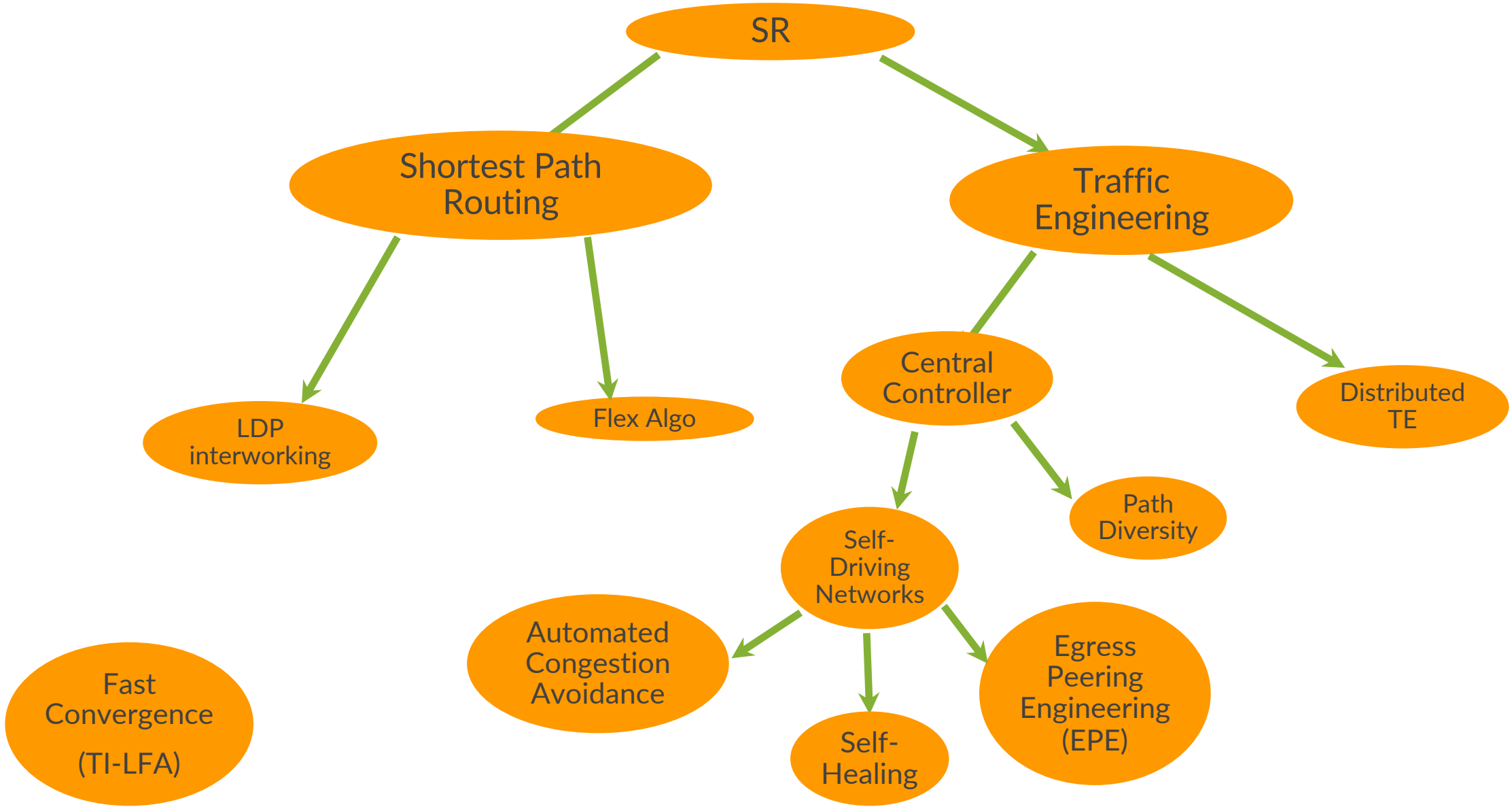
Segment Routing innovations: Flex Algo and BGP-CT

Julian Lucek

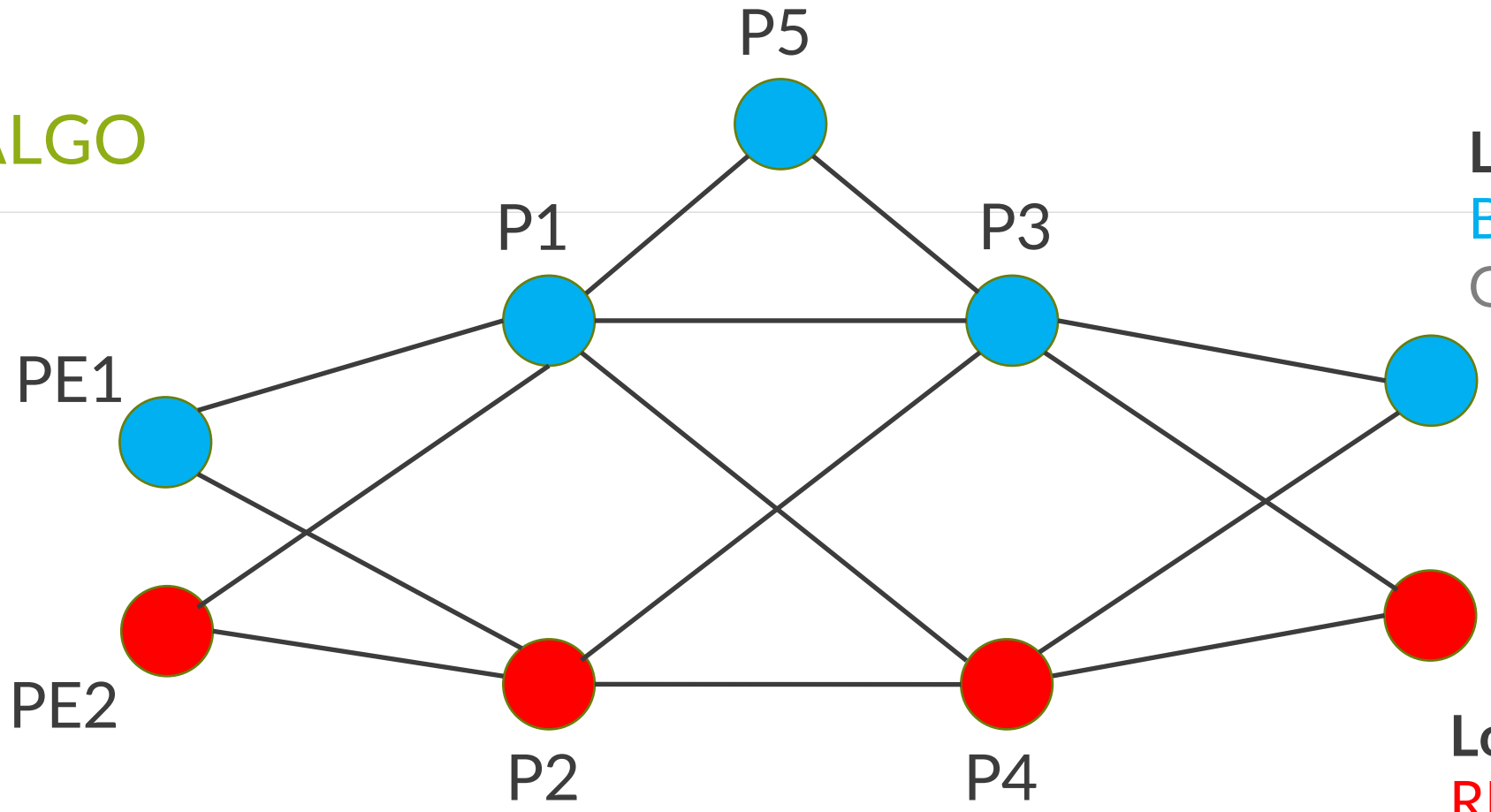
Juniper Networks | Engineering Simplicity

SEGMENT ROUTING (SR) LANDSCAPE



Flex Algo

FLEX-ALGO

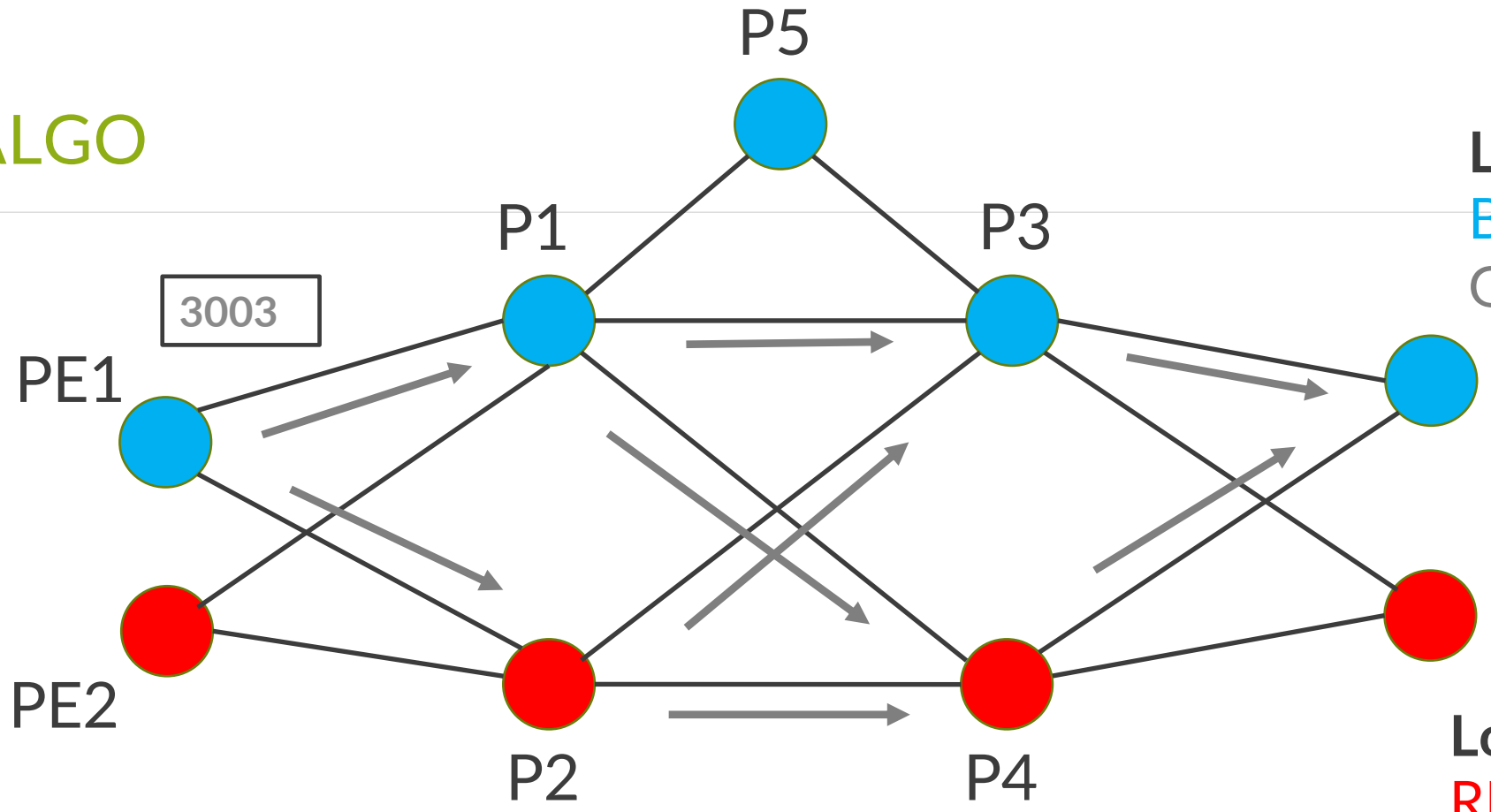


Lo0: 10.0.0.3
BLUE SID 1003
GREY SID 3003

Lo0: 10.0.0.4
RED SID 2004
GREY SID 3004

- A node can be a member of multiple algos.
- A node announces a different node SID for each algo that it is a member of.
- Separate Shortest Path First (SPF) calculations per algo.
- In the diagram, all nodes are members of the Grey topology. Nodes in one plane are members of Red algo, nodes in the other plane are members of the Blue algo.

FLEX-ALGO

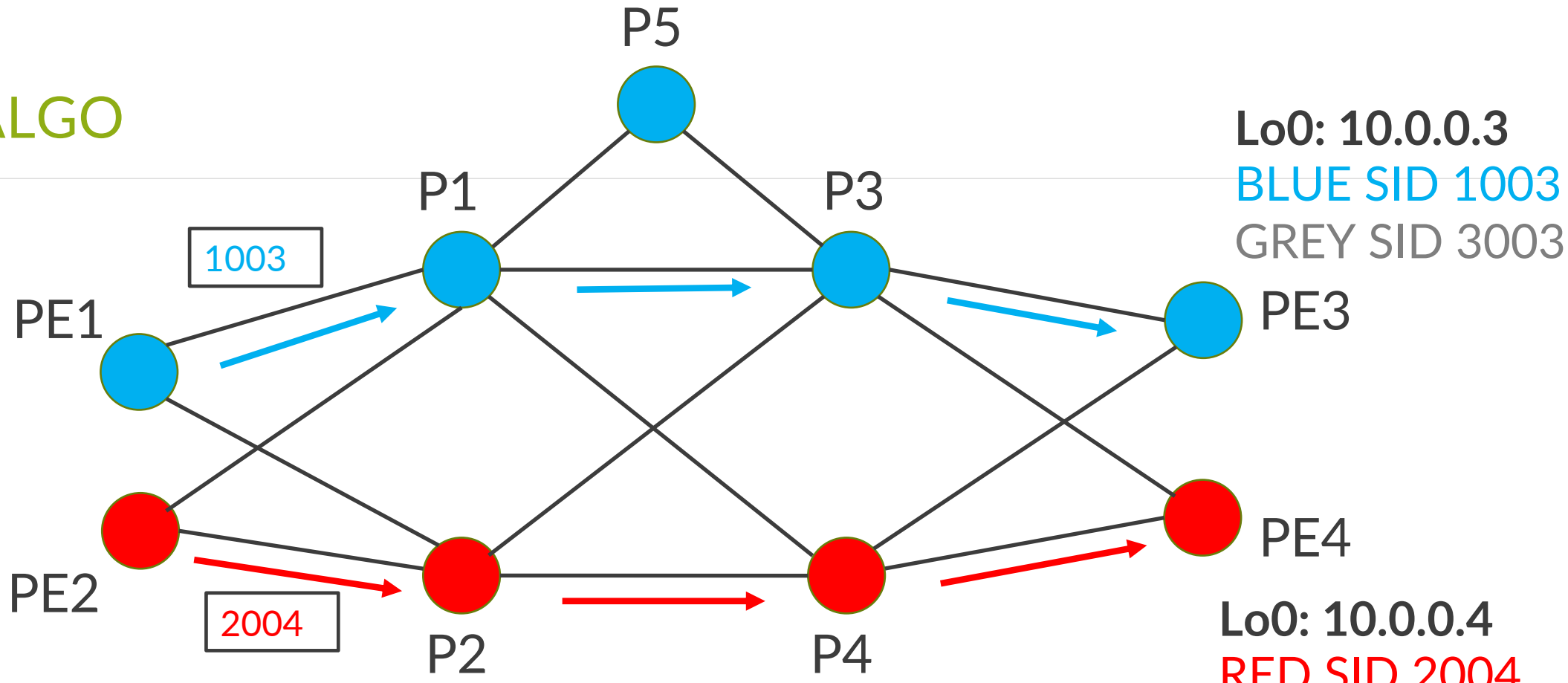


Lo0: 10.0.0.3
BLUE SID 1003
GREY SID 3003

Lo0: 10.0.0.4
RED SID 2004
GREY SID 3004

"Vanilla" traffic from PE1 is mapped to Grey topology, can go anywhere

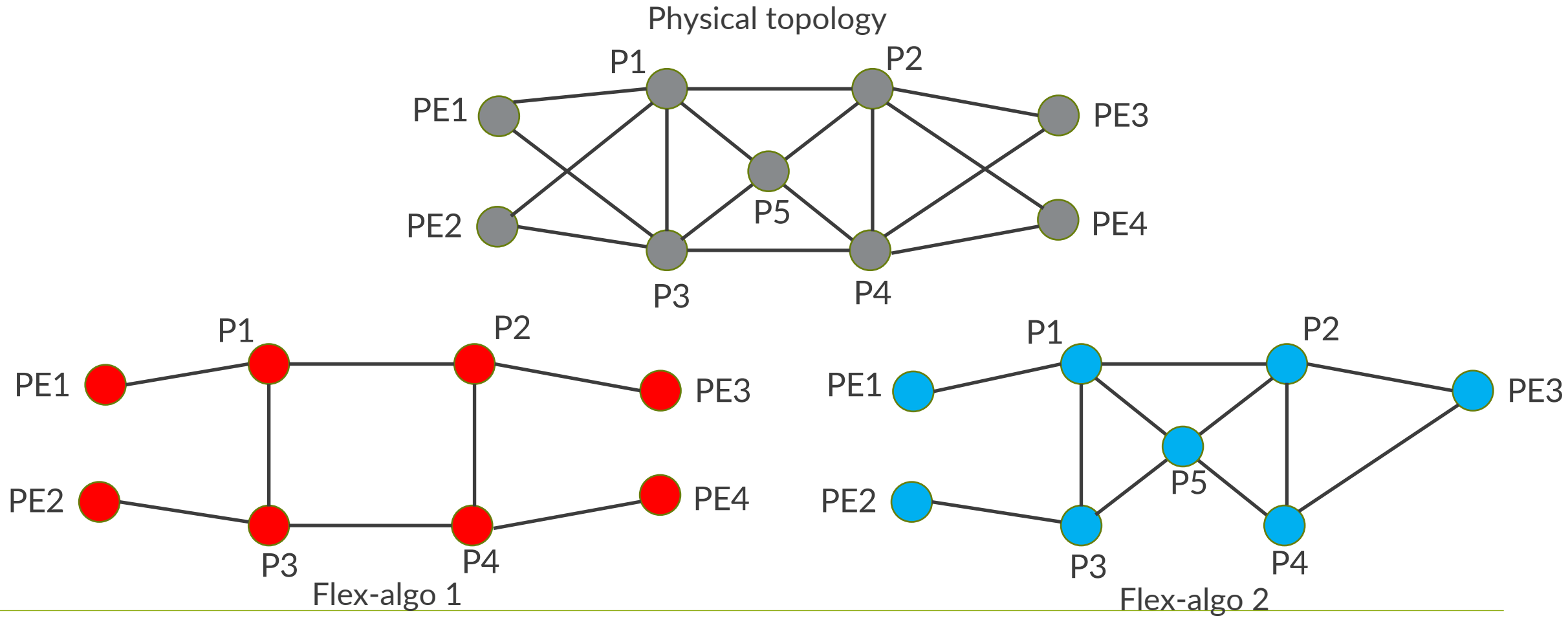
FLEX-ALGO



- We have some traffic from PE1 to PE3 that needs to be diversely routed from other traffic from PE2 to PE4
- Use Blue and Red algos respectively

ANOTHER FLEX-ALGO EXAMPLE

Nodes and links can be members of multiple flex-algos



METRIC ACCORDING TO TRAFFIC TYPE

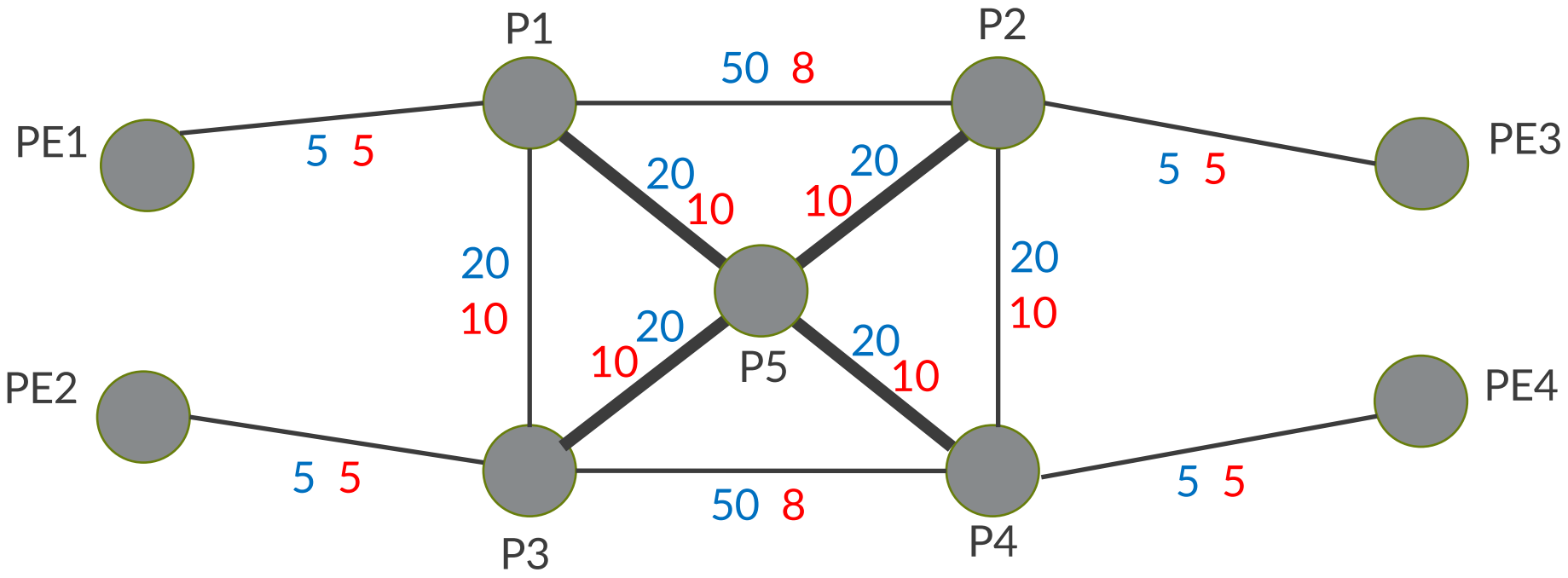
Another application for flex-algo is a situation where different traffic types need to use different link metrics.

For example, for delay sensitive traffic, you want a metric proportional to the latency of the link.

For bulk internet traffic, you want a metric inversely proportional to the bandwidth of the link.

In a given network, the two requirements might conflict!

METRIC ACCORDING TO TRAFFIC TYPE

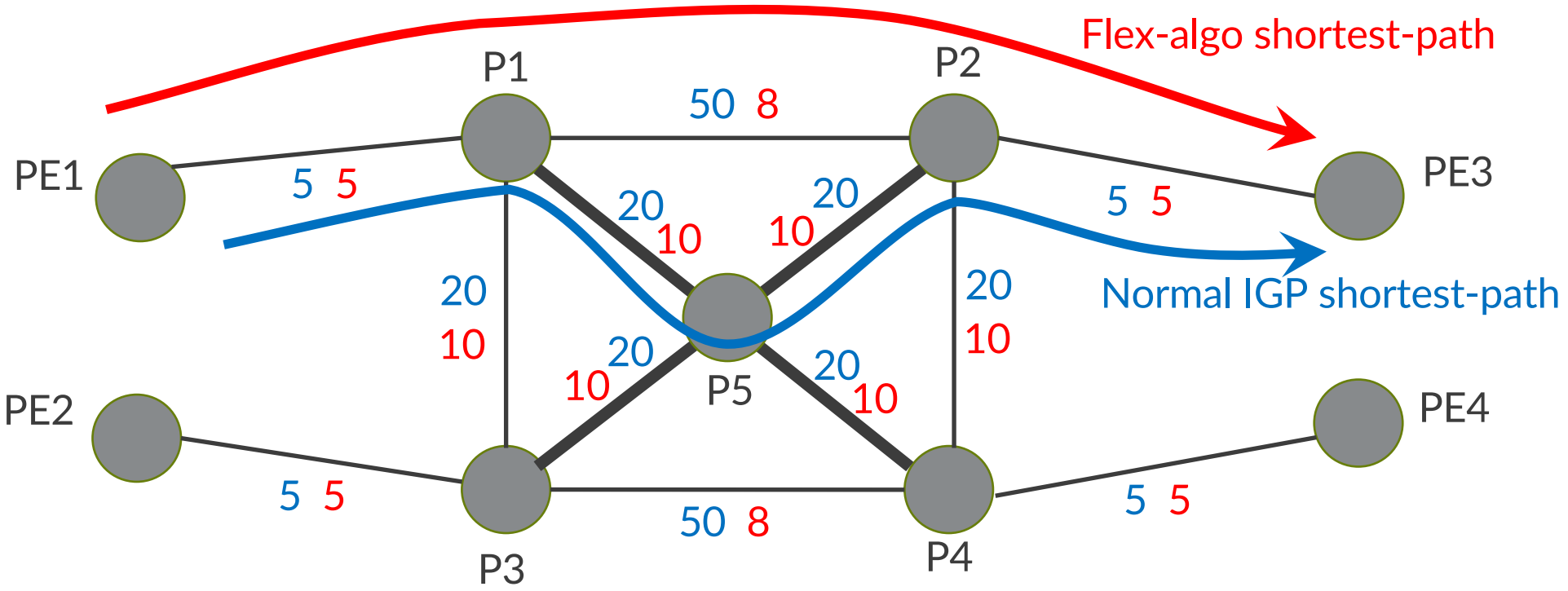


Normal IGP metric in blue font. Bulk traffic uses the normal IGP.

Red flex-algo metric in red font. Delay-sensitive traffic uses the red flex-algo.

In this example, all routers are members of the normal IGP and the red flex-algo.

METRIC ACCORDING TO TRAFFIC TYPE



FLEX-ALGO USING DELAY-METRIC

- In the IGP, on each interface configure a delay-metric (as well as the normal IGP metric).
  - The delay metric can either be static, or can be automatically derived from live TWAMP probe packet measurements.
- Configure the flex-algo to use the delay metric.

```
set routing-options flex-algorithm 129 definition priority 100

set routing-options flex-algorithm 129 definition metric-type delay-metric

set routing-options flex-algorithm 129 color 129
```

- In general, a flex-algo can be configured to use one of these metric-types:
  - normal IGP metric
  - TE-metric
  - delay-metric

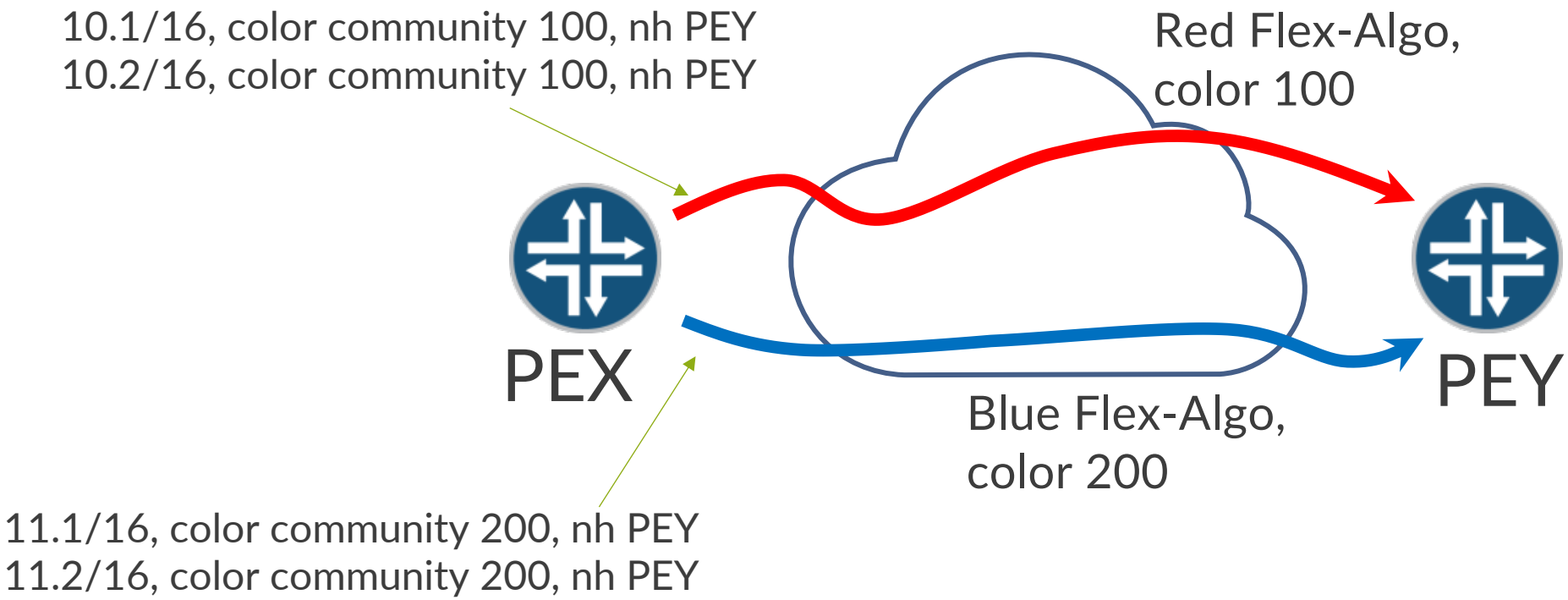
MAPPING TRAFFIC TO THE PREFERRED TRANSPORT

Colors provide a convenient way to auto-map prefixes/services to a particular transport

For example, mapping a particular VPN to a particular type of TE-tunnel or to a particular Flex-Algo.

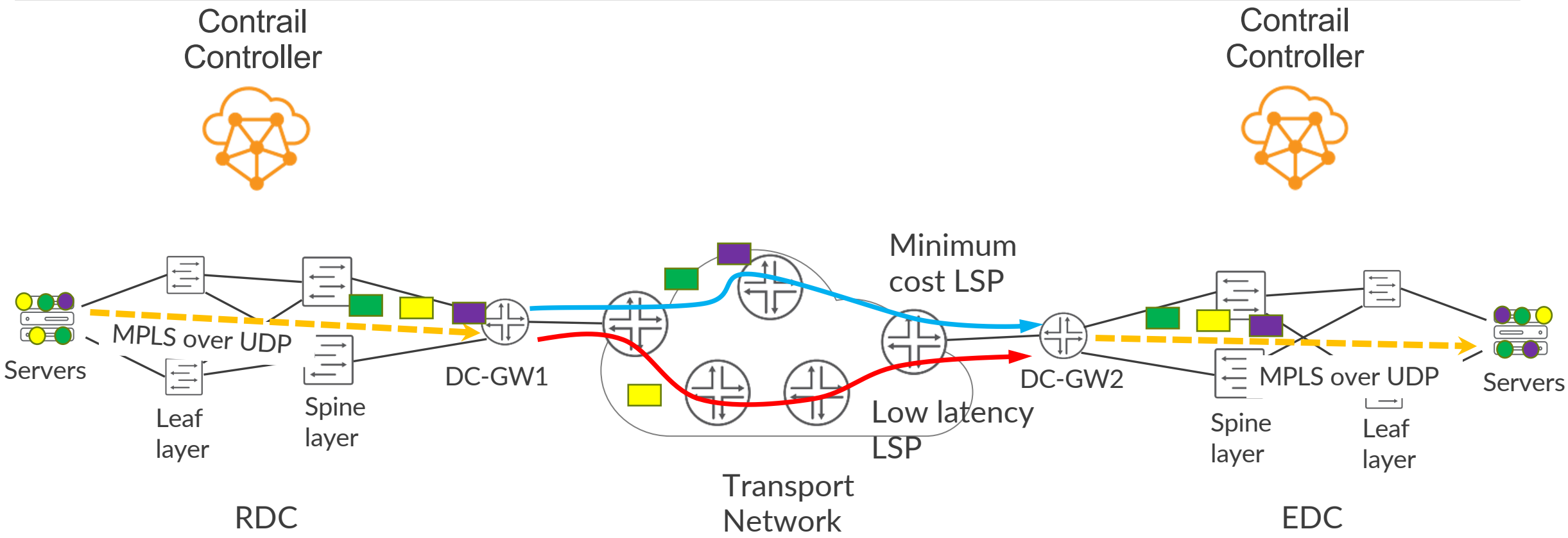
It's a very versatile approach, and also plays an important part in BGP-CT (see later)

AUTO-MAPPING OF COLORED PREFIXES ONTO COLORED TUNNELS

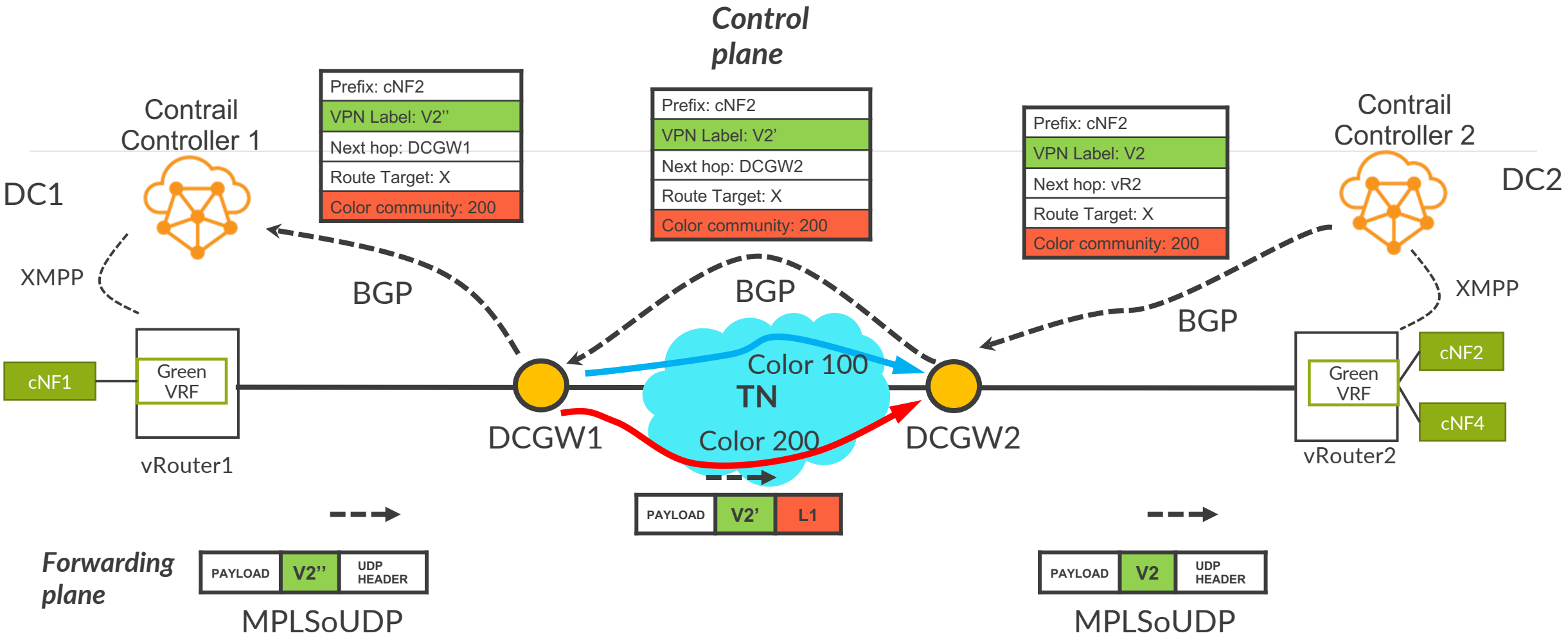


- A prefix (e.g. plain IP or VPN prefix) with a color community is automatically mapped onto the flex-algo with the matching color.
- Can also be used to map traffic onto a colored TE Tunnel (SR-TE or RSVP-TE).

SLICING ACROSS TELCO CLOUD AND TRANSPORT NETWORK



- Traffic from green and purple overlay networks (slices) need minimum cost transport.
- Traffic from yellow overlay network (slice) needs minimum cost transport.



Key point: DCGW1 puts the traffic for cNF2 onto the Flex-Algo or TE-LSP that has color 200, because the BGP prefix has color community C200.

BGP Classful Transport (BGP-CT)

BGP CLASSFUL TRANSPORT (BGP-CT)

A mechanism for extending color-mapping across multiple ASes.

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

RSVP-TE/SR-TE, with or without controller, Flex-Algo,

BGP-CT acts as the “glue” between domains

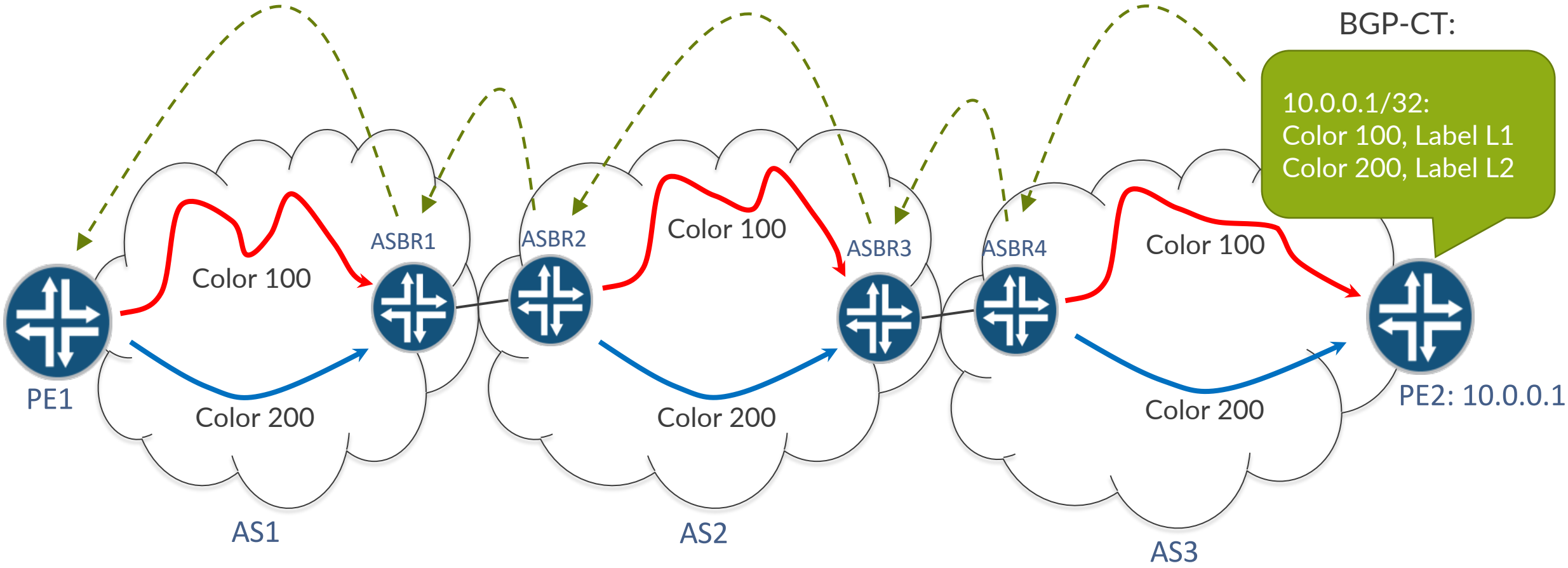
BGP-CT is similar to BGP-LU, except that it has {egress PE, color} granularity

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

FURTHER READING

Flex-Algo

https://datatracker.ietf.org/doc/draft-ietf-lsr-flex-algo/

BGP Classful Transport (BGP-CT)

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

Blog about Differentiated Transport for Cloud Overlay Networks

https://blogs.juniper.net/en-us/service-provider-transformation/differentiated-transport-across-the-wan-for-cloud-overlay-networks


Thank you
JUNIPER NETWORKS
Engineering
Simplicity