

BGP Classful Transport Planes

<https://tools.ietf.org/html/draft-kaliraj-idr-bgp-classful-transport-planes>

IETF 108

Kaliraj Vairavakkalai

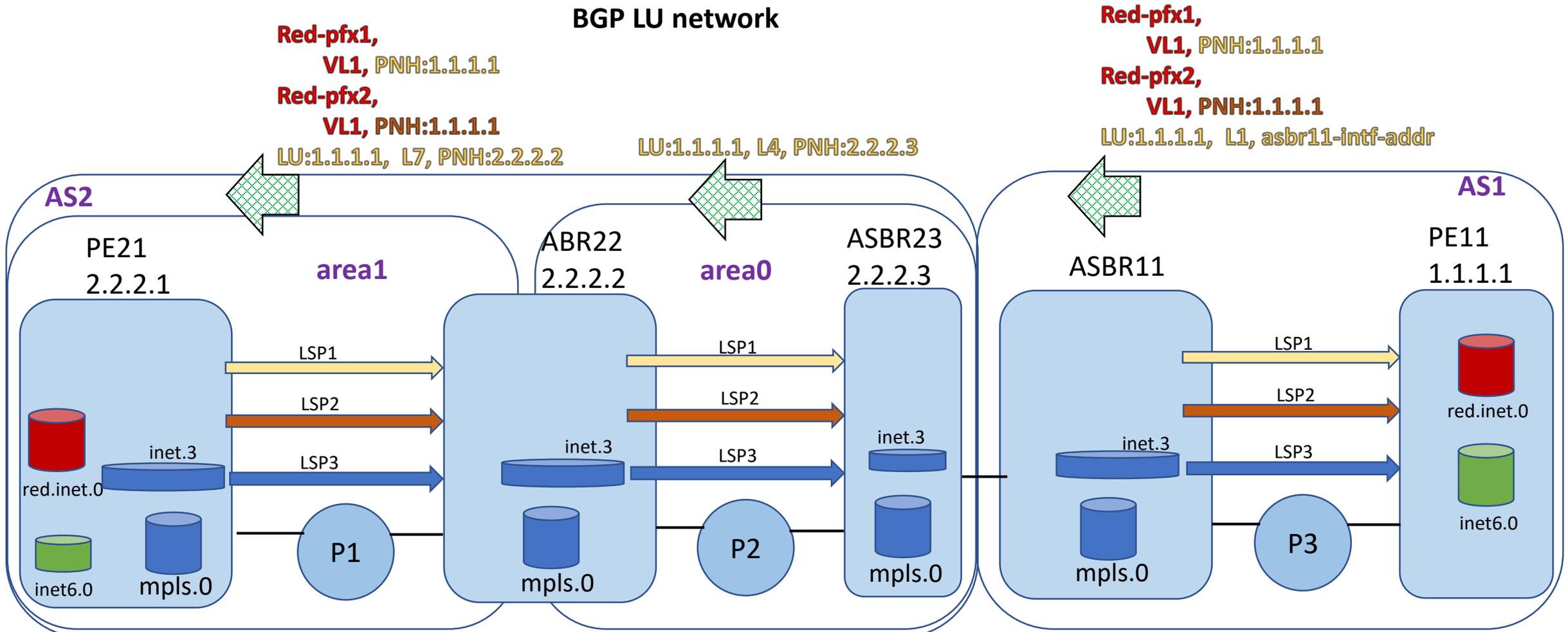
Juniper Networks

July 29, 2020

Problem

- A domain has intra-AS tunnels with varying TE characteristics (gold, silver, bronze).
- There could be multiple tunnels to the same destination. And different tunneling protocols creating those tunnels.
- These tunnels may need to be extended inter-domain, while preserving their TE characteristics end-to-end.
- Different Service routes want to resolve (put traffic) over intra/inter-domain tunnels of a certain TE characteristic, with an option to fallback on tunnels belonging to a different TE characteristic.
- How to extend BGP to signal these pieces of information, and get the job done.

BGP LU network



Red-pfx1,
VL1, PNH:1.1.1.1
Red-pfx2,
VL1, PNH:1.1.1.1
LU:1.1.1.1, L7, PNH:2.2.2.2

LU:1.1.1.1, L4, PNH:2.2.2.3

Red-pfx1,
VL1, PNH:1.1.1.1
Red-pfx2,
VL1, PNH:1.1.1.1
LU:1.1.1.1, L1, asbr11-intf-addr

PE21: red.inet.0
Pfx1 -> Push VL1, L7, LSP3
Pfx2 -> Push VL1, L7, LSP3
PE21:inet6.0
Pfx3 -> Push 2, L7, LSP3

ABR22: mpls.0
L7 -> Swap L4, Push LSP3

ASBR23: mpls.0
L4 -> Swap L1, asbr11-Intf

ASBR11: mpls.0
L1 -> Pop, Push LSP1 Labels

How? the constructs of BGP CT

- A domain has intra-AS tunnels with varying TE characteristics (Transport Class: gold, silver)
- Use “Transport Class Route Target” to signal transport class in BGP.
- There could be multiple tunnels to the same destination. Use “Route Distinguisher” to advertise them without path-hiding, and allow identifying originating PE.
- The tunnel may need to be extended inter-domain, while preserving the same Transport class end-to-end. Resolve BGP NH using tunnels belonging to the same Transport class. And follow RFC-4364 option-C style procedures, to create swap-routes on domain boundaries.
- New BGP transport layer address-family (SAFI: 76, “Classful Transport”) that follow RFC-4364 procedures.
- Service routes want to resolve using a Resolution scheme (*viz. use tunnels of a certain Transport class, with an option to fallback on other Transport classes*).
- Desired Resolution scheme is signaled via “Mapping community” which can be a function of transport-class.

BGP CT network

Red-pfx1, Comm-Gold,

VL1, PNH:1.1.1.1

Red-pfx2, Comm-Bronze,

VL1, PNH:1.1.1.1

RD1:1.1.1.1, RT-Gold,

L7, PNH:2.2.2.2

RD2:1.1.1.1, RT-Bronze,

L8, PNH:2.2.2.2

RD1:1.1.1.1, RT-Gold,

L4, PNH:2.2.2.3

RD2:1.1.1.1, RT-Bronze,

L5, PNH:2.2.2.3

Red-pfx1, Comm-Gold,

VL1, PNH:1.1.1.1

Red-pfx2, Comm-Bronze,

VL1, PNH:1.1.1.1

RD1:1.1.1.1, RT-Gold,

L1, asbr11-intf-addr

RD2:1.1.1.1, RT-Bronze,

L2, asbr11-intf-addr

AS2

PE21
2.2.2.1

area1

ABR22
2.2.2.2

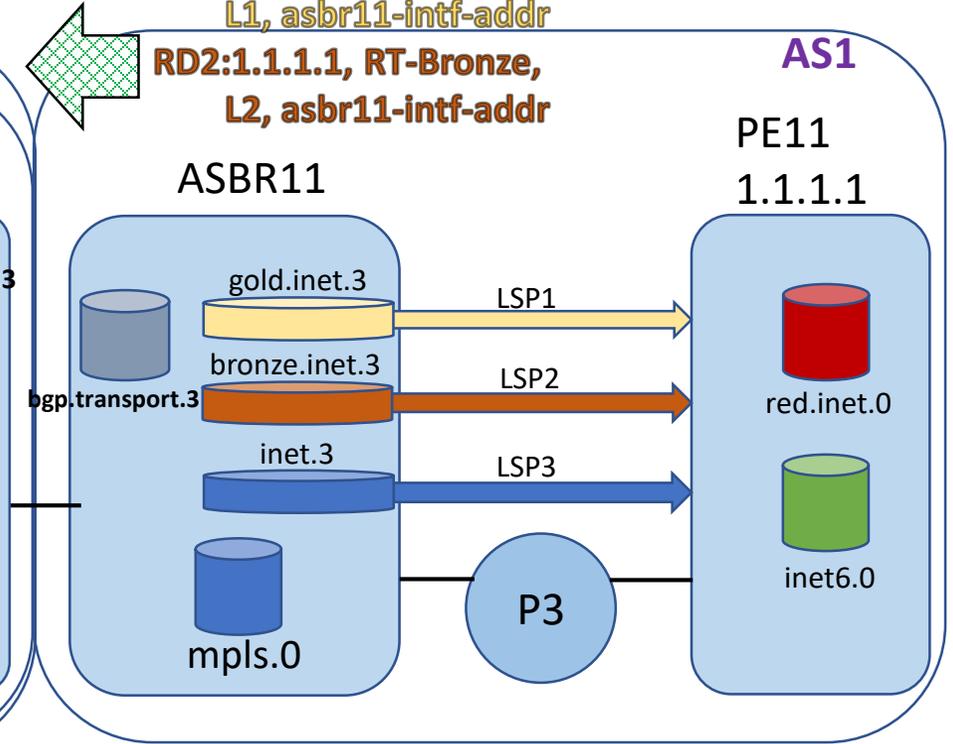
area0

ASBR23
2.2.2.3

AS1

PE11
1.1.1.1

ASBR11



PE21: red.inet.0

Pfx1 -> Push **VL1, L7, LSP1**

Pfx2 -> Push **VL1, L8, LSP2**

PE21:inet6.0

Pfx3 -> Push **2, L7, LSP1**

ABR22: mpls.0

L7 -> Swap **L4, Push LSP1**

L8 -> Swap **L5, Push LSP2**

L9 -> Swap **L6, Push LSP3**

ASBR23: mpls.0

L4 -> Swap **L1, asbr11-Intf**

L5 -> Swap **L2, asbr11-Intf**

L6 -> Swap **L3, asbr11-Intf**

ASBR11: mpls.0

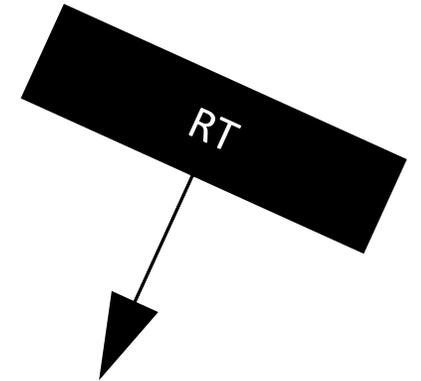
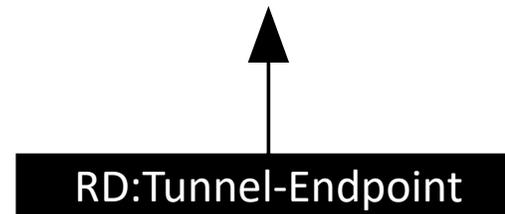
L1 -> Pop, Push **LSP1 Labels**

L2 -> Pop, Push **LSP2 Labels**

L3 -> Pop, Push **LSP3 Labels**

BGP CT – pcap sneak peak

```
Nov 10 22:00:51.708561 BGP SEND 13.21.0.13+65494 -> 13.21.0.21+179
Nov 10 22:00:51.708563 BGP SEND message type 2 (Update) length 98
Nov 10 22:00:51.708572 BGP SEND Update PDU length 98
Nov 10 22:00:51.708574 BGP SEND flags 0x40 code Origin(1): IGP
Nov 10 22:00:51.708580 BGP SEND flags 0x40 code ASPath(2) length 6: 1
Nov 10 22:00:51.708581 BGP SEND flags 0x80 code MultiExitDisc(4): 30
Nov 10 22:00:51.708596 BGP SEND flags 0xc0 code Extended Communities(16): transport-  
target:0:100
Nov 10 22:00:51.708605 BGP SEND flags 0x90 code MP_reach(14): AFI/SAFI 1/76
Nov 10 22:00:51.708611 BGP SEND      nhop 13.21.0.13 len 12
Nov 10 22:00:51.708631 BGP SEND      1.1.1.3:9:1.1.1.1/32 (label 299952)
```



Advantages

- Keep the heterogenous tunneling-domains (RSVP, SRTE, FlexAlgo, etc) loosely coupled and still preserve Transport-class end to end.
- Natural extension to BGP-LU RFC-4364 option-C deployments.
- Reuse of proven BGP-VPN technology at Transport layer.
 - `RD:TunnelEndpoint` takes care of path-hiding.
 - Transport class `RouteTarget` treats “Color” as an attribute (adjective), rather than part of NLRI (noun). Which is more appropriate.
- On-demand-NH comes for free, with RTC (RFC-4684) mechanisms for BGP-CT family
- New Route-target type avoids collision with existing service-routes RT namespace.
- Opens up new possibilities by extending applicability of time-tested RFC-4364 mechanisms at a new (transport) layer.

Why new address-family?

Why not re-use/hack existing families like LU, SRTE or L3VPN?

- Carrying 'Color' as attribute (RT) makes more sense, instead of in the NLRI.
- RD is the right distinguisher, end-to-end. Add-path-ID is per-session scope. Both are required, either one is not enough by itself.
- Use of RT allows for RTC like mechanisms, and the ODN. If we didn't use well-known RT ext-comm for route-leaking, this is not possible.
- Further overloading L3VPN (service family) with transport-routes is not good. As route-propagation path is different for service vs transport routes.
- Thus, new SAFI 76. A Transport family that can signal transport classes.