SEAMLESS SR ARCHITECTURE

Shraddha Hegde

Dec 2020

JUNIPER NETWORKS | Engineering Simplicity

WHY SEAMLESS SR?

- Widely Deployed Distributed network architecture such as Seamless MPLS.

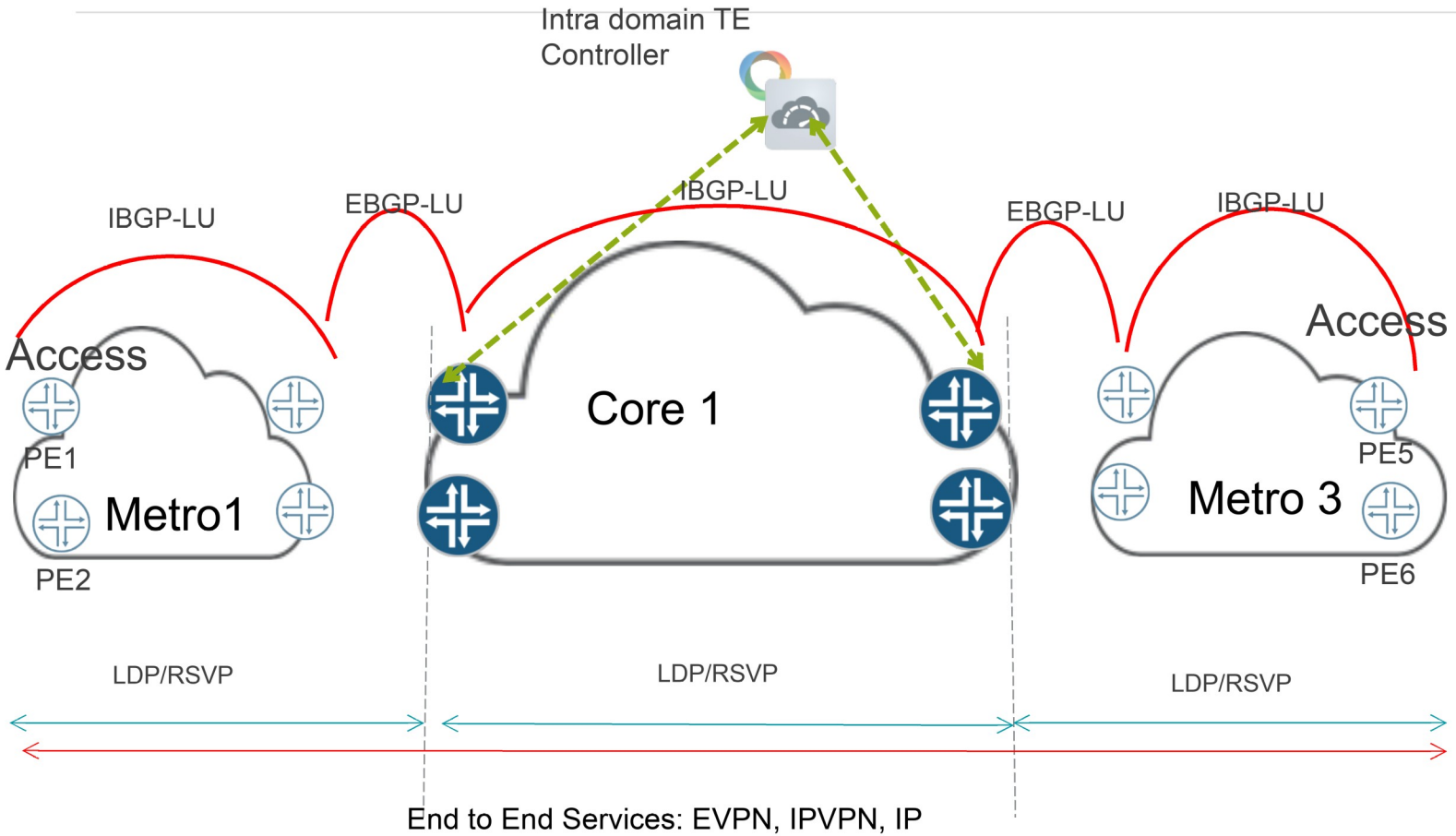
- 5G, IOT and Network traffic growth bringing Scaling and low latency requirements with E2E view

- Natural Upgrade needed from Seamless MPLS architecture to Segment Routing

- Lesser protocols and operationally easy

- E2E network slicing

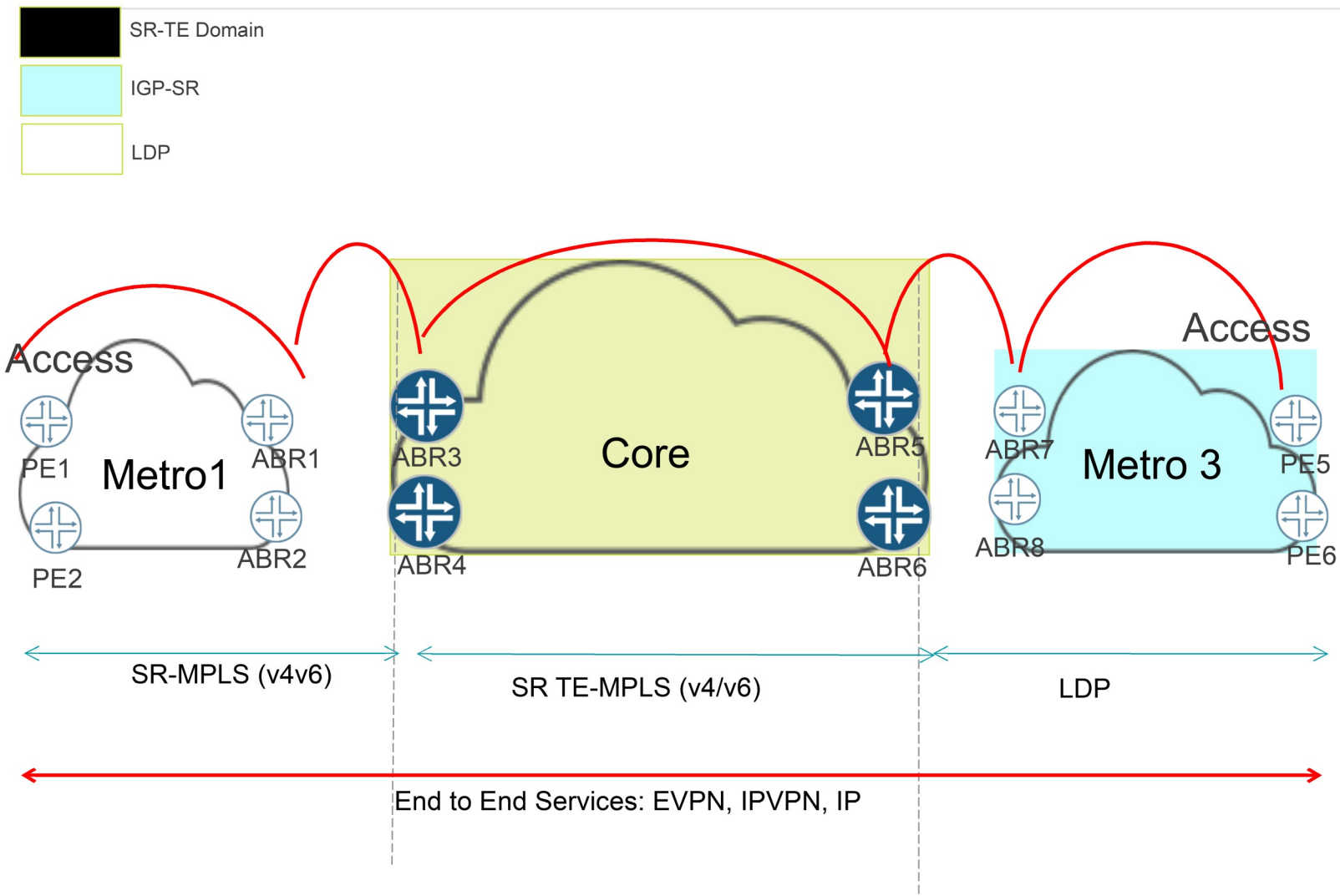
SEAMLESS MPLS ARCHITECTURE



- Large networks segregated into multiple IGP domains
- Core is in one IGP domain.
- Each of the metros in separate IGP domain
- BGP-LU provides E2E connectivity
- Services only on access node. Option C connectivity via BGP-LU
- Highly scalable architecture
- Widely deployed

SEAMLESS SR: INTRODUCTION

draft-hegde-spring-mpls-seamless-sr



- Independent SR migration in each domain
- BGP-LU for E-2-E connectivity
- Deterministic labels with BGP Prefix-SID (RFC 8669)

5G TRANSPORT REQUIREMENTS

- Massive bandwidth
- Timing and E2E Latency
- Operational Simplicity
- High availability
- High scalability
- E2E Service Differentiation
- Application aware routing

LIMITATIONS OF BGP-LU/BGP-PREFIX SID

- Provides a single E2E path
- The path selection is based on BGP best path selection, other parameters such as low latency cannot be used
- Multiple loopbacks with BGP-LU need to be used for multiple paths
- Service mapping need to be done at every border node
- Need a mechanism to simplify for 5G transport

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

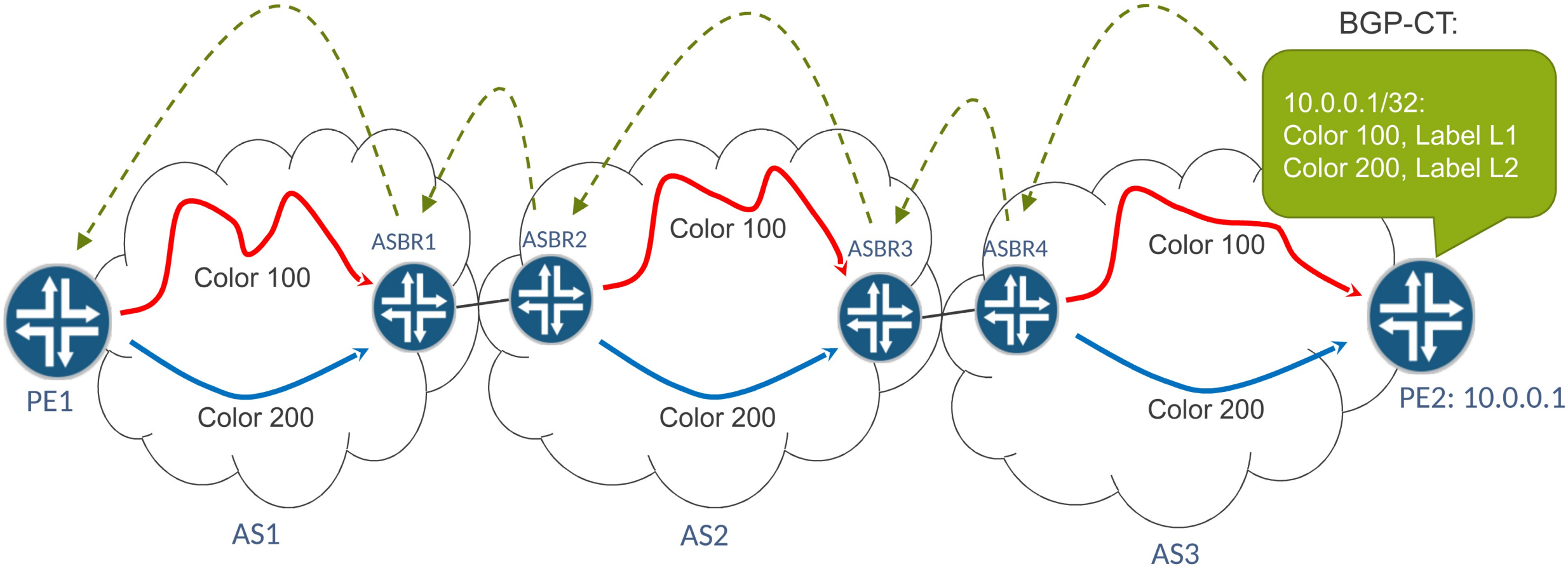
BGP-CT is similar to BGP-LU, except that it has {egress PE, color} granularity

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/

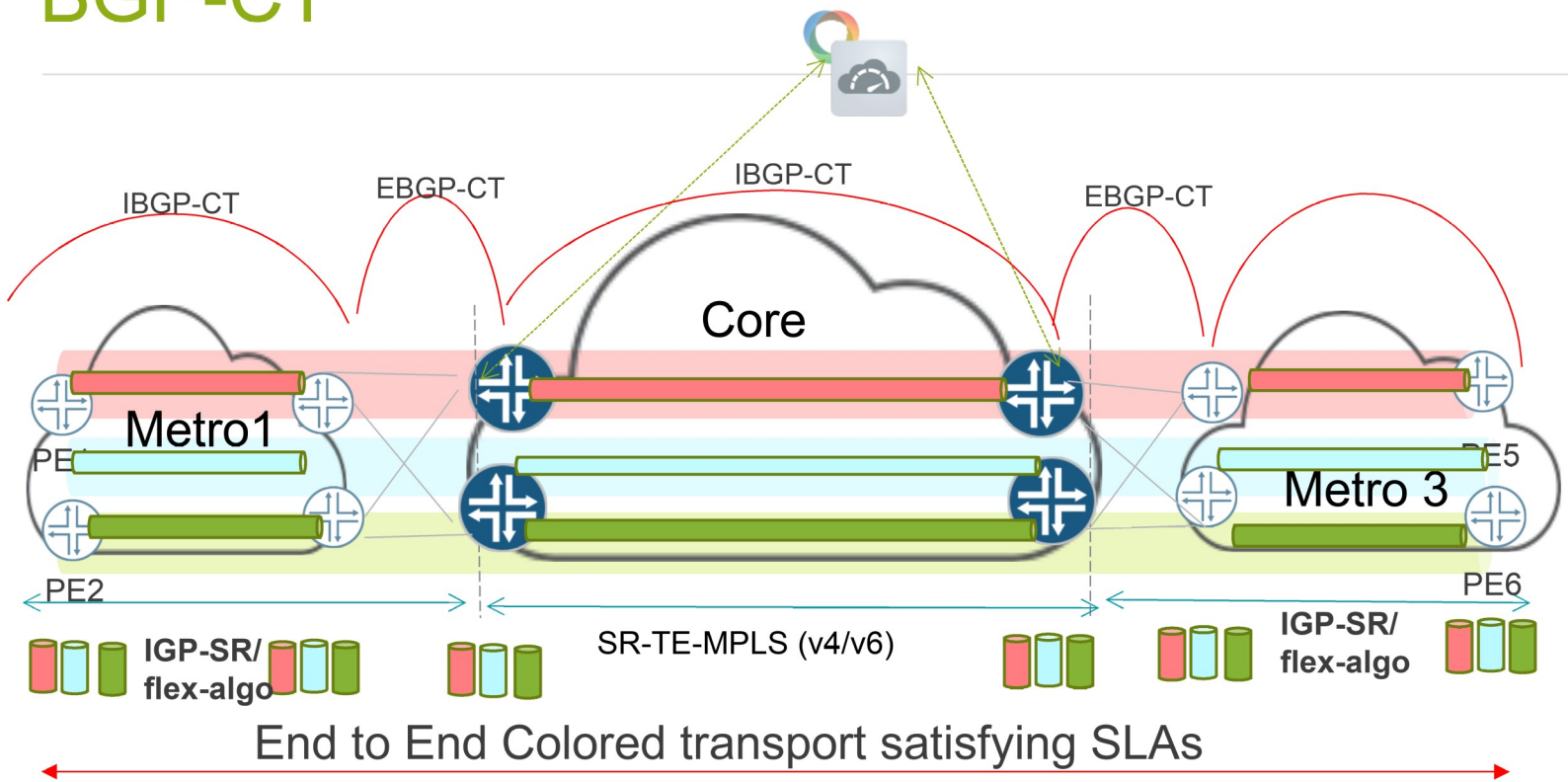
BGP-CT

- 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

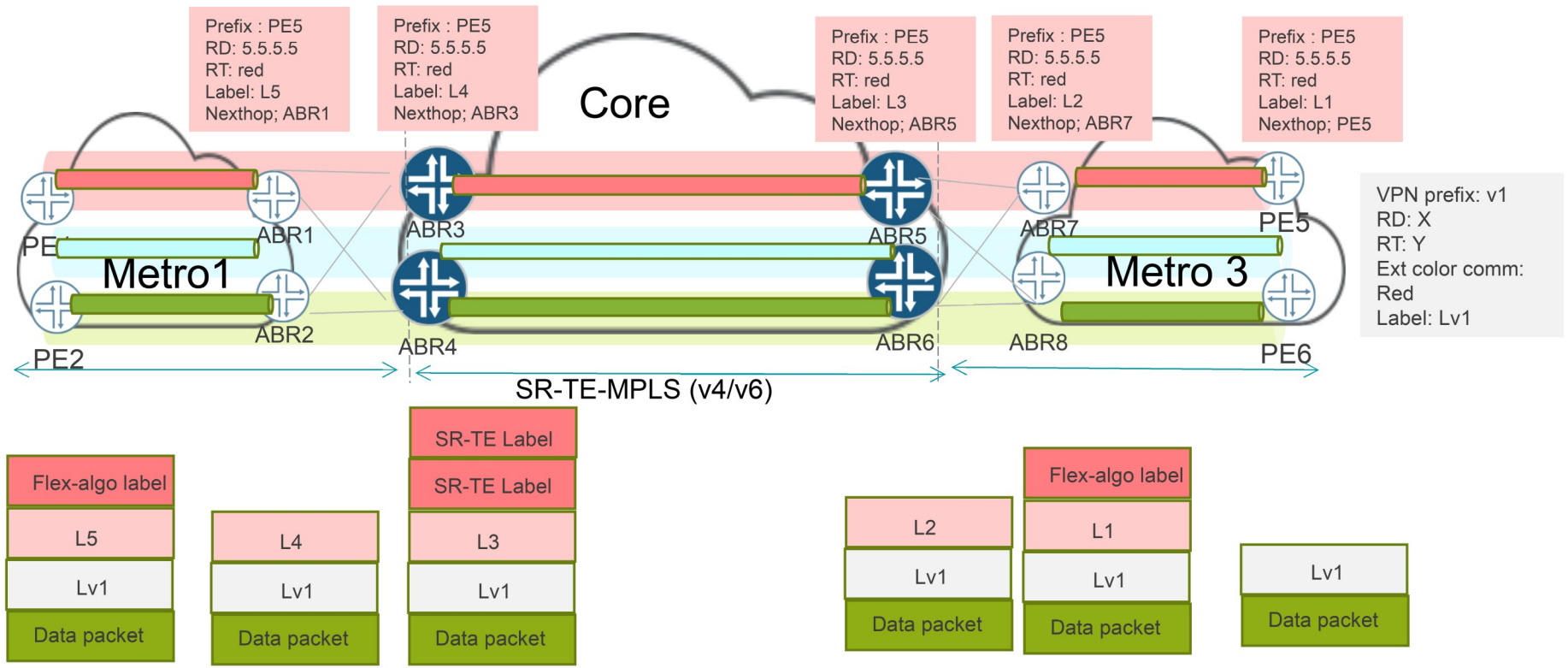
BGP-CT is independent of any VPN. It does not require VPN configuration.

RED Transport Class

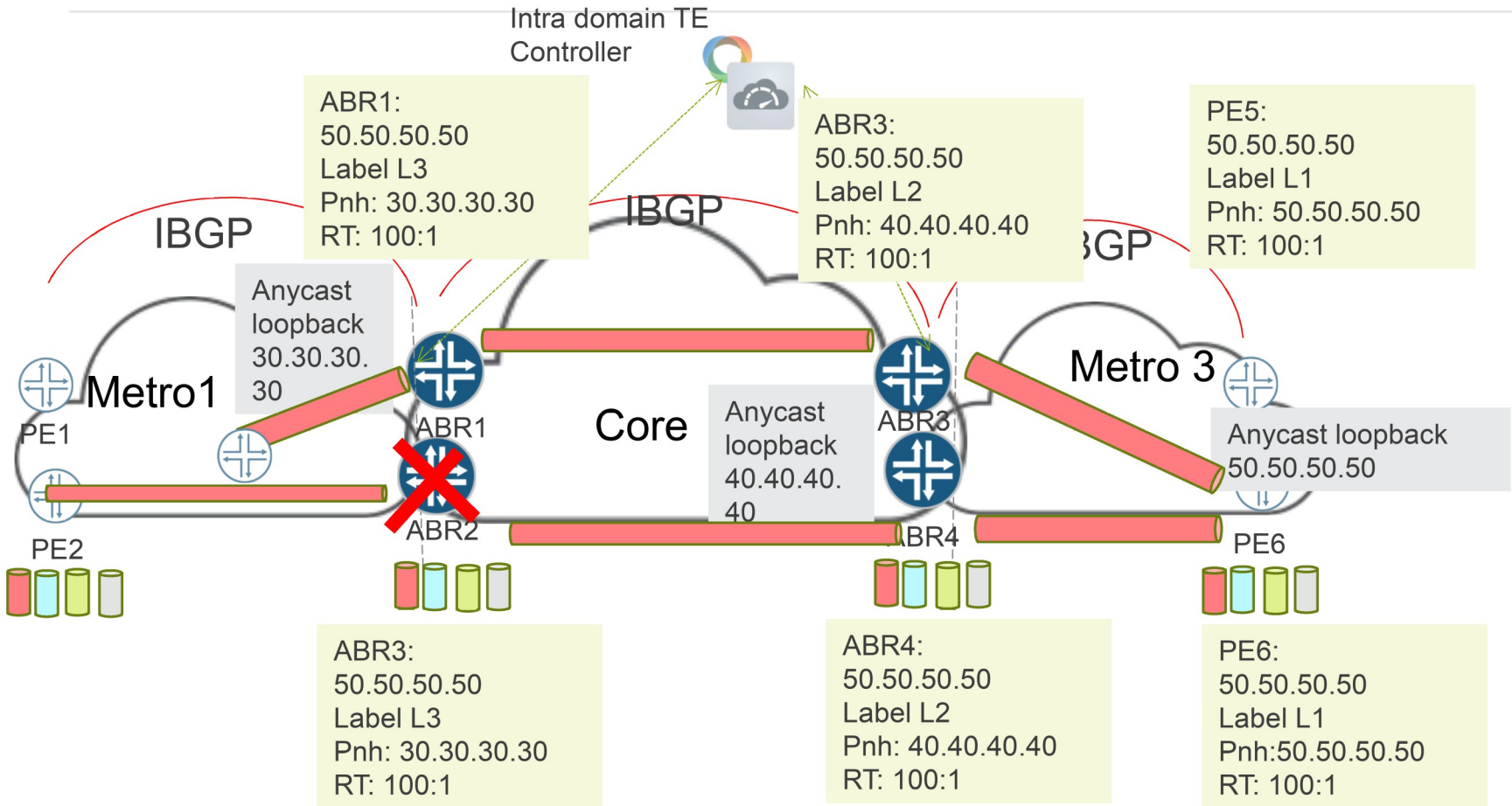
BLUE Transport Class

GREEN Transport Class

SEAMLESS SR : CLASSFUL TRANSPORT PACKET FORWARDING

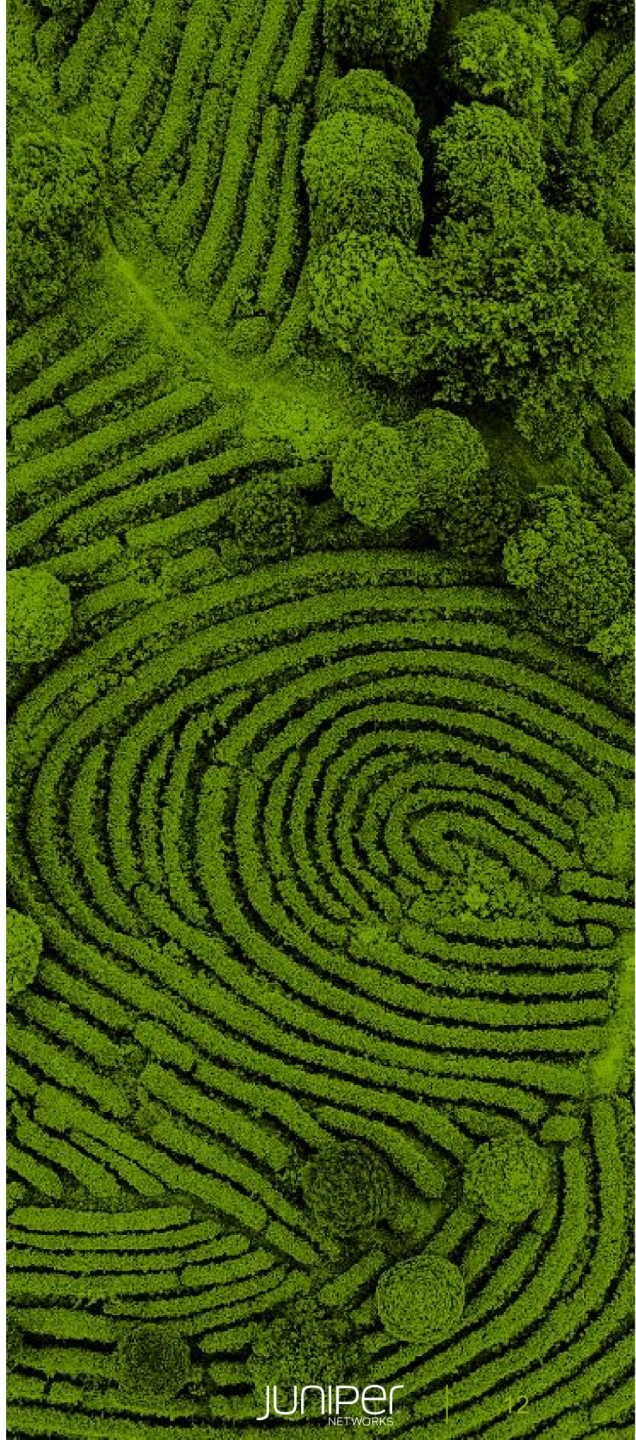


RESILIENCY

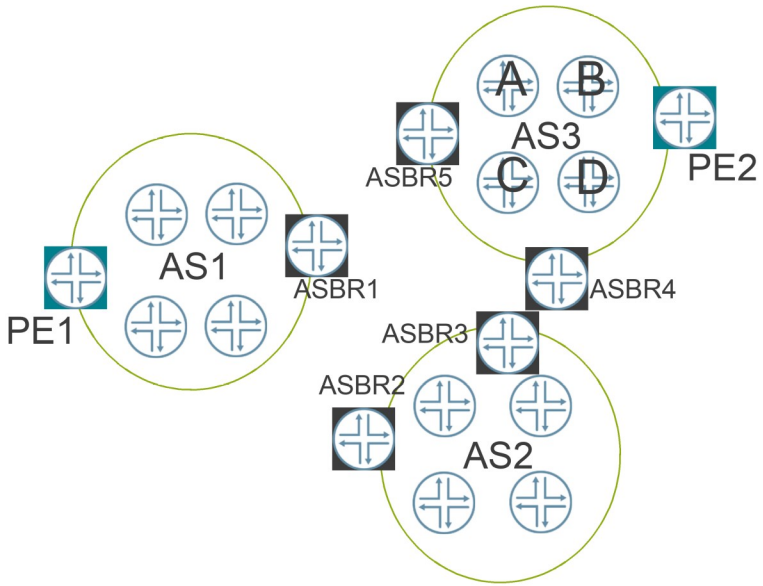


- Anycast loopbacks used as nexthops
- Same BGP-CT label on both ABRs
- TI-LFA for the anycast prefixes

USECASES



DATA SOVEREIGNTY USE CASE



Multiple AS

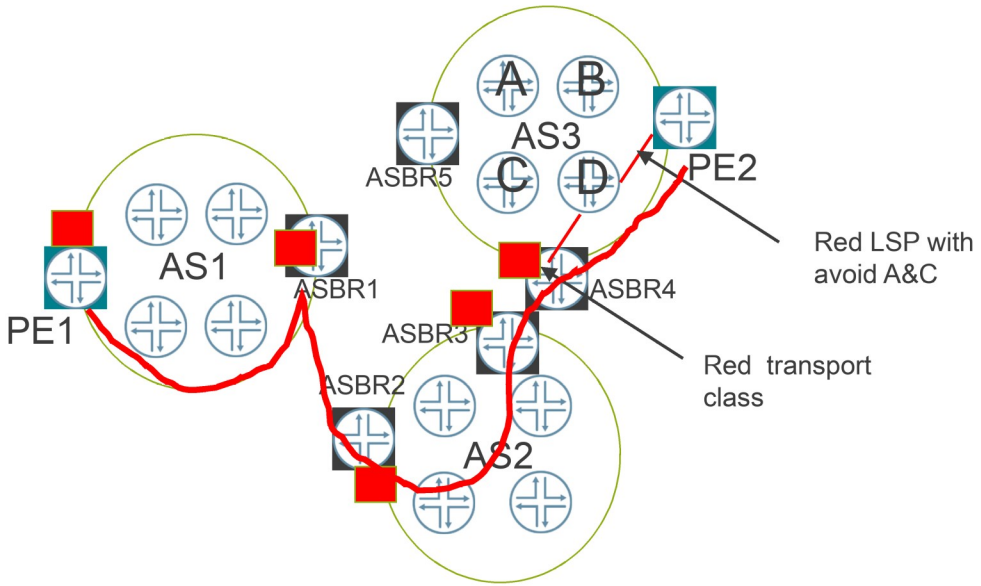
Each AS represents a continent

Data sovereignty

- Avoid node A and C

This “avoid node A& C” constraint is not applicable in AS1 and AS2

DATA SOVEREIGNTY USE CASE



Red LSP created on ASBR4

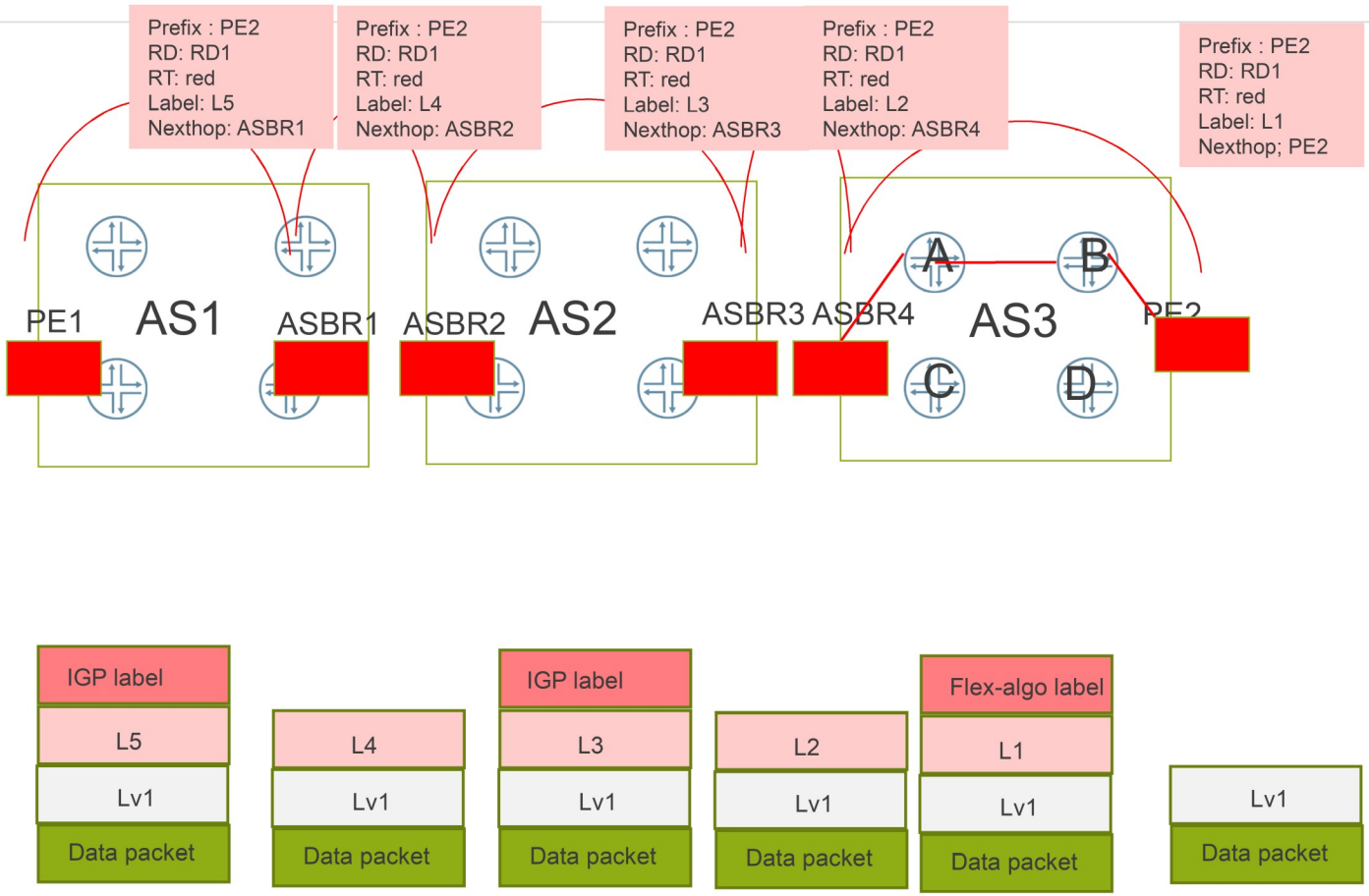
Since ASBR5 is connected to A & C, Red LSP isn't created

Red transport class created on all border nodes

Resolution

- AS3: Strict resolution
- AS1 &AS2: Resolve on Red fallback on best

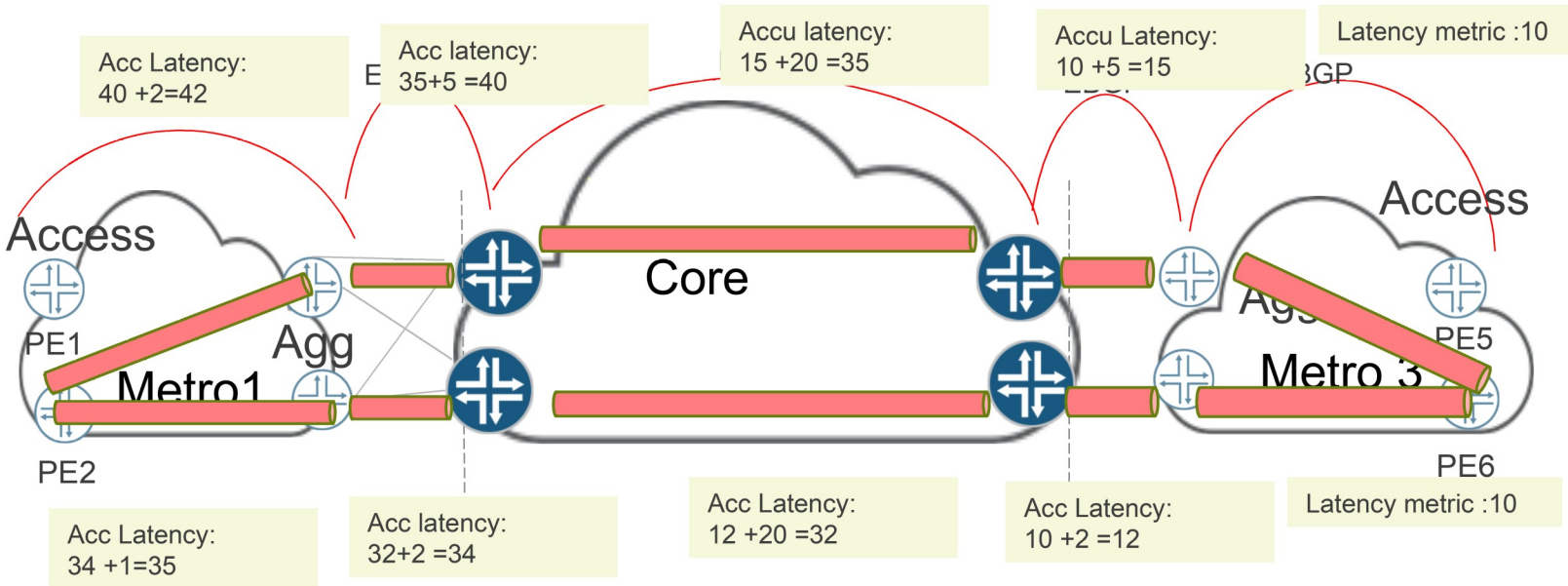
SOLUTION WITH BGP-CT



AS1 and AS2 do not need to create Red Transport Tunnel. BGP-CT will use best effort paths in AS1 and AS2

USECASES: E2E LOW LATENCY ROUTING

draft-ietf-idr-performance-routing



- Intra-domain latency accumulated
- BGP extension to carry accumulated latency metric in AIGP sub-TLV
- BGP best path selection at every border node based on accumulated latency metric

THANK YOU