

BGP FlowSpec extensions for Routing Policy Distribution(RPD) draft-li-idr-flowspec-rpd-01

Robin Li(lizhenbin@huawei.com)

Liang Ou(oul@gsta.com)

Yujia Luo(luoyuj@gsta.com)

Sujian Lu(jasonlu@tencent.com)

Vincent Zhuang(zhuangshunwan@huawei.com)

Eric Wu(eric.wu@huawei.com)

IETF94, Yokohama

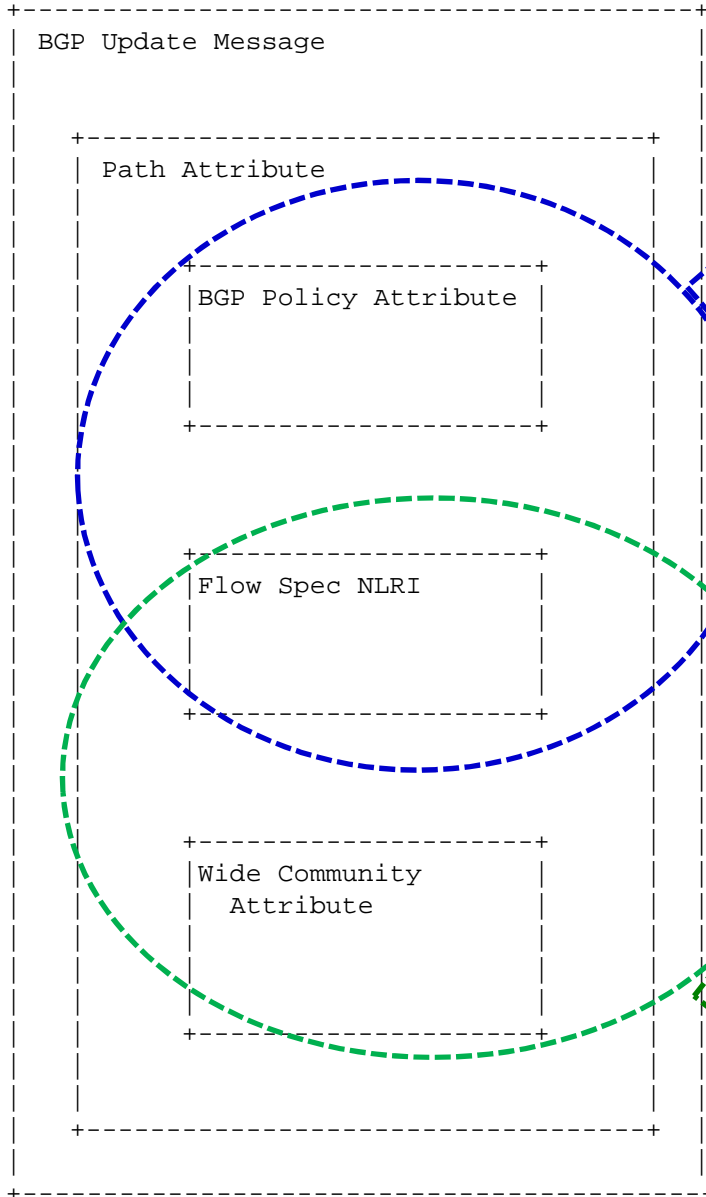
Changed from 00 version

□ Alternate protocol extensions using Wide Community

- Will discuss and hear suggestion from co-authors of Wide Community draft.
- Need WG's opinions about the more proper way.

□ One more operator has similar requirements and joined in. Maybe adding new use cases in next version.

RPD Mechanism in Summary



Option I:

1. Effective on which routes → Filtered by Flowspec NLRI
2. Effective on which peers → Filtered by BGP Policy Attribute
3. Take the action in BGP Policy Attribute

Option II:

1. Effective on which routes → Filtered by Flowspec NLRI
2. Effective on which peers → Filtered by Wide Community
3. Take the action in Wide Community

Protocol extensions option I(v00)

□ RPD protocol: BGP Flowspec

- Filtering rule: destination for prefix1/prefix2
- Action: R-bit introduced, more info carried in new attribute

```
+-----+
| reserved           | R | S | T |
+-----+
```

□ BGP Policy Attribute

• Attribute structure

```
+-----+
| Match fields (Variable)
+-----+
| Action fields (Variable)
+-----+
```

• Match field

```
+-----+
| Match Type (2 octets)
+-----+
| Number of Sub-TLVs (2 octets)
+-----+
| Sub-TLVs (Variable)
+-----+
```

• Action field

```
+-----+
| Action Type (2 octets)
+-----+
| Action Length (2 octets)
+-----+
| Action Values (Variable)
+-----+
```

- Action type 1: Route-Preference
- Action type 2: Route-Prepend-AS

□ Match type

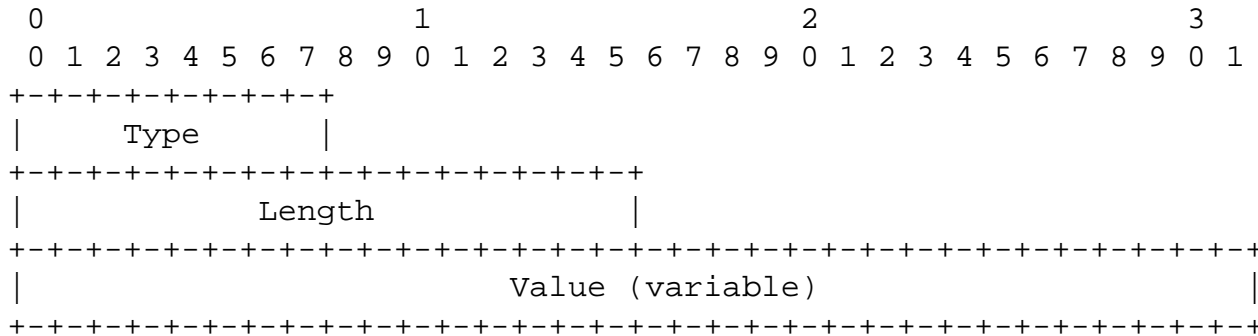
- Value 0: Permit, specifies the permit mode of a match rule
- Value 1: Deny, specifies the deny mode of a match rule.

□ Sub-TLVs

- Type 1: IPv4 Neighbor
- Type 2: IPv6 Neighbor
- Type 3: ASN list

Protocol extensions option II(v01)

□ New Wide Community Atoms



- Type 1: Autonomous System number list
- Type 2: IPv4 prefix (1 octet prefix length + prefix) list
- Type 3: IPv6 prefix (1 octet prefix length + prefix) list
- Type 4: Integer list
- Type 5: IEEE Floating Point Number list
- Type 6: Neighbor Class list
- Type 7: User-defined Class list
- Type 8: UTF-8 String
- Type TBD: BGP IPv4 neighbor --- Newly introduced in this draft
- Type TBD: BGP IPv6 neighbor --- Newly introduced in this draft

Protocol extensions option II(v01)

□ Inbound Traffic Control encoding example

```
0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Container Type 1 (1) |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|1 0 0 0 0 0 0 0|
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Hop Count: 0 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Length: 36 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Community: PREPEND N TIMES TO AS | 18 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Own ASN | 100 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Context ASN# | 100 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|ExcTargetTLV(2) | Length: 11 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| IPv4Neig(TBD) | Length: 8 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Local Speaker | #IGW2 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Remote Speaker | #Speaker1 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Param TLV (3) | Length: 7 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Integer (4) | Length: 4 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Prepend # | 5 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```

Protocol extensions option II(v01)

□ Outbound Traffic Control encoding example

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Container Type 1 (1)          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1 0 0 0 0 0 0 0|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Hop Count: 0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Length: 36 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Community: LOCAL PREFERENCE          20 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Own ASN          100 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Context ASN#          100 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|TargetTLV(1) | Length:          11 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| IPv4Neig(TBD) | Length:          8 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Local Speaker          #IGW2 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Remote Speaker          #Speaker1 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Param TLV (3) | Length:          7 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Integer (4) | Length:          4 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Increment #          100 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Next step

- ❑ Collect feedback and comments.
- ❑ Refine this draft according to comments.
- ❑ Adding new use cases from operators.

To Be Deleted

Motivation

□ Provider's requirements for traffic adjustment:

- Business development or network failure introduces link congestion and overload.
- Network transmission quality decreased as the result of delay, loss and need to adjust traffic to other paths.
- To control OPEX and CPEX, prefer the transit provider with lower price.

Motivation

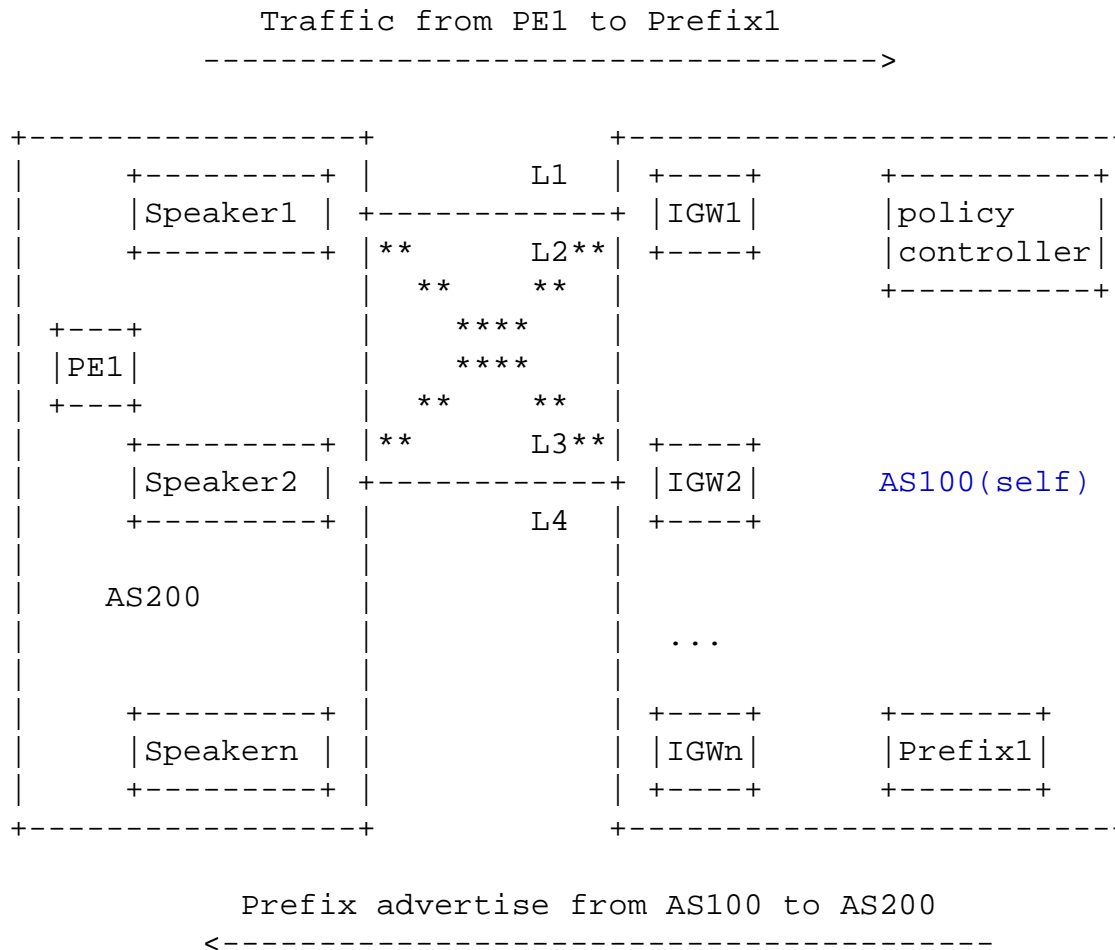
❑ Drawbacks using traditional routing policy:

- Device-based manual provisioning will cause configuration burden and misconfiguration.
- Complexity keeps increased gradually and difficulty to maintain.

Automatic provisioning mechanism is needed.

Application

□ Inbound traffic control



□ EBGP peering:

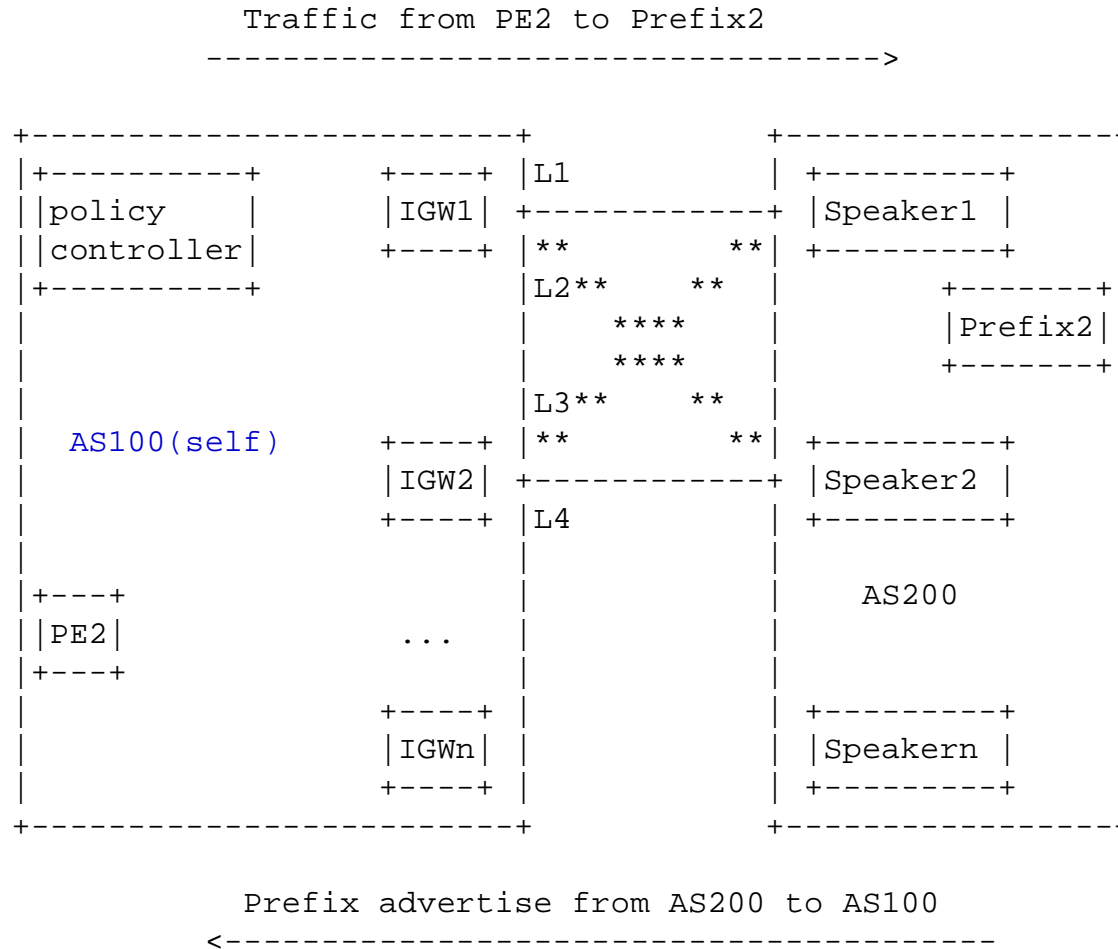
- Speaker1---L1---IGW1
- Speaker2---L2---IGW1
- Speaker1---L3---IGW2
- Speaker2---L4---IGW2

□ Requirement:

- Administration only on AS100
- Traffic enter AS100 through L3

Application

□ Outbound traffic control



□ EBGP peering:

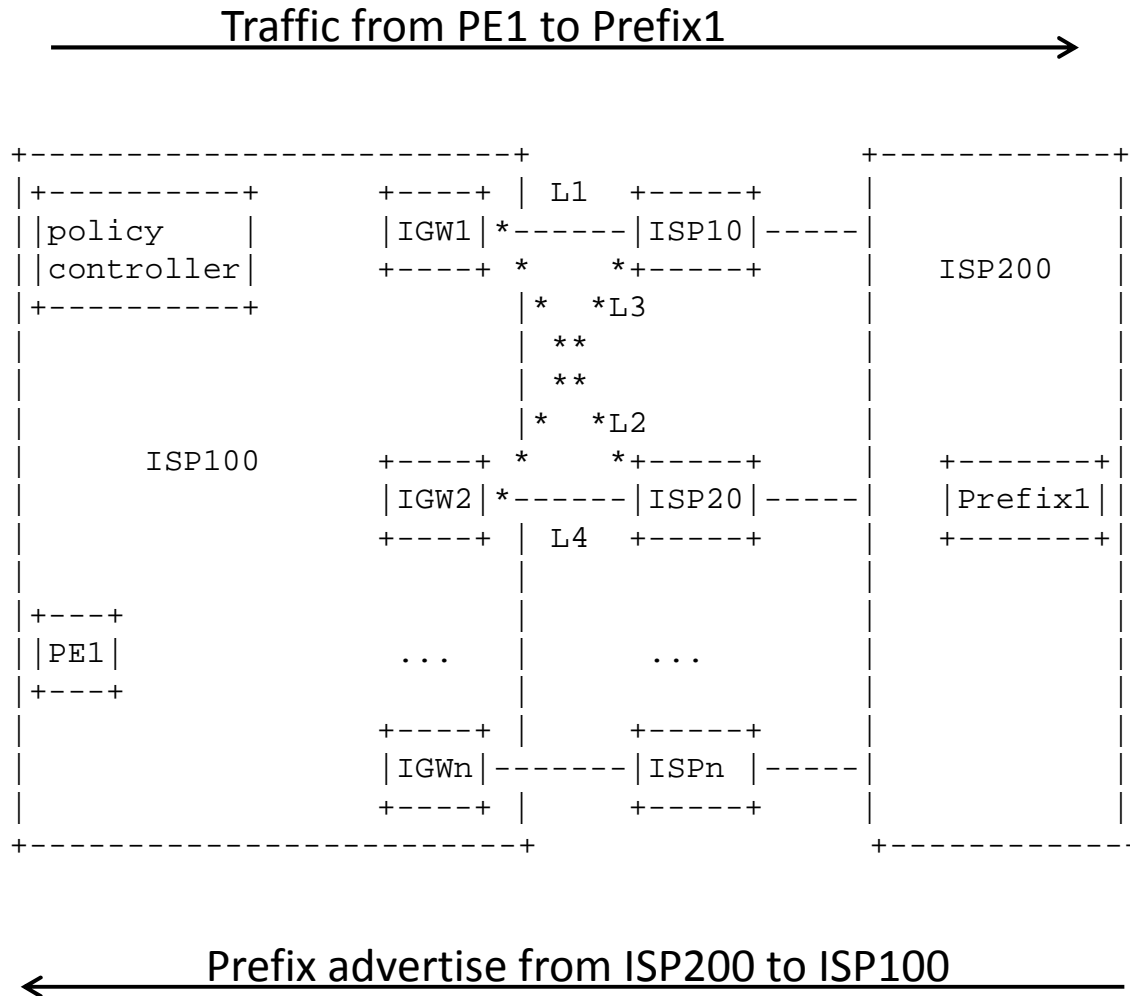
- IGW1---L1---Speaker1
- IGW1---L2---Speaker2
- IGW2---L3---Speaker1
- IGW2---L4---Speaker2

□ Requirement:

- Administration only on AS100
- Traffic exit through L3

Application

□ Outbound traffic control



□ EBGP peering:

- IGW1---L1---ISP10
- IGW1---L2---ISP20
- IGW2---L3---ISP10
- IGW2---L4---ISP20

□ Requirement:

- Traffic exit through L3
- Changing intra-AS route selection.

Solution

□ Routing Policy Distribution(RPD)

- Taking effect on control plane
- Impact decision on remote site

□ RPD protocol: BGP Flowspec

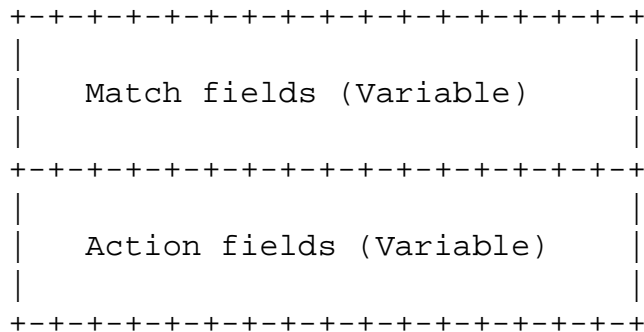
- Filtering rule: destination for prefix1/prefix2
- Action: R-bit introduced, more info carried in new attribute

```
+---+---+---+---+---+---+---+---+
| reserved           | R | S | T |
+---+---+---+---+---+---+---+---+
```

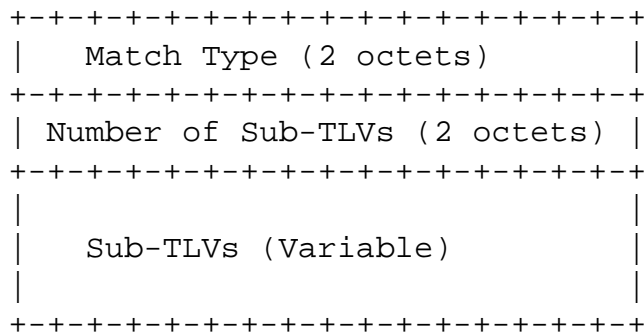
Solution

□ BGP Policy Attribute

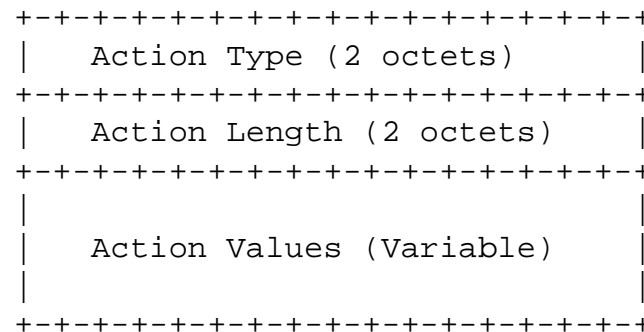
• Attribute structure



• Match field



• Action field



- Action type 1: Route-Preference
- Action type 2: Route-Prepend-AS

□ Match type

- Value 0: Permit, specifies the permit mode of a match rule
- Value 1: Deny, specifies the deny mode of a match rule.

□ Sub-TLVs

- Type 1: IPv4 Neighbor
- Type 2: IPv6 Neighbor
- Type 3: ASN list

Solution

□ Outbound traffic control

- Match type: permit
- IPv4 neighbor sub-TLV:
 - ✓ Local BGP Speaker IGW2
 - ✓ Remote BGP Peer Speaker1
- Action type: Route-Preference

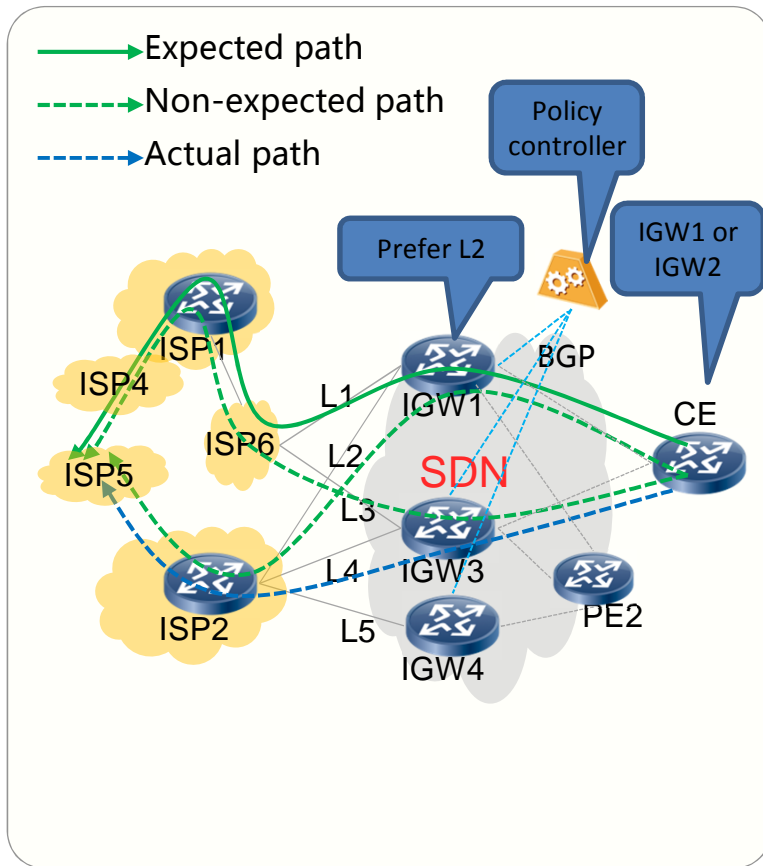
□ Inbound traffic control

- Match type: deny
- IPv4 neighbor sub-TLV:
 - ✓ Local BGP Speaker IGW2
 - ✓ Remote BGP Peer Speaker1
- Action type: Route-Prepend-AS
- Action value: Prepend-AS five times

Consideration

□ Current BGP-FS limitation

- Can't affect AS external device's decision.

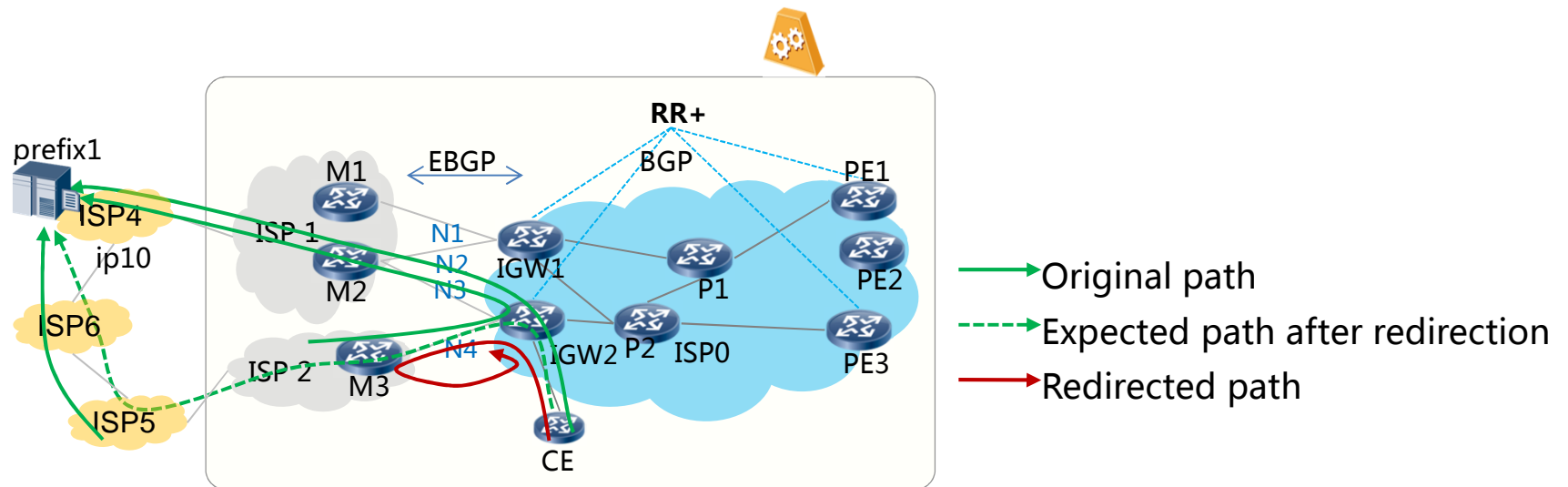


- On IGW1, L2 is preferred previously.
- BGP-Flowspec can redirect to L1 but CE may prefer IGW3, which makes IGW1's redirection useless.

Consideration

❑ Current BGP-FS limitation

- Changing decision in forwarding plane may introduce loop.



BGP-Flowspec only changes behavior in forwarding plane, no impact on control plane. IGW may still prefer M2 and advertise to M3. It is possible M3 still prefer IGW2's route so loop can happen between M3 and IGW2.