BGP FlowSpec extensions for Routing Policy Distribution(RPD)

draft-li-idr-flowspec-rpd-01

Zhenbin Li(lizhenbin@huawei.com)
Liang Ou(oul@gsta.com)
Yujia Luo(luoyuj@gsta.com)
Sujian Lu(jasonlu@tencent.com)
Shunwan Zhuang(zhuangshunwan@huawei.com)
Nan Wu(eric.wu@huawei.com)

IETF94, Yokohama

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.

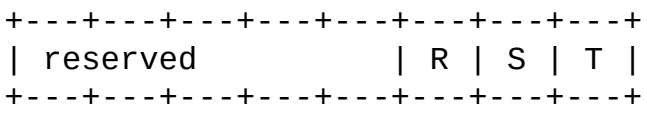
Solution

☐Routing Policy Distribution(RPD)

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

☐RPD protocol: BGP Flowspec

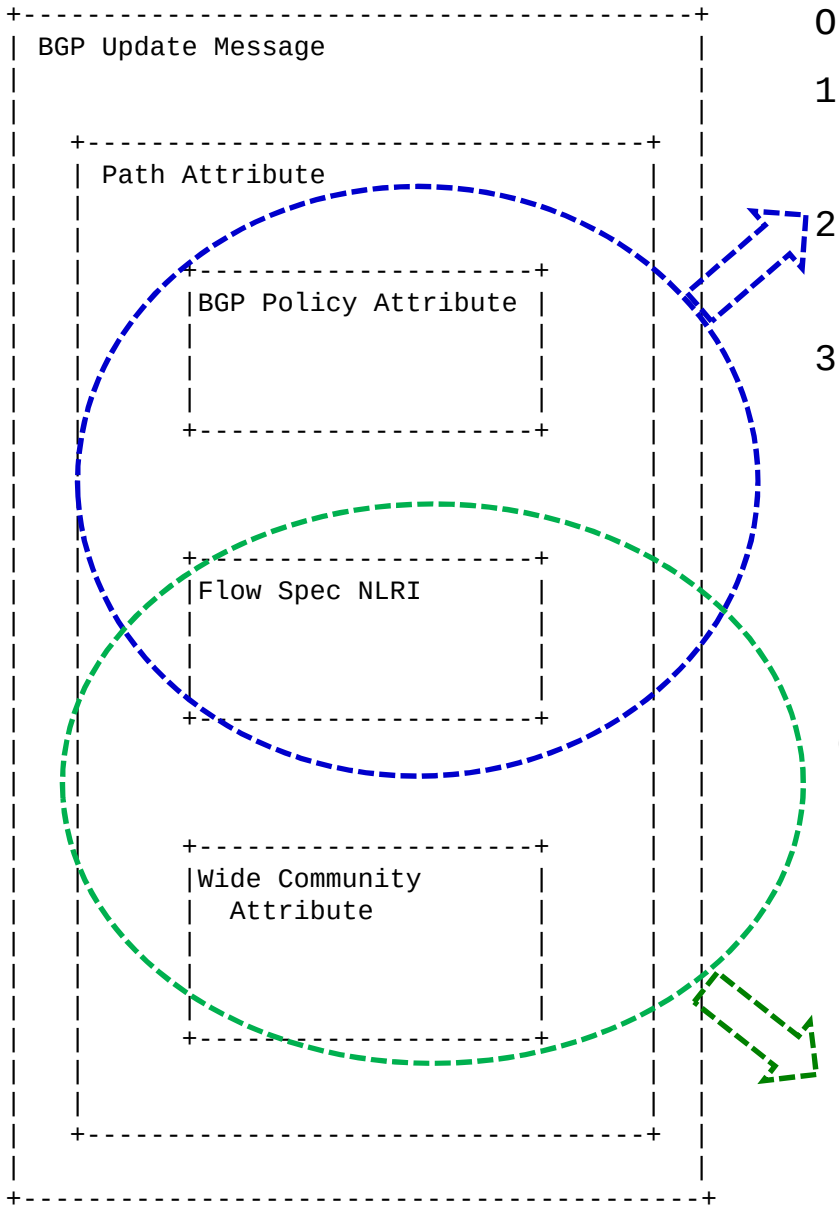
- Filtering rule: destination for prefix1/prefix2
- Action: R-bit introduced, more info carried in new attribute



Changed from 00 version

- Alternate protocol extensions using enhanced Wide Community
- One more operator, Tencent, 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

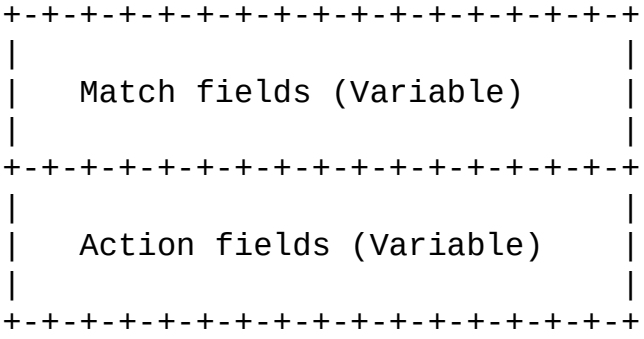
Option II:

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

Protocol extensions option I(v00)

☐BGP Policy Attribute

•Attribute structure

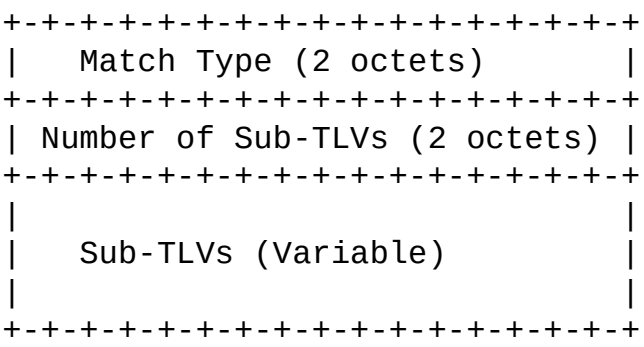


•Action field

•Action type 1: Route-Preference

•Action type 2: Route-Prepend-AS

•Match field



☐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

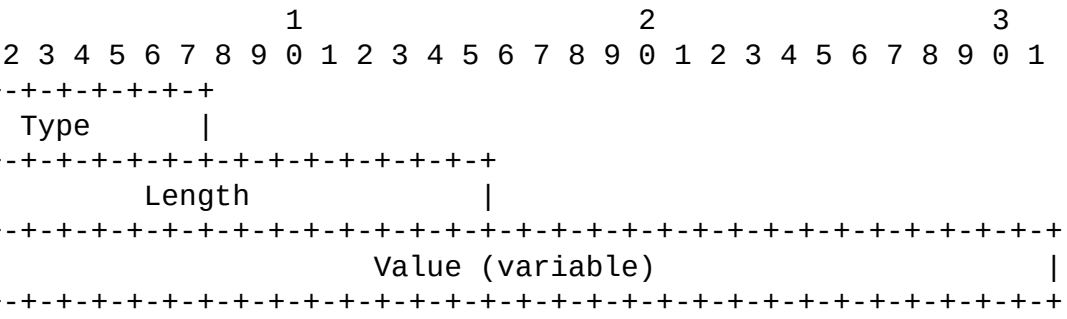
•Type 1: IPv4 Neighbor

•Type 2: IPv6 Neighbor

•Type 3: ASN list

Protocol extensions option II(v01)

- Wide Community is enhanced to filter a set of target routes to apply actions other than act as the attributes of advertised routes.
- New Wide Community Atoms

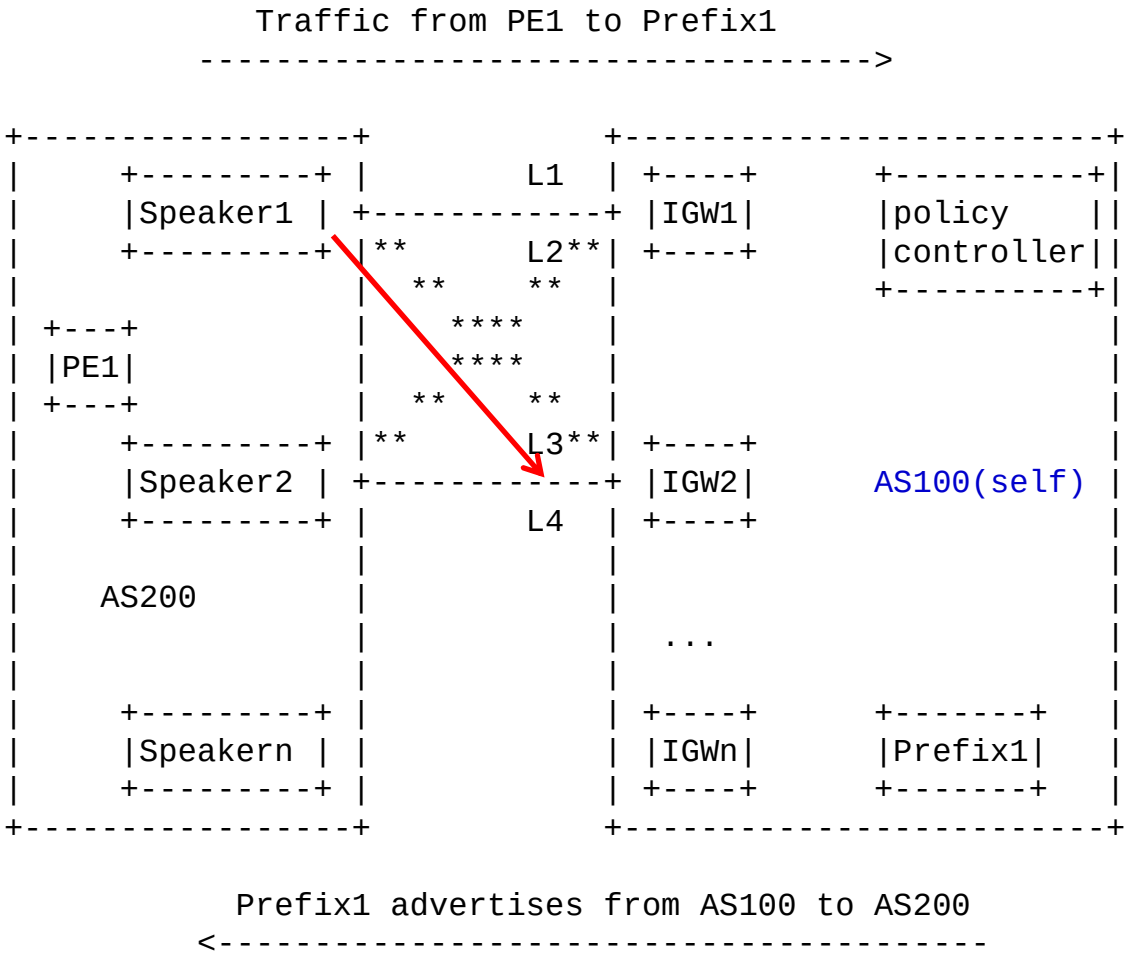


- Type 1: Autonomous System number list
- Type 2: IPv4 prefix (1 octet prefix length + prefix) list
- Type 3: IPv6 prefix (1 octet prefix length + prefix) list
- Type 4: Integer list
- Type 5: IEEE Floating Point Number list
- Type 6: Neighbor Class list
- Type 7: User-defined Class list7
- Type 8: UTF-8 String
- Type TBD: BGP IPv4 neighbor --- Newly introduced in this draft
- Type TBD: BGP IPv6 neighbor --- Newly introduced in this draft

- Actions of Wide Community can be reused and maybe enhanced in the future.

Application (1)

Inbound traffic control



- EBGP peering:
  - Speaker1---L1---IGW1
  - Speaker2---L2---IGW1
  - Speaker1---L3---IGW2
  - Speaker2---L4---IGW2

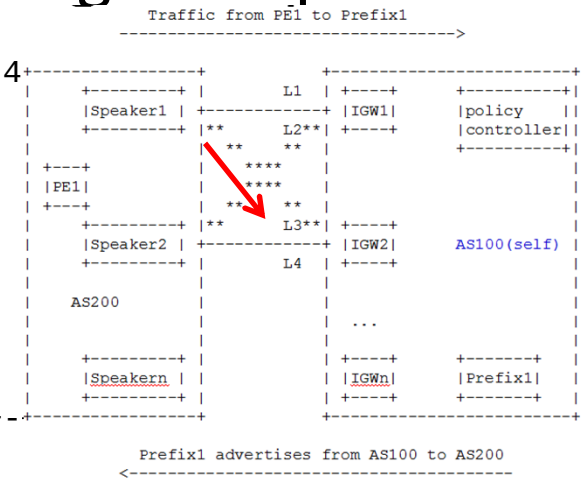
- Requirement:
  - Administration only on AS100
  - Traffic enter AS100 through L3

→ Traffic Direction

Encoding Example (1)

☐Inbound Traffic Control encoding example

```
 0                   1                   2
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 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 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```



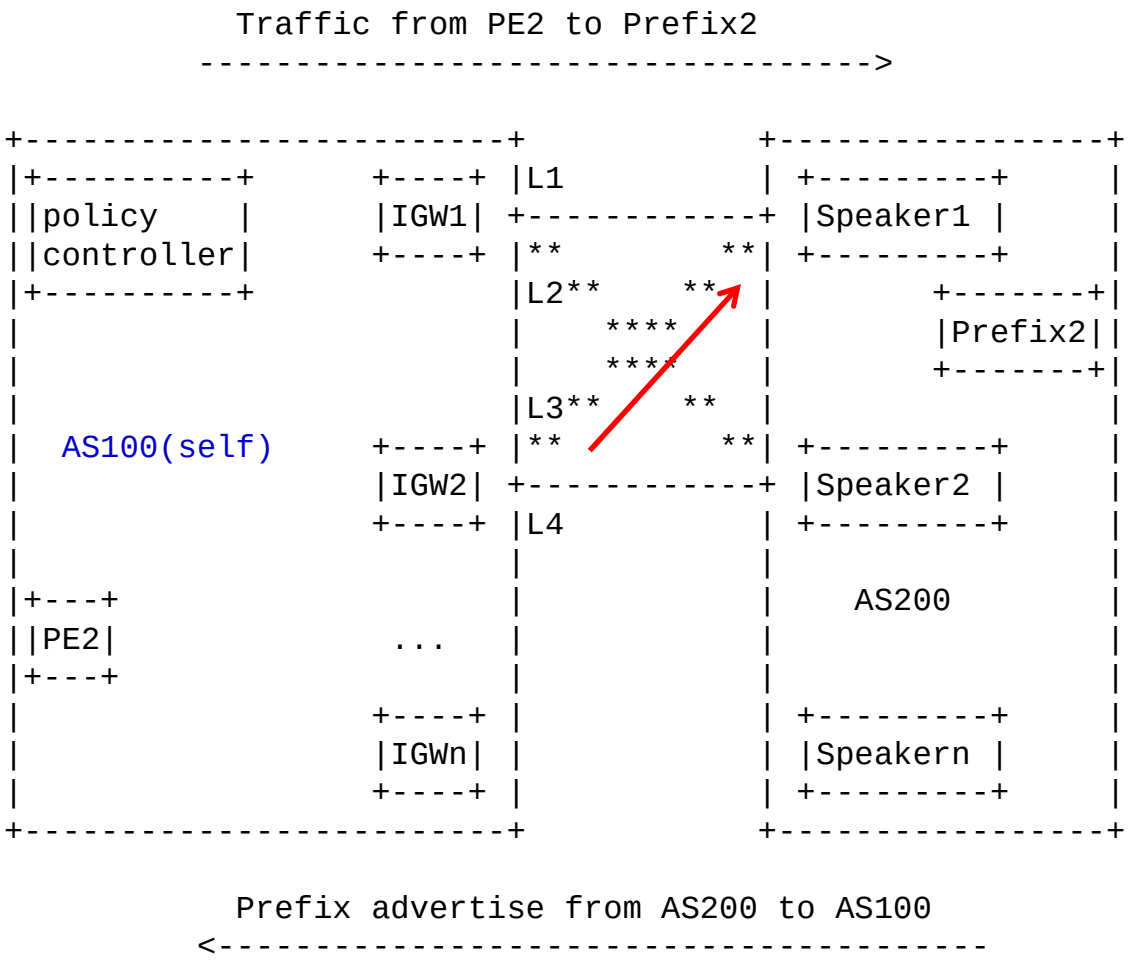
☐EBGP peering:
- •Speaker1---L1---IGW1
- •Speaker2---L2---IGW1
- •Speaker1---L3---IGW2
- •Speaker2---L4---IGW2

☐Requirement:
- •Administration only on AS100
- •Traffic enter AS100 through L3

- ☐ As required in the case, traffic from PE1 to Prefix1 need to enter through L3, so IGWs except IGW2 should prepend ASN list to Prefix1 when populating to AS100.
- ☐ As shown in left figure, community "PREPEND N TIMES TO AS" and "Exclude Target(s) TLV" are be used.

Application (2)

Outbound traffic control

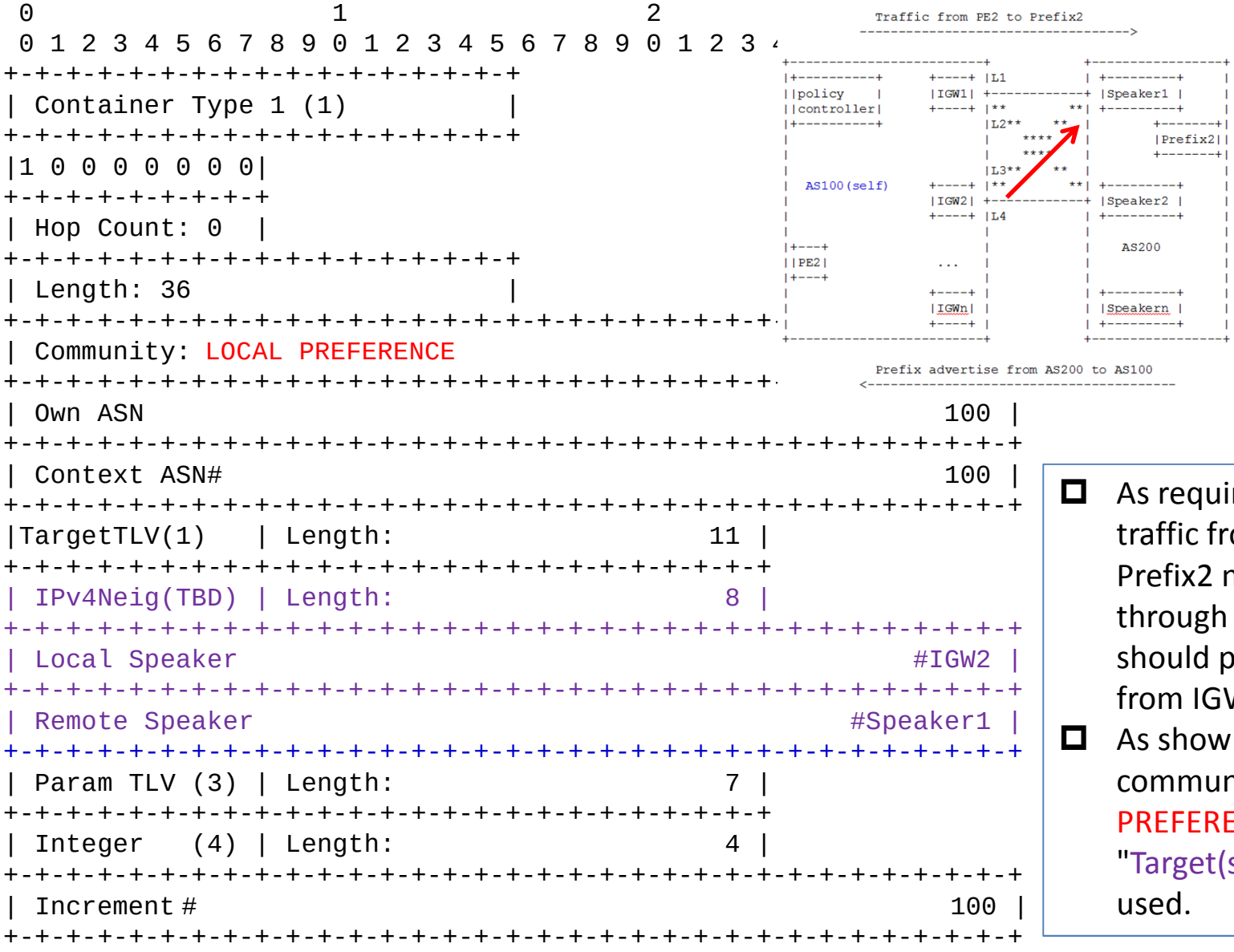


- EBGP peering:
  - IGW1---L1---Speaker1
  - IGW1---L2---Speaker2
  - IGW2---L3---Speaker1
  - IGW2---L4---Speaker2
- Requirement:
  - Administration only on AS100
  - Traffic exit through L3

Traffic Direction

Encoding Example (2)

Outbound Traffic Control encoding example



- EBGP peering:
  - IGW1---L1---Speaker1
  - IGW1---L2---Speaker2
  - IGW2---L3---Speaker1
  - IGW2---L4---Speaker2
- Requirement:
  - Administration only on AS100
  - Traffic exit through L3

- As required in the case, traffic from PE2 to Prefix2 need to exit through L3, so IGWs should prefer the route from IGW2 to Speaker1.
- As shown in left figure, community "LOCAL PREFERENCE" and "Target(s) TLV" are be used.

Next step

- Solicit comments on the alternative solutions.
- Refine this draft.
- Adding new use cases from operators.