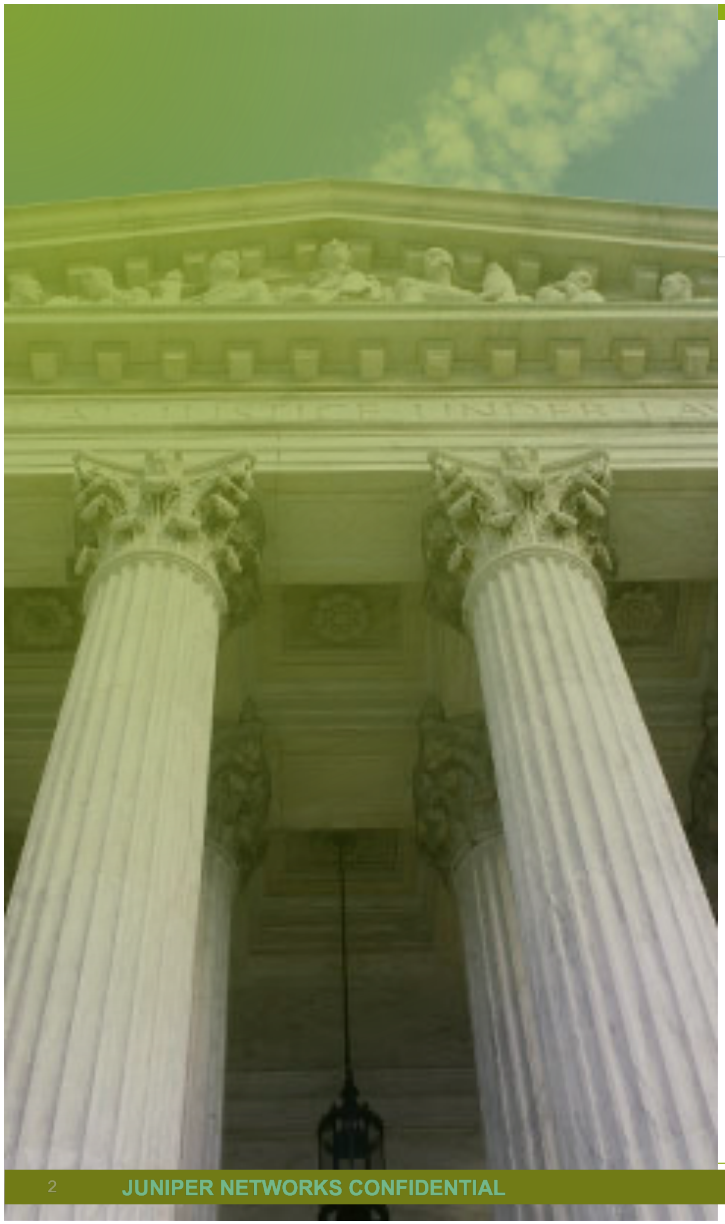
SEGMENT ROUTING

Shraddha Hegde

DENOG-13

Nov 7th 2021

JUNIPER NETWORKS | Engineering Simplicity



LEGAL STATEMENT

This statement of direction sets forth Juniper Networks' current intention and is subject to change at any time without notice.

No purchases are contingent upon Juniper Networks delivering any feature or functionality depicted in this presentation.

AGENDA

- Segment Routing Basics
  - SIDs and their uses
- Migrating to segment Routing
- Lab session

- Traffic engineering with Flex-algo
  - Flex-algo use cases
  - Flex-algo FAD advertisements
  - Flex-algo and TI-LFA
- Lab Session

- Seamless SR
- Lab session

SEGMENT ROUTING BASICS

WHY SEGMENT ROUTING?

- Reduces state on transit devices
- Reduces number of protocols on transit devices
- Native ECMP capabilities
- Simplifies Operations
- Programmable Architecture

SEGMENT ROUTING DATA PLANE

- SR-MPLS
  - No change to MPLS data plane
  - SIDs are represented with 20 bit labels
  - Explicit pathing with label stacking
  - Applicable to IPv4 and IPv6 networks
- SRv6
  - Applicable to IPv6 networks
  - New extension header SRH
  - SIDs 128 bit length
  - Hardware changes to support SRH processing

SEGMENT IDENTIFIERS

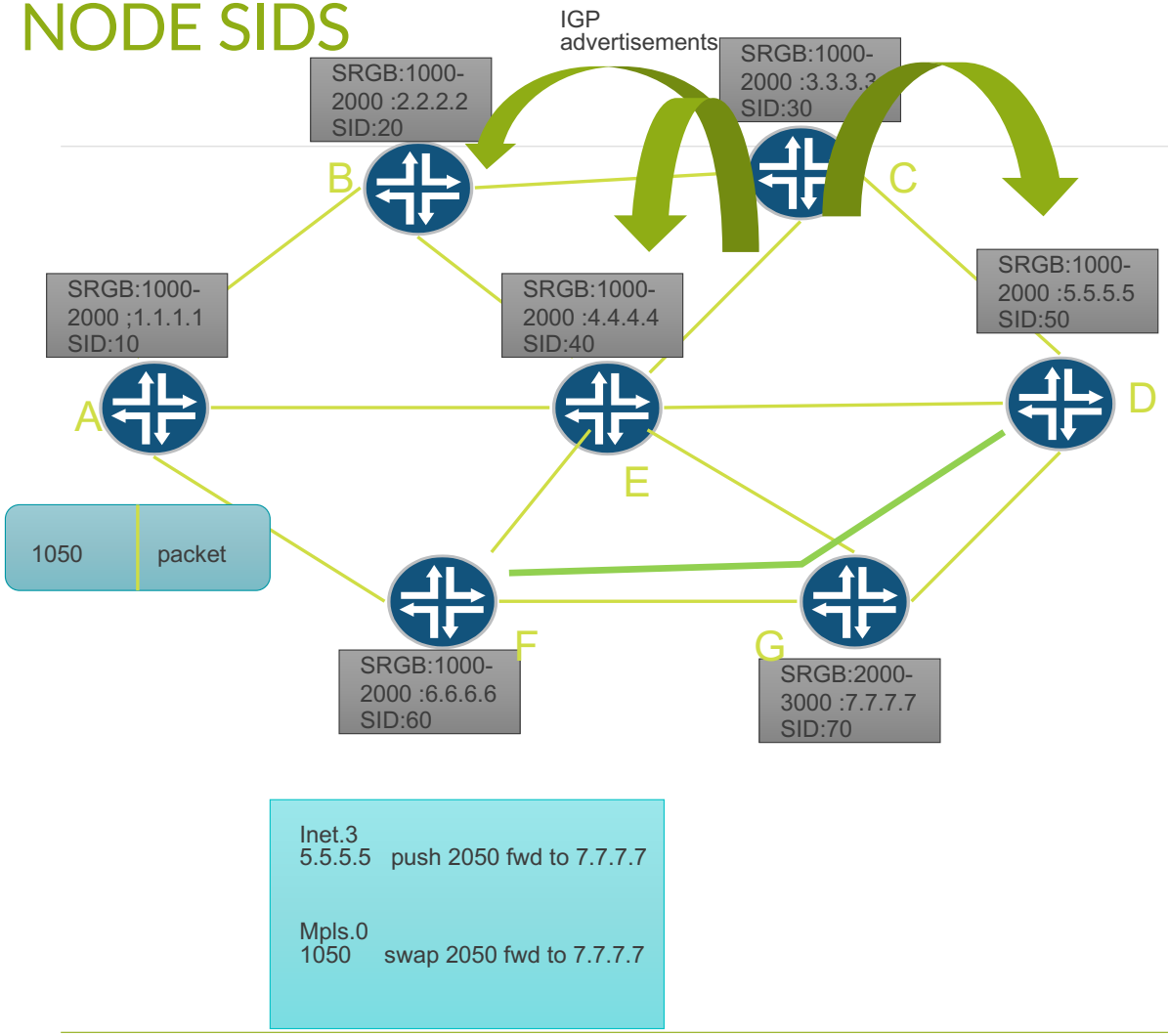
- Node-SIDs/Prefix-SIDs
  - Provide shortest path (like LDP)
- Adjacency SID/ Adjacency set SID
  - Forward along the link/ set of links
- Anycast SIDs
  - Similar to Prefix-SIDs allocated on more than one Node
- Binding-SIDs
  - forward along a particular path
- Mirror-SID
  - Used for Egress-protection

SRGB AND SRLB

- SRGB
  - Global label block allocated on every node
  - Has global significance
  - May be same or different on every node
  - Index allocated from the SRGB space
  - Labels derived on every node
  - Used by Node-SIDs/Prefix-SIDs/Anycast SIDs

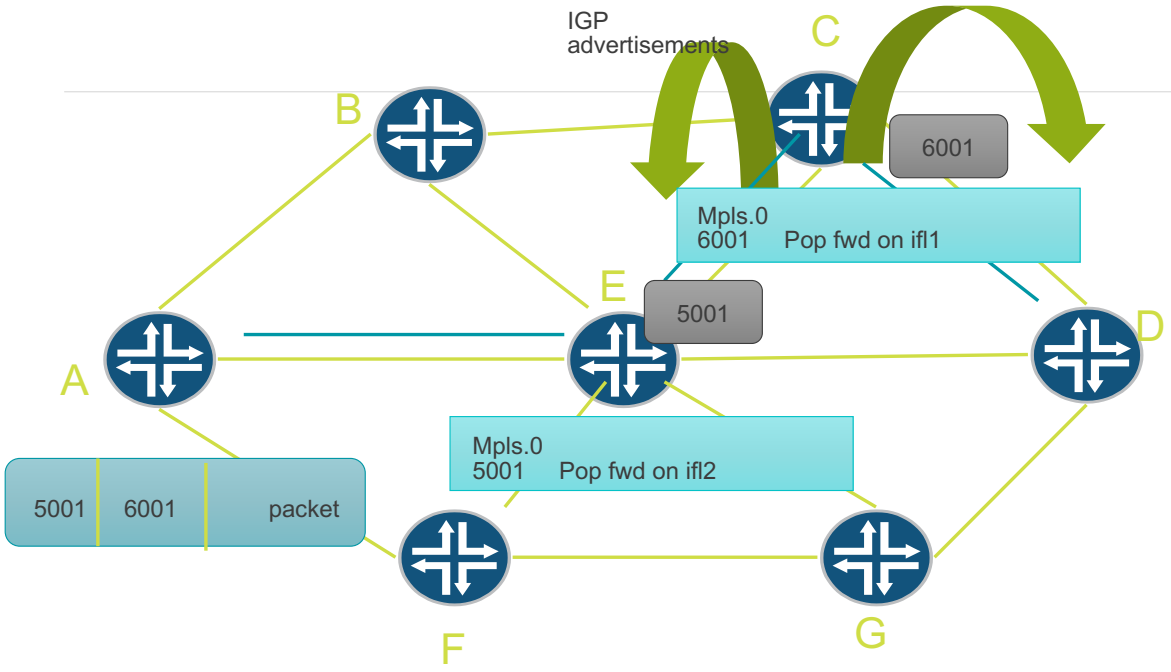
- SRLB
  - Local label block
  - No global significance
  - Static labels get allocated from this block

NODE SIDS



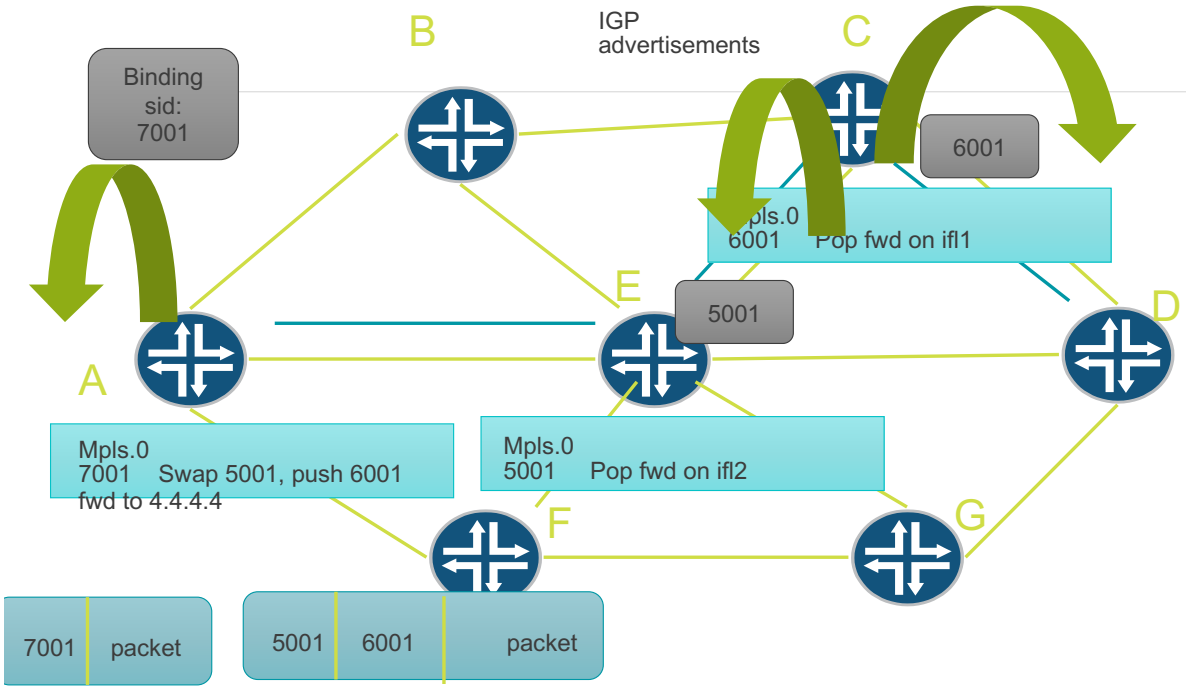
- The SRGB and node-sids are advertised in IGP
- IGP computes the nexthop for the node based on SPF
- Incoming labels are derived from Node's own SRGB and the destination's SID
- Outgoing labels are derived from the nexthop's SRGB and the destination SID

ADJACENCY SIDS



- Dynamic labels vs static labels
- POP label and forward on specific interface
- Adjacency hold timers

BINDING-SIDS



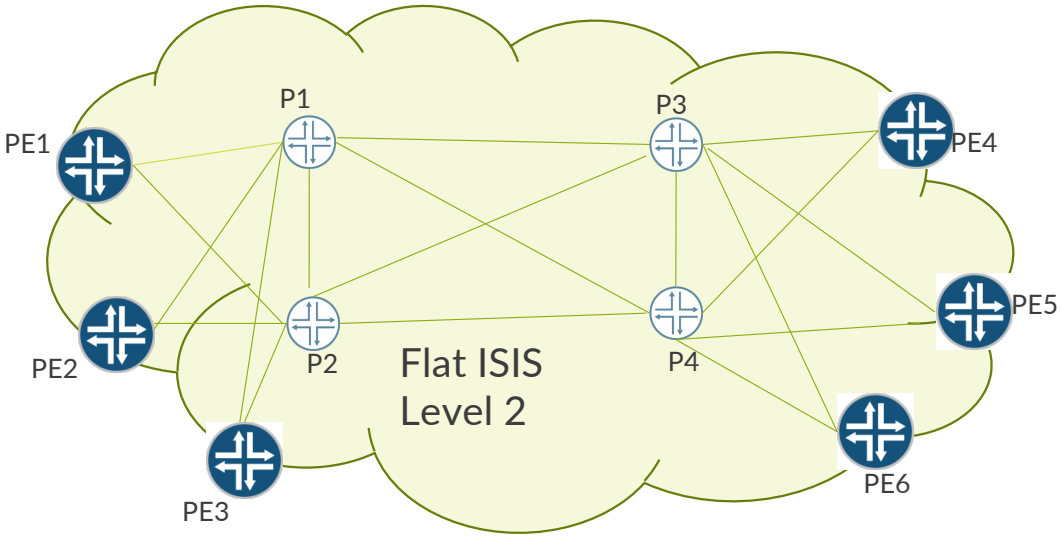
- They represent a path , instead of interfaces
- When a packet arrives with binding-label on top of stack, it gets steered into the tunnel represented by the label-stack
- Binding-sids could be dynamically assigned by a node or get statically assigned from controller.

MIGRATIONS

MIGRATIONS

- Decide migration usecase
- Decide the size of the SRGB/SRLB
- Decide Traffic migration mechanisms
- Decide protection mechanisms
- Traffic engineering options
  - Flex-algo/SR-TE

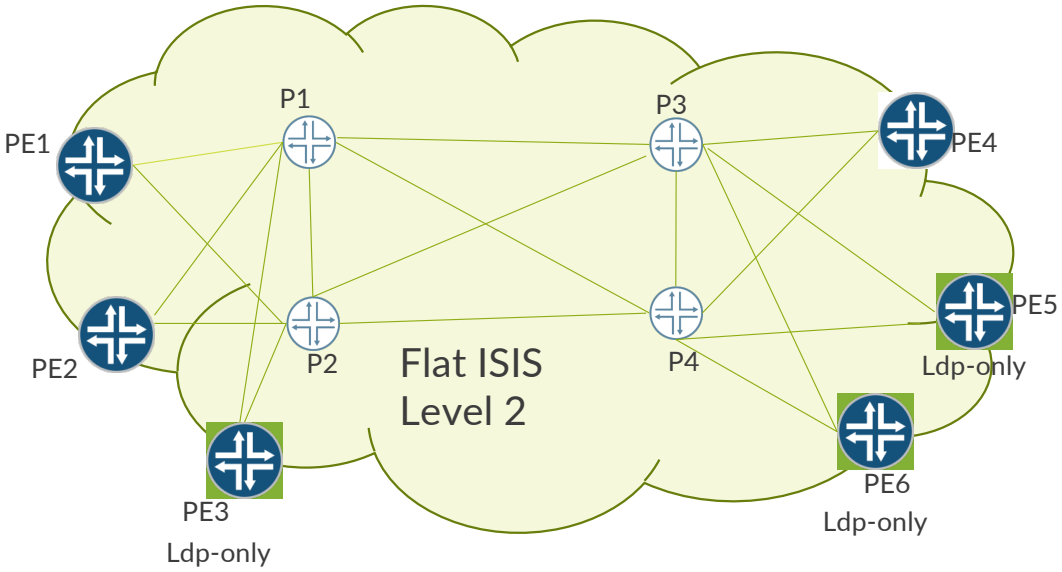
LDP TO SR MIGRATION



- Upgrade all nodes
- Configure SRGB
- Enable SR
- Verify all paths

- Shifting traffic
  - Ships in the night
    - Separate loopback for SR
  - Preference based

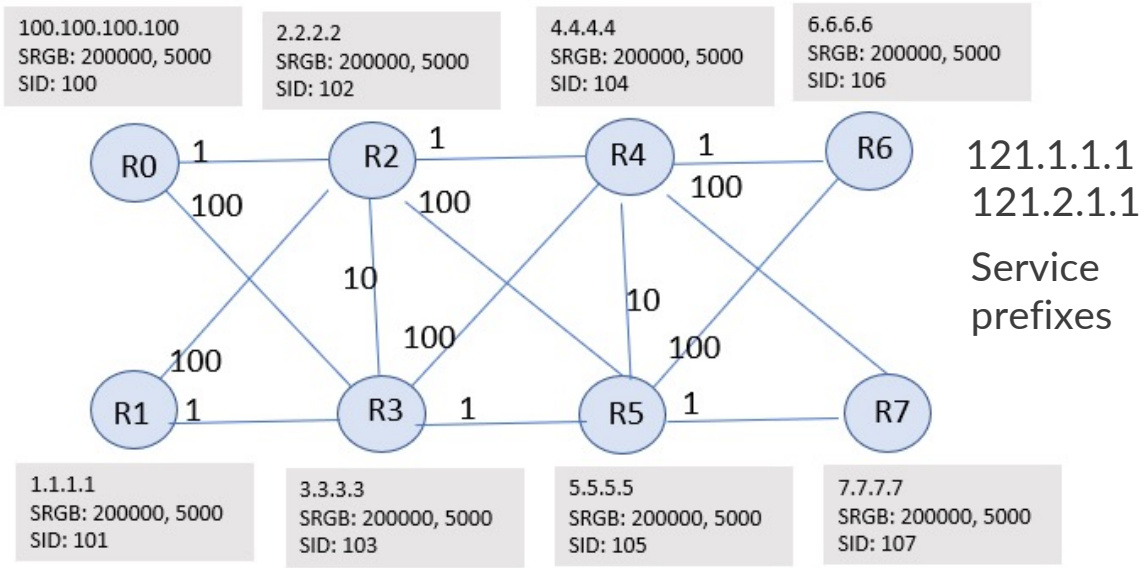
LDP TO SR MIGRATION



- Ships in the night
  - LDP and SR co-exist
  - LDP on all nodes
  - SR on upgraded nodes
  - LDP used for ldp-only destinations
  - SR paths used for SR enabled destination
- Mapping server
  - Mapping server SIDs for LDP-Only
  - Mapping client stitching at border

LAB SESSION

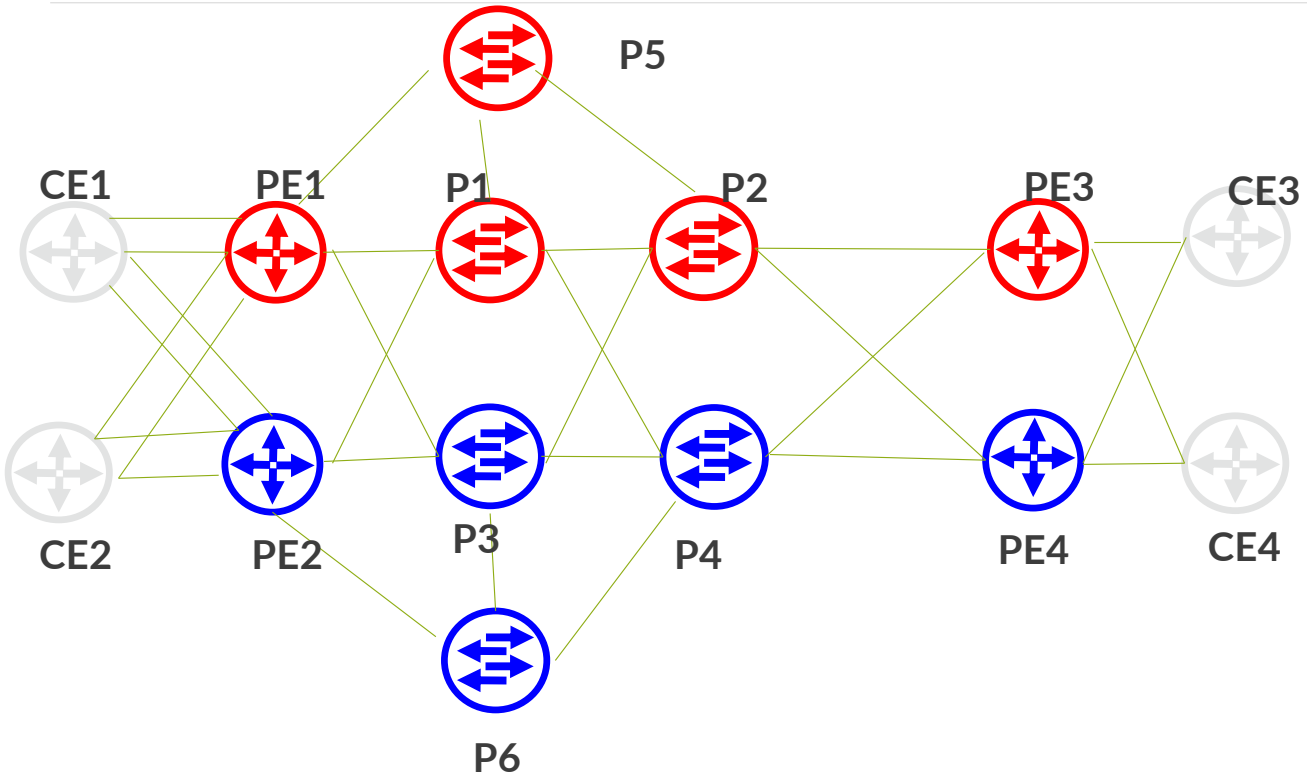
LAB TOPOLOGY



- ISIS Flat L2 domain
- LDP enabled
- IP Service prefixes attached to R6
- Service prefixes on LDP
- Enable segment routing
  - Configure SRGB
  - Configure Node-Sids/Adj-SIDs
  - Reboot
  - Verify E2E SR paths
  - Switch the traffic to SR

TRAFFIC ENGINEERING WITH FLEX-ALGO

FLEX-ALGO



Usecases:

- Dual plane networks with strict traffic isolation requirements
- Low latency Routing
- Data sovereignty
- Combining two different networks

Scaling:

- 4-8 Flex-algo

FAD ADVERTISEMENTS

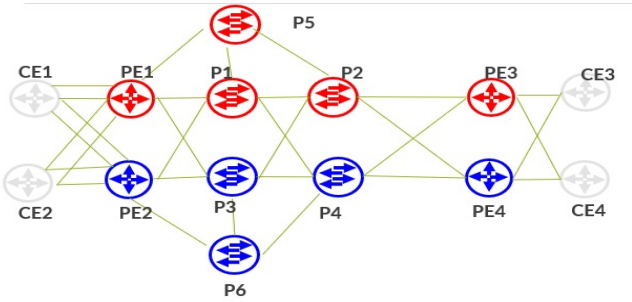
- Every node FAD can be configured
- For easy operations, only one or two nodes can be FAD servers
- FAD election between FAD advertisements and selection of winner
- Changing FAD
  - Update the FAD definition
  - Assign higher priority

Ordered rule for selection

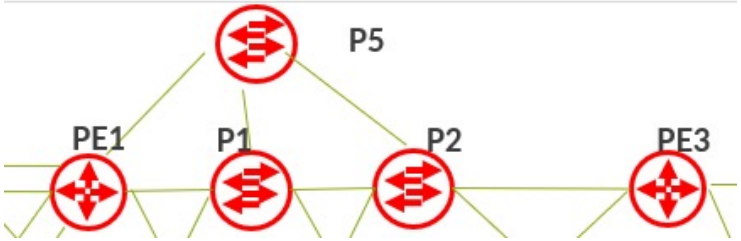
1. Highest priority
2. Highest system ID

SPF AND ROUTE PROCESSING

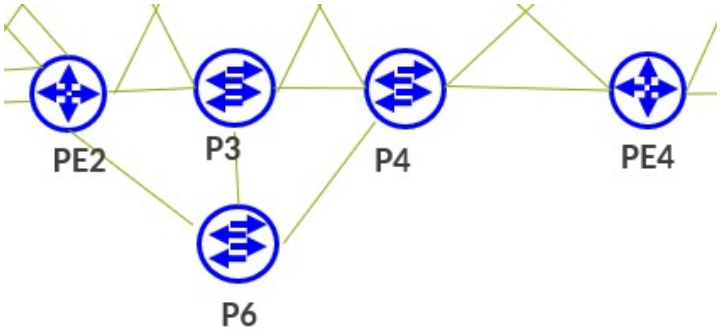
- Topology Filtering
- SPF
  - SPF per Flex-algo
- Route processing
  - Per Flex-algo SIDs
  - Nexthop processing
  - ECMP processing
- Route download



Flex-algo 128 topology



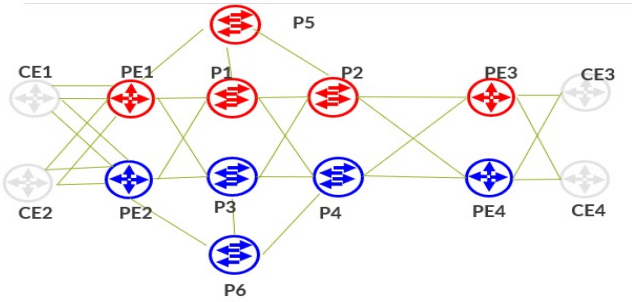
Flex-algo 129 topology



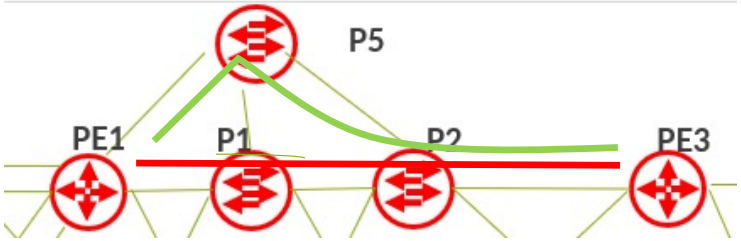
FLEX-ALGO & TI-LFA

- Topology Filtering

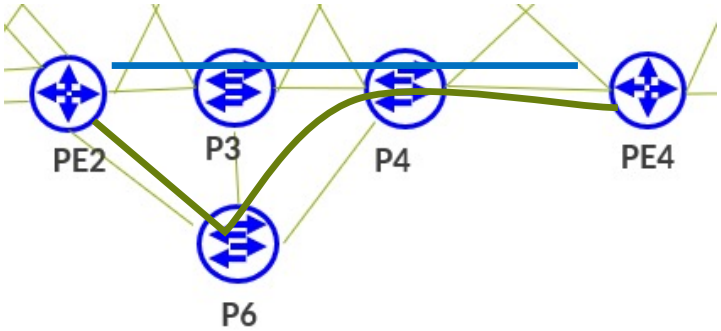
- SPF
  - SPF per Flex-algo
  - TI-LFA computation within flex-algo topology



Flex-algo 128 topology



Flex-algo 129 topology



FLEX-ALGO CONFIG

PE1 is a FAD server for Flex-algo 128.

Other nodes take the definition from the server

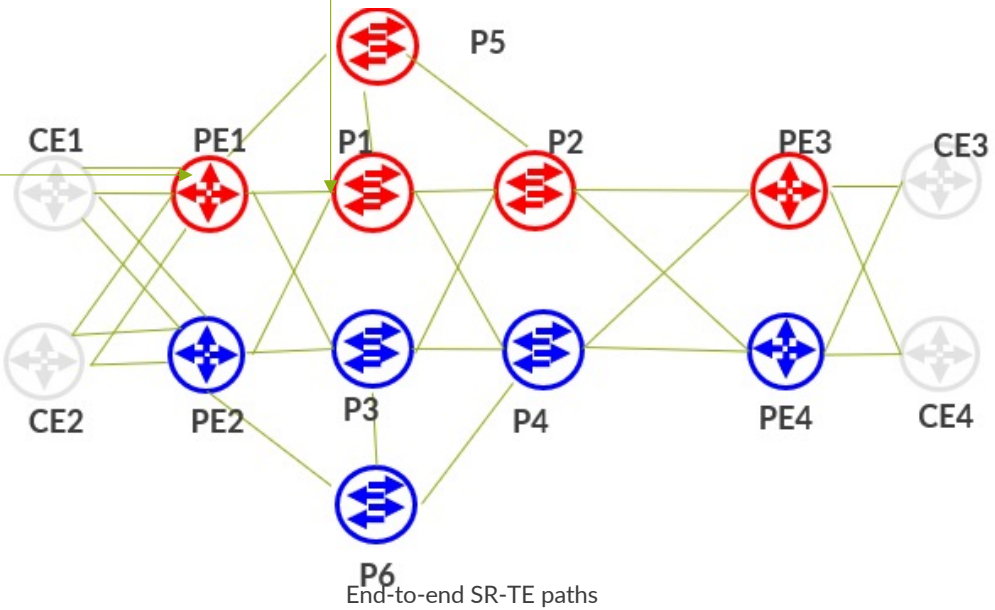
```
 protocols isis source-packet-routing
{
            flex-algorithm  128;
}
 protocols isis export flex-sid

 policy-options policy-statement
flex-sid {
             term one {
             from route-filter
3.3.3.3/32 exact
                                then
prefix-segment {

algorithm 128 index  310
}
```

```
set routing-options {
             flex-algorithm 128  {
            definition {
            metric-type te-metric
                                spf;
}
 protocols isis source-packet-routing {
            flex-algorithm  128;
}
 protocols isis export flex-sid

 policy-options policy-statement flex-sid
{
             term one {
             from route-filter
100.100.100.100/32 exact
                                then
prefix-segment {
```



End-to-end SR-TE paths

PUTTING TRAFFIC ON FLEX-ALGO

Service prefix associated with ext color community on egress.

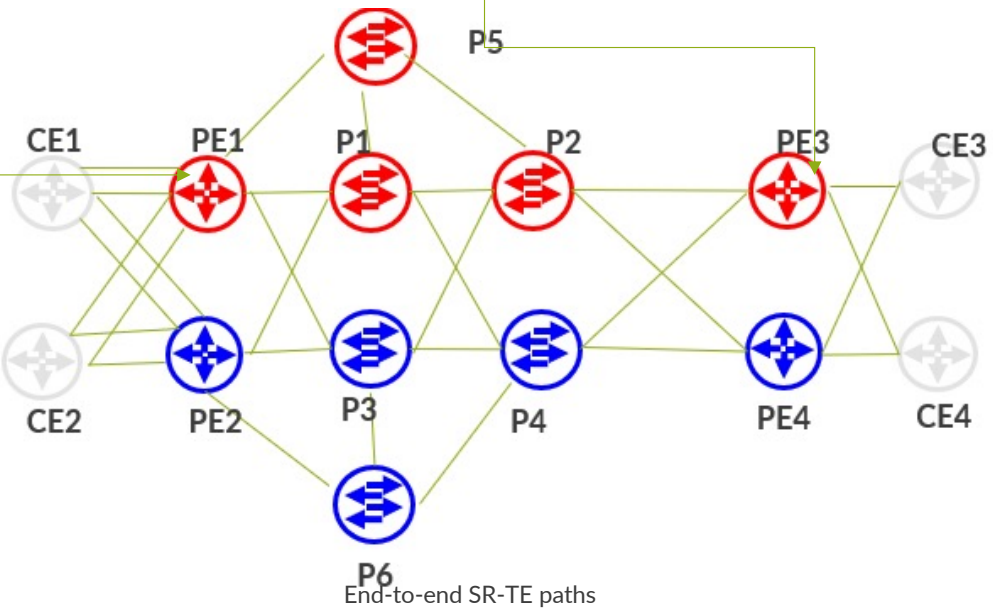
On ingress, Flex-algo routes created in color table

Service prefix automatically resolves on

corresponding flex-algo route

```
 policy-options policy-statement
flex-sid {
            term one {
            from route-filter
192.1.1.1/32 exact
                                then
community add com1;
}
set policy-options community com1
members color:0:128;
```

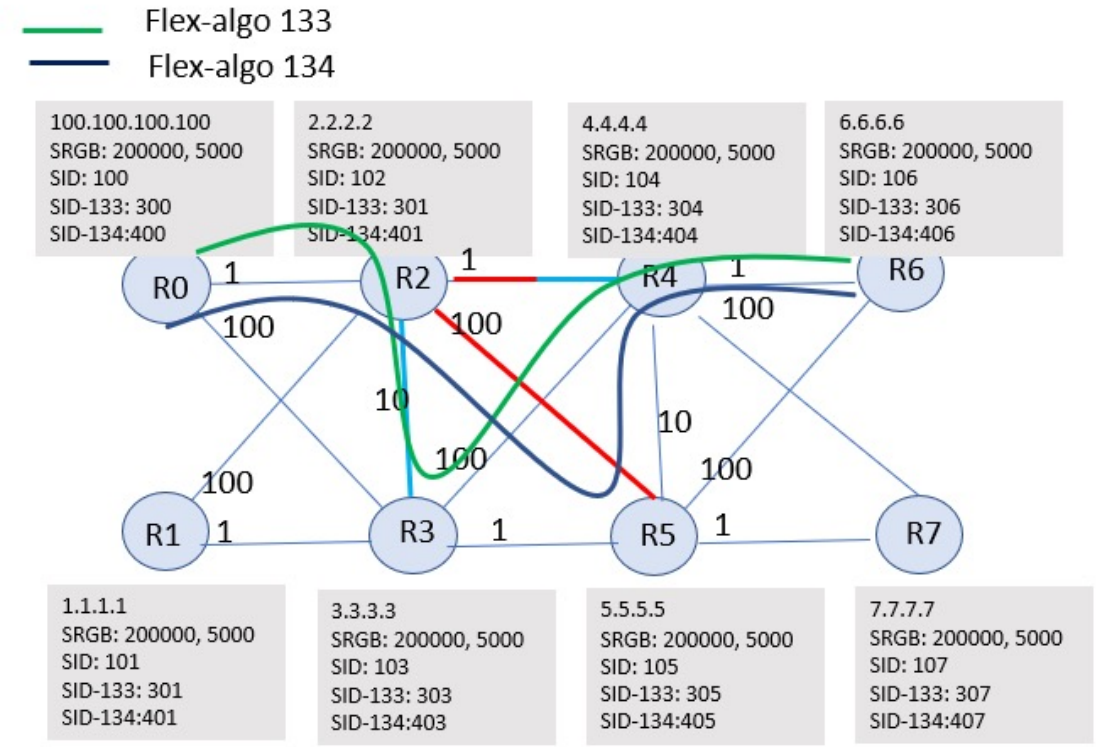
```
3.3.3.3-32
            *[L-ISIS/14] 00:14:39, metric 2
             >  to 21.0.2.2 via ge-0/0/1.0, Push
200100
              to 21.0.3.2 via ge-0/0/2.0, Push
200100
3.3.3.3/32
            *[L-ISIS/14] 00:14:39, metric 1
             >  to 21.0.2.2 via ge-0/0/1.0
```



End-to-end SR-TE paths

LAB SESSION

LAB TOPOLOGY



- Flex-alg0 133
  - Igp-metric, SPF, exclude red
- Flex-algo 134
  - Igp-metric, SPF, exclude blue

- Flex-algo
  - Configure FAD
  - Configure Flex-algo SIDs
  - Verify E2E flex-algo paths
  - Switch the traffic to flex-algo

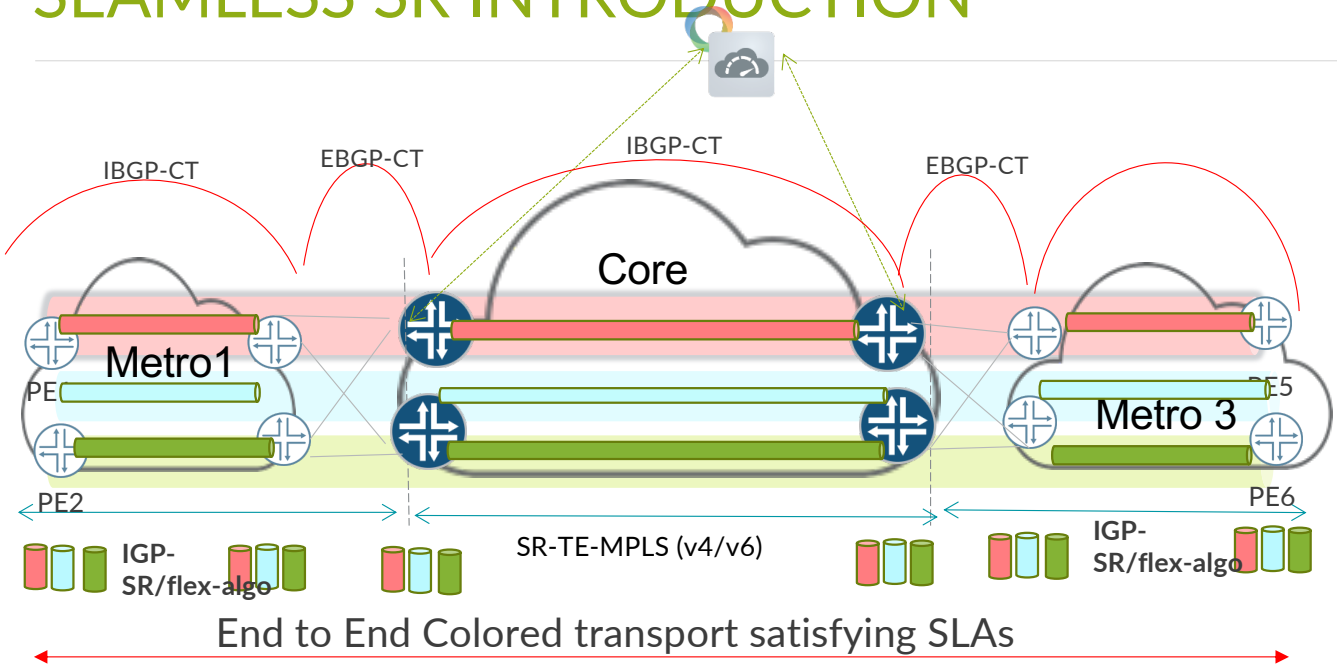
SEAMLESS SR FOR INTENT AWARE ROUTING

WHY SEAMLESS SR?

- Widely Deployed Distributed network architecture such as Seamless MPLS.
- 5G, IOT and Network traffic growth bringing Scaling and low latency requirements with E2E view
- Natural Upgrade needed from Seamless MPLS architecture to Segment Routing
- Lesser protocols and operationally easy
- E2E network slicing

draft-hegde-spring-mpls-seamless-sr

SEAMLESS SR INTRODUCTION



- Distributed Routing architecture
- Service routes only on service end-points
- Independent IGP domains
- Stitched together with BGP
- BGP-CT for E2E intent aware paths

RED Transport Class

BLUE Transport Class

GREEN Transport Class

BGP CLASSFUL TRANSPORT (BGP-CT)

A mechanism for extending color-mapping across multiple ASes.

- No need for a controller, but can be used if desired
- No need to expose internal topology of a domain to any other domain

Each domain can make its own choice of transport technology independently of what other domains are using

- SR-TE, Flex-Algo, RSVP…
- BGP-CT acts as the “glue” between domains

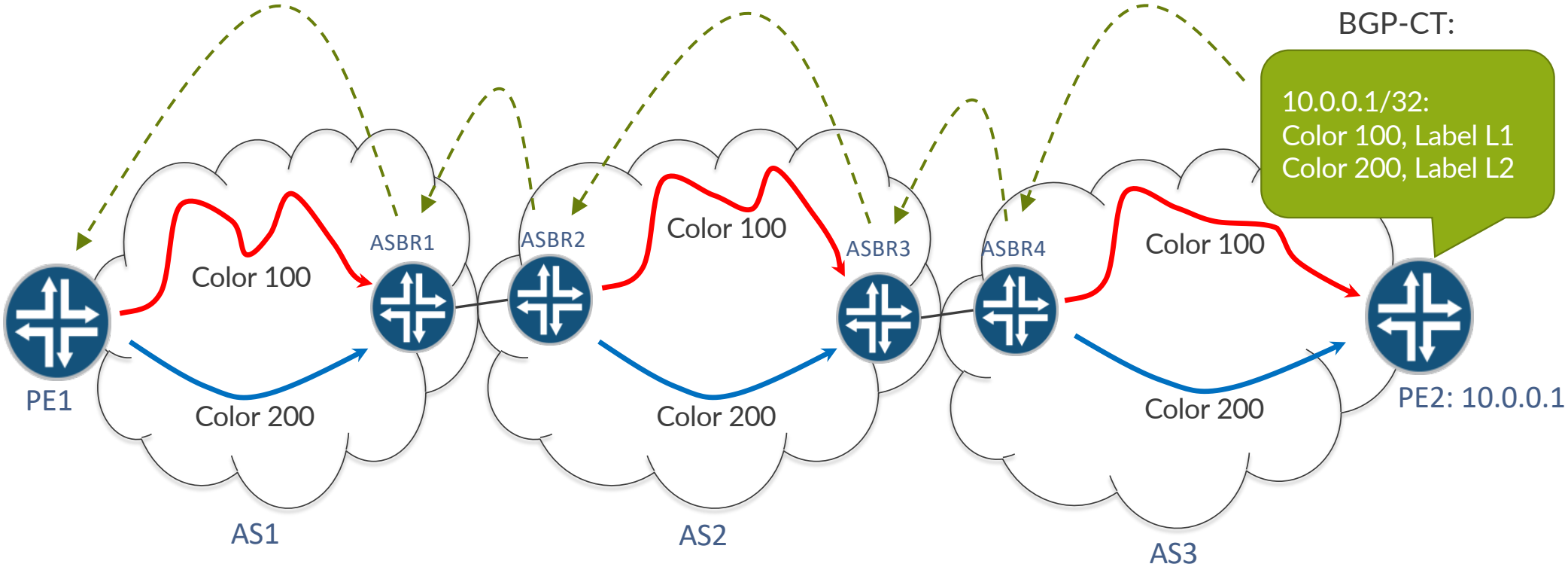
BGP-CT is similar to BGP-LU, except that it has {egress PE, color} granularity

Color denotes the “flavor” of the transport e.g. minimum latency, cheapest monetary cost.

See https://datatracker.ietf.org/doc/draft-kaliraj-idr-bgp-classful-transport-planes/

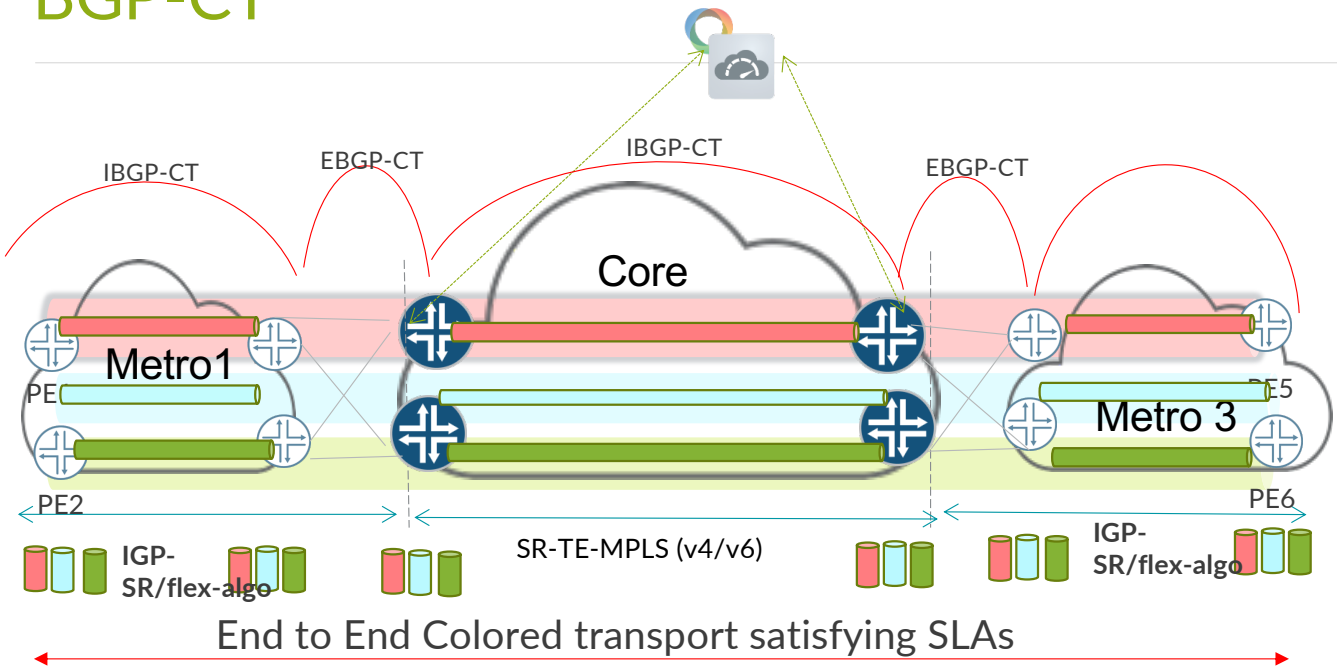
BGP-CT

- PE1 maps prefixes (according to color community) to the matching color BGP-CT label and local tunnel or flex-algo to ASBR1.
- In turn, ASBR2 maps traffic to tunnel or flex-algo to ASBR3 according to the color of the incoming BGP-CT label.



e.g. Color 100 = cheapest cost
Color 200 = minimum latency

BGP-CT



BGP-CT NLRI:

Prefix : PE loopback

RD : Distinguisher

Route Target community:
Corresponds to Transport Class

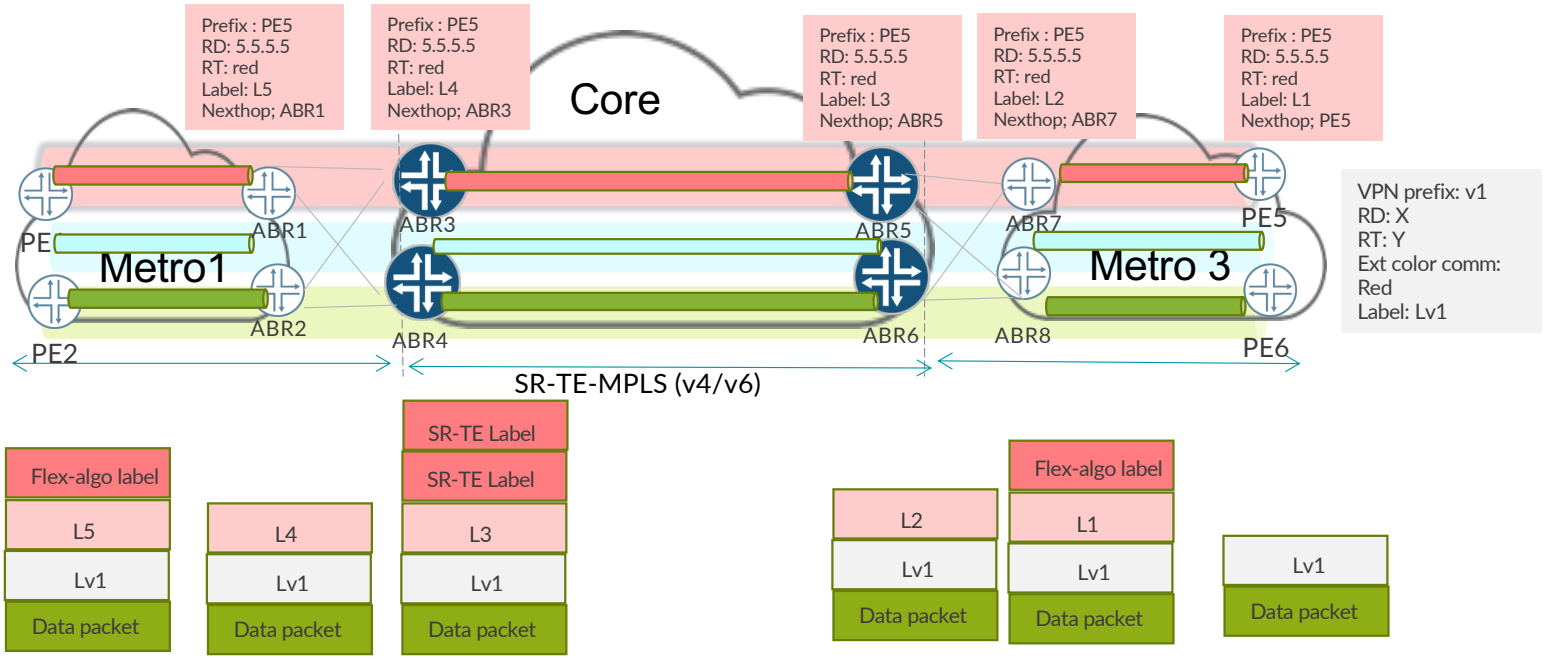
BGP-CT is independent of any VPN. It does not require VPN configuration.

RED Transport Class

BLUE Transport Class

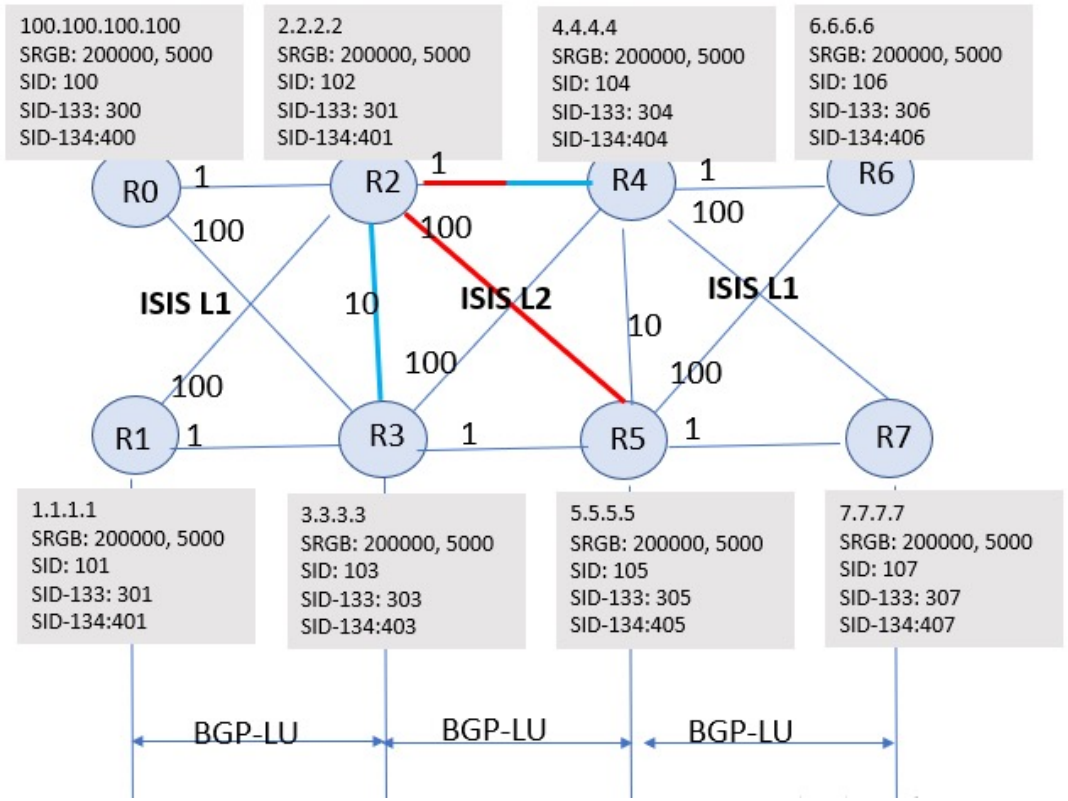
GREEN Transport Class

SEAMLESS SR : CLASSFUL TRANSPORT PACKET FORWARDING



LAB SESSION

LAB TOPOLOGY



- ISIS L1 and ISIS L2 domains
  - No leaking between levels
  - BGP-LU to provide E2E connectivity
- BGP-CT
  - Enable BGP-CT family
  - E2E intent aware paths

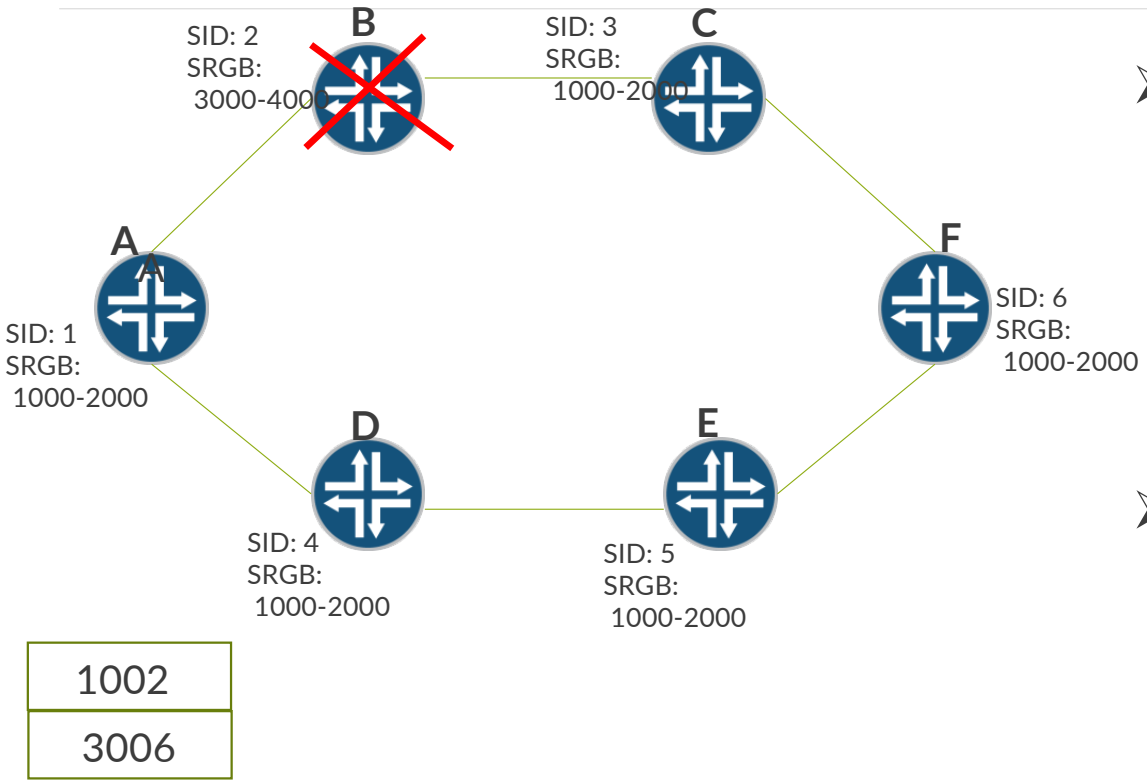
THANK YOU

TI-LFA

- Post convergence loop free paths used as backup
- Near 100% backup coverage
- Backup paths computed automatically, no need to provision backup paths
- No separate capacity planning for backup paths as post convergence path used for backup
- Ability to account for local and remote SRLG failures

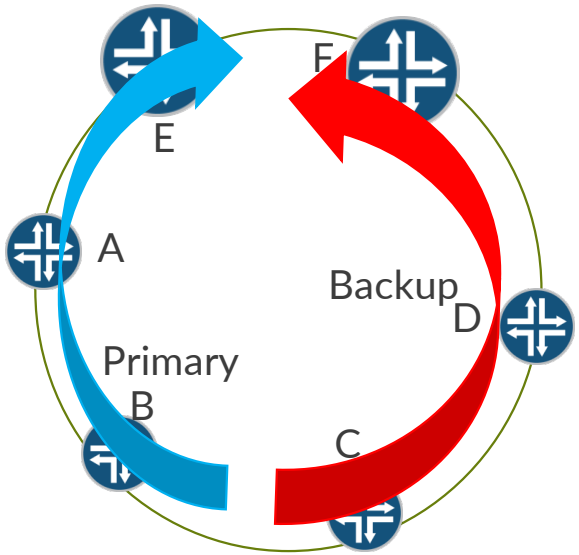
```
protocols isis interface ge-0/0/0.0 {
    level 2 {
        post-convergence-lfa;
    }
protocols isis backup-spf-options {
    use-post-convergence-lfa;
}
```

NODE PROTECTION FOR SR-TE PATHS

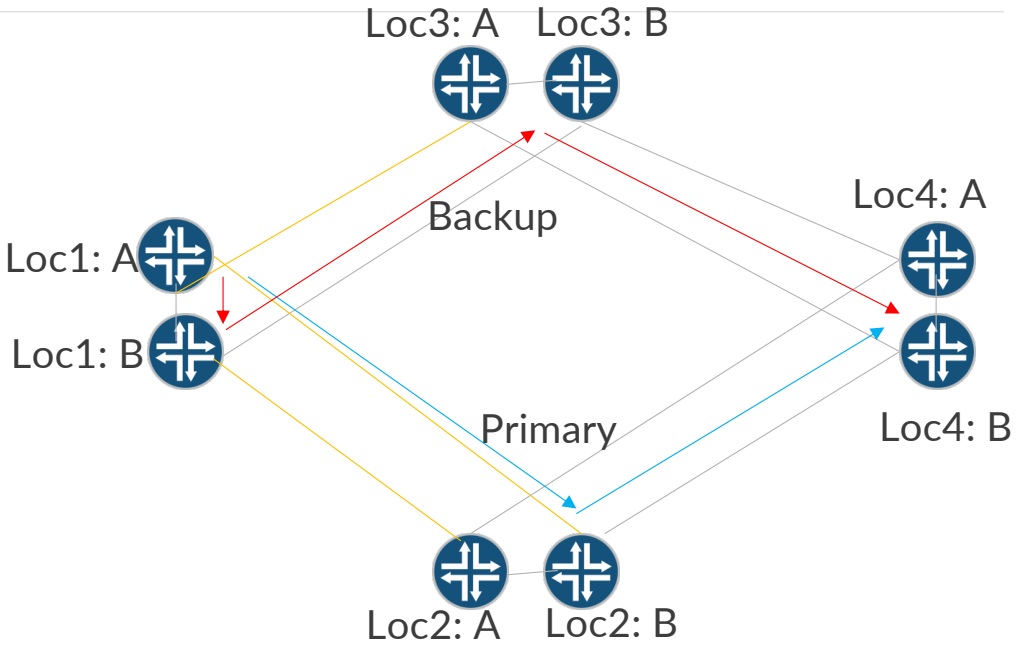


- **Context table based solution**
  - Build context table for every neighbor
  - Program paths for labels from neighbors' perspective
  - On failure
    - lookup next label in context table
- **Optimized solution**
  - Uniform SRGB
  - Global Adj-SIDs
  - No context table
  - On Failure
    - lookup next label in Global table

TI-LFA WITH SRLG PROTECTION



- Congruent to post convergence path
- Provides better coverage



- Remote SRLG consideration for TI-LFA
  - Avoid backup path along paths with same SRLG