen, inter-AS flow-specification may greatly improve the ability to successfully mitigate DDoS attacks. Filtering on a broader scale reduces traffic hotspots (and congestion) closer to the downstream networks. Flow-specification is a possible approach to allow a more distributed filtering throughout large sections of the Internet not limited to a single carrier-network itself.

[4]After a BGP notification message the session needs to be terminated and restarted, there is no graceful recovery from such situation.

2 Methods

A router lab consisting of four routers and additional management devices has been setup for the inter-AS flow-specification verification. For verification/monitoring of the BGP messages all direct connections between the routers were wired via a switch with a separate vlan for each point-2-point connection. Port-mirroring towards a network capturing device was configured on this layer-2 switch for each vlan to capture the entire BGP traffic. Figure 1 shows the resulting network but without the intermediate switch (for clarity reasons).

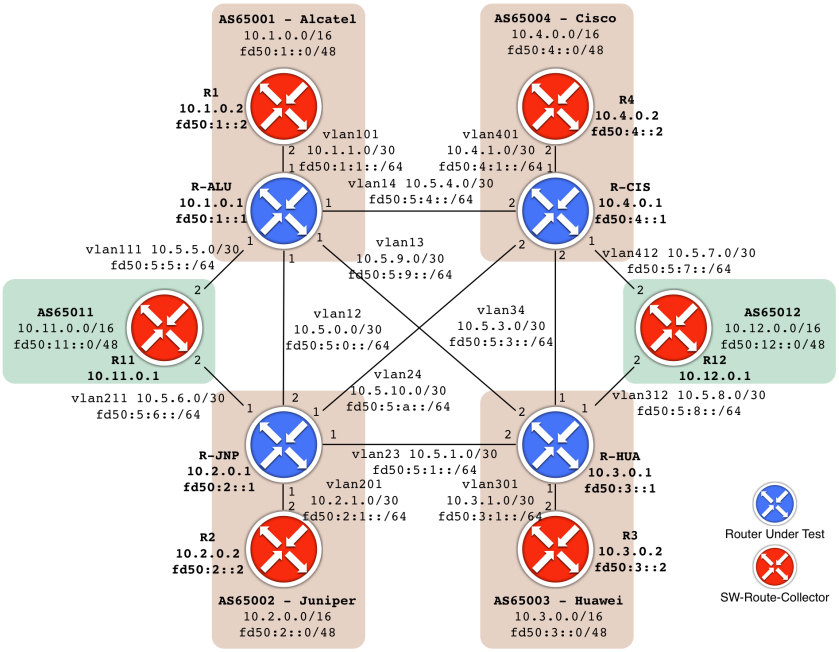

Figure 1: Lab network diagram

The entire verification of the behaviour in the lab was limited to control-plane behaviour no actual data-frames were sent during the tests, except from occasional ICMP-echos or UDP-traceroutes to verify the IPv4 routing within the lab. All the test verification was based on the output of *show* commands on the routers and the stability of the BGP sessions themselves.

Neither the configuration of the lab, nor the design of the test-cases were meant to do a complete functional analysis of the vendor's flow-specification implementation or RFC conformance tests of the devices under test and thus are not suitable for vendor or product selection. Additionally some of the test-cases were designed to max out limits in the implementations or are border cases not particular useful in production environments.

2.1 Router Hardware

The lab consisted of four routers manufactured by different vendors[5] that are frequently seen in the Austrian Internet carrier landscape.

Nokia/Alcatel and Cisco offered suitable hardware from their demo stock including support for configuration and debugging. Juniper suggested a firmware for the lab MX480 in next layer's lab. The Huawei router could be taken from T-Mobile's lab.

Lab-ID	**Manufacturer**	**Type**	**Firmware**
R-ALU	Nokia/Alcatel	7750 SR-c4	TiMOS-B-14.0.R3
R-CIS	Cisco	ASR 9001-S	IOS-XR 5.3.2
R-HUA	Huawei	NE40E-X8A	V800R007C00SPC100
R-JNP	Juniper	MX480	15.1F5.15

Table 1: Router Hardware/Firmware

All routers except for the Huawei router, switches and packet capture PC were installed in a single rack in next layer's datacenter. The Huawei router was remotely connected via a layer-2 transparent ethernet service from T-Mobile's PoP.

2.1.1 Flowspec related `show` commands

The CLIs of four different vendors required different commands for verifying the behaviour. Sometimes it was possible to verify the behaviour based on the received BGP messages on R1-4. Table 2 gives an overview of the required commands on the different platforms. Detailed documentation on the specific commands can be found in the vendor's documentation.

2.2 BGP Daemon (SW-Route-Collector)

In order to inject flow-specification routes into the test-network and verify the distribution of these routes multiple instances of the BGP-daemon ExaBGP were added to the network. The documentation and source-code can be found on Github:

`https://github.com/mshahbaz/exabgp`

As a base for the tests version 3.4.11 of ExaBGP has been used. ExaBGP is a very flexible BGP implementation written entirely in Python. It can send

[5]The vendors have been ask to suggest a hardware and software for the tests to be carried out.

Vendor	Command
Alcatel	show router bgp neighbor <ip> ipv4 advertised-routes show router bgp neighbor <ip> ipv4 received-routes show router bgp neighbor <ip> flow-ipv4 advertised-routes brief show router bgp neighbor <ip> ipv4 advertised-routes show router bgp neighbor <ip> ipv4 received-routes show router bgp neighbor <ip> flow-ipv4 advertised-routes show filter ip "fSpec-0" detail
Cisco	show flowspec vrf default ipv4 detail show flowspec vrf default ipv4 internal show bgp ipv4 flowspec show bgp trace flowspec show flowspec trace client event error show flowspec trace manager event error debug bgp update debug flowspec client debug flowspec error debug flowspec all
Huawei	disp bgp peer <ip> disp bgp flow peer <ip> disp bgp flow routing-table disp bgp routing-table disp bgp flow routing-table peer <ip> received-routes disp bgp flow routing-table peer <ip> advertised-routes disp bgp routing-table peer <ip> received-routes disp bgp routing-table peer <ip> advertised-routes disp flowspec statistics <idx-from-routing-table>
Juniper	show route table inetflow.0 all show route receive-protocol bgp <ip> table inetflow.0 all (extensive) show route table inet.0 all show route receive-protocol bgp <ip> table inet.0 all (extensive) show firewall filter __flowspec_default_inet__

Table 2: CLI commands overview

and receive routes but does not implement a RIB. It can easily be modified to send any arbitrary announcement and log all the received BGP messages (if decodable).

The tests required ExaBGP to announce IPv4 unicast and IPv4 flow-specification routes and receive/log the messages from the routers. The json interface was used to manual insert routes to simulate a dynamic behaviour (ie. announce and withdraw routes while the BGP daemon was running). Listing 1 shows

the ExaBGP configuration that was used in order to have ExaBGP listen on TCP port 1234 for a connection to issue json commands. The connection to port 1234 was established using `telnet` and the command were manually pasted into the terminal.

Listing 1: ExaBGP json configuration

```
process stio {
       run /usr/bin/nc -l 1234;
       encoder json;
       receive {
           parsed;
           update;
           neighbor-changes;
       }
}
```

2.3 Packet Capture / Analysis (tcpdump, Wireshark)

During the tests all packets were captured via `tcpdump`[6] and written to the disk of the Lab PC for later analysis via `Wireshark`[7]. Since the purpose of the tests did not require to generate any traffic except for network control, packet capture performance was not an issue. The size of the capture files where moderately (usually even below 1Mbyte).

To save the raw packets (in tcpdump's own pcap format) to the disk the following command was used:

```
tcpdump -ni em1 -s0 -w <filename>.pcap
```

The resulting files were then copied to the laptops for analysis via `Wireshark`.

2.4 Lab Services

Additional network services were needed for a consistent view over the network and for improved management:

- NTP Server (for consistent timestamps in logs)
- Syslog (remote logging service)
- DNS

[6] tcpdump is standard unix/linux command-line packet capture tool.
[7] Wireshark protocol analyser can be found at https://www.wireshark.org/

2.5 Testcases

This section gives only a brief overview of the intended test cases and why this particular test was selected. More in depth details on the test-configurations and how the tests were performed can be found in Section 3.

2.5.1 General Match Pattern

Announce a flow specification containing most special cases of flow type components and operators as specified by RFC5575 Section 4 and verify the behaviour and stability of the network. See Section 3.1 for details and results.

2.5.2 Action Extended Communities / Community rewriting

Announce a flow specification with the action communities defined in Section 7 of RFC5575 and verify if they are implemented as transitive or non-transitive extended community. RFC5575 defines the traffic-rate community explicitly as non-transitive, but as of RFC4360 [7] this is actually assigned as a transitive community by IANA[8]. Furthermore this test should verify if the implementation of the action community handling is consistent over all vendors.

Since one of the goals of this work is to evaluate the readiness of flow specification implementations in inter AS deployments community/action-policy rewriting is required at AS-borders to reflect the local AS policies. Such BGP-policies should be able to remove or replace actions (extended communities) received by a neighbouring BGP speaker and filter announcements based on the NLRI and received (extended) communities.

See Section 3.3 for details and results.

2.5.3 Flow Specification Validation

Flow specifications may get valid or invalid (in the manner of Section 6 of RFC5575) over time when IPv4 routing changes. These changes need to be reflected in the flow specification filters selected by the routers. This test verifies if flow specifications that should get invalid because the best match

[8] As all the other communities defined in RFC5575, where the RFC5575 is unclear and missing information about transitivity of all the other defined action communities

IPv4 route has changed over time are actually removed from the flow filters of the routers. See Section 3.5 for details and results.

2.5.4 Other

Some more test cases were planned but since implementations are still lacking many features these have been skipped until there are proper implementations available. See Section 3.6 for missing features that may be required in an inter provider flow specification setting.

3 Results

The lab described in Section 2 was constantly reconfigured during tests to reflect the actual test-cases. The base router configurations can be found in appendix 5 including the interface configurations and BGP peer configuration. The configuration of ExaBGP R11, R12 is shown alongside the test-cases since those routers have been heavily reconfigured during the tests.

3.1 General Match Pattern

This test was designed to verify if all the flow component types specified in the RFC5575 are supported by the routers and could be correctly dissected by the firmware. A rather complex flow NLRI was configured and announced by R11, R12. The announcements were such that the verification (Section 6 of RFC5575) should pass those NLRIs as valid. It was not the aim of the test to see a actual firewall rule being produced for that NLRI on each plattform because semantically the announced rule would never match a packet.

Section 4 of RFC5575 - type 3 defines a operator value encodings for comparison operations[9]. These operator-value pairs can be chained together to produce comparison operations like the following. This example encoding table is only given for type 3 flow component (IP-Protocol) to demonstrate why this complex operator was used (see table 3) for this test.

ip-protocol (Type-3 flow component) = 0,1,3,5,6,7,10-12,13-15,17-19,255

All flow components announced during this test were constructed accordingly. The complete NLRI consisted of all possible flow components in one NLRI. Table 4 shows the resulting flowspec NLRI that was announced from R11 and R12.

The configuration of ExaBGP for announcing the NLRI (table 4) from R11 is shown in Listing 2.

Listing 2: IPv4 and Flow-Route configuration of R11

```
static {
   route 10.11.0.0/16 self;
}
flow {
  route {
    match {
         destination 10.11.255.1/32;
         source 10.12.255.0/24;
         protocol =0 =1 =3 =5 =6 =7 >=10&<=12 >=13&<=15 >=17&<=19 =255;
         port =0 =21 =23 =25 =26 =27 >=30&<=32 >=33&<=35 >=37&<=39 =65535;
```

[9]This operator value encoding is used for all of the value comparison operations used for the IPv4-Field-Type matching.

Operator	Value	Description
00000001b	0	eq operator
00000001b	1	eq operator
00000001b	3	eq operator
00000001b	5	eq operator (may be aggregated by the implementation with 6 and 7)
00000001b	6	eq operator (may be aggregated by the implementation with 5 and 7)
00000001b	7	eq operator (may be aggregated by the implementation with 5 and 6)
00000011b	10	gt, eq operator (may get aggregated with 13-15)
01000101b	12	AND lt, eq operator (may get aggregated with 13-15)
00000011b	13	gt, eq operator (may get aggregated with 10-12)
01000101b	15	AND, lt, eq operator (may get aggregated with 10-12)
10000001b	255	End-of-list, eq operator (maximum value possible for this type)

Table 3: Type 3 encoding of the test-pattern

```
        destination-port =0 =41 =43 =45 =46 =47 >=50&<=52 >=53&<=55 >=57&<=59
            =65535;
        source-port =0 =61 =63 =65 =66 =67 >=70&<=72 >=73&<=75 >=77&<=79
            =65535;
        icmp-type =0 =1 =3 =5 =6 =7 >=10&<=12 >=13&<=15 >=17&<=19 =255;
        icmp-code =0 =10 =21 =23 =25 =26 =27 >=30&<=32 >=33&<=35 >=37&<=39
            =255;
        tcp-flags [fin syn rst push ack urgent];
        packet-length =0 =40 =46 =201 =203 =205 =206 =207 >=300&<=302
            >=303&<=305 >=307&<=309 =65535;
        dscp =0 =1 =3 =5 =6 =7 >=10&<=12 >=13&<=15 >=17&<=19 =48 =63;
        fragment [ not-a-fragment dont-fragment is-fragment first-fragment
            last-fragment ];
    }
    then {
      accept;
    }
  }
}
```

After starting the R11 ExaBGP instance it was planned to check all the router's BGP tables and verify the received flow specifications and routes. However, after R11 started to announce the configured route and flow specification it was immediately noticeable (from the log entries of the attached syslog server) that some BGP sessions in the Lab started to flap endlessly. As long as R11 was announcing its the test flow-specification no stable state of the network could be reached. The analysis showed that the following BGP sessions where unstable:

- R-JNP to R11
- R-JNP to R-ALU

Type	Name	Matching operator, value
1	Destination prefix	10.[local-as-specific].255.0/24
2	Source prefix	10.12.255.0/24
3	IP protocol	0,1,3,5,6,7,10-12,13-15,17-19,255
4	Port	0,21,23,25,26,27,30-32,33-35,37-39,65535
5	Destination port	0,41,43,45,46,47,50-52,53-55,57-59,65535
6	Source port	0,61,63,65,66,67,70-72,73-75,77-79,65535
7	ICMP type	0,1,3,5,6,7,10-12,13-15,17-19,42, 255
8	ICMP code	0,10,21,23,25,26,27,30-32,33-35,37-39, 255
9	TCP flags	ack, fin, push, rst, syn, urgent
10	Packet length	0, 40, 46, 201,203,205,206,207,300-302,303-305,307-309, 65535
11	DSCP	0,1,3,5,6,7,10-12,13-15,17-19, 48,63
12	Fragment	dont-fragment, is-fragment, first-fragment, last-fragment

Table 4: NLRI Testpattern

- R-JNP to R-CIS
- R-CIS to R4
- R-CIS to R-HUA
- R-CIS to R12

We decided to further investigate the misbehaviour and tried to reduce the complexity of the network for a root cause analysis.

3.1.1 R-JNP BGP Flap Cause Analysis

From the log messages on R-JNP it was clear that R-JNP was not able to decode the received flow-specification from R11. We deactivated all other BGP sessions between R-JNP and other routers to eliminate possible other influences and were able to reproduce the constant BGP flaps even in a resulting network where only R11 and R-JNP are involved. The log entries observed on R-JNP are shown in listing 3.

Listing 3: Juniper syslog messages, route flaps

```
Jun 28 10:41:58 <daemon.warn> r-jnp mx480-01-re1 rpd[14661]:
    RPD_BGP_NEIGHBOR_STATE_CHANGED: BGP peer 10.5.6.2 (External AS 65011) changed
    state from Established to Idle (event RecvUpdate) (instance master)
Jun 28 10:41:58 <daemon.warn> r-jnp mx480-01-re1 rpd[14661]: bgp_rcv_nlri:9989:
    NOTIFICATION sent to 10.5.6.2 (External AS 65011): code 3 (Update Message
    Error) subcode 10 (bad address/prefix field), Reason: peer 10.5.6.2 (External
    AS 65011) update included invalid route zero-len/0 (0 of 47)
Jun 28 10:41:58 <daemon.err> r-jnp mx480-01-re1 rpd[14661]:
    bgp_inetflow_get_prefix: can't resolve inetflow prefix range
Jun 28 10:41:58 <daemon.err> r-jnp mx480-01-re1 rpd[14661]: Received malformed
    update from 10.5.6.2 (External AS 65011)
```

```
Jun 28 10:41:58 <daemon.err> r-jnp mx480-01-re1 rpd[14661]: Family inet-flow,
     prefix zero-len/368
```

The cause could either be a bug in ExaBGP sending out a malformed flow-specification NLRI or a problem with Juniper's firmware. So we used Wireshark to analyse the packet captures recorded on the wire R11 to R-JNP and noticed that not even Wireshark was able to dissect the BGP UPDATE sent by ExaBGP[10]. This is the reason why we initially thought that ExaBGP is sending a malformed NLRI. However, why did R-ALU not complain about that particular UPDATE[11]? We needed to manually dissect the UPDATE in order to find out that the UPDATE sent by ExaBGP indeed was correctly formatted. This lead to the following conclusions:

1. Juniper's firmware has problems decoding certain NLRIs.
2. Wireshark has a bug in the BGP, flow-specification dissector.

This section will continue with analysis of the NLRI decoding issue observed on R-JNP. See the paragraph below on Wireshark's BGP dissector for further analysis of the Wireshark issue.

Since the Juniper implementation was known to usually correctly decode flow-specification NLRIs we modified the ExaBGP configuration and removed one flow-component-type after the other and checked if Juniper still sends out BGP NOTIFICATIONS when receiving the resulting NLRI. We started with removing the type-12 flow component. Only after removing all components (type-12 up to type-5) the NOTIFICATION messages disappeared. Further tests showed that the following flow-component combinations in a single NLRI cause Juniper's implementation to send a BGP NOTIFICATION:

```
Port (type-4) + Destination-port (type-5) (+ any other type)
Port (type-4) + Source-port (type-6) (+ any other type)
```

While from a semantical point of view such an NLRI may not make any sense, the NLRI is supposed to be treated as opaque to BGP and thus should not trigger a BGP NOTIFICATION as long as the NLRI is correct from the syntactical point of view (correct encoding of the flow components into the NLRI byte-string).

The behaviour of R-JNP is unexpected. A flow specification NLRI may traverse multiple routers until it is received by the first Juniper implementation that suffers from this bug and then trigger a BGP NOTIFICATION.

[10] Wireshark marked the BGP UPDATE in question as malformed.

[11] The session between R11 and R-ALU was stable during the initial test.

In that case such a BGP UPDATE may lead to BGP sessions flapping not only on adjacent routers but also on remote routers not directly connected with the router that originates that NLRI as long as it is part of that same flow-specification domain (independent of the AS).

Comment on `show route table inetflow.0` command output:

During debugging the command `show route table inetflow.0 (detail/extensive)` was regularly used and it seems that output clipping occurs when it needs to display a large flow-specification NLRI. Listing 4 demonstrates the output of a large flow filter which is clipped after a view lines (it is missing the "3". The end should output as "=03". We did not find a way to output very long flow filters using the CLI.

Listing 4: Juniper CLI clipping very long flow-specifications

```
inetflow.0: 8 destinations, 15 routes (7 active, 0 holddown, 8 hidden)
10.11.255/24,10.12.255/24,proto=0,=1,=3,=5,=6,=7,>=10&<=12,>=13&<=15,>=17&
<=19,=255,dstport=0,=41,=43,=45,=46,=47,>=50&<=52,>=53&<=55,>=57&<=59,=655
35,srcport=0,=61,=63,=65,=66,=67,>=70&<=72,>=73&<=75,>=77&<=79,=65535,icmp
-type=0,=1,=3,=5,=6,=7,>=10&<=12,>=13&<=15,>=17&<=19,=255,icmp-code=0,=10,
=21,=23,=25,=26,=27,>=30&<=32,>=33&<=35,>=37&<=39,=255,len=0,=40,=46,=201,
=203,=205,=206,=207,>=300&<=302,>=303&<=305,>=307&<=309,=65535,dscp=0,=1,=
3,=5,=6,=7,>=10&<=12,>=13&<=15,>=17&<=19,=48,=63,frag=00,=01,=02,=0/term:6
 (1 entry, 1 announced)
```

3.1.2 R-CIS BGP Flap Cause Analysis

After having found the root cause of R-JNP BGP session flapping the configuration of R11 was changed to announce a NLRI without type-4 and type-5 components. See table 5 for the resulting flow-specification NLRI.

Listing 5: Modified IPv4 and Flow-Route configuration of R11

```
static {
   route 10.11.0.0/16 self;
}
flow {
  route {
    match {
         destination 10.11.255.1/32;
         source 10.12.255.0/24;
         protocol =0 =1 =3 =5 =6 =7 >=10&<=12 >=13&<=15 >=17&<=19 =255;
         source-port =0 =61 =63 =65 =66 =67 >=70&<=72 >=73&<=75 >=77&<=79
              =65535;
         icmp-type =0 =1 =3 =5 =6 =7 >=10&<=12 >=13&<=15 >=17&<=19 =255;
         icmp-code =0 =10 =21 =23 =25 =26 =27 >=30&<=32 >=33&<=35 >=37&<=39
              =255;
         tcp-flags [fin syn rst push ack urgent];
         packet-length =0 =40 =46 =201 =203 =205 =206 =207 >=300&<=302
              >=303&<=305 >=307&<=309 =65535;
         dscp =0 =1 =3 =5 =6 =7 >=10&<=12 >=13&<=15 >=17&<=19 =48 =63;
```

```
            fragment [ not-a-fragment dont-fragment is-fragment first-fragment
                last-fragment ];
        }
        then {
          accept;
        }
      }
    }
```

After starting up ExaBGP the R-JNP sessions seemed stable. However we still observed most of the sessions towards R-CIS randomly flapping. It looked like there are other reasons for these BGP flaps.

During debugging all BGP sessions on R-CIS but the one to R-JNP were disabled. The remaining session between R-CIS and R-JNP was stable and the received flow-specification on R-CIS seemed correct (verified with `show bgp ipv4 flowspec`). For further debugging we enabled the BGP session to R4 and observed constant BGP flapping between R-CIS and R4 again. In this case R-CIS was receiving the NOTIFICATION from ExaBGP and ExaBGP was claiming to be unable to parse the NLRI received from R-CIS. Again we defined two potential reasons for the NOTIFICATION:

1. ExaBGP is unable to parse a correct NLRI.
2. Cisco's firmware is sending out a incorrect flow-specification.

We gave Wireshark a try but since we did not have a fix for the bug in Wireshark (see below) indeed it was unable to parse the offending UPDATE.

Again we decided to remove one component type after the other. After removing the first component (type-12) the session between R-CIS and R4 was stable again. However, we could not find any type combination triggering the problem.

We decided to manually compare the byte-string of the NLRI we manually dissected earlier to the NLRI sent by R-JNP and received by R-CIS and that sent from R-CIS towards R4. The NLRI received from R11 and sent by R-JNP to R-CIS were equal. However the NLRI sent by R-CIS to R4 did not match the original NLRI.

Manual dissection of this NLRI showed that R-CIS seems to wrongly encode the length field within the flow-specification NLRI larger then 239 byte. NLRI encoded by R-CIS:

```
0xf2 0x01 0x18 0x0a 0x0b 0xff ...
--1- --2- --3- --4- --5- --6- --7-
```

1. length = 242

2. type = 1 (destination IP)
3. CIDR-Length = 24
4. IP Byte 1
5. IP Byte 2
6. IP Byte 3
7. ...

However, Section 4 of RFC5575 defines a 2-byte length encoding for flow-specification NLRIs larger then 239 byte:

```
If the NLRI length value is smaller than 240 (0xf0 hex), the length
field can be encoded as a single octet.  Otherwise, it is encoded as
an extended-length 2-octet value in which the most significant nibble
of the first byte is all ones.
```

The correct encoding of the NLRI in question should be:

```
0xf0 0xf2 0x01 0x18 0x0a 0x0b ...
--1- --2- --3- --4- --5- --6- --7-
```

1. extended length byte 1
2. extended length byte 2 = 242
3. type = 1 (destination IP)
4. CIDR-Length = 24
5. IP Byte 1
6. IP Byte 2
7. ...

Cisco accepted the bug report that we filed as CSCva38418 ("BGP flowspec incorrectly encodes length of NLRI") and supplied a software revision that fixes this problem. While writing this article this fix is already generally available.

3.1.3 Wireshark BGP Dissector Crash Analysis

Since Wireshark was the only tool available to us to assist dissecting the BGP UPDATE messages we used it very frequently and we noticed very early in the process that it was not always able to correctly dissect all BGP messages that we recorded.

Debugging the previous issues showed, that we could not trust the router's firmwares either, thus needed to manually dissect many BGP updates until we noticed that Wireshark was suffering a very similar problem as Cisco's implementation: Whenever a flow-specification UPDATE message was larger then 239 byte it could not be dissected correctly. See figure 2.

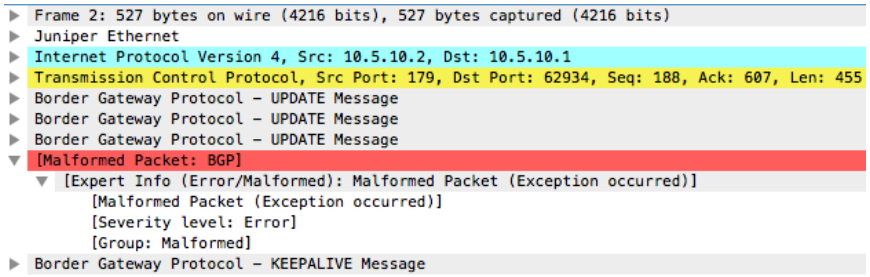

Figure 2: Wireshark dissector error

Wireshark is an open source implementation so we did not need to rely on blackbox-testing principles, but were able to actually dig into the BGP dissector source code and verify the behaviour according to the source code. The code had a special case for extended length flow-specification NLRI which was good, but it manipulated the length field in a wrong way and thus was not able to determine the correct length of the NLRI.

The bug was filed as Wireshark-Bug 12568 and fixed within a few days. While writing this article the fix is already in the current stable Wireshark versions (2.2.1).

3.2 Simplified Match Pattern

Since we observed the problems with R-CIS and R-JNP we simplified the match pattern in order to get a stable network. Required changes in the match pattern were the following:

- Do not use type-3 and (type-4 or type-5) flow components in a single NLRI (R-JNP).
- Keep the total size of the NLRI under 240 byte (R-CIS).

We decided to entirely remove the type-3 component (but keep type-4 and type-5) and remove the last two match conditions on type-12 (fragment-bits) from the original NLRI. This gives a total length of the flow-specification NLRI of 238 byte.

After restarting ExaBGP R11 again, the BGP sessions were stable. We waited some time for the network to converge and verified the announcements of the flow-specifications on all the routers using the appropriate commands (see section 2.1.1).

We quickly noticed that neither R-ALU nor R-HUA were propagating our flow-specification to their iBGP neighbours (R1, R3) nor to their eBGP neighbours. Additional investigation for the causes was necessary.

3.2.1 R-ALU not Propagating Flow Announcement Analysis

Since the BGP sessions were all stable `show` commands on R-ALU were used for further investigation. Immediately we could see that the flow-specification that we announced was correctly parsed and could be seen in R-ALU's RIB. See listing 6.

Listing 6: Flow-specification on R-ALU not validated correctly

```
*A:R-ALU# show router bgp neighbor 10.5.5.2 flow-ipv4 received-routes
===============================================================================
 BGP Router ID:10.1.0.1      AS:65001      Local AS:65001
===============================================================================
 Legend -
 Status codes : u - used, s - suppressed, h - history, d - decayed, * - valid
                l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes : i - IGP, e - EGP, ? - incomplete


===============================================================================
BGP FLOW IPV4 Routes
===============================================================================
Flag  Network           Nexthop               LocalPref      MED
      As-Path
-------------------------------------------------------------------------------
i     --                0.0.0.0               n/a            None
      65011

      NLRI Subcomponents:
      Dest Pref : 10.11.255.255/32
      Src Pref : 10.12.255.0/24
      ....
      ...
```

The console output shows that the route is not active and thus not propagated to any neighbour. A reason for a flow route not being active is, that it has failed the route validation (Section 6 of RFC 5575). We reconfigured the router and deactivated flow-specification validation for that particular eBGP neighbour and **after** re-announcing the flow route it became active

and was propagated to its neighbours according to BGP router propagation principles.

The packet captures of the announcements from R11 to R-ALU showed that after starting the ExaBGP process the flow-specification UPDATE was sent before the associated IPv4 announcement that is required for successful flow-specification validation (see table 5).

t	**Update-Type NLRI**
1	MP-REACH-NLRI (afi=1(IPv4), safi=133(Flow-specification)) dst-prefix=10.11.255.255/32 ...
2	NLRI 10.11.0.0/16

Table 5: BGP UPDATES order, ExaBGP R11 to R-ALU after startup

Since the IPv4-NLRI (10.11.0.0/16) was being announced **after** the flow-specification NLRI. Indeed the initial validation of the flow-specification is supposed to fail, because the the associated IPv4 prefix has not been announced yet. However after the announcement of the IPv4-NLRI the flow-specification filters should become active because all requirements for successful validation are now satisfied. We reconfigured the R11 to R-ALU session again to reenable flow validation for that session and manually fed the announcements into ExaBGP to have a defined order of IPv4-NLRI first and flow-specification NLRI second. The result was, that the flow-specification could successfully be validated. We could prove that flow-specification was not revalidated in case of routing table changes in IPv4. This lead to additional test-cases we performed for all routers (Section 3.5).

3.2.2 R-HUA not Propagating Flow Announcement Analysis

On R-HUA we were unable to see our flow-specification route with CLI commands. From the packet captures we saw the flow-specification being announced but it somehow disappeared. Again we decided to remove flow component after flow component to see if this problem was triggered by a flow component (type). After removing the first flow component (type-12 fragment) the flow specification announcement could seen on R-HUA and was propagated to its peers.

R-HUA is not propagating flow-specification filters when they contain certain types. This is not according to RFC5575 where the NLRI is defined as an opaque string for BGP.

We isolated R-HUA from the remaining infrastructure and reconfigured R12 to directly announce multiple combinations of flow-specification NLRI type-components to R-HUA while having R3 listen to all propagated flow-specifications. We tested for the following combinations:

- type-12(fragment) + all other types
- type-4(port) + type-5(dst-port) / type-6(src-port)
- type-3(protocol) + type-7 + type-8, type-9 (icmp-type, icmp-code, tcp-flags)

We picked those combination because the first was known to be dropped by R-HUA, the second lead to problems on R-JNP and the last does not match any packet. It turned out that the first two announcements are perfectly valid but are never propagated. The third will never match any packet but is always (independent on the actual values used for the type-3 protocol matching) propagated.

From our blackbox testing we concluded that Huawei's implementation does not propagate flow-specification filters when it is unable to perform packet matching for that particular combination of types. The examples that we found are (not necessarily a complete list):

- type-12(fragment) (+ any other type)
- type-4(port) + (type-5 and/or type-6)

Since R-HUA is not propagating some flow-specification filters (that completely comply with the standard and should be successfully verified), routers behind R-HUA, that may be able to act on such filters, may never receive the filters nor further propagate the filters towards other networks or routers.

3.3 Action Extended Communities Transitivity

To verify the transitivity of the action communities R11 and R12 were configured to announce a simple flow-specification towards their neighbours. To this simple match criteria all of the defined traffic action communities were attached. The ExaBGP configuration of R11, R12 is shown in listing 7.

Listing 7: ExaBGP configuration for traffic action transitivity tests

```
# R11
static {
   route 10.11.0.0/16 self;
}
flow {
  route {
     match {
        source 10.255.255.11/32;
        destination 10.11.255.255/32;
        protocol tcp;
     }
```

```
          then {
             discard;
             redirect 30740:12345;
             mark 12;
             action sample;
          }
       }
    }

    # R12
    static {
      route 10.12.0.0/16 next-hop self;
    }
    flow {
       route {
          match {
             source 10.255.255.11/32;
             destination 10.12.255.255/32;
             protocol tcp;
          }
          then {
             discard;
             redirect 30740:12345;
             mark 12;
             action sample;
          }
       }
    }
```

The test routers should receive the flow-specifications not only from the originating neighbours (R11, R12) but also from their neighbouring routers. In this case the flow-specifications announcement already traversed another AS. If the implementation of that particular neighbour filters certain communities it should be possible to see that some action communities are missing after the flow-specification is traversing a particular AS.

The listings 8 to 11 show the output of the BGP tables on all of the routers. The relevant parts are highlighted in boxes.

Listing 8: Action community transitivity verification on R-ALU

```
A:R-ALU>show>router>bgp#      routes flow-ipv4
===============================================================================
 BGP Router ID:10.1.0.1      AS:65001      Local AS:65001
===============================================================================
 Legend -
 Status codes : u - used, s - suppressed, h - history, d - decayed, * - valid
                l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes : i - IGP, e - EGP, ? - incomplete


===============================================================================
BGP FLOW IPV4 Routes
===============================================================================
Flag  Network           Nexthop               LocalPref      MED
      As-Path
-------------------------------------------------------------------------------
u*>i  --                0.0.0.0               n/a            None
      65011
```

```
      Community Action:  rate-limit:  0 kbps
      Community Action:  redirect-to-vrf:30740:12345
      Community Action:  mark-dscp:  12
      Community Action:  sample-log:  0:2
      NLRI Subcomponents:
      Dest Pref : 10.11.255.255/32
      Src Pref : 10.255.255.11/32
      Ip Proto : [ == 6 ]
*i    --                0.0.0.0              n/a           None
      65002 65011
      Community Action:  rate-limit:  0 kbps
      Community Action:  sample-log:  0:2
      Community Action:  redirect-to-vrf:30740:12345
      Community Action:  mark-dscp:  12
      NLRI Subcomponents:
      Dest Pref : 10.11.255.255/32
      Src Pref : 10.255.255.11/32
      Ip Proto : [ == 6 ]
u*>i  --                0.0.0.0              n/a           None
      65002 65003 65012
      Community Action:  rate-limit:  0 kbps
      Community Action:  sample-log:  0:2
      Community Action:  redirect-to-vrf:30740:12345
      Community Action:  mark-dscp:  12
      NLRI Subcomponents:
      Dest Pref : 10.12.255.255/32
      Src Pref : 10.255.255.11/32
      Ip Proto : [ == 6 ]
------------------------------------------------------------------------------
Routes : 3
==============================================================================
A:R-ALU>show>router>bgp#
```

Listing 9: Action community transitivity verification on R-CIS

```
RP/0/RSP0/CPU0:R-CIS#show bgp ipv4 flowspec
Wed Aug 3 14:36:14.422 CEST
BGP router identifier 10.4.0.1, local AS number 65004
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0x0  RD version: 2270
BGP main routing table version 2270
BGP NSR Initial initsync version 1 (Reached)
BGP NSR/ISSU Sync-Group versions 0/0
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
*> Dest:10.11.255.255/32,Source:10.255.255.11/32,Proto:=6/120
                    0.0.0.0                              0 65002 65011 i
*> Dest:10.12.255.255/32,Source:10.255.255.11/32,Proto:=6/120
                    0.0.0.0                              0 65012 i
*                   0.0.0.0                              0 65002 65003 65012 i


Processed 2 prefixes, 3 paths
```

```
RP/0/RSP0/CPU0:R-CIS#show bgp ipv4 flowspec
     Dest:10.11.255.255/32,Source:10.255.255.11/32,Proto:=6/120
Wed Aug 3 14:36:29.395 CEST
BGP routing table entry for
     Dest:10.11.255.255/32,Source:10.255.255.11/32,Proto:=6/120
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker             2267       2267
Last Modified: Aug 3 14:11:48.810 for 00:24:40
Paths: (1 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.1
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.1
  65002 65011
    0.0.0.0 from 10.5.10.1 (10.2.0.1)
      Origin IGP, localpref 100, valid, external, best, group-best
      Received Path ID 0, Local Path ID 1, version 2267
      Extended community:  FLOWSPEC Traffic-rate:0,0 FLOWSPEC Traffic-action:2,0
      FLOWSPEC Redirect-RT:30740:12345 FLOWSPEC Traffic-mark:0x0c

RP/0/RSP0/CPU0:R-CIS#show bgp ipv4 flowspec
     Dest:10.12.255.255/32,Source:10.255.255.11/32,Proto:=6/120
Wed Aug 3 14:36:34.380 CEST
BGP routing table entry for
     Dest:10.12.255.255/32,Source:10.255.255.11/32,Proto:=6/120
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker             2270       2270
Last Modified: Aug 3 14:22:48.810 for 00:13:45
Paths: (2 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.1
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.1
  65012, (received & used)
    0.0.0.0 from 10.5.7.2 (10.12.0.1)
      Origin IGP, localpref 100, valid, external, best, group-best
      Received Path ID 0, Local Path ID 1, version 2270
      Extended community:  FLOWSPEC Traffic-rate:0,0 FLOWSPEC Traffic-action:2,0
      FLOWSPEC Redirect-RT:30740:12345 FLOWSPEC Traffic-mark:0x0c
  Path #2: Received by speaker 0
  Not advertised to any peer
  65002 65003 65012
    0.0.0.0 from 10.5.10.1 (10.2.0.1)
      Origin IGP, localpref 100, valid, external, invalid flowspec-path
      Received Path ID 0, Local Path ID 0, version 0
      Extended community:  FLOWSPEC Traffic-rate:0,0 FLOWSPEC Traffic-action:2,0
      FLOWSPEC Redirect-RT:30740:12345 FLOWSPEC Traffic-mark:0x0c

RP/0/RSP0/CPU0:R-CIS#
```

Listing 10: Action community transitivity verification on R-HUA

```
<R-HUA>disp bgp flow routing-table

 BGP Local router ID is 10.3.0.1
```

```
 Status codes: * - valid, > - best, d - damped, x - best external, a - add path,
               h - history, i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
 RPKI validation codes: V - valid, I - invalid, N - not-found



 Total Number of Routes: 2
 * >  ReIndex : 4994
      Dissemination Rules:
       Destination IP : 10.12.255.255/32
       Source IP     : 10.255.255.11/32
       Protocol      : eq 6
       MED     :                   PrefVal : 0
       LocalPref:
       Path/Ogn : 65012i
 * >  ReIndex : 5057
      Dissemination Rules:
       Destination IP : 10.11.255.255/32
       Source IP     : 10.255.255.11/32
       Protocol      : eq 6
       MED     :                   PrefVal : 0
       LocalPref:
       Path/Ogn : 65002 65011i

<R-HUA>disp bgp flow routing-table 4994

 BGP local router ID : 10.3.0.1
 Local AS number : 65003
 Paths:  1 available, 1 best
 ReIndex : 4994
 Order  : 3221225471
 Dissemination Rules :
   Destination IP : 10.12.255.255/32
   Source IP     : 10.255.255.11/32
   Protocol      : eq 6

 BGP flow-ipv4 routing table entry information of 4994:
 Match action :
    apply deny
    apply sample
    apply remark-dscp 12
    apply redirect vpn-target 30740:12345
 From: 10.5.8.2 (10.12.0.1)
 Route Duration: 0d00h16m20s
 AS-path 65012, origin igp, pref-val 0, valid, external, best, pre 255
 Advertised to such 2 peers:
    10.5.1.1
    10.5.8.2

<R-HUA>disp bgp flow routing-table 5057

 BGP local router ID : 10.3.0.1
 Local AS number : 65003
 Paths:  1 available, 1 best
 ReIndex : 5057
 Order  : 1610612735
 Dissemination Rules :
   Destination IP : 10.11.255.255/32
   Source IP     : 10.255.255.11/32
```

```
  Protocol      : eq 6

 BGP flow-ipv4 routing table entry information of 5057:
 Match action :
   apply deny
   apply sample
   apply remark-dscp 12
   apply redirect vpn-target 30740:12345
 From: 10.5.1.1 (10.2.0.1)
 Route Duration: 0d00h26m36s
 AS-path 65002 65011, origin igp, pref-val 0, valid, external, best, pre 255
 Advertised to such 2 peers:
    10.5.1.1
    10.5.8.2

<R-HUA>
```

Listing 11: Action community transitivity verification on R-JNP

```
flow@mx480-01-re1> show route table inetflow.0 detail all


inetflow.0: 2 destinations, 4 routes (2 active, 0 holddown, 2 hidden)
10.11.255.255,10.255.255.11,proto=6/term:1 (2 entries, 1 announced)
       *BGP    Preference: 170/-101
               Next hop type: Fictitious, Next hop index: 0
               Address: 0x95af704
               Next-hop reference count: 4
               State: <Active Ext>
               Local AS: 65002 Peer AS: 65011
               Age: 3:29
               Validation State: unverified
               Task: BGP_65011.10.5.6.2
               Announcement bits (2): 0-Flow 1-BGP_RT_Background
               AS path: 65011 I
               Communities:  traffic-rate:0:0 traffic-action:  sample redirect:30740:12345
               traffic-marking:12
               Accepted
               Validation state: Accept, Originator: 10.5.6.2, Nbr AS: 65011
               Via: 10.11.0.0/16, Active
               Localpref: 100
               Router ID: 10.11.0.1
        BGP                /-101
               Next hop type: Fictitious, Next hop index: 0
               Address: 0x95af704
               Next-hop reference count: 4
               State: <Hidden Ext>
               Inactive reason: Unusable path
               Local AS: 65002 Peer AS: 65001
               Age: 3:08
               Validation State: unverified
               Task: BGP_65001.10.5.0.1
               AS path: 65001 65011 I
               Communities:  traffic-rate:0:0 traffic-action:  sample redirect:30740:12345
               traffic-marking:12
               Accepted
               Validation state: Reject, Originator: 10.5.0.1, Nbr AS: 65001
               Via: 10.11.0.0/16, Active
               Localpref: 100
               Router ID: 10.1.0.1
```

```
            Hidden reason: Flow-route fails validation


10.12.255.255,10.255.255.11,proto=6/term:2 (2 entries, 1 announced)
       *BGP   Preference: 170/-101
              Next hop type: Fictitious, Next hop index: 0
              Address: 0x95af704
              Next-hop reference count: 4
              State: <Active Ext>
              Local AS: 65002 Peer AS: 65003
              Age: 16:14
              Validation State: unverified
              Task: BGP_65003.10.5.1.2
              Announcement bits (2): 0-Flow 1-BGP_RT_Background
              AS path: 65003 65012 I
              Communities:  traffic-rate:0:0 traffic-action:  sample redirect:30740:12345
              traffic-marking:12
              Accepted
              Validation state: Accept, Originator: 10.5.1.2, Nbr AS: 65003
              Via: 10.12.0.0/16, Active
              Localpref: 100
              Router ID: 10.3.0.1
        BGP               /-101
              Next hop type: Fictitious, Next hop index: 0
              Address: 0x95af704
              Next-hop reference count: 4
              State: <Hidden Ext>
              Inactive reason: Unusable path
              Local AS: 65002 Peer AS: 65004
              Age: 2:31
              Validation State: unverified
              Task: BGP_65004.10.5.10.2
              AS path: 65004 65012 I
              Communities:  traffic-rate:0:0 traffic-action:  sample redirect:30740:12345
              traffic-marking:12
              Accepted
              Validation state: Reject, Originator: 10.5.10.2, Nbr AS: 65004
              Via: 10.12.0.0/16, Active
              Localpref: 100
              Router ID: 10.4.0.1
              Hidden reason: Flow-route fails validation

{master}
flow@mx480-01-re1>
```

All flow-specifications received by all routers seem to have all action communities attached. None of the four RFC 5575 defined communities are treated as non-transitive on any of the platforms. This violates Section 7 of RFC 5575, since the `traffic-rate` action is explicitly defined as non transitive there. For all the other traffic action communities the transitivity property is not defined in the RFC but all implementations decided to treat these remaining communities as transitive. See Section 4 for suggested revised behaviour.

3.4 Path Attribute Modification / Policies

In an inter AS setting BGP policies are heavily used to influence the propagation of routing information to reflect the contracts between organisations and for traffic engineering. Such policies usually involve setting, deleting BGP communities modifying LOCAL_PREF, MULTI_EXIT_DISC or entirely filtering the BGP UPDATE on ingress (upon receiving an UPDATE from a neighbour) or egress (while sending an UPDATE to a neighbour).

Filtering based on the NLRI (ie. flow-specification type values) was not possible on any of the tested implementations. Furthermore it was not even possible on most of the implementations to create a BGP policy that filters/modifies flow-specification UPDATES based on action communities[12] or any other community attached to a flow-specification NLRI.

Specially the fact that it is not possible to filter on traffic action communities rises many security concerns:

- The `redirect` action is most likely not useful in an inter AS setting. A third party may redirect traffic into any arbitrary MPLS VPN if policies cannot filter such updates.
- The `traffic-marking` action may allow a third party to map certain traffic[13] into any arbitrary forwarding-class and thus override a carriers QoS constrains and policies[14].
- The `traffic-action` action may allow a third party to divert certain traffic to router's control-plane and overwhelm the control-plane with packets.

During the testing we where unable to get any information if these problems regarding filtering are likely to be addressed in future firmware updates. These limitations are likely to exist for a longer time.

3.5 Flow Specification Validation

During the initial flow tests it could be shown that at least one implementation does only validate of the flow-specification NLRI once it is received (Section 3.2.1). Later routing table updates do not trigger a flow-specification NLRI to be revalidated. This leads to an unpredictable race condition in the

[12] Juniper's policy-statement implementation allows to match on communities/extended communities but does not allow to add a match for the address-family "flow-specification" and thus is not very useful for flow-spec matching.

[13] Matched by the flow-specification NLRI

[14] This may even allow some traffic to go into network-control forwarding-classes.

flow specification validation since the arrival time of IPv4 BGP UPDATES versus flow-sepecification UPDATES may lead to different validation results.

RFC 5575 does not explicitly require revalidation of flow-specification on IPv4 routing updates, however it compares the validation with Section 9.1.2 of RFC 4271 [6] behaviour in case of NEXT_HOP route being unresolvable. When a NEXT_HOP is resolvable again a BGP route immediately gets feasible again and thus is a candidate route for the BGP route selection process. Thus some kind of *revalidation* of the NEXT_HOP is taking place hence flow-specification NLRIs need to get revalidated when IPv4 updates occur.

In order to test the implementations for such a race condition all the BGP sessions between R-ALU, R-CIS, R-HUA and R-JNP have been deactivated and the only remaining sessions were those to R11, R12 and to R1-4. ExaBGP R11, R12 was reconfigured in order not to announce any route but to listen on its json interface for commands.

Listing 12: ExaBGP configuration for validation race testing R-ALU

```
group IPV4 {
    hold-time 180;
    graceful-restart 1200;
    family {
       ipv4 unicast;
       ipv4 flow;
    }

   process stio {
        run /usr/bin/nc -l 1234;
        encoder json;
        receive {
            parsed;
            update;
            neighbor-changes;
        }
   }

   # iBGP neighbor
   neighbor 10.1.0.1 {
      router-id 10.1.0.2;
      local-address 10.1.0.2;
      local-as 65001;
      peer-as 65001;
      static {
        route 10.1.0.0/16 next-hop 10.1.0.2;
      }
   }

   # eBGP neighbor
   neighbor 10.5.5.1 {
      router-id 10.11.0.1;
      local-address 10.5.5.2;
      local-as 65011;
      peer-as 65001;
   }
}
```

The following steps where performed (in chronological order) during the testing (the IP addresses below show the test case for R-ALU) to test for a possible race condition in the flow-specification validation:

1. **Startup ExaBGP:** no routes announced yet.
2. **Verify:** no routes should be received.
3. **Announce via ExaBGP:** `neighbor 10.5.5.1 announce route 10.11.0.0/16 next-hop self`
4. **Verify:** IPv4 prefix 10.11.0.0/16 should be accepted.
5. **Announnce via ExaBGP:** `neighbor 10.5.5.1 announce flow route {\n match {\n source 10.0.0.1/32;\n destination 10.11.255.255/32;\n}\n then {\n accept;\n}\n}\n\n}`
6. **Verify:** The flow-specification should be accepted.
7. **Withdraw via ExaBGP:** `neighbor 10.5.5.1 withdraw route 10.11.0.0/16 next-hop self`
8. **Verify:** The IPv4 prefix should disappear.
9. **Verify:** The flow-specification should become unfeasible (validation fail).
10. **Withdraw via ExaBGP:** `neighbor 10.5.5.1 withdraw flow route {\n match {\n source 10.0.0.1/32;\n destination 10.11.255.255/32;\n}\n then {\n accept;\n} \n}\n\n}`
11. **Verify:** The flow-specification should entirely disappear.
12. **Announnce via ExaBGP:** `neighbor 10.5.5.1 announce flow route {\n match {\n source 10.0.0.1/32;\n destination 10.11.255.255/32;\n}\n then {\n accept;\n}\n}\n\n}`
13. **Verify:** The flow-specification should be announced but unfeasible (validation fail).
14. **Announce via ExaBGP:** `neighbor 10.5.5.1 announce route 10.11.0.0/16 next-hop self`
15. **Verify:** IPv4 prefix 10.11.0.0/16 should be accepted.
16. **Verify:** The flow-specification should become feasible (correctly validated).

All routers except for R-ALU passed the above verification. On R-ALU in step 9 the flow-specification did not become unfeasible and in step 16 the flow-specification did not become active (after the announcement of a matching IPv4 prefix). Since the internal design of this feature in that firmware is unknown to us we waited 5 minutes after the announcements/withdraws to see if after some time a revalidation is performed, but this was not the case.

A bug with Nokia/Alcatel was filed and confirmed. The latest available software release[15] lists *"Installed validated FlowSpec routes do not disappear when next-hop disappears"* as a current known limitation.

3.6 Missing Features

During lab setup we noticed that many features, that are known as best practice in inter AS BGP setups in other address families, are either completely missing or not supported by some of the vendors. The following list shows features that that are vital for many operators but poorly available in current firmwares:

- **BGP import/export Policies:** Matching upon flow-specification components and traffic action communities and modifying action communities, filter updates or modify other BGP path attributes (LOCAL_PREF, MULTI_EXIT_DISC).
- **Flowspec in a VRF:** For security reasons more and more operators move their entire Internet routing into a VRF. Such designs require flow-specification BGP sessions to operate within a VRF.
- **IPv6 Flowspec:** There is no current RFC that defines flow-specification for IPv6 filtering. All implementations are based on a expired IETF IDR working-group document `I-D.ietf-idr-flow-spec-v6` [5].

There may be many more desired features and addons for flow-specification. Based on the current activities and Internet drafts regarding flow-specification being submitted to the IETF this is a very active field of current development[16].

3.7 ExaBGP IPv6 Flow NLRI Parsing Bug

While in the end IPv6 tests were entirely skipped due to feature availability on the platforms (Section 3.6), the initial configuration of the lab contained

[15] SROS 14.0.R6

[16] See the search option on the IETF data-tracker `https://datatracker.ietf.org`

IPv6 flow specification and routing where it was supported. Later this configuration has been removed or not maintained. However, during the initial setup a ExaBGP bug was discovered that lead to ExaBGP not being able to parse certain IPv6 flow components and issuing a BGP NOTIFICATION message when it received certain IPv6 flow components.

Since ExaBGP is a open source BGP implementation, it was possible to locate the problem within the source code of ExaBGP and show that all IPv6 flow component-types that exist in IPv4 flow-specification as well, could not be correctly parsed. Table 6 gives a list of flow-components that triggered a NOTIFICATION in the lab.

Type	**IPv4 Type Name**	**IPv6 Type Name**	**triggers bug**
1	IPv4 dest. prefix	IPv6 dest. prefix	no
2	IPv4 source prefix	IPv6 source prefix	no
3	IP protocol	Next header	no
4	Port	same as IPv4	yes
5	Destination port	same as IPv4	yes
6	Source port	same as IPv4	yes
7	ICMP type	same as IPv4	yes
8	ICMP code	same as IPv4	yes
9	TCP flags	same as IPv4	yes
10	Packet length	same as IPv4	yes
11	DSCP	Traffic class	no
12	Fragment	Fragment (different encoding)	yes
13	n/a	Flow label	no

Table 6: IPv6 flow types that trigger a BGP NOTIFICATION

After opening a case with ExaBGP (Bug #436) Thomas Mangin was able to supply a patch after a few hours. Releases 3.4.17 include a fix for that issue.

4 Conclusion

The goal of this work was to produce a *known to be working* set of configuration suitable for a robust inter AS flow-specification implementation. Very soon it was clear that we are unable to produce such a configuration because of multiple limitations in the implementations. We decided to demonstrate the issues/bugs and limitations that we identified.

The following bugs where found during the setup of the lab:

- Juniper's implementation terminates the BGP session (NOTIFICATION) on certain flow-specification type combinations (Section 3.1.1).
- Cisco's implementation sent corrupt BGP flow-specification UPDATE messages (peers to respond with terminating the BGP sessions with BGP NOTIFICATION) (Section 3.1.2).
- Wireshark was unable to parse flow-specification NLRIs larger than 139 bytes (Section 3.1.3).
- Alcatel's implementation did not revalidate flow-specifications in case of routing table changes (Sections 3.2.1 and 3.5).
- Huawei's implementation did not propagate certain flow-specification type combinations (Section 3.2.2).
- ExaBGP was not able to parse certain IPv6 flow types (Section 3.7).
- Transitivity of action communities violating RFC 5575 (Section 3.3).

Our test cases were not targeted to find bugs in a systematic way. The bugs were incidentally triggered during configuring the lab. If systematic test methods were applied we think that one could find more bugs in the implementations.

Not only bugs, but multiple missing features were identified. We think that the following, currently missing, additional features are required to allow a safe inter provider flow specification implementation. The missing features above may introduce serious security problems in case of inter provider flow specification:

1. Proper BGP filtering and rewriting of action-communities to secure eBGP sessions against unwanted actions and filters (Section 3.4).
2. Possibility to limit the resources for flow specification (like ie. max-prefix): Exhausting resources (like tcam resources or memory) on the router platforms may lead to performance degradation or network outages.

3. Proper feature testing / yet undiscovered bugs in the implementations (see above).

We could also show that the different implementations sometimes treat NLRIs in a different way and that in a multi vendor environment that could lead to unpredictable filter propagation and constant BGP session flaps. While we already listed the inconsistent behaviours within the list of bugs above, we think that what we listed as bugs sometimes is a result of unclear sections and definitions in RFC 5575. As a result of this work we published a IETF draft to improve flow specification interoperability and consistent behaviour over different implementations:

`I-D.draft-loibl-bacher-idr-flowspec-clarification` [3]

This published draft updates and clarifies the following definitions of RFC 5575:

- Clarification of the comparison operator
- Clarification of the component type length
- (Re-)Validation of the flow specification NLRI
- Transitivity of traffic filtering actions
- Clarification of flowspec NLRI parsing and validation

While this draft has already been published we noticed that independently to our work, there were ongoing efforts regarding an updated RFC 5575 in order to clarify possible traffic action community interference mainly by S. Hares and R. Raszuk. Together we decided to come up with an RFC 5575bis draft that contains our proposed changes, the clarification of traffic action interference and a cleanup of the entire RFC 5575 wherever we thought that the current text needed additional clarifications:

`I-D.hr-idr-rfc5575bis` [2]

With this update we think that a proper interoperable implementation should be possible and unambiguous sections have been improved.

Given the current bugs, interoperability issues and missing features we do not recommend flow specification BGP sessions between different carriers without an additional network element that could act as a flow specification NLRI and action-community screening device. Such a device may be based on ExaBGP plus some additional software to perform the actual filter screening and potential rewrite (based on operator's decisions). Invalid flow specification NLRIs or action filters have the potential to remotely trigger a complete network failure.

5 Acknowledgements

The authors would like to thank Nokia/Alcatel Austria and Cisco for supplying their lab equipment and support. Alexander Mayrhofer, Nicolas Fevrier and Robert Raszuk for their comments and support. Susan Hares for her comments and support regarding IETF procedures.

References

[1] Martin Bacher. Addressing ddos attacks with bgp flowspec. Master's thesis, Fachhochschule Technikum Wien, Hochstaedtplatz 5, 1200 Wien, 2016.

[2] Susan Hares, Robert Raszuk, Danny McPherson, Christoph Loibl, and Martin Bacher. Dissemination of flow specification rules. Internet-Draft draft-hr-idr-rfc5575bis-02, IETF Secretariat, November 2016. `http://www.ietf.org/internet-drafts/draft-hr-idr-rfc5575bis-02.txt`.

[3] Christoph Loibl and Martin Bacher. Flowspec clarification. Internet-Draft draft-loibl-bacher-idr-flowspec-clarification-00, IETF Secretariat, August 2016. `http://www.ietf.org/internet-drafts/draft-loibl-bacher-idr-flowspec-clarification-00.txt`.

[4] P. Marques, N. Sheth, R. Raszuk, B. Greene, J. Mauch, and D. McPherson. Dissemination of Flow Specification Rules. RFC 5575 (Proposed Standard), August 2009. Updated by RFC 7674.

[5] Danny McPherson, Robert Raszuk, Burjiz Pithawala, Andy, and Susan Hares. Dissemination of flow specification rules for ipv6. Internet-Draft draft-ietf-idr-flow-spec-v6-07, IETF Secretariat, March 2016. `http://www.ietf.org/internet-drafts/draft-ietf-idr-flow-spec-v6-07.txt`.

[6] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). RFC 4271 (Draft Standard), January 2006. Updated by RFCs 6286, 6608, 6793, 7606, 7607, 7705.

[7] S. Sangli, D. Tappan, and Y. Rekhter. BGP Extended Communities Attribute. RFC 4360 (Proposed Standard), February 2006. Updated by RFCs 7153, 7606.

A Router Base Configurations

A.1 R-ALU Alcatel/Nokia

```
A:R-ALU# admin display-config
# TiMOS-B-14.0.R3 both/hops ALCATEL SR 7750 Copyright (c) 2000-2016 Alcatel-Lucent.
# All rights reserved. All use subject to applicable license agreements.
# Built on Wed May 25 17:42:59 PDT 2016 by builder in /rel14.0/b1/R3/panos/main


# Generated TUE AUG 30 07:42:59 2016 UTC


exit all
configure
#--------------------------------------------------
echo "System Configuration"
#--------------------------------------------------
    system
        name "R-ALU"
        dns
        exit
        snmp
            shutdown
        exit
        time
            ntp
                server 192.168.0.250
                no shutdown
            exit
            sntp
                shutdown
            exit
            dst-zone CEST
                start last sunday march 01:00
                end last sunday october 01:00
            exit
            zone CET
        exit
        thresholds
            rmon
            exit
        exit
    exit
#--------------------------------------------------
echo "System Security Configuration"
#--------------------------------------------------
    system
        security
            telnet-server
            user "flow"
                password "XXXX"
                access console
                console
                    no member "default"
                    member "administrative"
                exit
            exit
            ssh
                preserve-key
            exit
```

```
            per-peer-queuing
        exit
    exit
#--------------------------------------------------
echo "System Login Control Configuration"
#--------------------------------------------------
    system
        login-control
            ssh
                inbound-max-sessions 10
            exit
            idle-timeout 1440
        exit
    exit
#--------------------------------------------------
echo "Log Configuration"
#--------------------------------------------------
    log
        syslog 1
            address 192.168.0.250
            no log-prefix
        exit
        log-id 10
            from main
            to syslog 1
            no shutdown
        exit
        log-id 11
            from change
            to syslog 1
            no shutdown
        exit
        log-id 12
            from debug-trace
            to syslog 1
            no shutdown
        exit
    exit
#--------------------------------------------------
echo "System Security Cpm Hw Filters and PKI Configuration"
#--------------------------------------------------
    system
        security
        exit
    exit
#--------------------------------------------------
echo "QoS Policy Configuration"
#--------------------------------------------------
    qos
    exit
#--------------------------------------------------
echo "Card Configuration"
#--------------------------------------------------
    card 1
        card-type iom-c4-xp
        mcm 1
            mcm-type mcm-xp
            no shutdown
        exit
        mda 1
            mda-type m2-10gb-xp-xfp
            no shutdown
```

```
        exit
        mda 3
            mda-type c1-1gb-sfp
            no shutdown
        exit
        mda 4
            mda-type c1-1gb-sfp
            no shutdown
        exit
        mda 5
            no shutdown
        exit
        no shutdown
    exit
#--------------------------------------------------
echo "Port Configuration"
#--------------------------------------------------
    port 1/1/1
        ethernet
            mode hybrid
            encap-type dot1q
        exit
        no shutdown
    exit
    port 1/1/2
        shutdown
        ethernet
            mode hybrid
            encap-type dot1q
        exit
    exit
    port 1/3/1
        shutdown
        ethernet
            mode hybrid
            encap-type dot1q
        exit
    exit
    port 1/4/1
        shutdown
        ethernet
            mode hybrid
            encap-type dot1q
        exit
    exit
    port 1/5/1
        shutdown
        ethernet
            mode hybrid
            encap-type dot1q
        exit
    exit
    port 1/5/2
        shutdown
        ethernet
            mode hybrid
            encap-type dot1q
        exit
    exit
#--------------------------------------------------
echo "System Sync-If-Timing Configuration"
#--------------------------------------------------
```

```
    system
        sync-if-timing
            begin
            commit
        exit
    exit
#--------------------------------------------------
echo "Management Router Configuration"
#--------------------------------------------------
    router management
    exit


#--------------------------------------------------
echo "Router (Network Side) Configuration"
#--------------------------------------------------
    router Base
        network-domains
            network-domain "INTERNET"
            exit
        exit
        interface "intAS"
            address 10.1.1.1/30
            port 1/1/1:101
            ipv6
                address fd50:1:1::1/64
            exit
            no shutdown
        exit
        interface "system"
            address 10.1.0.1/32
            ipv6
                address fd50:1::1/128
            exit
            no shutdown
        exit
        interface "toCISCO"
            address 10.5.4.1/30
            port 1/1/1:14
            ipv6
                address fd50:5:4::1/64
            exit
            no shutdown
        exit
        interface "toHUA"
            address 10.5.9.1/30
            port 1/1/1:13
            ipv6
                address fd50:5:9::1/64
            exit
            no shutdown
        exit
        interface "toJNP"
            address 10.5.0.1/30
            port 1/1/1:12
            ipv6
                address fd50:5::1/64
            exit
            no shutdown
        exit
        interface "toR11"
            address 10.5.5.1/30
```

```
            port 1/1/1:111
            ipv6
                address fd50:5:5::1/64
            exit
            no shutdown
        exit
        autonomous-system 65001
#--------------------------------------------------
echo "Static Route Configuration"
#--------------------------------------------------
        static-route-entry 10.1.0.2/32
            next-hop 10.1.1.2
                no shutdown
            exit
        exit
        static-route-entry fd50:1::2/128
            next-hop fd50:1:1::2
                no shutdown
            exit
        exit
    exit


#--------------------------------------------------
echo "Service Configuration"
#--------------------------------------------------
    service
        customer 1 create
            description "Default customer"
        exit
        customer 100 create
            description "INTERNET"
        exit
        ies 1 customer 1 create
            interface "mgtitf" create
            exit
        exit
        vprn 100 customer 100 create
            interface "Loopback" create
            exit
            interface "r-cis" create
            exit
            interface "r-hua" create
            exit
            interface "r-jnp" create
            exit
            interface "r11" create
            exit
            interface "r1" create
            exit
        exit
        ies 1 customer 1 create
            interface "mgtitf" create
                address 192.168.0.1/24
                sap 1/1/1:66 create
                exit
            exit
            no shutdown
        exit
        vprn 100 customer 100 create
            shutdown
            description "INTERNET"
```

```
autonomous-system 65001
route-distinguisher 65001:100
vrf-target target:65001:100
interface "Loopback" create
    address 10.10.0.1/32
    ipv6
        address fd50:10::1/128
    exit
exit
interface "r-cis" create
    address 10.50.4.1/30
    ipv6
        address fd50:50:4::1/64
    exit
    sap 1/1/1:1014 create
    exit
exit
interface "r-hua" create
    address 10.50.9.1/30
    ipv6
        address fd50:50:9::1/64
    exit
    sap 1/1/1:1013 create
    exit
exit
interface "r-jnp" create
    address 10.50.0.1/30
    ipv6
        address fd50:50::1/64
    exit
    sap 1/1/1:1012 create
    exit
exit
interface "r11" create
    address 10.50.5.1/30
    ipv6
        address fd50:50:5::1/64
    exit
    sap 1/1/1:1111 create
    exit
exit
interface "r1" create
    address 10.10.1.1/30
    ipv6
        address fd50:10:1::1/64
    exit
    sap 1/1/1:1101 create
    exit
exit
static-route-entry 10.10.0.2/32
    next-hop 10.10.1.2
        shutdown
        description "R1-Loopback"
    exit
exit
static-route-entry fd50:10::2/128
    next-hop "fd50:10.1::2"
        shutdown
        description "R1-Loopback"
    exit
exit
bgp
```

```
local-as 65001
router-id 10.1.0.1
group "eBGPv4"
    type external
    flowspec-validate
    neighbor 10.50.0.2
        description "R-JNP"
        family ipv4 flow-ipv4
        peer-as 65002
    exit
    neighbor 10.50.4.2
        description "R-CIS"
        family ipv4 flow-ipv4
        peer-as 65004
    exit
    neighbor 10.50.5.2
        description "R11"
        family ipv4 flow-ipv4
        peer-as 65011
    exit
    neighbor 10.50.9.2
        description "R-HUA"
        family ipv4 flow-ipv4
        peer-as 65003
    exit
exit
group "eBGPv6"
    type external
    flowspec-validate
    neighbor fd50:50::2
        description "R-JNP"
        family ipv6 flow-ipv6
        peer-as 65002
    exit
    neighbor fd50:50:4::2
        description "R-CIS"
        family ipv6 flow-ipv6
        peer-as 65004
    exit
    neighbor fd50:50:5::2
        description "R11"
        family ipv6 flow-ipv6
        peer-as 65011
    exit
    neighbor fd50:50:9::2
        description "R-HUA"
        family ipv6 flow-ipv6
        peer-as 65003
    exit
exit
group "iBGPv4"
    shutdown
    next-hop-self
    type internal
    peer-as 65001
    local-address 10.10.0.1
    neighbor 10.10.0.2
        description "R1"
        family ipv4 flow-ipv4
    exit
exit
group "iBGPv6"
```

```
                shutdown
                next-hop-self
                type internal
                peer-as 65001
                local-address fd50:10::1
                neighbor fd50:10::2
                    description "R1"
                    family ipv6 flow-ipv6
                exit
            exit
            no shutdown
        exit
    exit
#--------------------------------------------------
echo "Router (Service Side) Configuration"
#--------------------------------------------------
    router Base
#--------------------------------------------------
echo "Policy Configuration"
#--------------------------------------------------
        policy-options
            begin
            policy-statement "staticToBGP"
                entry 10
                    from
                        protocol static
                    exit
                    action accept
                    exit
                exit
            exit
            commit
        exit
#--------------------------------------------------
echo "BGP Configuration"
#--------------------------------------------------
        bgp
            local-as 65001
            router-id 10.1.0.1
            rib-management
                ipv4
                    leak-import "back"
                exit
            exit
            group "eBGPv4"
                type external
                neighbor 10.5.0.2
                    description "R-JNP"
                    family ipv4 flow-ipv4
                    peer-as 65002
                exit
                neighbor 10.5.4.2
                    description "R-CIS"
                    family ipv4 flow-ipv4
                    peer-as 65004
                exit
                neighbor 10.5.5.2
                    description "R11"
                    family ipv4 flow-ipv4
                    peer-as 65011
                exit
```

```
            neighbor 10.5.9.2
                shutdown
                description "R-HUA"
                family ipv4 flow-ipv4
                peer-as 65003
            exit
        exit
        group "eBGPv6"
            shutdown
            type external
            flowspec-validate
            neighbor fd50:5::2
                description "R-JNP"
                family ipv6 flow-ipv6
                peer-as 65002
            exit
            neighbor fd50:5:4::2
                description "R-CIS"
                family ipv6 flow-ipv6
                peer-as 65004
            exit
            neighbor fd50:5:5::2
                description "R11"
                family ipv6 flow-ipv6
                peer-as 65011
            exit
            neighbor fd50:5:9::2
                description "R-HUA"
                family ipv6 flow-ipv6
                peer-as 65003
            exit
        exit
        group "iBGPv4"
            next-hop-self
            type internal
            peer-as 65001
            local-address 10.1.0.1
            neighbor 10.1.0.2
                description "R1"
                family ipv4 flow-ipv4
            exit
        exit
        group "iBGPv6"
            shutdown
            next-hop-self
            type internal
            peer-as 65001
            local-address fd50:1::1
            neighbor fd50:1::2
                description "R1"
                family ipv6 flow-ipv6
            exit
        exit
        no shutdown
    exit
exit


#--------------------------------------------------
echo "Source IP Address Configuration"
#--------------------------------------------------
    system
```

```
        security
            source-address
                application syslog "mgtitf"
            exit
        exit
    exit
#--------------------------------------------------
echo "System Time NTP Configuration"
#--------------------------------------------------
    system
        time
            ntp
            exit
        exit
    exit


exit all


# Finished TUE AUG 30 07:43:00 2016 UTC
```

A.2 R-CIS Cisco

```
!! IOS XR Configuration 5.3.3
!! Last configuration change at Thu Aug 25 14:29:55 2016 by flow
!
hostname R-CIS
clock timezone CET 1
clock summer-time CEST date march 5 2000 02:00 october 5 2035 03:00 60
logging console disable
logging monitor disable
logging buffered 100000000
logging buffered informational
logging 192.168.0.250 vrf default severity info port default
domain name lab.local
vrf TEST
 address-family ipv4 unicast
 !
!
vrf INTERNET
 address-family ipv4 unicast
 !
 address-family ipv4 flowspec
 !
 address-family ipv6 unicast
 !
 address-family ipv6 flowspec
 !
!
ftp client passive
ftp client source-interface TenGigE0/0/2/0.66
ntp
 server 192.168.0.250 minpoll 8 maxpoll 12 iburst
!
!
class-map type traffic match-all fs_tuple
 match destination-address ipv4 10.4.255.255 255.255.255.255
 match source-address ipv4 10.3.255.255 255.255.255.255
 match destination-port 1000-1003 1005-1006 1008-1009 1010-1013 1015-1017
```

```
 match source-port 1000-1003 1005-1006 1008-1009 1010-1013 1015-1017
 match protocol 1-2 3-4 5-6 7-8 9-10
 end-class-map
!
policy-map type pbr fs_table_default
 class type traffic fs_tuple
  drop
 !
 class type traffic class-default
 !
 end-policy-map
!
interface Loopback0
 description iBGP-Loopback
 ipv4 address 10.4.0.1 255.255.255.255
 ipv6 address fd50:4::1/128
 ipv6 enable
!
interface Loopback10
 description iBGP-Loopback - VRF INTERNET
 vrf INTERNET
 ipv4 address 10.40.0.1 255.255.255.255
 ipv6 address fd50:40::1/128
 ipv6 enable
!
interface MgmtEth0/RSP0/CPU0/0
 shutdown
!
interface MgmtEth0/RSP0/CPU0/1
 shutdown
!
interface TenGigE0/0/2/0
!
interface TenGigE0/0/2/0.14
 description R-ALU
 ipv4 address 10.5.4.2 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:5:4::2/64
 ipv6 enable
 encapsulation dot1q 14
!
interface TenGigE0/0/2/0.24
 description R-JNP
 ipv4 address 10.5.10.2 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:5:a::2/64
 ipv6 enable
 encapsulation dot1q 24
!
interface TenGigE0/0/2/0.34
 description R-HUA
 ipv4 address 10.5.3.2 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:5:3::2/64
 ipv6 enable
 encapsulation dot1q 34
!
interface TenGigE0/0/2/0.66
 description LAB-MGMT
 ipv4 address 192.168.0.4 255.255.255.0
 encapsulation dot1q 66
!
```

```
interface TenGigE0/0/2/0.401
 description R4
 ipv4 address 10.4.1.1 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:4:1::1/64
 ipv6 enable
 encapsulation dot1q 401
!
interface TenGigE0/0/2/0.412
 description R12
 ipv4 address 10.5.7.1 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:5:7::1/64
 ipv6 enable
 encapsulation dot1q 412
!
interface TenGigE0/0/2/0.1014
 description R-ALU - VRF INTERNET
 vrf INTERNET
 ipv4 address 10.50.4.2 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:50:4::2/64
 ipv6 enable
 encapsulation dot1q 1014
!
interface TenGigE0/0/2/0.1024
 description R-JNP - VRF INTERNET
 vrf INTERNET
 ipv4 address 10.50.10.2 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:50:a::2/64
 ipv6 enable
 encapsulation dot1q 1024
!
interface TenGigE0/0/2/0.1034
 description R-HUA - VRF INTERNET
 vrf INTERNET
 ipv4 address 10.50.3.2 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:50:3::2/64
 ipv6 enable
 encapsulation dot1q 1034
!
interface TenGigE0/0/2/0.1401
 description R4 - VRF INTERNET
 vrf INTERNET
 ipv4 address 10.40.1.1 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:40:1::1/64
 ipv6 enable
 encapsulation dot1q 1401
!
interface TenGigE0/0/2/0.1412
 description R12 - VRF INTERNET
 vrf INTERNET
 ipv4 address 10.50.7.1 255.255.255.252
 ipv6 nd suppress-ra
 ipv6 address fd50:50:7::1/64
 ipv6 enable
 encapsulation dot1q 1412
!
interface TenGigE0/0/2/1
```

```
 shutdown
!
interface TenGigE0/0/2/1.3000
!
interface TenGigE0/0/2/2
 shutdown
!
interface TenGigE0/0/2/3
 shutdown
!
route-policy RMAP-PERMIT
  done
end-policy
!
route-policy RMAP-FS-PERMIT
  done
end-policy
!
router static
 address-family ipv4 unicast
  10.4.0.2/32 10.4.1.2 description R4-Lo
 !
 address-family ipv6 unicast
  fd50:4::2/128 fd50:4:1::2 description R4-Lo
 !
!
router bgp 65004
 bgp router-id 10.4.0.1
 address-family ipv4 unicast
 !
 address-family vpnv4 unicast
 !
 address-family ipv6 unicast
 !
 address-family vpnv6 unicast
 !
 address-family ipv4 flowspec
 !
 address-family ipv6 flowspec
 !
 address-family vpnv4 flowspec
 !
 address-family vpnv6 flowspec
 !
 neighbor-group EBGP
  address-family ipv4 unicast
   route-policy RMAP-PERMIT in
   route-policy RMAP-PERMIT out
  !
  address-family ipv4 flowspec
   route-policy RMAP-PERMIT in
   route-policy RMAP-PERMIT out
   soft-reconfiguration inbound
  !
 !
 neighbor-group iBGP
  local address 10.4.0.1
  address-family ipv4 unicast
   route-policy RMAP-PERMIT in
   route-policy RMAP-PERMIT out
   next-hop-self
  !
```

```
 address-family ipv4 flowspec
  route-policy RMAP-PERMIT in
  route-policy RMAP-PERMIT out
  validation disable
  soft-reconfiguration inbound
 !
!
neighbor-group EBGPv6
 shutdown
 address-family ipv6 unicast
  route-policy RMAP-PERMIT in
  route-policy RMAP-PERMIT out
 !
 address-family ipv6 flowspec
  route-policy RMAP-PERMIT in
  route-policy RMAP-PERMIT out
  soft-reconfiguration inbound
 !
!
neighbor-group iBGPv6
 shutdown
 local address fd50:4::1
 address-family ipv6 unicast
  route-policy RMAP-PERMIT in
  route-policy RMAP-PERMIT out
  next-hop-self
 !
 address-family ipv6 flowspec
  route-policy RMAP-PERMIT in
  route-policy RMAP-PERMIT out
  validation disable
  soft-reconfiguration inbound
 !
!
neighbor 10.4.0.2
 remote-as 65004
 use neighbor-group iBGP
 description R4
!
neighbor 10.5.3.1
 remote-as 65003
 use neighbor-group EBGP
 shutdown
 description R-HUA
!
neighbor 10.5.4.1
 remote-as 65001
 use neighbor-group EBGP
 shutdown
 description R-ALU
!
neighbor 10.5.7.2
 remote-as 65012
 use neighbor-group EBGP
 description R12
!
neighbor 10.5.10.1
 remote-as 65002
 use neighbor-group EBGP
 description R-JNP
!
neighbor fd50:4::2
```

```
  remote-as 65004
  use neighbor-group iBGPv6
  description R4
 !
 neighbor fd50:5:3::1
  remote-as 65003
  use neighbor-group EBGPv6
  description R-HUA
 !
 neighbor fd50:5:4::1
  remote-as 65001
  use neighbor-group EBGPv6
  description R-ALU
 !
 neighbor fd50:5:7::2
  remote-as 65012
  use neighbor-group EBGPv6
  description R12
 !
 neighbor fd50:5:a::1
  remote-as 65002
  use neighbor-group EBGPv6
  description R-JNP
 !
!
generic-interface-list double1
 interface TenGigE0/0/2/0
!
flowspec
 local-install interface-all
 address-family ipv4
  local-install interface-all
 !
!
ssh server v2
ssh server vrf default
ssh timeout 120
end
```

A.3 R-HUA Huawei

```
!Software Version V800R007C00SPC100
!Last configuration was updated at 2016-08-25 14:49:02+02:00 DST by bacherm
!Last configuration was saved at 2016-06-27 14:14:04+02:00 DST by flow
#
clock timezone CET add 01:00:00
clock daylight-saving-time CEST repeating 02:00 last Sun Mar 02:00 last Sun Oct
     01:00
#
sysname R-HUA
#
set neid 2fc2d7
#
FTP server enable
#
undo info-center source default channel 4
info-center loghost 192.168.0.250 level informational
info-center timestamp debugging short-date precision-time tenth-second
info-center timestamp log short-date precision-time millisecond
#
```

```
fan speed auto
#
undo user-security-policy enable
#
service-template template-default0
#
service-template template-default1
#
service-template template-default2
#
service-template template-default3
#
service-template template-default4
#
ntp-service server disable
ntp-service ipv6 server disable
ntp-service unicast-server 192.168.0.250
#
undo telnet server enable
undo telnet ipv6 server enable
#
diffserv domain default
#
diffserv domain 5p3d
#
soc
#
ip vpn-instance INTERNET
 ipv4-family
  route-distinguisher 65003:100
 ipv6-family
  route-distinguisher 65003:100
#
ip vpn-instance mgmt
 ipv4-family
  route-distinguisher 1:1
  vpn-target 1:1 export-extcommunity
  vpn-target 1:1 import-extcommunity
#
bfd
#
acl name AL-FS-TEST basic
#
aaa
 user-password min-len 3
 local-user flow password irreversible-cipher $1a$XXXXX$
 local-user flow service-type ftp ssh
 local-user flow level 15
 local-user flow ftp-directory cfcard:
 #
 authentication-scheme default0
 #
 authentication-scheme default1
 #
 authentication-scheme default
  authentication-mode local radius
 #
 authorization-scheme default
 #
 accounting-scheme default0
 #
 accounting-scheme default1
```

```
 #
 domain default0
 #
 domain default1
 #
 domain default_admin
#
license
#
interface Eth-Trunk4
#
interface GigabitEthernet1/0/1
 description next_layer
 undo shutdown
 undo dcn
 statistic enable
#
interface GigabitEthernet1/0/1.13
 vlan-type dot1q 13
 description R-ALU
 ipv6 enable
 ip address 10.5.9.2 255.255.255.252
 ipv6 address FD50:5:9::2/64
 statistic enable
#
interface GigabitEthernet1/0/1.23
 vlan-type dot1q 23
 description R-JNP
 ipv6 enable
 ip address 10.5.1.2 255.255.255.252
 ipv6 address FD50:5:1::2/64
 statistic enable
#
interface GigabitEthernet1/0/1.34
 vlan-type dot1q 34
 description R-CIS
 ipv6 enable
 ip address 10.5.3.1 255.255.255.252
 ipv6 address FD50:5:3::1/64
 statistic enable
#
interface GigabitEthernet1/0/1.66
 vlan-type dot1q 66
 description LAB-MGMT
 ip address 192.168.0.3 255.255.255.0
#
interface GigabitEthernet1/0/1.301
 vlan-type dot1q 301
 description R3
 ipv6 enable
 ip address 10.3.1.1 255.255.255.252
 ipv6 address FD50:3:1::1/64
 statistic enable
#
interface GigabitEthernet1/0/1.312
 vlan-type dot1q 312
 description R12
 ipv6 enable
 ip address 10.5.8.1 255.255.255.252
 ipv6 address FD50:5:8::1/64
 statistic enable
#
```

```
interface GigabitEthernet1/0/1.1013
 vlan-type dot1q 1013
 description R-ALU - VRF INTERNET
 ip binding vpn-instance INTERNET
 ipv6 enable
 ip address 10.50.9.2 255.255.255.252
 ipv6 address FD50:50:9::2/64
 statistic enable
#
interface GigabitEthernet1/0/1.1023
 vlan-type dot1q 1023
 description R-JNP - VRF INTERNET
 ip binding vpn-instance INTERNET
 ipv6 enable
 ip address 10.50.1.2 255.255.255.252
 ipv6 address FD50:50:1::2/64
 statistic enable
#
interface GigabitEthernet1/0/1.1034
 vlan-type dot1q 1034
 description R-CIS - VRF INTERNET
 ip binding vpn-instance INTERNET
 ipv6 enable
 ip address 10.50.3.1 255.255.255.252
 ipv6 address FD50:50:3::1/64
 statistic enable
#
interface GigabitEthernet1/0/1.1301
 vlan-type dot1q 1301
 description R3 - VRF INTERNET
 ip binding vpn-instance INTERNET
 ipv6 enable
 ip address 10.30.1.1 255.255.255.252
 ipv6 address FD50:30:1::1/64
 statistic enable
#
interface GigabitEthernet1/0/1.1312
 vlan-type dot1q 1312
 description R12 - VRF INTERNET
 ip binding vpn-instance INTERNET
 ipv6 enable
 ip address 10.50.8.1 255.255.255.252
 ipv6 address FD50:50:8::1/64
 statistic enable
#
interface GigabitEthernet1/0/2
 undo shutdown
 undo dcn
#
interface GigabitEthernet1/0/3
 undo shutdown
 undo dcn
#
interface GigabitEthernet1/0/4
 undo shutdown
 undo dcn
#
interface GigabitEthernet1/0/5
 undo shutdown
 undo dcn
#
interface GigabitEthernet1/0/6
```

```
 undo shutdown
 undo dcn
#
interface GigabitEthernet1/0/7
 undo shutdown
 undo dcn
#
interface GigabitEthernet1/0/8
 undo shutdown
#
interface GigabitEthernet1/0/9
 undo shutdown
#
interface GigabitEthernet1/0/10
 undo shutdown
#
interface GigabitEthernet1/0/11
 shutdown
#
interface LoopBack0
 ipv6 enable
 ip address 10.3.0.1 255.255.255.255
 ipv6 address FD50:3::1/128
#
interface NULL0
#
bgp 65003
 router-id 10.3.0.1
 undo default ipv4-unicast
 graceful-restart
 graceful-restart timer wait-for-rib 360
 peer 10.3.0.2 as-number 65003
 peer 10.3.0.2 description R3
 peer 10.5.1.1 as-number 65002
 peer 10.5.1.1 description R-JNP
 peer 10.5.3.2 as-number 65004
 peer 10.5.3.2 description R-CIS
 peer 10.5.8.2 as-number 65012
 peer 10.5.8.2 description R12
 peer 10.5.9.1 as-number 65001
 peer 10.5.9.1 description R-ALU
 peer FD50:3::2 as-number 65003
 peer FD50:3::2 description R3
 peer FD50:5:1::1 as-number 65002
 peer FD50:5:1::1 description R-JNP
 peer FD50:5:3::2 as-number 65004
 peer FD50:5:3::2 description R-CIS
 peer FD50:5:8::2 as-number 65012
 peer FD50:5:8::2 description R12
 peer FD50:5:9::1 as-number 65001
 peer FD50:5:9::1 description R-ALU
 #
 ipv4-family unicast
  undo synchronization
  peer 10.3.0.2 enable
  peer 10.3.0.2 next-hop-local
  peer 10.5.1.1 enable
  peer 10.5.3.2 enable
  peer 10.5.8.2 enable
  peer 10.5.9.1 enable
 #
 ipv4-family flow
```

```
  peer 10.3.0.2 enable
  peer 10.3.0.2 validation-disable
  peer 10.5.1.1 enable
  peer 10.5.3.2 enable
  peer 10.5.8.2 enable
  peer 10.5.9.1 enable
 #
 ipv6-family unicast
  undo synchronization
  peer FD50:3::2 enable
  peer FD50:3::2 next-hop-local
  peer FD50:5:1::1 enable
  peer FD50:5:3::2 enable
  peer FD50:5:8::2 enable
  peer FD50:5:9::1 enable
 #
 ipv4-family vpn-instance INTERNET
#
undo dcn
#
route-policy RMAP-DENY-FLOWSPEC deny node 10
 if-match ip-prefix PL-FS-TEST
#
route-policy RMAP-DENY-FLOWSPEC permit node 20
#
ip ip-prefix PL-FS-TEST index 10 permit 20.0.0.0 30
ip community-filter advanced COM-DENY-FLOWSPEC-ACTION permit ^0:0
ip community-filter advanced COM-DENY-FLOWSPEC-ACTION permit ^8006:0:0
ip community-filter advanced COM-DENY-FLOWSPEC-ACTION permit ^0.*$
ip community-filter advanced COM-DENY-FLOWSPEC-ACTION permit ^64.*$
ip community-filter advanced COM-DENY-FLOWSPEC-ACTION permit 8006:0:0
ip extcommunity-list soo basic CL-DENY-FLOWSPEC-ACTION index 10 permit 0:0
ip extcommunity-list soo advanced CL-DENY-FLOWSPEC-ACTION-2 index 10 permit 10 ^0:0
#
ip route-static 10.3.0.2 255.255.255.255 10.3.1.2 description R3
#
ipv6 route-static FD50:3::2 128 FD50:3:1::2 description R3
#
#
snmp-agent trap type base-trap
#
lldp enable
#
stelnet server enable
scp server enable
#
ssh client first-time enable
#
#
user-interface con 0
 authentication-mode password
 set authentication password cipher $1a$XXXXX$
#
user-interface vty 0 4
 authentication-mode aaa
 user privilege level 15
 idle-timeout 60 0
 protocol inbound ssh
#
local-aaa-server
#
return
```

<R-HUA>

A.4 R-JNP juniper

```
## Last commit: 2016-06-27 21:59:46 CEST by flow
version 15.1F5.15;
system {
    domain-name nextlayer.at;
    time-zone Europe/Zurich;
    no-redirects;
    no-redirects-ipv6;
    authentication-order [ tacplus password ];
    root-authentication {
        encrypted-password "$1$XXXXX"; ## SECRET-DATA
    }
    name-server {
        10.255.0.4;
    }
    tacplus-server {
        10.255.0.4 {
            secret "$9$XXXXX"; ## SECRET-DATA
            timeout 1;
        }
    }
    accounting {
        events [ login change-log interactive-commands ];
        destination {
            tacplus {
                server {
                    10.255.0.4 {
                        secret "$9$XXXXX"; ## SECRET-DATA
                        timeout 1;
                    }
                }
            }
        }
    }
    login {
        user admin-template {
            uid 2002;
            class super-user;
        }
        user flow {
            uid 2003;
            class super-user;
            authentication {
                encrypted-password "$1$XXXXX"; ## SECRET-DATA
            }
        }
        user lab {
            uid 2000;
            class super-user;
            authentication {
                encrypted-password "$1$XXXXX"; ## SECRET-DATA
            }
        }
        user remote {
            uid 2001;
            class super-user;
        }
```

```
    }
    services {
        ssh {
            max-sessions-per-connection 32;
        }
        netconf {
            ssh;
        }
    }
    syslog {
        user * {
            any emergency;
        }
        host 192.168.0.250 {
            any any;
            source-address 192.168.0.2;
        }
        file messages {
            any notice;
            authorization info;
        }
        file interactive-commands {
            interactive-commands any;
        }
        file default-log-messages {
            any info;
            match "(requested 'commit' operation)|(copying configuration to
                 juniper.save)|(commit complete)|ifAdminStatus|(FRU power)|(FRU
                 removal)|(FRU insertion)|(link
                 UP)|transitioned|Transferred|transfer-file|(license add)|(license
                 delete)|(package -X update)|(package -X delete)|(FRU Online)|(FRU
                 Offline)|(plugged in)|(unplugged)|CFMD_CCM_DEFECT| LFMD_3AH |
                 RPD_MPLS_PATH_BFD|(Master Unchanged, Members Changed)|(Master
                 Changed, Members Changed)|(Master Detected, Members Changed)|(vc
                 add)|(vc delete)|(Master detected)|(Master changed)|(Backup
                 detected)|(Backup changed)|(interface vcp-)";
            structured-data;
        }
        time-format year millisecond;
        source-address 10.255.0.17;
    }
    commit synchronize;
    ntp {
        server 192.168.0.250;
    }
}
chassis {
    redundancy {
        routing-engine 0 master;
        routing-engine 1 backup;
        failover {
            on-loss-of-keepalives;
            on-disk-failure;
        }
        graceful-switchover;
    }
    fpc 1 {
        pic 1 {
            tunnel-services;
        }
    }
    fpc 2 {
```

```
        power off;
    }
    fpc 3 {
        power off;
    }
    fpc 4 {
        power off;
    }
    fpc 5 {
        power off;
    }
    network-services enhanced-ip;
}
interfaces {
    xe-0/2/1 {
        flexible-vlan-tagging;
        encapsulation flexible-ethernet-services;
        unit 12 {
            description "LINK TO AS65001 - ALCATEL";
            vlan-id 12;
            family inet {
                address 10.5.0.2/30;
            }
            family inet6 {
                address fd50:5:0::2/64;
            }
        }
        unit 23 {
            description "LINK TO AS65003 - HUAWEI";
            vlan-id 23;
            family inet {
                address 10.5.1.1/30;
            }
            family inet6 {
                address fd50:5:1::1/64;
            }
        }
        unit 24 {
            description "LINK TO AS65004 - CISCO";
            vlan-id 24;
            family inet {
                address 10.5.10.1/30;
            }
            family inet6 {
                address fd50:5:a::1/64;
            }
        }
        unit 66 {
            description "LINK TO MGMT";
            vlan-id 66;
            family inet {
                address 192.168.0.2/24;
            }
        }
        unit 201 {
            description "LINK TO iBGP ROUTE-COLLECTOR";
            vlan-id 201;
            family inet {
                address 10.2.1.1/30;
            }
            family inet6 {
                address fd50:2:1::1/64;
```

```
    }
}
unit 211 {
    description "LINK TO AS65011 ROUTE-COLLECTOR";
    vlan-id 211;
    family inet {
        address 10.5.6.1/30;
    }
    family inet6 {
        address fd50:5:6::1/64;
    }
}
unit 1012 {
    description R-ALU;
    vlan-id 1012;
    family inet {
        address 10.50.0.2/30;
    }
    family inet6 {
        address fd50:50:0::2/64;
    }
}
unit 1023 {
    description R-HUA;
    vlan-id 1023;
    family inet {
        address 10.50.1.1/30;
    }
    family inet6 {
        address fd50:50:1::1/64;
    }
}
unit 1024 {
    description R-CIS;
    vlan-id 1024;
    family inet {
        address 10.50.10.1/30;
    }
    family inet6 {
        address fd50:50:a::1/64;
    }
}
unit 1201 {
    description R2;
    vlan-id 1201;
    family inet {
        address 10.20.1.1/30;
    }
    family inet6 {
        address fd50:20:1::1/64;
    }
}
unit 1211 {
    description R11;
    vlan-id 1211;
    family inet {
        address 10.50.6.1/30;
    }
    family inet6 {
        address fd50:50:6::1/64;
    }
}
```

```
    }
    lo0 {
        unit 0 {
            family inet {
                address 10.2.0.1/32;
            }
            family inet6 {
                address fd50:2::1/128;
            }
        }
        unit 1000 {
            family inet {
                address 10.20.0.1/32;
            }
            family inet6 {
                address fd50:2::1/128;
            }
        }
    }
}
routing-options {
    nonstop-routing;
    rib inet6.0 {
        static {
            route fd50:2::2/128 next-hop fd50:2:1::2;
        }
    }
    static {
        route 10.2.0.2/32 next-hop 10.2.1.2;
    }
    flow {
        term-order standard;
    }
    router-id 10.2.0.1;
    autonomous-system 65002;
    forwarding-table {
        export ECMP;
    }
}
protocols {
    bgp {
        log-updown;
        group EBGP {
            family inet {
                unicast;
                flow;
            }
            family inet6 {
                unicast;
                flow;
            }
            neighbor 10.5.6.2 {
                peer-as 65011;
            }
            neighbor 10.5.0.1 {
                peer-as 65001;
            }
            neighbor 10.5.1.2 {
                peer-as 65003;
            }
            neighbor 10.5.10.2 {
                peer-as 65004;
```

```
            }
        }
        group IBGP {
            local-address 10.2.0.1;
            family inet {
                unicast;
                flow {
                    no-validate THEN-ACCEPT;
                }
            }
            export NEXT-HOP-SELF;
            neighbor 10.2.0.2 {
                local-address 10.2.0.1;
                peer-as 65002;
            }
        }
        group EBGP6 {
            family inet6 {
                unicast;
                flow;
            }
            neighbor fd50:5:6::2 {
                peer-as 65011;
            }
            neighbor fd50:5:0::1 {
                peer-as 65001;
            }
            neighbor fd50:5:1::2 {
                peer-as 65003;
            }
            neighbor fd50:5:a::2 {
                peer-as 65004;
            }
        }
        group IBGP6 {
            local-address fd50:2::1;
            family inet6 {
                unicast;
                flow {
                    no-validate THEN-ACCEPT;
                }
            }
            export NEXT-HOP-SELF;
            neighbor fd50:2::2 {
                peer-as 65002;
            }
        }
    }
}
policy-options {
    policy-statement ECMP {
        term ECMP {
            then {
                load-balance per-packet;
            }
        }
    }
    policy-statement NEXT-HOP-SELF {
        from protocol bgp;
        then {
            next-hop self;
        }
```

```
    }
    policy-statement THEN-ACCEPT {
        then accept;
    }
}
routing-instances {
    INTERNET {
        instance-type vrf;
        interface xe-0/2/1.1012;
        interface xe-0/2/1.1023;
        interface xe-0/2/1.1024;
        interface xe-0/2/1.1201;
        interface xe-0/2/1.1211;
        interface lo0.1000;
        route-distinguisher 65002:100;
        vrf-target target:65002:100;
        routing-options {
            rib INTERNET.inet6.0 {
                static {
                    route fd50:2::2/128 next-hop fd50:2:1::2;
                }
            }
            static {
                route 10.20.0.2/32 next-hop 10.20.1.2;
            }
        }
        protocols {
            bgp {
                disable;
                group EBGP {
                    family inet {
                        unicast;
                        flow;
                    }
                    family inet6 {
                        unicast;
                        flow;
                    }
                    neighbor 10.50.6.2 {
                        peer-as 65011;
                    }
                    neighbor 10.50.0.1 {
                        peer-as 65001;
                    }
                    neighbor 10.50.1.2 {
                        peer-as 65003;
                    }
                    neighbor 10.50.10.2 {
                        peer-as 65004;
                    }
                }
                group IBGP {
                    local-address 10.20.0.1;
                    family inet {
                        unicast;
                        flow {
                            no-validate THEN-ACCEPT;
                        }
                    }
                    export NEXT-HOP-SELF;
                    neighbor 10.20.0.2 {
                        local-address 10.20.0.1;
```

```
                    peer-as 65002;
                }
            }
            group EBGP6 {
                family inet6 {
                    unicast;
                    flow;
                }
                neighbor fd50:50:6::2 {
                    peer-as 65011;
                }
                neighbor fd50:50:0::1 {
                    peer-as 65001;
                }
                neighbor fd50:50:1::2 {
                    peer-as 65003;
                }
                neighbor fd50:50:a::2 {
                    peer-as 65004;
                }
            }
            group IBGP6 {
                local-address fd50:20::1;
                family inet6 {
                    unicast;
                    flow {
                        no-validate THEN-ACCEPT;
                    }
                }
                export NEXT-HOP-SELF;
                neighbor fd50:20::2 {
                    peer-as 65002;
                }
            }
        }
    }
}

{master}
```