



Note Well

Any submission to the IETF intended by the Contributor for publication as all or part of an IETF Internet-Draft or RFC and any statement made within the context of an IETF activity is considered an "IETF Contribution". Such statements include oral statements in IETF sessions, as well as written and electronic communications made at any time or place, which are addressed to:

- The IETF plenary session
- The IESG, or any member thereof on behalf of the IESG
- Any IETF mailing list, including the IETF list itself, any working group or design team list, or any other list functioning under IETF auspices
- Any IETF working group or portion thereof
- Any Birds of a Feather (BOF) session
- The IAB or any member thereof on behalf of the IAB
- The RFC Editor or the Internet-Drafts function

All IETF Contributions are subject to the rules of [RFC 5378](#) and [RFC 8179](#).

Statements made outside of an IETF session, mailing list or other function, that are clearly not intended to be input to an IETF activity, group or function, are not IETF Contributions in the context of this notice. Please consult [RFC 5378](#) and [RFC 8179](#) for details.

A participant in any IETF activity is deemed to accept all IETF rules of process, as documented in Best Current Practices RFCs and IESG Statements.

A participant in any IETF activity acknowledges that written, audio and video records of meetings may be made and may be available to the public.



I E T F



Autodiscovery

- Clear interest in WG to work on this topic.
- No clear consensus on a specific approach.
- Proposals (four of them) progressively closer semantically, seems likely convergence can happen.
- But, important differences remain.
 - At least: transport (L2, L3), liveness, security, maybe multihop.
- Chairs propose to charter a design team to close on a single proposal, by next meeting.
 - Might build on one of the existing drafts, might be a new draft, up to the design team.
 - Emphasis on pragmatism, OK to limit applicability (for example, to a single administrative domain).
- Questions? Comments? Volunteers?

draft-hujun-idr-bgp-ipsec-01

draft-hujun-idr-bgp-ipsec- transport-mode-00

Hu Jun, Nokia

IETF 106

Updates in draft-hujun-idr-bgp-ipsec-01

- replaces color sub-TLV with a new IPsec configuration tag sub-TLV
- add rule on selecting TLV when there multiple feasible TLVs in section (#operation)
- change crypto used in example of section (#operation)
- change title from "BGP Signaled IPsec Tunnel Configuration" to "BGP Provisioned IPsec Tunnel Configuration"
- Add a section (#operationspecifics) on some operation specifics
- add more content in (#security)
- add specification of number of time each new sub-TLV allowed in a given tunnel TLV
- add clarification in section (#intro) to clarify IPsec tunnel means IPsec tunnel mode
- traffic selector protocol and port range now come from tag mapped configuration

draft-hujun-idr-bgp-ipsec-transport-mode-00

- This draft defines a method to advertise IP tunnel encapsulation with IPsec transport mode protection in BGP; e.g GRE with IPsec transport mode, VXLAN with IPsec transport mode ..etc

```
-----  
|IPv4 header  | ESP | GRE | Payload |   ESP   | ESP|  
|(any options)| Hdr | Hdr | Packet  | Trailer | ICV|  
-----
```

```
          |<-- encryption --->|
```

```
          |<----- integrity ---->|
```

Example: IPv4 GRE tunnel packet with ESP transport protection

IP Tunnel with IPsec Transport Mode

- Unlike IPsec tunnel mode, which is essentially encapsulate whole IP packet as an payload of a new IPsec tunnel packet, IPsec transport mode does not introduce any new IP header, so it is not a tunnel stack as in “X in Y” type;
- Due to this is the reason, IP tunnel with IPsec transport mode doesn't fit in current spec of ietf-idr-tunnel-encaps, an extension is needed, draft-hujun-idr-bgp-ipsec-transport-mode-00 is proposed to address such use case;

How does it work?

- A new IPsec Transport Protected sub-TLV is introduced, its value its value is a IPsec configuration tag as defined in hujun-idr-bgp-ipsec.
- When an IP tunnel encapsulation TLV include this new sub-TLV, it means advertising router requires IPsec transport mode protection for the corresponding IP tunnel, using the IPsec config as following:
 - ESP transport mode
 - private and public routing instance is same as routing instance in which the packet to be forwarded
 - peer tunnel address is same as indicated by Remote Endpoint sub-TLV
 - local traffic selector:
 - address range: local tunnel endpoint address
 - protocol: tag mapped configuration
 - port range: tag mapped configuration
 - remote traffic selector:
 - address range: address in Remote Endpoint sub-TLV of selected tunnel encapsulation TLV
 - protocol: tag mapped configuration
 - port range: tag mapped configuration
- its transform and other configuration maps to the tag indicated in the IPsec configuration tag sub-TLV

WG Adoption

As extensions of WG draft ietf-idr-tunnel-encaps, I propose to adopt both draft-hujun-idr-bgp-ipsec and draft-hujun-idr-bgp-ipsec-transport-mode

draft-dunbar-idr-sdwan-port-safi-05

Linda Dunbar

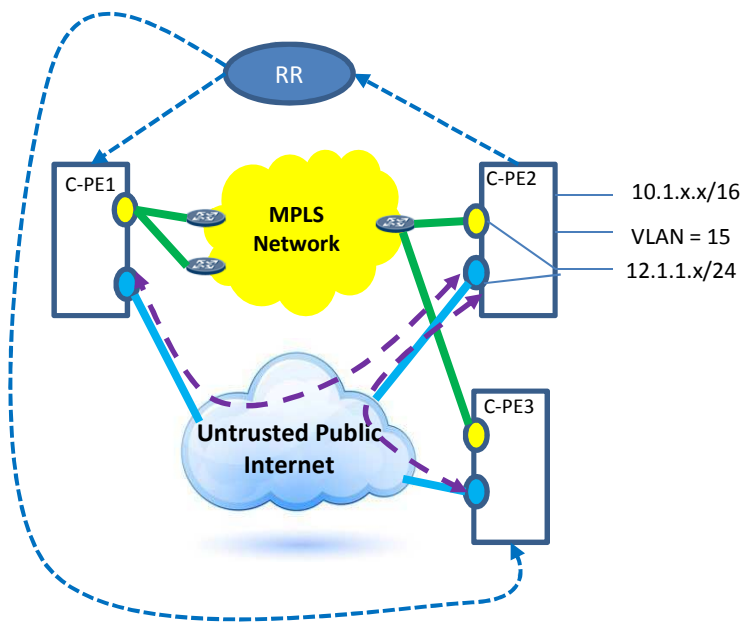
Sue Hares

IETF 106 Nov 2019

Intention of the draft

- Informational draft
- Purpose:
 - We applied a Port-SAFI in the first come first serve (FCFS) category (SAFI =74)
 - Intended to inform how the Port-SAFI is used for SDWAN overlay network
- We would like to hear your feedback.

Port Based IPsec Tunnel Confederated IPsec via RR



Regular MPLS BGP Routes Update

BGP UPDATE Messages from C-PE2 to announce all the routes attached

- MP-NLRI Path Attribute
- Nexthop (C-PE2)
- NLRI
 - 10.1.x.x.
 - VLAN 15
 - 12.1.1x

BGP UPDATE Messages from C-PE2 to RR for WAN port properties:

- MP-NLRI Path Attribute:
 - Port Identifier encoding
- Tunnel-Encap Path Attribute:
 - NAT for the WAN Port
 - IPsec SA-12 (C-PE1->C-PE2 via the specific port)
 - IPsec SA-32 (C-PE3->C-PE2 via the specific port)

New NLRI for the WAN Port

N subTLVs in the Tunnel Encap Path Attribute

Attributes for End Point Identity

NLRI Length	1 octet	SDWAN Can have different TYPE
Network-Type	2 Octets	
Port-Distinguisher	4 octets	Locally significant within the node
SDWAN-Site-ID	4 octets	
SDWAN-Node-ID	4 or 16 octets	routable address across WAN

- NLRI Length: expressed in bits as defined in [RFC4760].
- Network-Type: SDWAN
- Port Distinguisher: Locally significant Port identifier.
- SDWAN-Site-ID: Globally unique site identifier.
- SDWAN Node ID: Locally significant node identifier (system ID or the loopback address (IPv4 or IPv6)).

Advantage of new NLRI: to represent different address space than client routes: SDWAN WAN port; similarr approach as the new NLRI used for SR Policy

Disadvantage of new NLRI: intermediate Routers can drop the UPDATE due to not recognizing the new NLRI.

Not applicable to SDWAN overlay, as the UPDATE to RR is simple IP forwarding, not terminated by any routers/switches in between

SubTLV for the NAT Property of the WAN Port

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|EncapExt Type | EncapExt subTLV Length |I|O|R|R|R|R|R|R|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| NAT Type      | Encap-Type   |Trans networkID| RD ID   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
|           Local  IP Address
|           32-bits for IPv4, 128-bits for Ipv6
|           ~~~~~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
|           Local  Port
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
|           Public IP
|           32-bits for IPv4, 128-bits for Ipv6
|           ~~~~~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
|           Public Port
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Flags:

-I bit (CPE port address or Inner address scheme)

If =0 → inner addr is IPv4.

If =1 → inner address is IPv6.

-O bit (Outer address scheme):

If =0 → the public (outer) address is IPv4.

If =1 → the public (outer) address is IPv6.

-R bits: reserved for future use.

Must be set to 0 now.

NAT Type: without NAT; 1:1 static NAT; Full Cone; Restricted Cone; Port Restricted Cone; Symmetric; or Unknown (i.e. no response from the STUN server).

Encap Type : the supported encap types for the port facing public network, such as IPsec+GRE, IPsec+VxLAN, IPsec without GRE, GRE (when packets don't need encryption)

Transport Network ID: Central Controller assign a global unique ID to each transport network ;

RD ID: Routing Domain ID , Need to be global unique.

Local IP: The local (or private) IP address of the port ;

Local Port: used by Remote SDWAN node for establishing IPsec to this specific port.

Public IP: The IP address after the NAT. If NAT is not used, this field is set to NULL.

Public Port: The Port after the NAT. If NAT is not used, this field is set to NULL.

SubTLV for the Port Based IPsec

The IPsecSA sub-TLV is for the SDWAN edge node to establish IPsec security association with their peers via the port that face untrusted network:

[illegible]

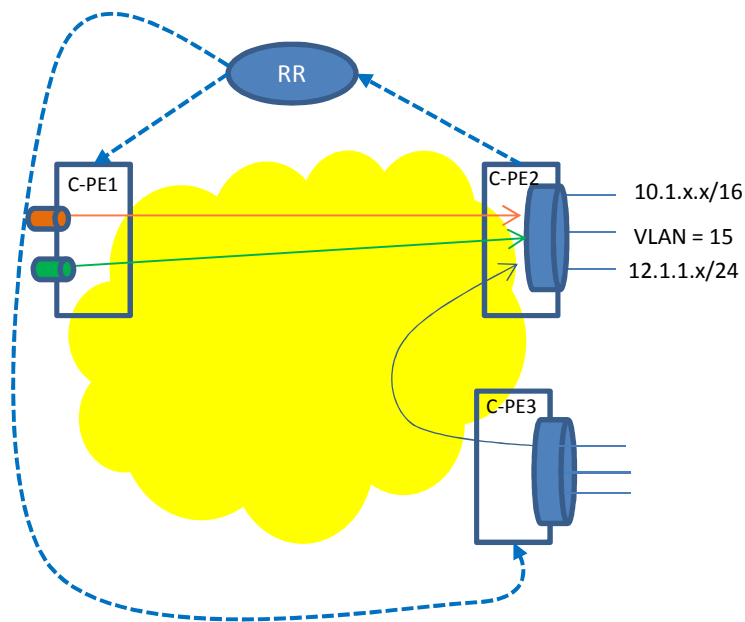
Device Information Message (DIM) are derived from
draft-carrel-ipsecme-controller-ike-01

Next Step

- Call for WG adoption
- Why?
 - Demonstrate how BGP is used in Port Based IPsec to scale SDWAN overlay

BACKUP SLIDES

Recap of BESS' presentation on BGP for Homogeneous SDWAN



One BGP UPDATE Message from C-PE2 to RR:

- multiple routes encoded in the MP-NLRI Path Attribute
 - 10.1.x.x/16
 - VLAN #15
 - 12.1.1.x/24
- IPsec attributes are encoded in the Tunnel-Encap Path Attribute
 - IPsec attributes for all possible remote nodes, or
 - IPsec attributes for specific remote nodes, or
 - IPsec attributes for specific remote subnets
-

WHY BGP

- here are some of the Compelling reasons of using BGP to distribute SDWAN edge properties among peers that might be spread across the globe:
- (note: the BGP for SDWAN Edges is running at different layers than the BGP for underlay networks, i.e. not “FLAT” BGP. They are among SDWAN edges, not for exposing to underlay provider as you stated EBGp. When the underlay network service providers use SDWAN to temporarily expand bandwidth in some segments, they have more reason to use BGP to minimize amount of learning & configuration of introducing new protocols in their environment)
- – BGP already widely deployed as sole protocol (see RFC 7938). Even if not for this purpose of propagating SDWAN WAN port properties, the BGP base protocol implementation is supported by virtually all switches/routers (virtual & physical). Even AWS VPC export the BGP routes.
- – Wide acceptance – minimal learning (which is very important requirement for operations)
- – Robust and simple implementation,
- – Reliable transport
- – Guaranteed in-order delivery
- – Incremental updates
- – No flooding and selective filtering
- – RR already has the capability to apply policies to communications among peers.
- Bottom line: It is much easier to add one function than adding a brand-new protocol stack.
- Alternative: extending LISP, NHRP, DSVPN/DMVPN
- – In addition to more proposal changes needed, NHRP/DSVPN/DMVPN don't scale well.
- – More learning, more barrier to be deployed, just think how many decades of painful journey deploying IPv6.
- –
- Prior extension of BGP for non-client routes reachability: Flowspec, BGP LS, Segment routing policies, etc

Deprecation of AS_SET and AS_CONFED_SET

<https://tools.ietf.org/html/draft-ietf-idr-deprecate-as-set-confed-set-02>

K. Sriram, Warren Kumari, Lilia Hannachi, Jeff Haas

IETF IDR WG Meeting

IETF 106

November 2019

Interest in deprecation of AS_SET and AS_CONFED_SET

- WG seems to have strong motivation to eliminate the use of these Attributes

Analysis of AS_SETs in BGP (IPv4)

Unique prefixes (with or without AS_SET) : 826535

Total # routes with AS_SETs : 477

routes with only one AS in AS_SET : 383

routes that are /24 prefix (aggregate) announcements : 239

Total # routes that seem meaningless or malformed : 456

Total # routes that seem meaningful : 21

Details: https://www.nist.gov/sites/default/files/documents/2019/10/23/detailed-as_set-analysis.txt

Analysis of AGGREGATOR, ATOMIC_AGGREGATE

*** When there is AGGREGATOR without AS_SET ***

Unique prefixes (with or without AS_SET) : 826535

Unique prefixes without AS_SET but with AGGREGATOR: 75698 (9.2%)

Unique prefixes with ATOMIC_AGGREGATE: 47258

Unique prefixes with AGGREGATOR and ATOMIC_AGGREGATE: 44971

Unique prefixes with AGGREGATOR and without ATOMIC_AGGREGATE: 31769

https://www.nist.gov/sites/default/files/documents/2019/10/23/detailed-as_set-analysis.txt

Source of some unusual aggregated AS_PATHs (jhaas)

41.196.34.0/23 701 174 8452 24863 {37069}

RFC 4271 compliant implementations of aggregation can yield an AS_SET of length one under the following conditions:

1. One or more contributing routes that are completely internal. (NULL AS_PATH.)
2. One or more contributing routes with the same single AS number.

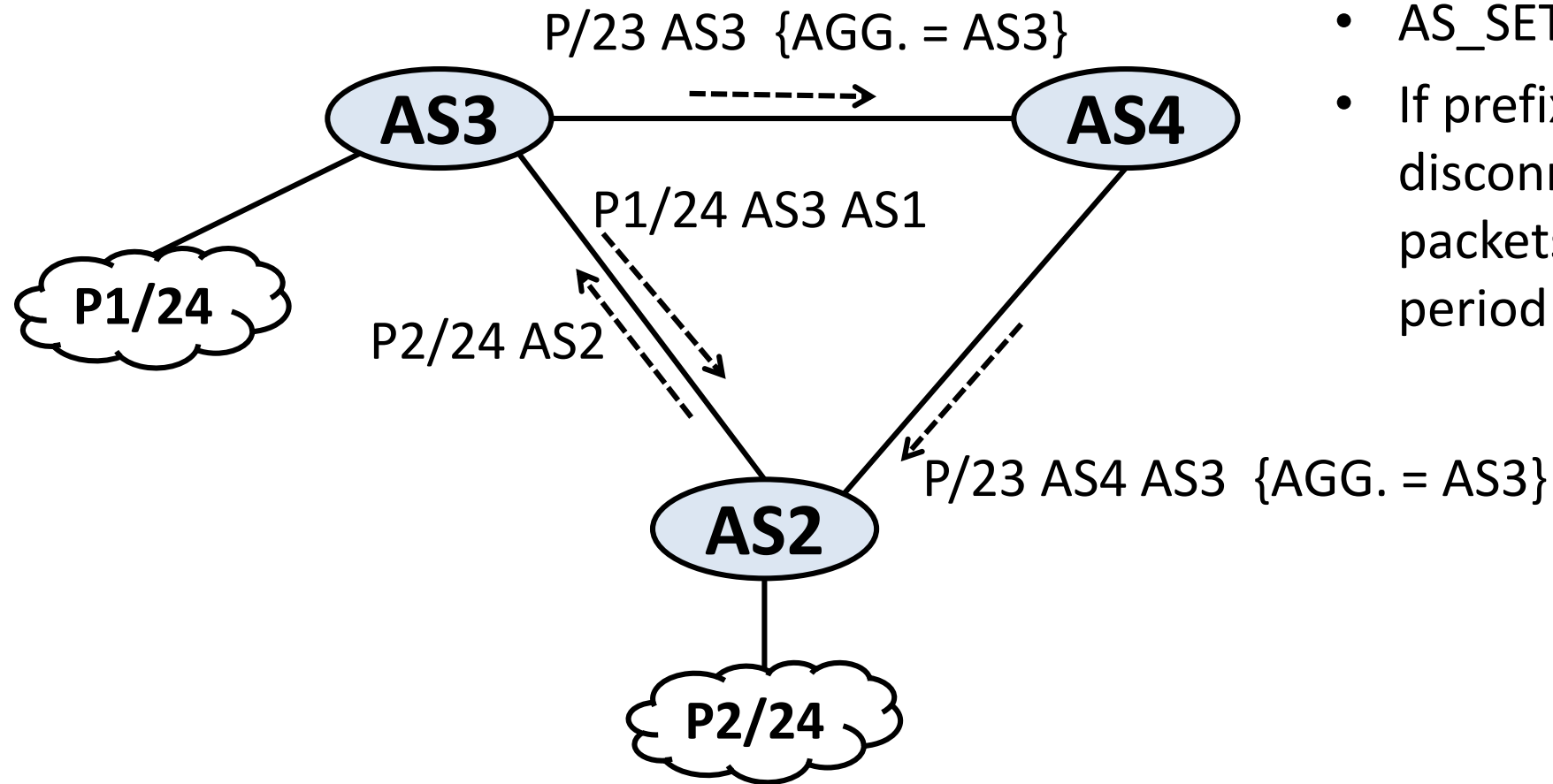
The longest common AS_PATH per the rules is NULL. Putting all additional ASes in a set yields an AS_SET of length one.

Source of some unusual aggregated AS_PATHs (jhaas) (2)

It is possible to alter the code for such cases of a single AS in the AS_SET.

- In such a case, it could be merged into the adjacent AS_SEQ.
- The path length is preserved.
- However, does this properly preserve the origin intent? This may be arguable. “brief” style aggregation discarding the AS_SET may be the right thing to do here.

Common Scenario: AGGREGATOR without AS_SET



- AS_SET not really required
- If prefix P2/24 gets disconnected, then data packets loop for a brief period

$$P/23 = P1/24 + P2/24$$

AGG. = AGGREGATOR

MUST / SHOULD Question

- Conformant BGP speakers MUST NOT locally generate BGP UPDATE messages containing *SET
- Upon receipt of messages with *SET, conformant BGP speakers SHOULD use the "Treat-as-withdraw" error handling behavior
 - SHOULD → MUST ?

Updating RFCs 4271, 5065, 6793 – Level of Detail?

- RFC 4271 has 26 mentions of AS_SET
- RFC 5065 has 11 mentions of AS_CONFED_SET
- RFC 6793 has 1 mention of AS_SET and 10 mentions of AS_CONFED_SET

Strategy for making necessary updates to these RFCs?

RFC 5065: Autonomous System Confederations for BGP

RFC 6793: BGP Support for Four-Octet Autonomous System (AS) Number Space

Alternative path for standardization (jhaas)

- Most (all?) implementations of BGP should be able to support “brief” style aggregation already. No new code need be deployed to change how aggregation works.
 - And no need to intrusively change several RFCs.
 - RFC 6472 already covers this requirement.
- Implementations should be asked to add a policy element that permits AS_SETS to be detected.
 - Having done so, it is possible to implement policy to discard routes having AS_SETs.
 - In the absence of operators cleaning up routes that have sets, RPKI filtering will eventually provide them “incentive” to clean up.

Alternative path for standardization (jhaas) (2)

- An Operational Considerations section for this document should be added that covers the issues with not using sets:
 - The aggregator must supply the more specific contributors to the contributing ASes.
 - The aggregator should not supply the aggregate route to the contributing ASes.
 - ASes that have reachability that is being aggregated should likely reject routes that contain their reachability to prevent forwarding loops.
 - Potentially enshrine the practice of internally advertising a discard route for the destination addresses belonging to one's subnet to prevent in-AS traffic from being sent off-AS. (However, see AS-bridging scenarios.)

Questions / Discussion



November 2019
IDR Working Group – IETF 106 Singapore

draft-ietf-idr-segment-routing-te-policy

Stefano Previdi

Clarence Filsfils, Ketan Talaulikar (Cisco Systems)

Paul Mattes (Microsoft)

Eric Rosen (Juniper)

Dhanendra Jain, Steven Lin (Google)

Overview & Status

- BGP SR Policy SAFI introduced for signaling of SR Policies from controllers to headend(s) via BGP
 - It is the companion document of draft-ietf-spring-segment-routing-policy and covers the BGP protocol encoding part only
- Stable draft with multiple implementations and deployments
 - First individual draft version posted in 2016
 - WG adopted in 2017
 - Now prepare for WGLC ...
- The SR Policy Architecture (draft-ietf-spring-segment-routing-policy) is also stable with multiple implementations and likely to start WGLC soon in SPRING WG

Overview of Updates in v08

- Error Handling aspects consolidated
- IANA Considerations updated
- Security Considerations added
- Updated text to reflect the discussion on use of RT on the list
- Align with Segment Types updated in SR Policy draft to avoid mix-up with codepoints
- Updated codepoints allocated via Early Allocation process
- Minor editorial updates

Error Handling

- Consolidated all Error Handling procedures in a new Sec 5 that covers these aspects
- Treat-as-withdraw for all errors related to TLVs/sub-TLVs and procedures defined in this document
- AFI/SAFI disable or session-reset for errors when the NLRIs cannot be parsed
- Clarified that Semantic Validation of information in TLVs/sub-TLVs that are consumed by SR Policy processing (i.e. not used by BGP) is not performed by BGP.

IANA Consideration

- Add request for new Registry for Color Extended Community Field for the 2 CO (color-only) bits defined in this document
- Changed allocation policy to Specification Required since the flags and some other fields are very small space for FCFS policy
- Added guidance for DE

Next Step – Preparing for WGLC

- Request WG review & inputs
- Update implementation report on IDR wiki



November 2019
IDR Working Group – IETF 106 Singapore

draft-ietf-idr-bgp-ls-app-specific-attr

Ketan Talaulikar, Peter Psenak (Cisco Systems)
Jeff Tantsura (Apstra)

Overview

- BGP LS extension for IGP drafts:
 - ietf-isis-te-app
 - ietf-ospf-te-link-attr-reuse
- Allows link attributes to be signaled on a per application basis
- Applications:
 - RSVP TE, SRTE, LFA, Flex-Algo, ...
- ASLA (application Specific Link Attributes) TLV defined in both IGPs

Progression Updates

- First presented at IDR WG Interim before IETF 103 Bangkok
- Adopted as WG document after IETF 104 Prague
- Codepoints allocated via Early Allocation process

Summary of Updates in v01

- Aligned with the changes to underlying IGP Specifications
- Removed Max Reservable BW and Unreserved BW attributes since they are RSVP-TE specific
- Clarified text on processing of ISIS ASLA sub-TLV
- Introduced text to clarify scenarios where existing deployments of SRTE can continue using existing BGP-LS top-level TLVs originally introduced for RSVP-TE
- Simplified the backward compatibility considerations in BGP-LS by leveraging the backward compatibility aspects of the underlying IGPs

Next Step

- Request WG review and inputs
- IGP drafts are post WGLC status and this BGP-LS document can progress to WGLC once we have implementations

BGP Flexible Color-Based Tunnel Selection

draft-shen-idr-flexible-color-tunnel-selection-00

Yimin Shen (yshen@juniper.net)

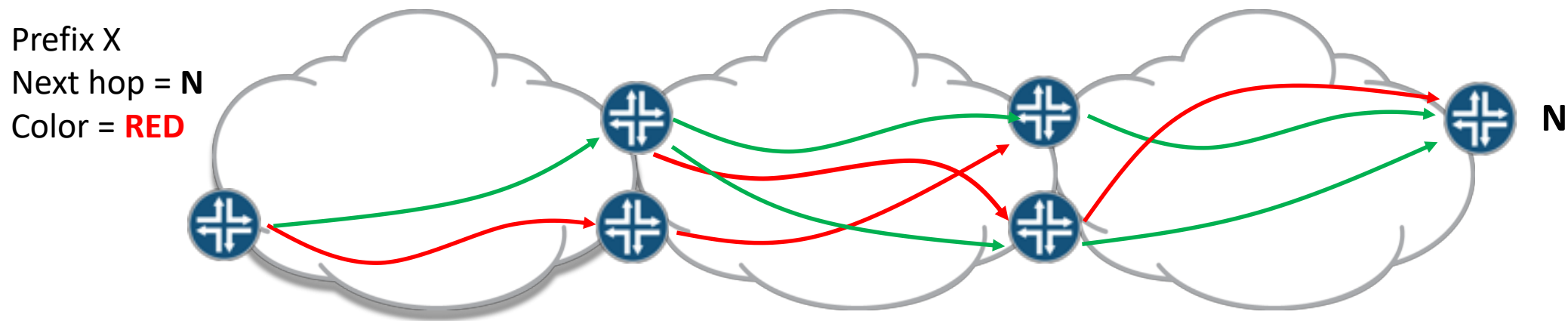
Ravi Singh (ravi.singh.ietf@gmail.com)

IETF 106, Singapore

Nov 2019

Color as Criterion in Tunnel Selection

- Payload prefix with BGP next hop **N** and color **RED** → **RED** tunnel to **N**
- Use cases: inter-domain TE, SR, etc.



Color-Based Tunnel Selection

- Color provides a generic notion for network attributes:
 - TE characteristic
 - Virtual topology
 - Network slice
 - Path computation algorithm
- Color Extended Community can be signaled across AS, and from controller to routers.
- A generic service mapping mechanism based on TE, slicing, virtualization, flex-algo, etc., suitable for inter-domain and controller-driven environments.

Extended Mapping Modes and Flexible Selection Scheme

Extended mapping modes

- IP-color, with optional fallback colors
- Color-only, with optional fallback colors
- IP-any-color
- IP-only
- Converted-IPv6
- Converted-IPv6-color, with optional fallback colors
- Converted-IPv6-any-color
- Color-profile

Flexible selection scheme

- Consists of a sequence of extended mapping modes
- Falls back from one mode to next mode in the specified order

Example: payload prefix with nexthop **N** and color **RED**.

Flexible selection scheme:

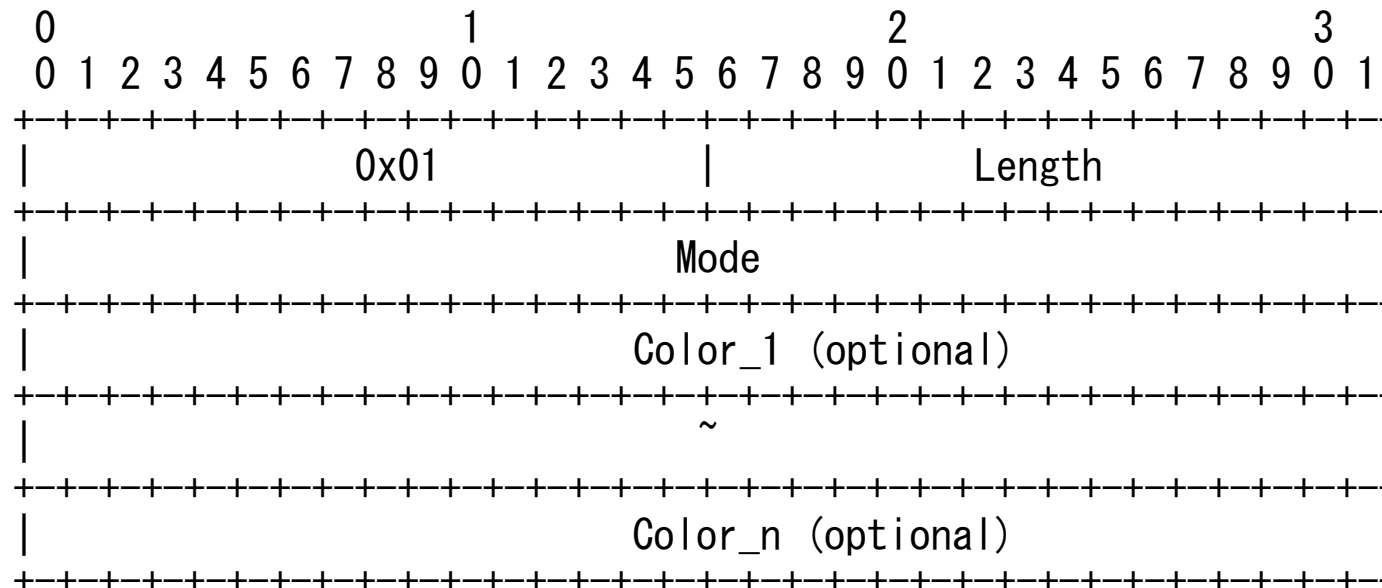
- (1) IP-color, fallback = {**BLUE**, **GREEN**}
 - Attempt **RED** tunnel to **N**
 - Fall back to **BLUE** tunnel to **N**
 - Fall back to **GREEN** tunnel to **N**
- (2) Converted-IPv6-color, fallback = {**BLUE**}
 - Attempt **RED** tunnel to **N'**
 - Fall back to **BLUE** tunnel to **N'**
- (3) IP-only
 - Attempt uncolored tunnel to **N**

Provisioning Modes of Flexible Selection Scheme

- Ingress router
 - Configure as a policy
 - Apply to payload prefixes
- Egress router
 - Signal via a new **Flexible Color Tunnel Selection** Path Attribute to ingress routers
 - Apply to payload prefixes on each ingress router
- Controller
 - Signal via a new **Flexible Color Tunnel Selection** Path Attribute to ingress routers
 - Apply to payload prefixes on each ingress router

BGP Flexible Color Tunnel Selection Path Attribute

- Carries the information of a tunnel selection scheme.
- Comprises a sequence (in the fallback order) of Extended Mapping Mode TLVs.



Relationship with Color-Only Bits of Color Extended Community

- *draft-filsfils-spring-segment-routing-policy* and *draft-previdi-idr-segment-routing-te-policy* defines two “Color-Only” bits in the Color Extended Community, and three fallback modes.
- This draft supports user-defined flexible selection schemes
 - Fully supports the fallback modes pre-defined by the CO bits
- If flexible selection scheme and the CO bits co-exist:
 - Local policy > received CO bits > received Flexible Color Tunnel Path Attribute

Next Steps

- Request review
- Welcome comments and suggestions

Destination-IP-Origin-AS Filter for BGP Flow Specification

draft-wang-idr-flowspec-dip-origin-as-filter

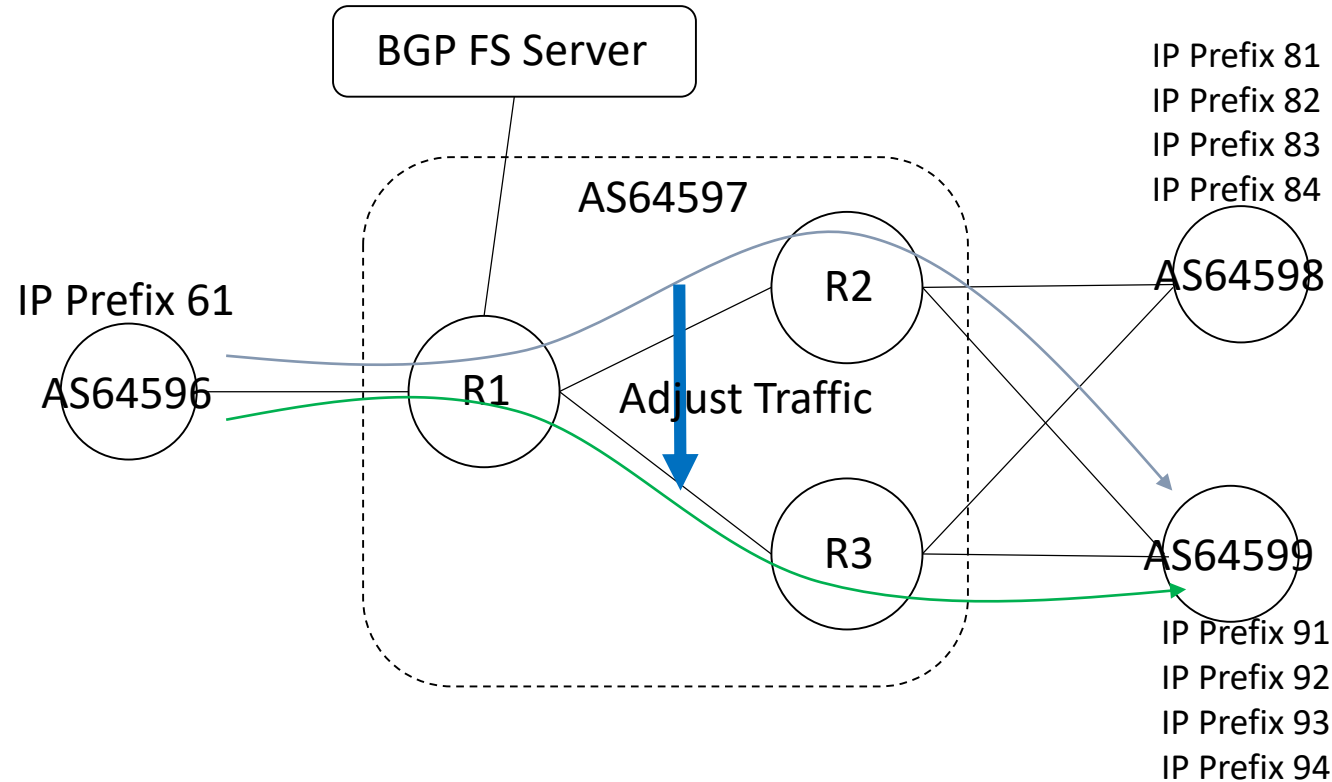
IETF 106 Singapore

Wang Haibo, Huawei

Wang Aijun, China Telecom

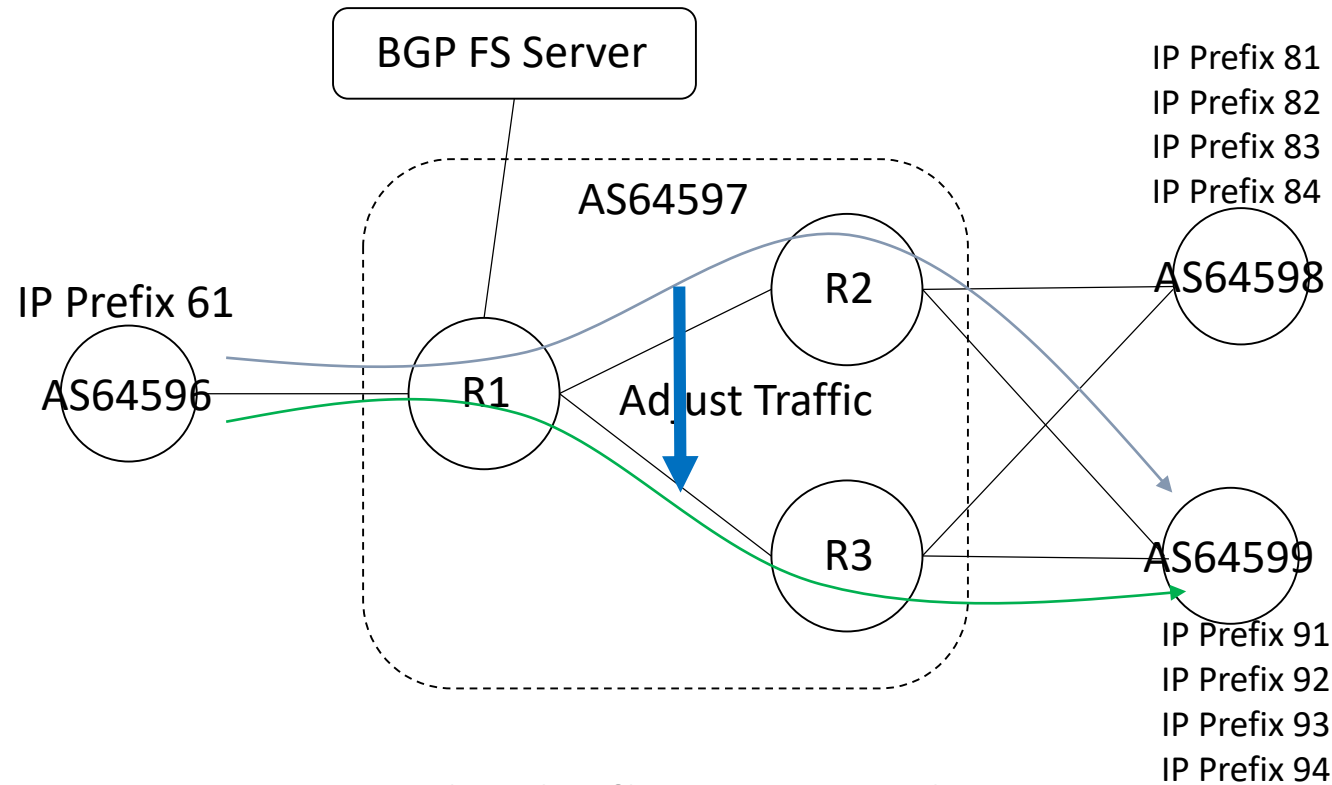
Zhuang Shunwan, Huawei

Why need Dest-IP-Origin-AS rule



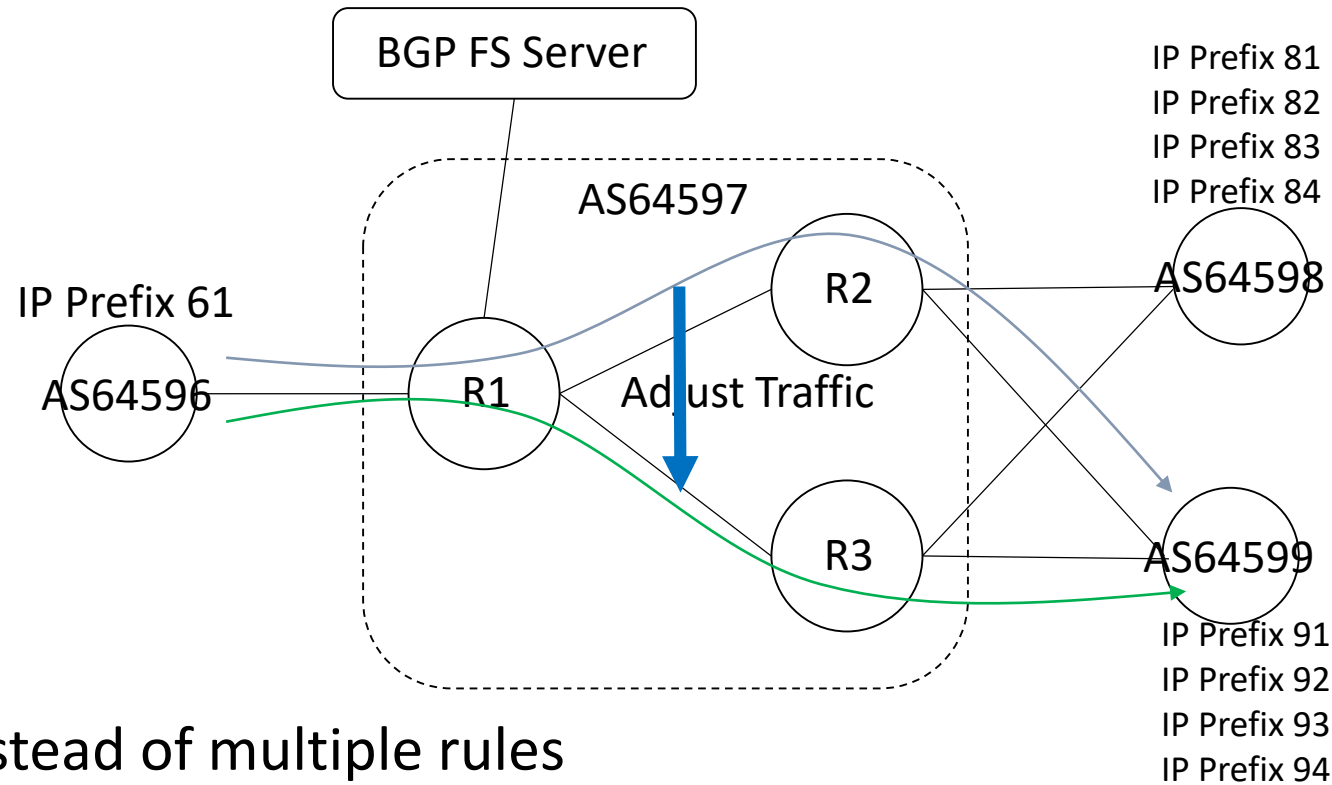
- Purpose: Adjust traffic(AS64596 IP61 to AS64599), from R2 to R3
 - Requirement received from some operators

Current method



- Current: FS server send multiple flowspec rules to R1
 - Rule 1: SIP IP61, DIP IP91, Action: redirect to R3
 - Rule 2: SIP IP61, DIP IP92, Action: redirect to R3
 - ... and continue for all IP belong to AS64599

Proposed method



- One rule instead of multiple rules
 - Rule : SIP IP61, DAS AS64599, Action: redirect to R3
- Benefit:
 - Easier for operation
 - Reduce messages from server to router
 - May reduce numbers of flowspec table entries (based on device implementation)

New component type for DAS

- Add new component for NLRI:
 - Type TBD1 - Destination-IP-Origin-AS
 - Encoding: <type (1 octet), [op, value]+>
 - Op: 0 1 2 3 4 5 6 7
 - +---+---+---+---+---+---+---+---+
 - | e | a | len | 0 | lt | gt | eq |
 - +---+---+---+---+---+---+---+---+
 - Value : four octets ASN
- Suggested method:
 - Lookup FIB to get the DIP's Origin AS(called DAS)
 - Use the DAS to match this rule
- Another option:
 - Expand DAS to multiple DIP based rules

Application and evolution

- Combine with existing flowspec components for traffic matching
- More new types could be introduced
 - Source IP origin AS
 - Destination IP Community
 - Source IP Community
 - and others...

Comments received

- There was some discussion about this usage:
 - <https://mailarchive.ietf.org/arch/msg/idr/x6xoTdJ9vBVDaqh2dv5NshMJ8Go>
- Security Considerations
 - Capability negotiation for directly connection between routers and flowspec servers
 - May use BMP to collect the routers capability, whether support this component type
 - May use node target to control which router to use(draft-dong-idr-node-target-ext-comm-01)
- Whether use new SAFI for this component type
 - The new component type also provides a match criteria under flowspec architecture
 - The new component type may be used together with existing component types

Next step

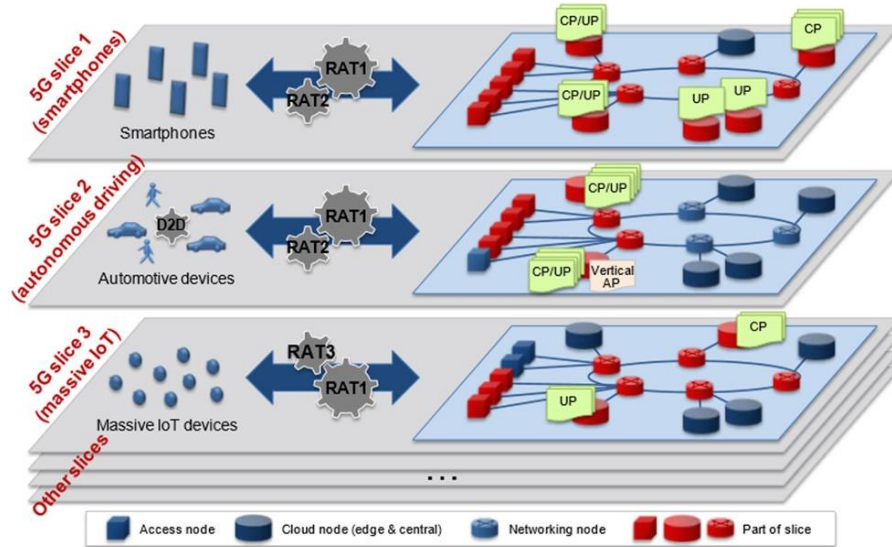
- Keep on updating and welcome comments
- Thank you😊

draft-chan-bgp-lu2

Color Operation with BGP-LU (Slicing by color)

Louis Chan
Juniper Network
louisc@juniper.net
Nov 2019

5G/Metro Network Slicing



- New proposal via color
- Mix of RSVP, SR, LDP
- Controller-less
- Transport service across domain
- Very deterministic path possible
 - e.g. RSVP-RSVP-RSVP



Issues: SR-TE E2E solution with controller

- How to scale the controller, if network size > 10K
 - Up to 300K routers in size
- Remote failure detection
- Label depth at first ingress router
 - Cannot avoid Binding-SID (or alike) in the middle of the network
 - But then, problem is the node failure of the router holding B-SID

Two drafts submitted

- BGP-LCU
 - <https://tools.ietf.org/html/draft-szarecki-idr-bgp-lcu-traffic-steering-00>
 - New SAFI to include color in NLRI
 - New length field
- BGP-LU2
 - https://datatracker.ietf.org/doc/draft-chan-idr-bgp-lu2/?include_text=1
 - Use of color extended community for slicing
 - Conserve the original NLRI and hence the length field
 - Reuse RFC8277 (BGP-LU) and extend the support for color
 - Need new BGP capability code

Motivation

- Multiple Transport Service by **Slicing BGP** by CO₂LR
 - Low latency, diversity, BE....
- No need to run SR in all domains – RSVP, LDP, MPLSoUDP compatible
- Predictable label depth
- Remote Failure Signaling
- Scalable to >100k routers in size
- **Controller-less!** But the architecture is still SR controller compatible
- Design for BGP-LU backward compatibility

BGP-LU and BGP-LU2

BGP-LU

Prefix	Label	- OR -
--------	-------	--------

Prefix	Label		Label
--------	-------	---	-------


Single label

Multiple labels

BGP-LU2

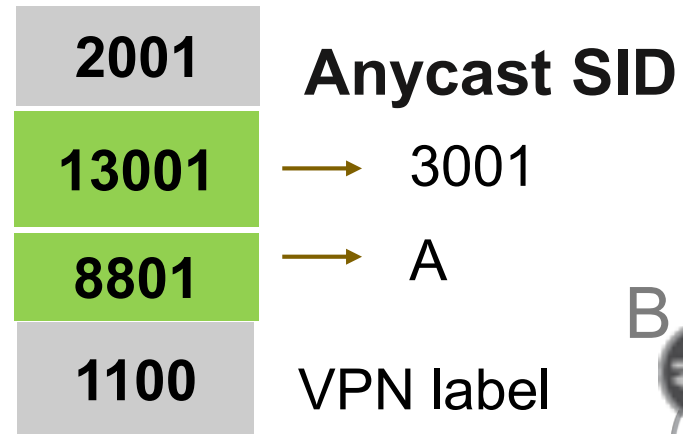
Prefix	Label	- OR -
--------	-------	--------

Prefix	Label		Label
--------	-------	--	-------

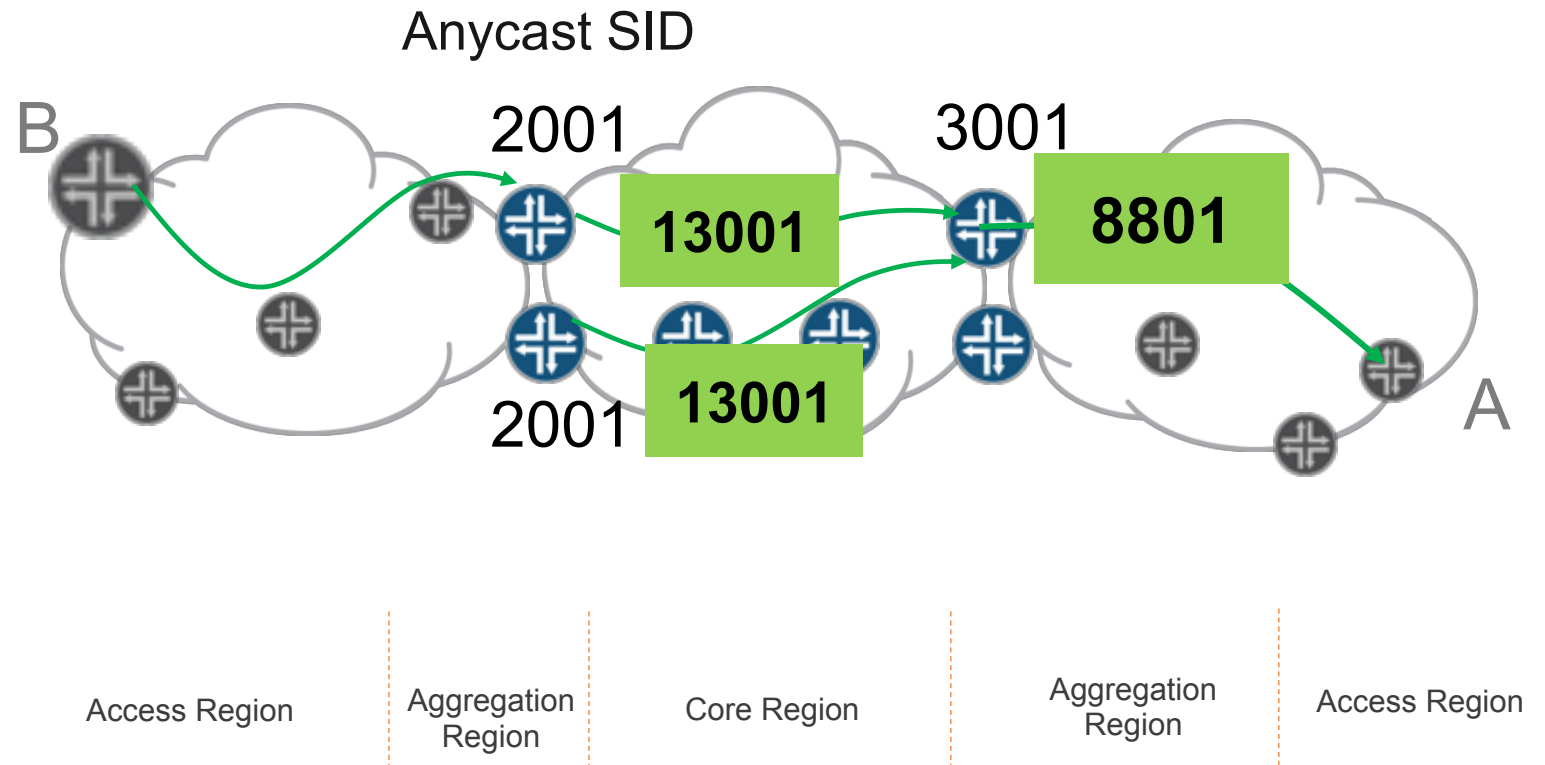
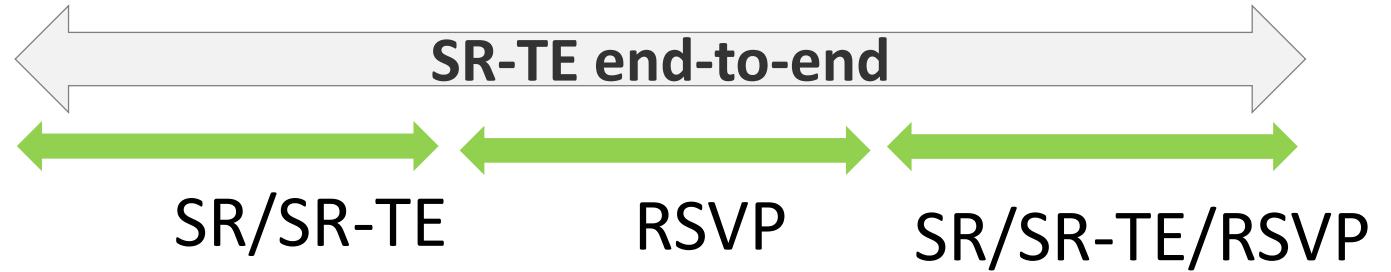
Prefix	Label		Label
--------	-------	---	-------

Prefix	Label		Label
--------	-------	---	-------

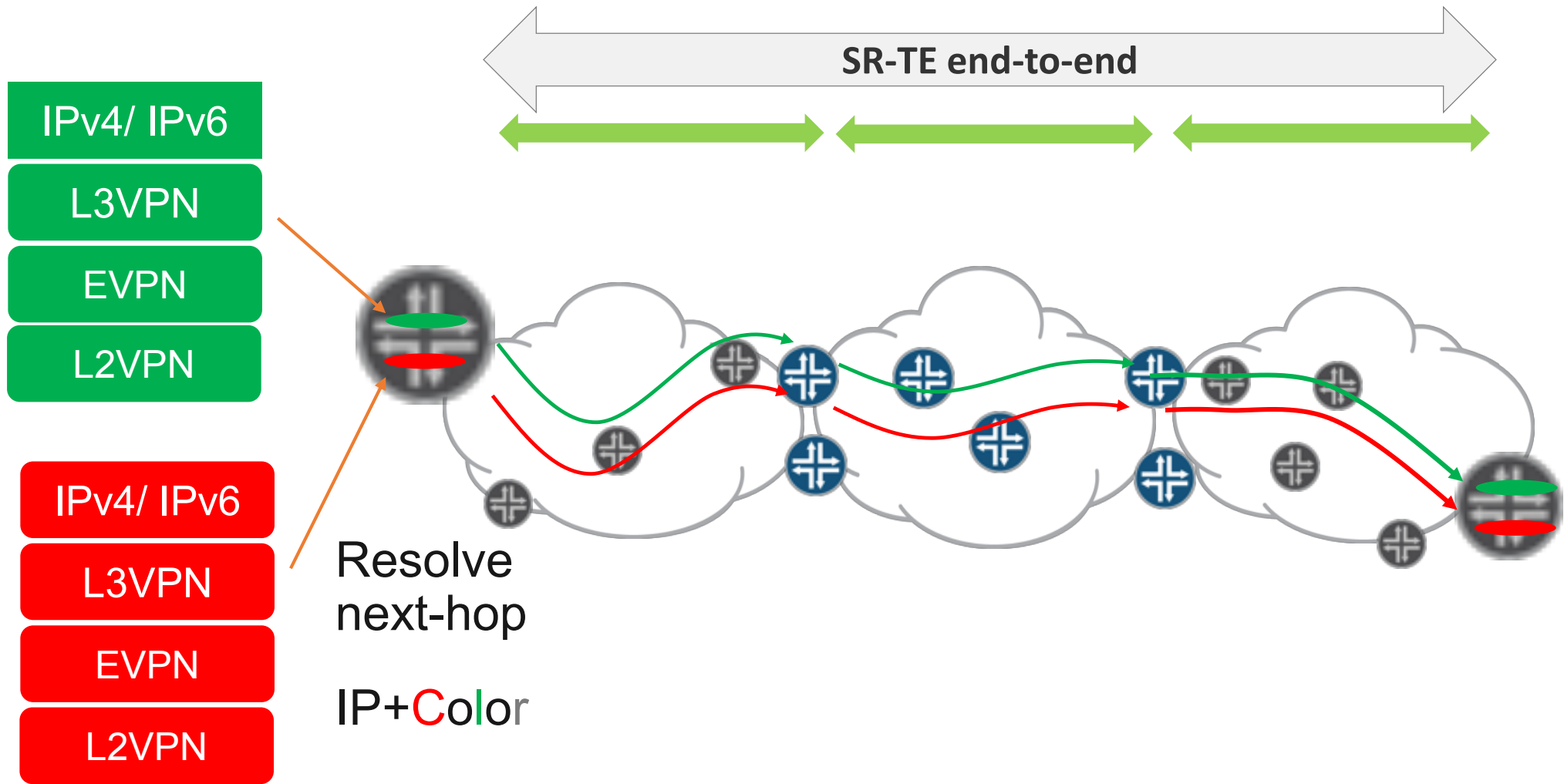
E2E – A BGP proposal (aka LU2)



Similar to
BGP-LU
BUT with
COLOR



Service Mapping Method



Based on “color” attribute embedded

Choices of color encoding

- Use of Path Attribute for color – as in the current draft
 - Save byte count for the same advertisement
 - Not normal BGP operation
- Use of new Attribute to prepend to NLRI
 - Just like add path
 - Prepend 4+2 byte (color + 2 reserved byte)
 - Repeat the same info
 - Additional 6KB for 1000 prefixes

Alternatives: Prepend Color info

Similar to Add-path

