SDN Controllers in the WAN: protocols and applications

Julian Lucek

jlucek@juniper.net

@julianlucek

JUNIPER NETWORKS | Engineering Simplicity

Agenda

SDN Controller for WAN concept

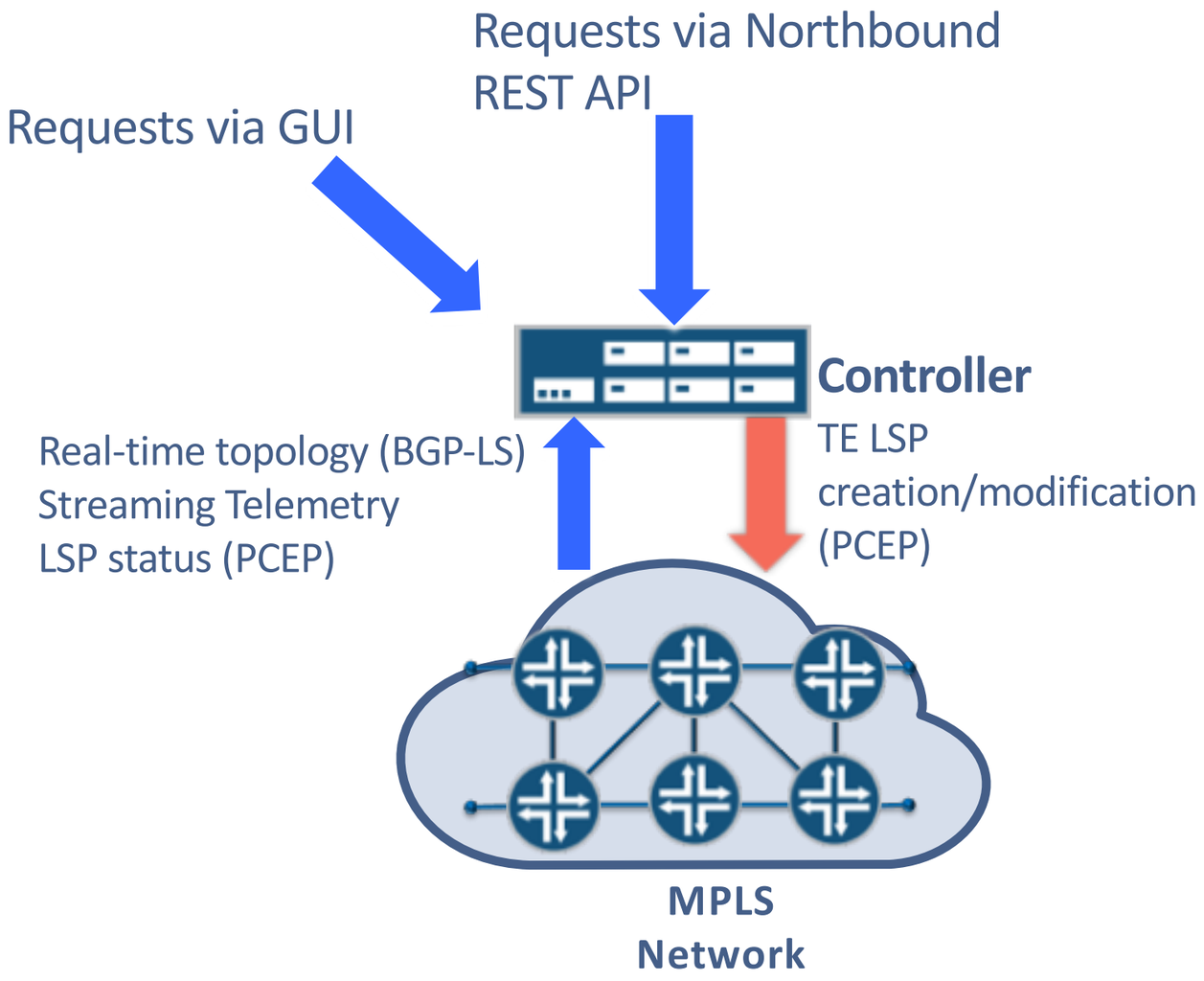
Enabling protocols

- BGP Link State (BGP-LS)
- Path Computation Element Protocol (PCEP)

Applications

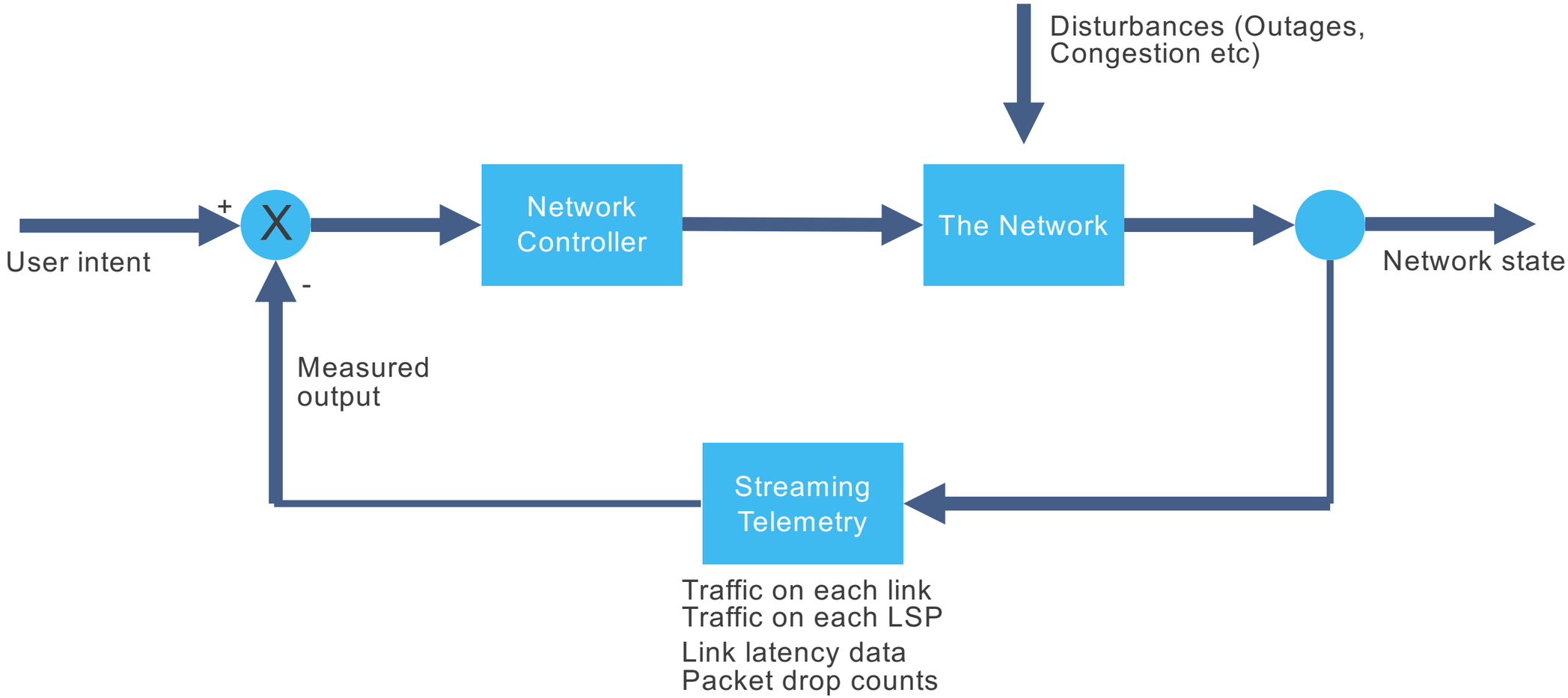
- Path Diversity
- Telemetry-driven congestion avoidance
- Dynamic minimum-latency path

WAN SDN Controller Concept



- The Controller is a real-time controller, directly coupled to the live network (rather than being a passive observer or an offline planning tool)
- It is part of the *control plane* of the network
- It receives input from the network itself via control plane protocols and Streaming Telemetry
- It receives input from a human operator via a GUI, and from orchestrators/OSS via a northbound REST API
  - E.g. requests to set up LSPs with particular attributes
- Shipping for 3 years, now deployed in Telcos, ISPs, Transit Providers, high-end Enterprises

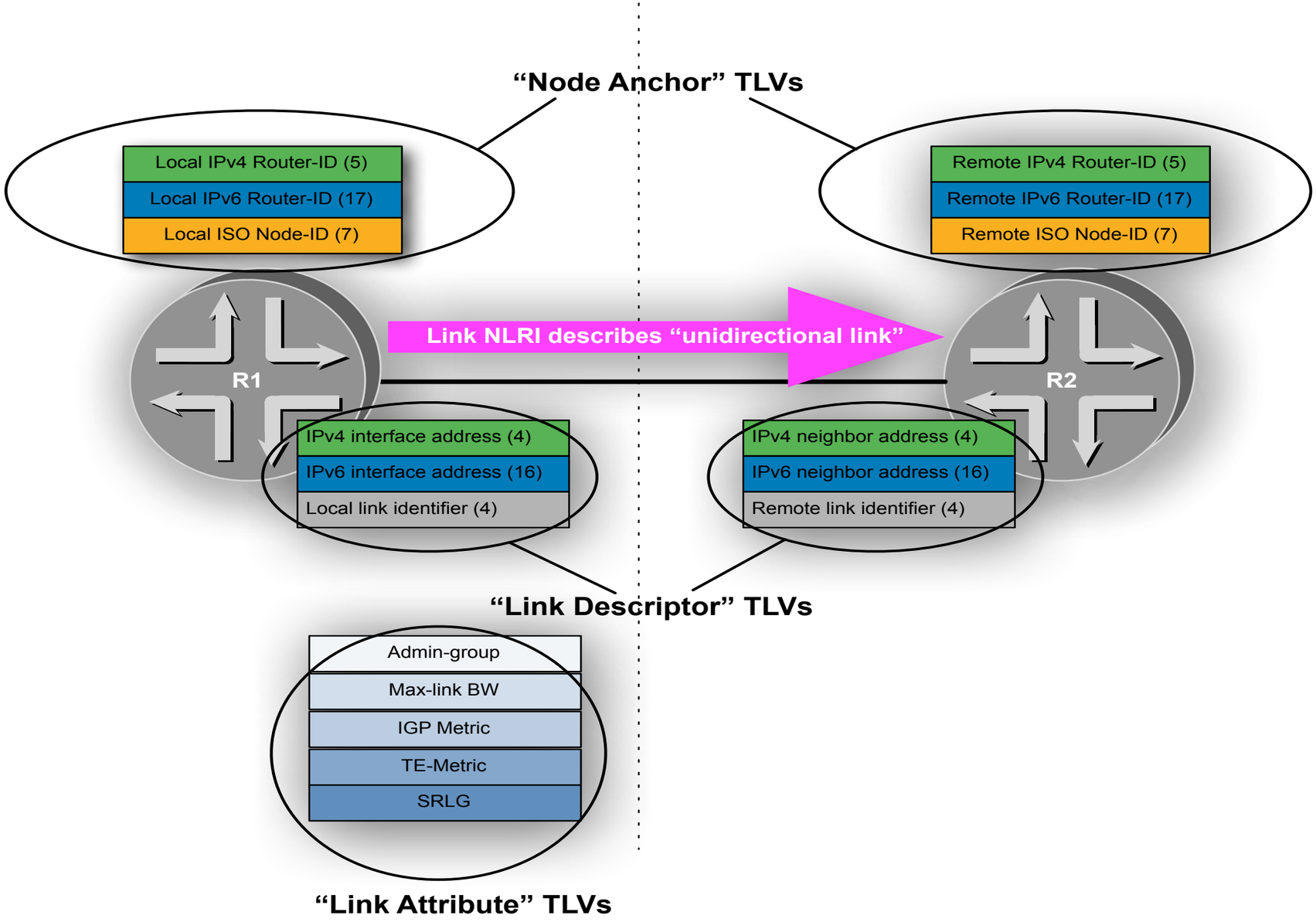
The Power of Telemetry in Adaptive Control Systems



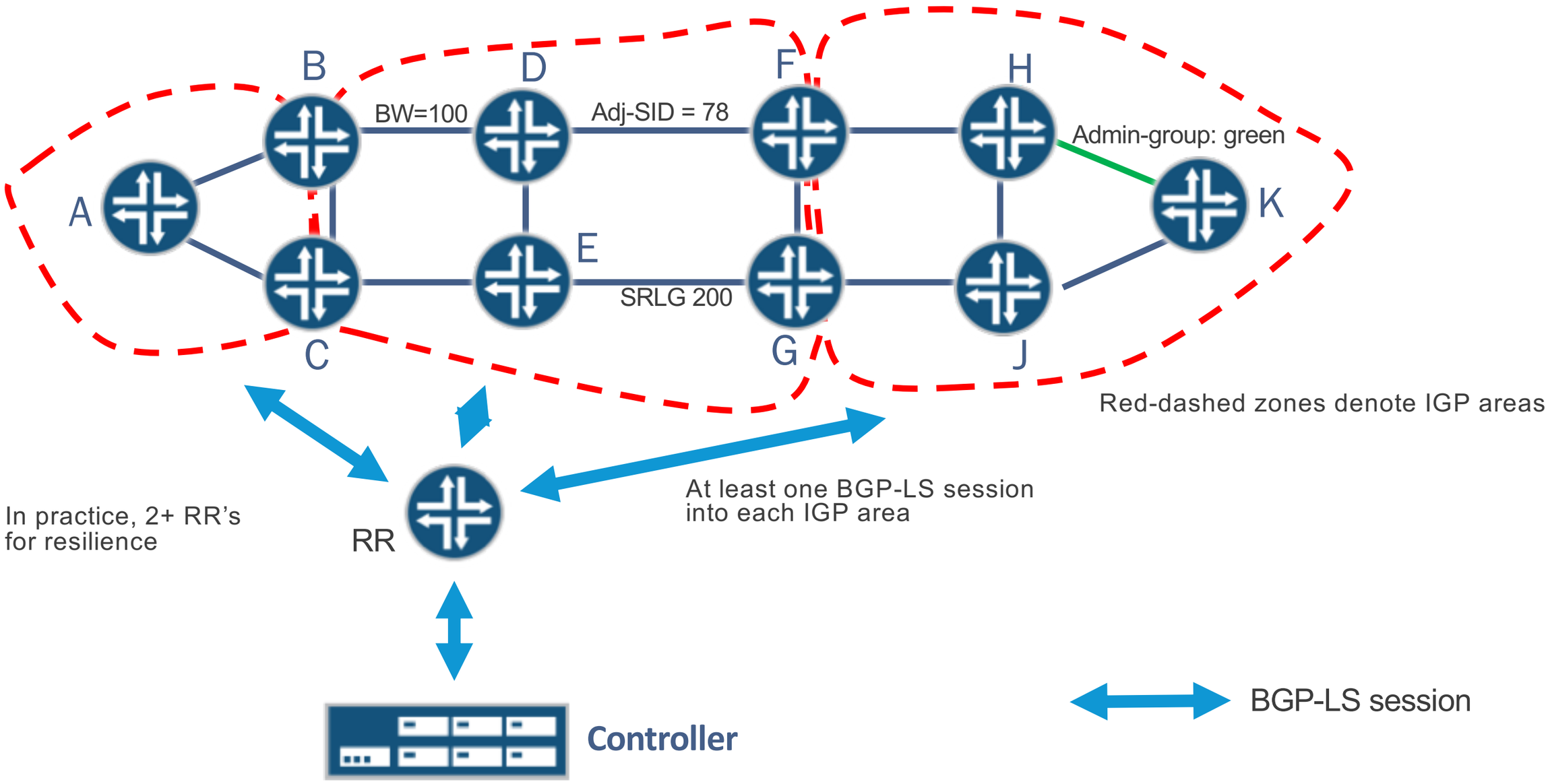
Streaming Telemetry allows much more frequent updates than SNMP. Push paradigm, rather than request/response. Stats collected on linecard are sent directly from there without passing through control processor.

BGP Link State (BGP-LS)

BGP-LS TLVs


"Node Anchor" TLVs
Local IPv4 Router-ID (5)
Local IPv6 Router-ID (17)
Local ISO Node-ID (7)
Remote IPv4 Router-ID (5)
Remote IPv6 Router-ID (17)
Remote ISO Node-ID (7)
Link NLRI describes "unidirectional link"
R1
R2
IPv4 interface address (4)
IPv6 interface address (16)
Local link identifier (4)
IPv4 neighbor address (4)
IPv6 neighbor address (16)
Remote link identifier (4)
"Link Descriptor" TLVs
Admin-group
Max-link BW
IGP Metric
TE-Metric
SRLG
"Link Attribute" TLVs

BGP-LS in action



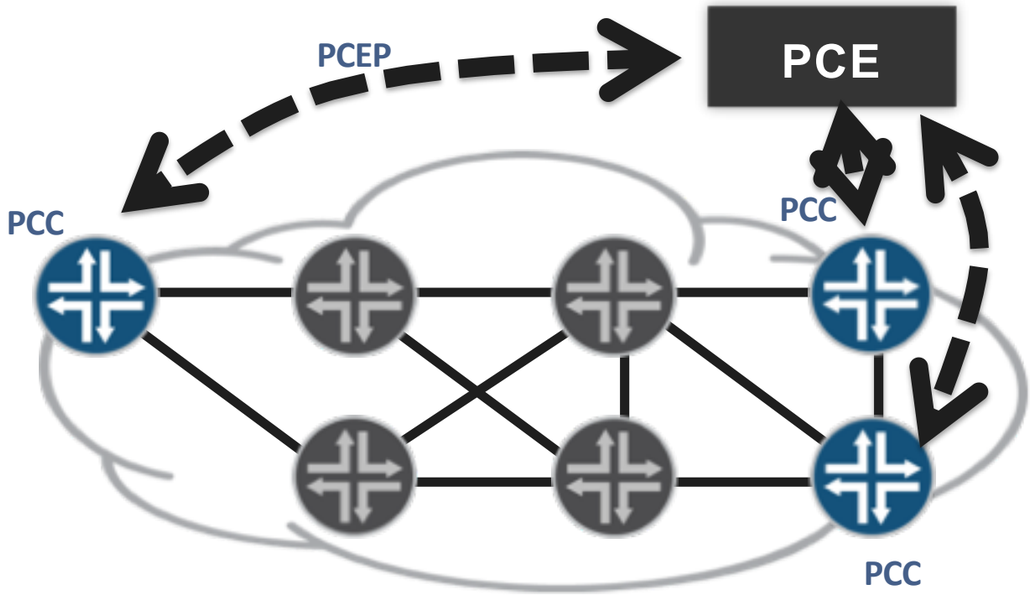
Path Computation Element Protocol (PCEP)

PCE: A standards-based approach

What is it?

- PCE: Path Computation Element (RFC 4655)

  An entity that can calculate paths in the network



Components

- PCE: Path Computation Element

  Computes the path

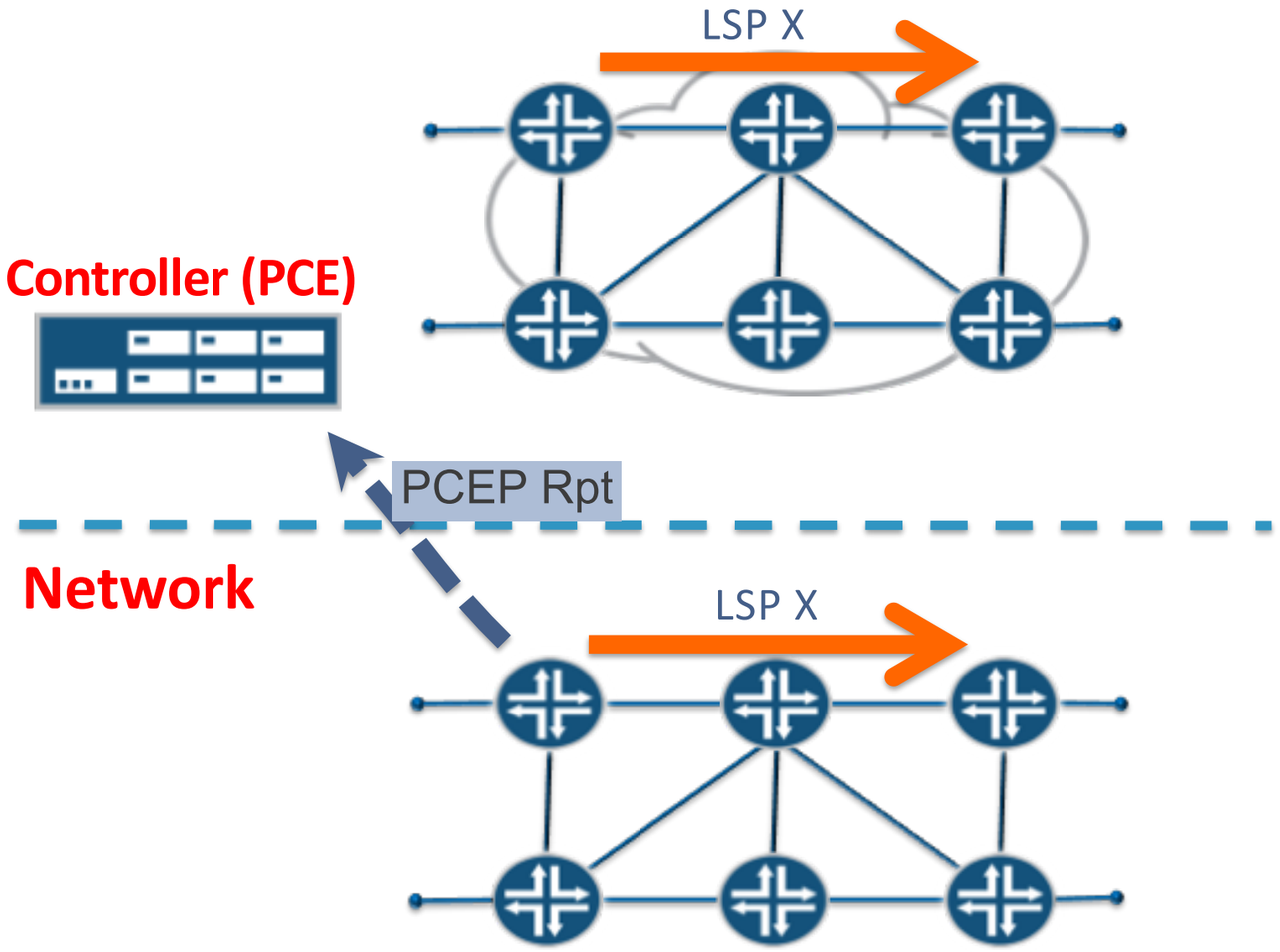
- PCC: Path Computation Client

  Receives the path. Sets up LSP using RSVP or SPRING.

- PCEP: PCE protocol (RFC 5440)

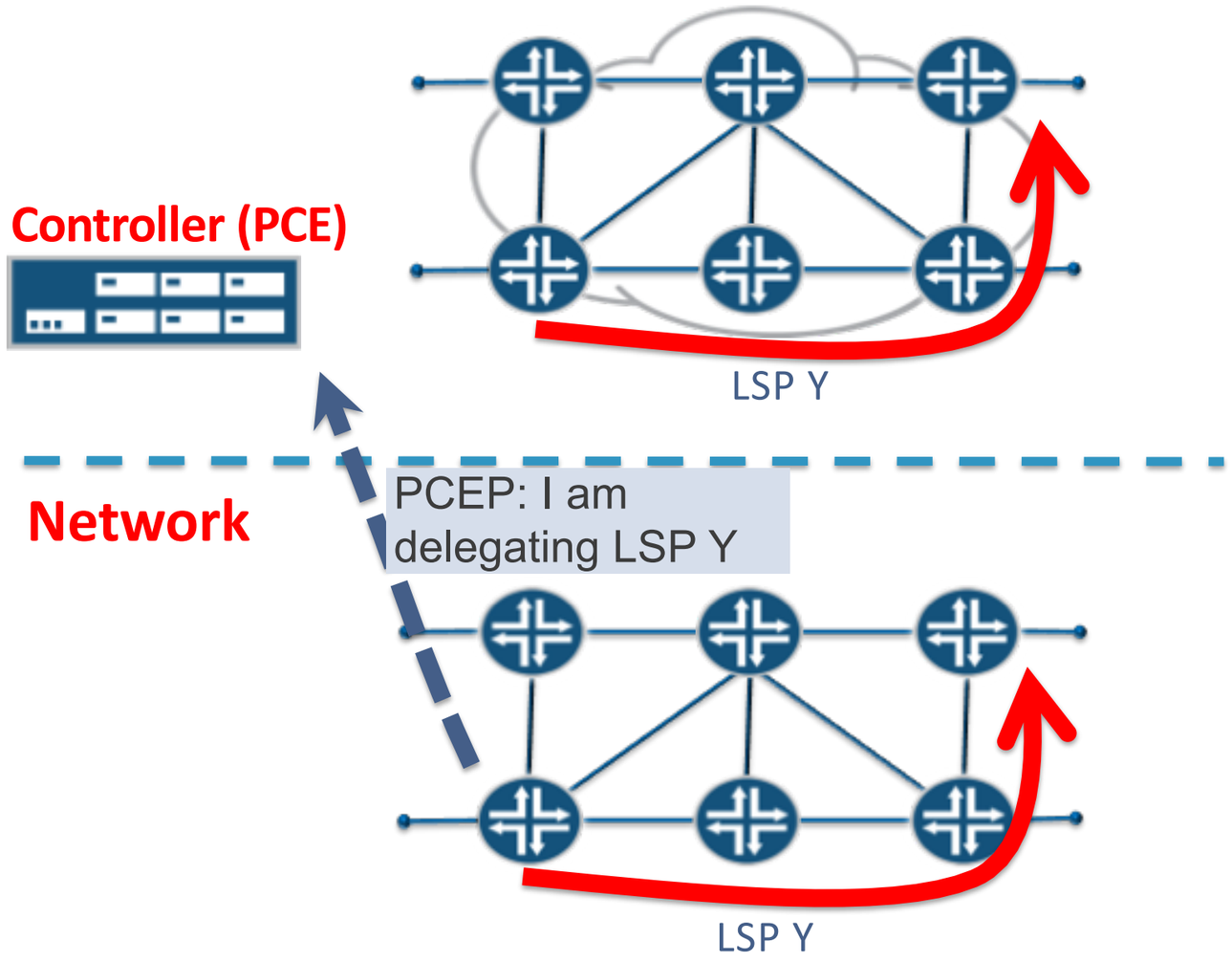
  For PCE/PCC communication

“Vanilla” LSP



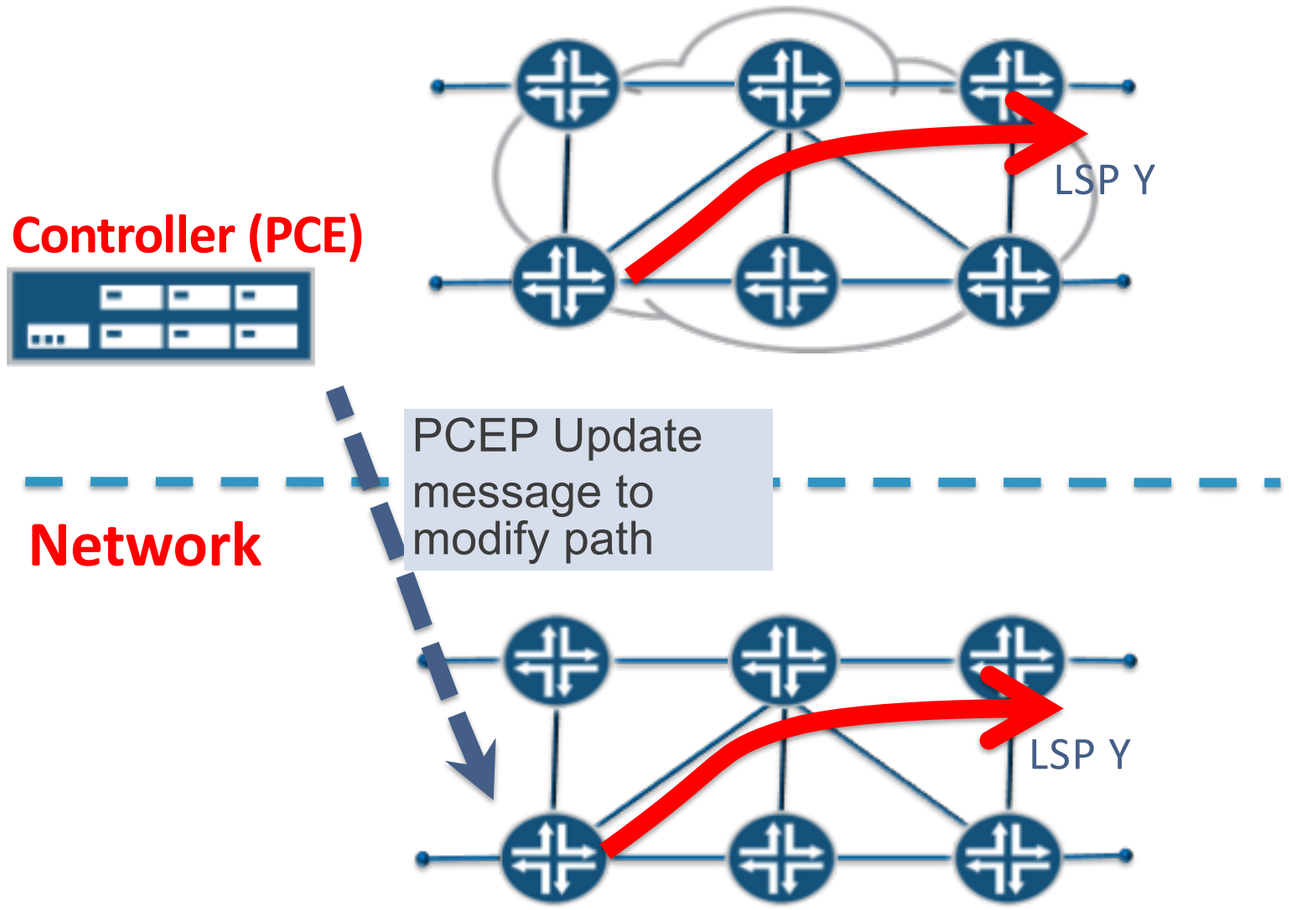
- Vanilla LSP is configured on ingress router
- Ingress router reports parameters of LSP to the PCE via PCEP, e.g.
  - Full Path
  - Bandwidth reservation
  - Priority
  - Status (Up/Down)
- PCE is *not* authorized to modify the LSP
- Nevertheless, PCEP very useful for visualizing status/path of LSP, and its history

Delegated LSP



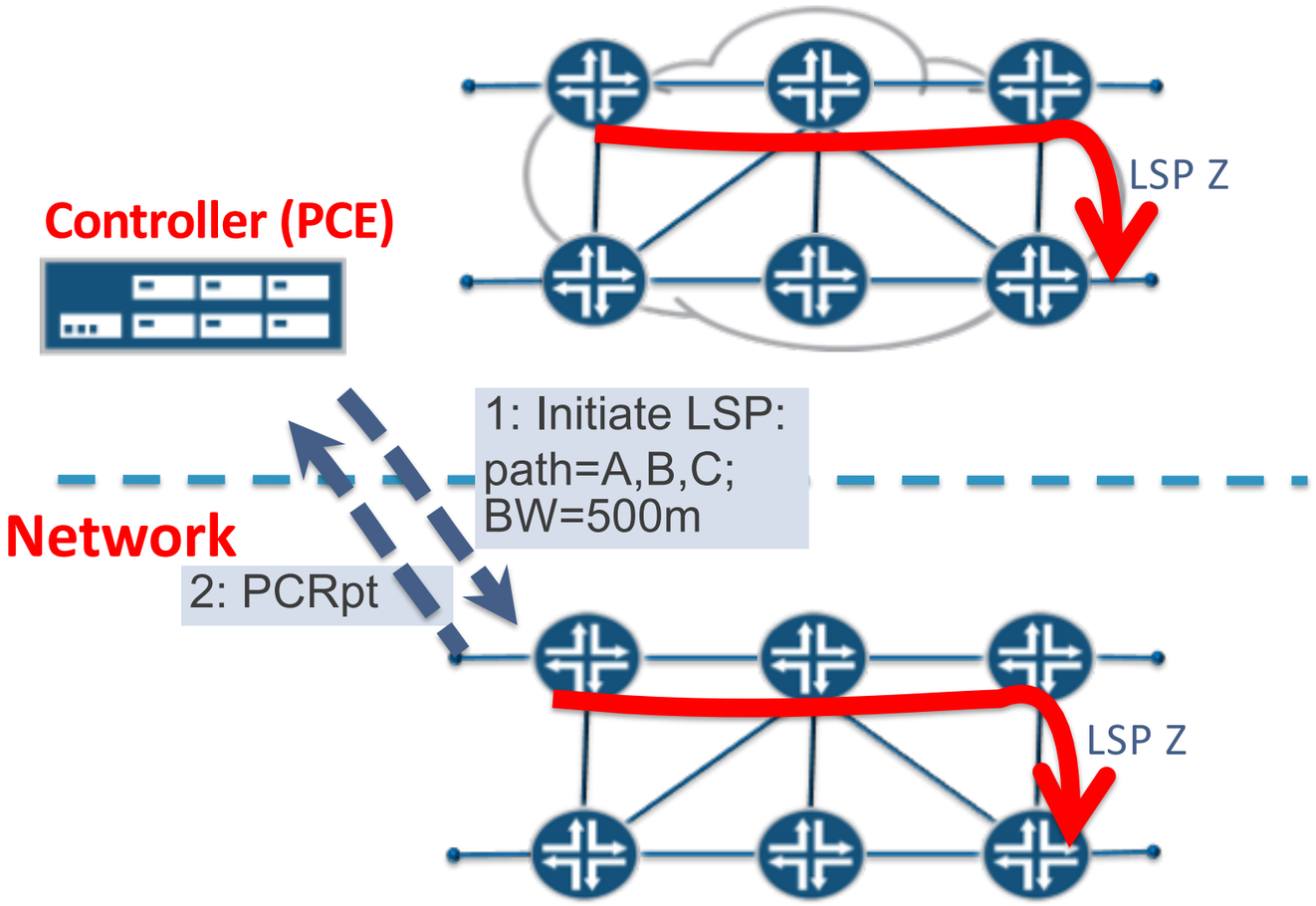
- LSP is configured on ingress router
- LSP can be delegated to controller (via CLI config), either at the same time that the LSP is configured, or sometime later

Delegated LSP (cont'd)



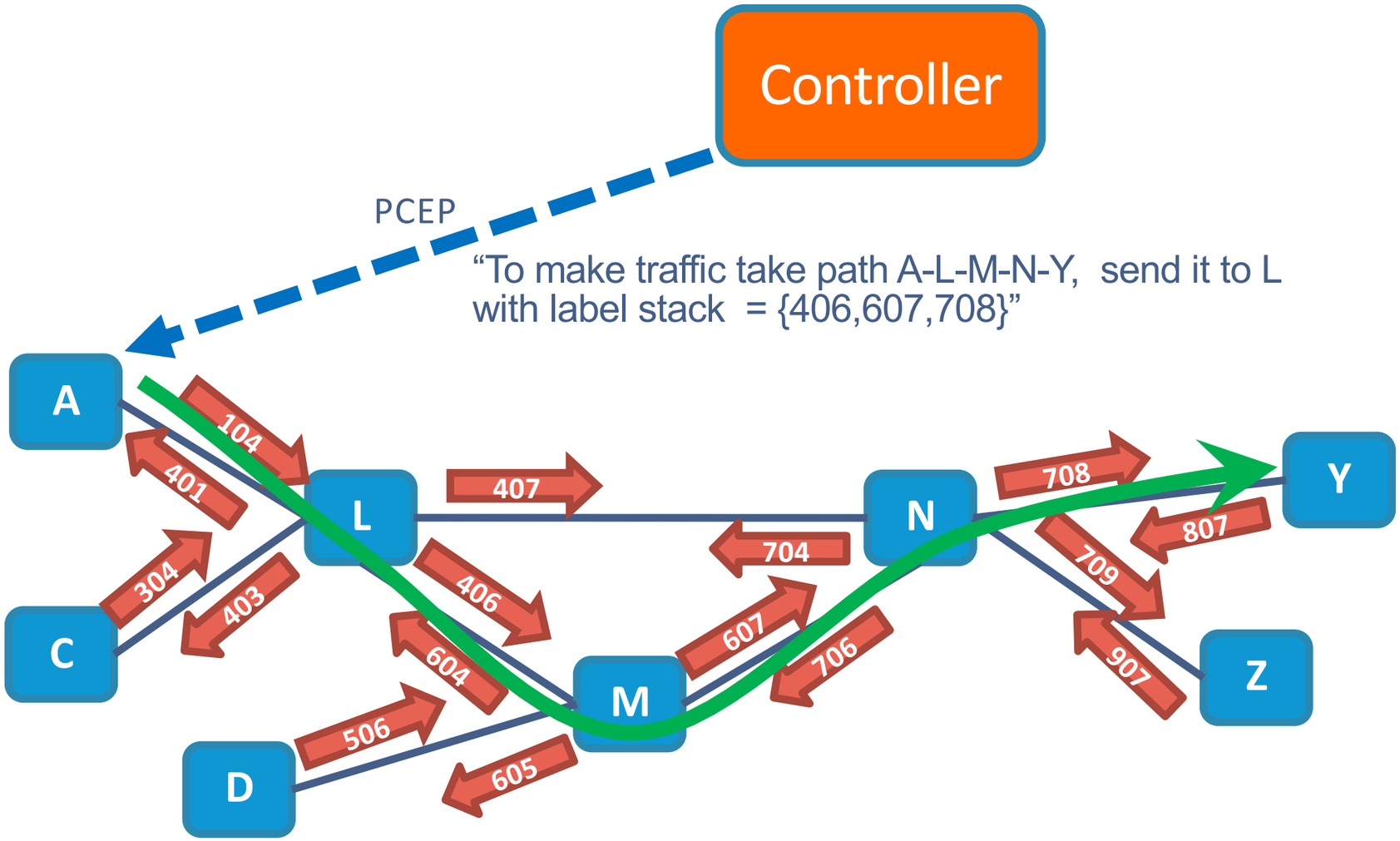
- Once the LSP is delegated to the PCE, the PCE can modify parameters of the LSP at any time, if needed

PCE-initiated LSP



- PCE computes LSP path and sends LSP set-up request via PCEP to ingress router
- Ingress router signals LSP through network, and sends confirmation via PCEP to PCE
- PCE can subsequently modify the LSP as required. PCE can request tear down of LSP when no longer needed.

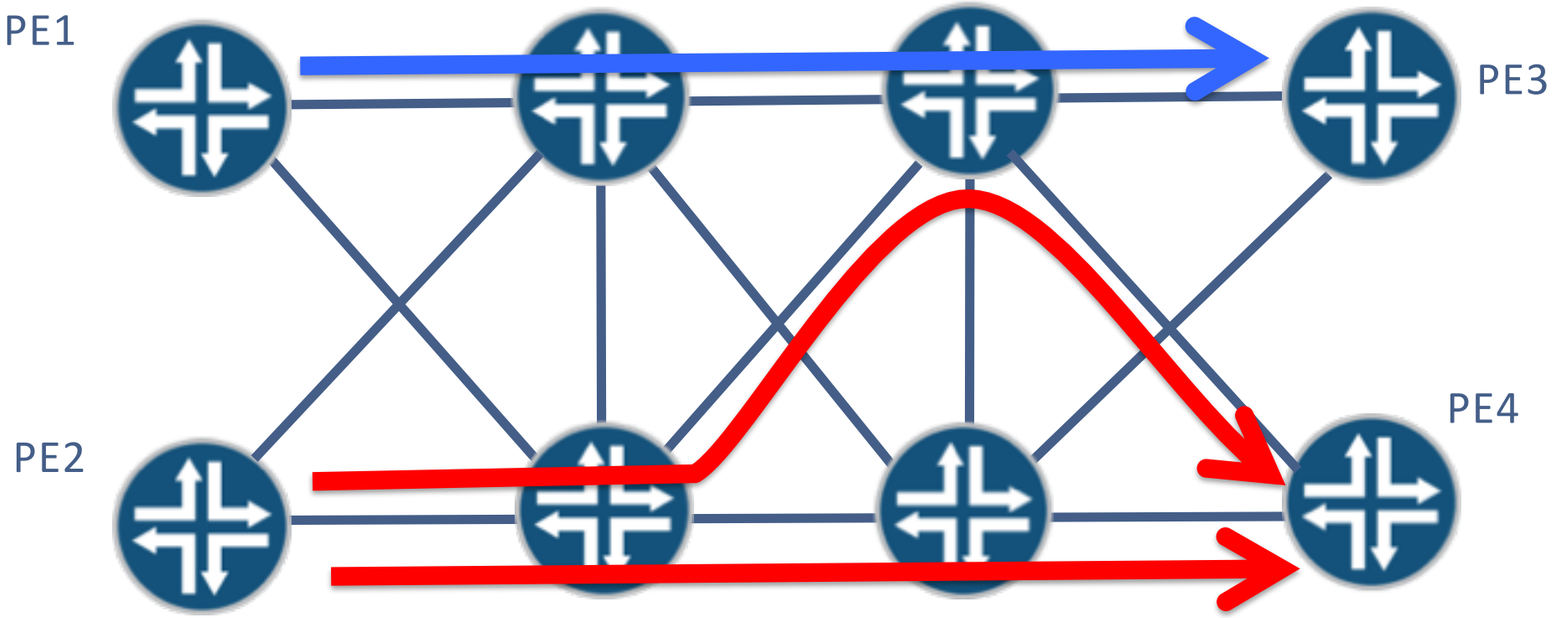
PCEP for Segment Routing



Orange arrows show Adj-SIDs

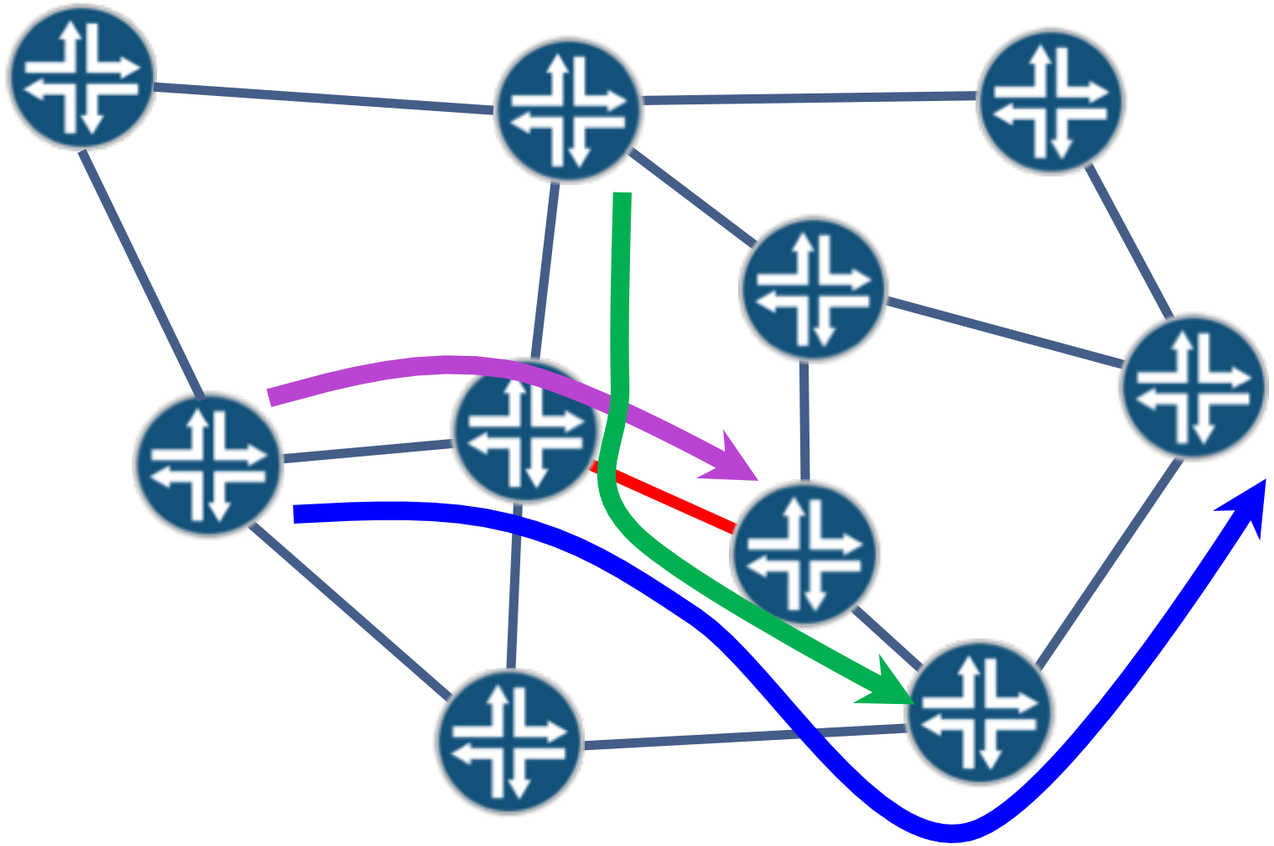
Applications of SDN Controllers in the WAN

LSP Path Diversity



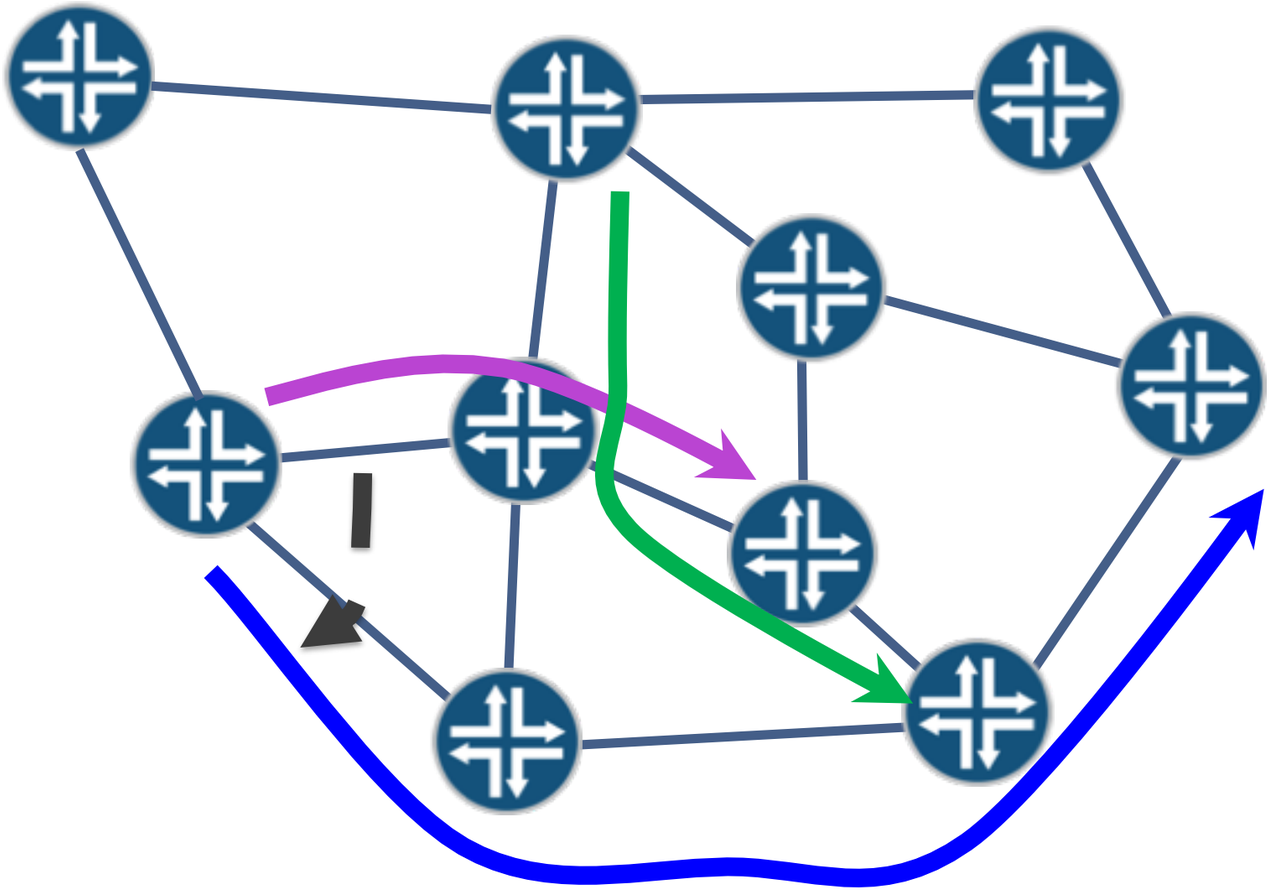
- The two paths must not have any nodes or links in common – including the PEs
- Difficult to achieve path diversity when each ingress node calculates its own path – probability of no fate-sharing at all is only 1 in 16!
- Central controller co-computes both paths to ensure diversity.
- Applications: path-diverse pseudowires, signaling traffic in mobile networks, broadcast TV

Automated actions, based on link utilisation (1)



Controller knows via Streaming Telemetry that red link is currently experiencing high utilization

Automated actions, based on link utilisation (2)



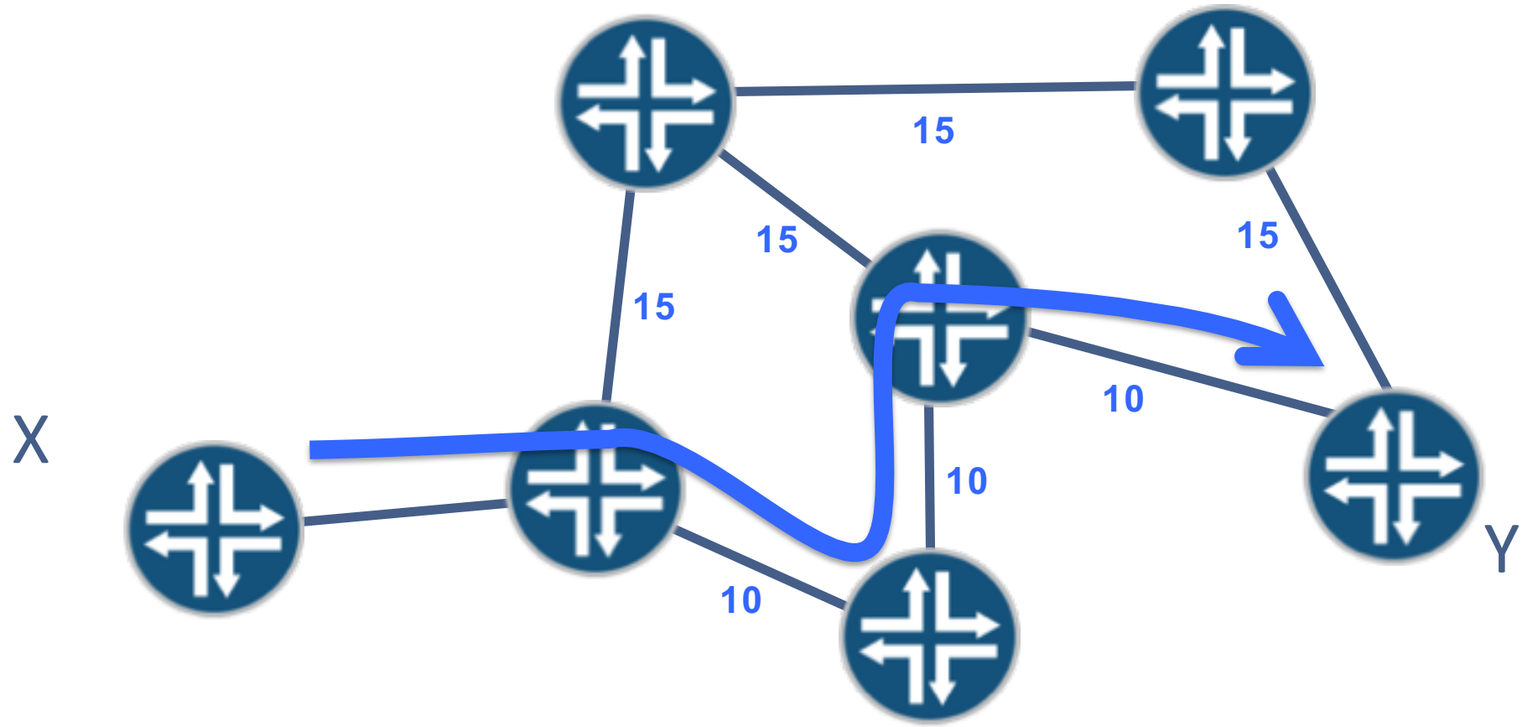
Controller also knows via Streaming Telemetry how much traffic is travelling on each LSP

So it automatically moves away some LSPs from the congested link

Other triggers for moving LSPs: interface error counts, queue high-water marks, planned maintenance window

Programmable cost function

Cost function = lowest IGP metric path that meets the required path constraints (BW etc)

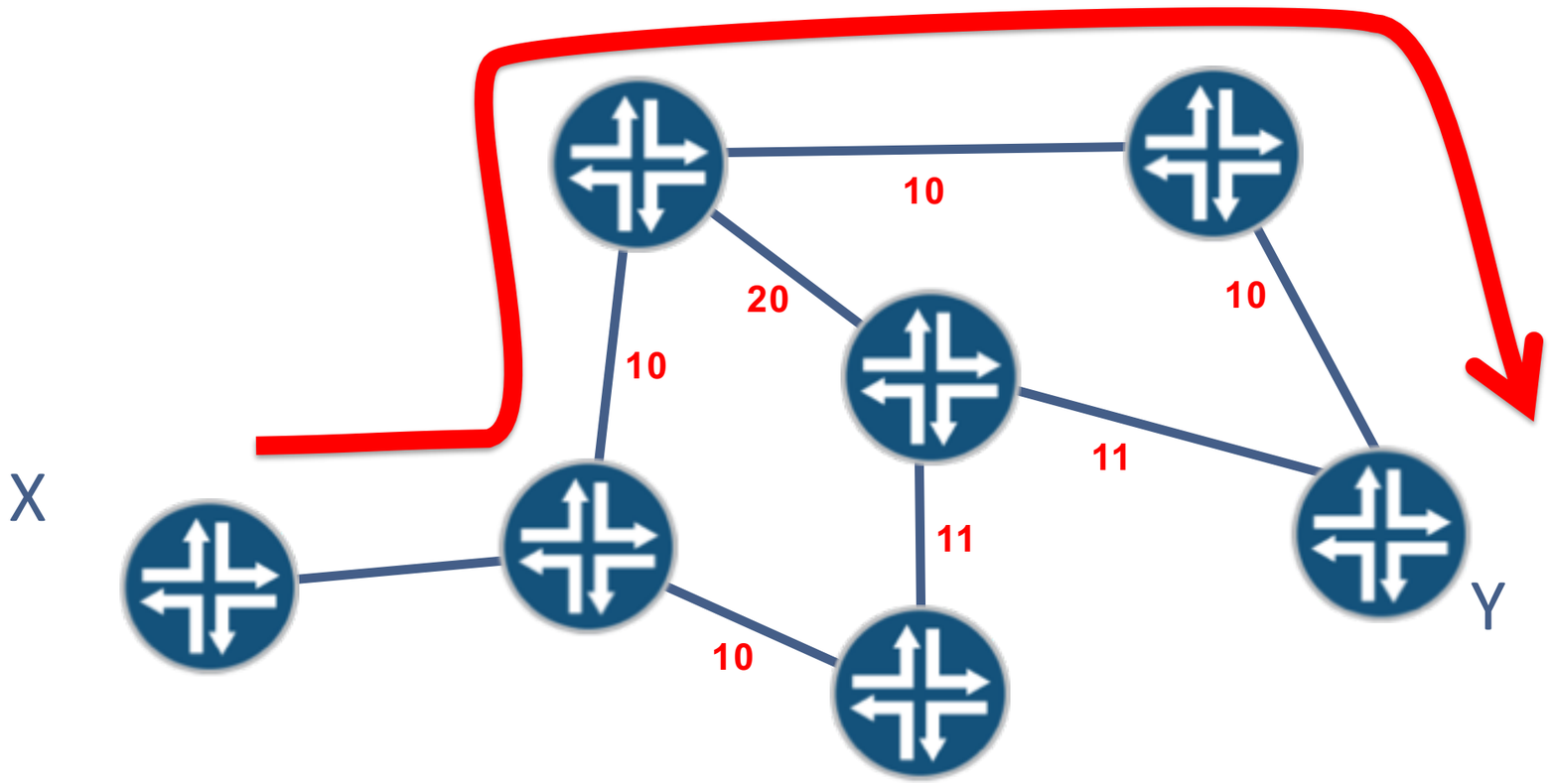


Lowest IGP metric path

Blue numbers show IGP metric

Programmable cost function (cont'd)

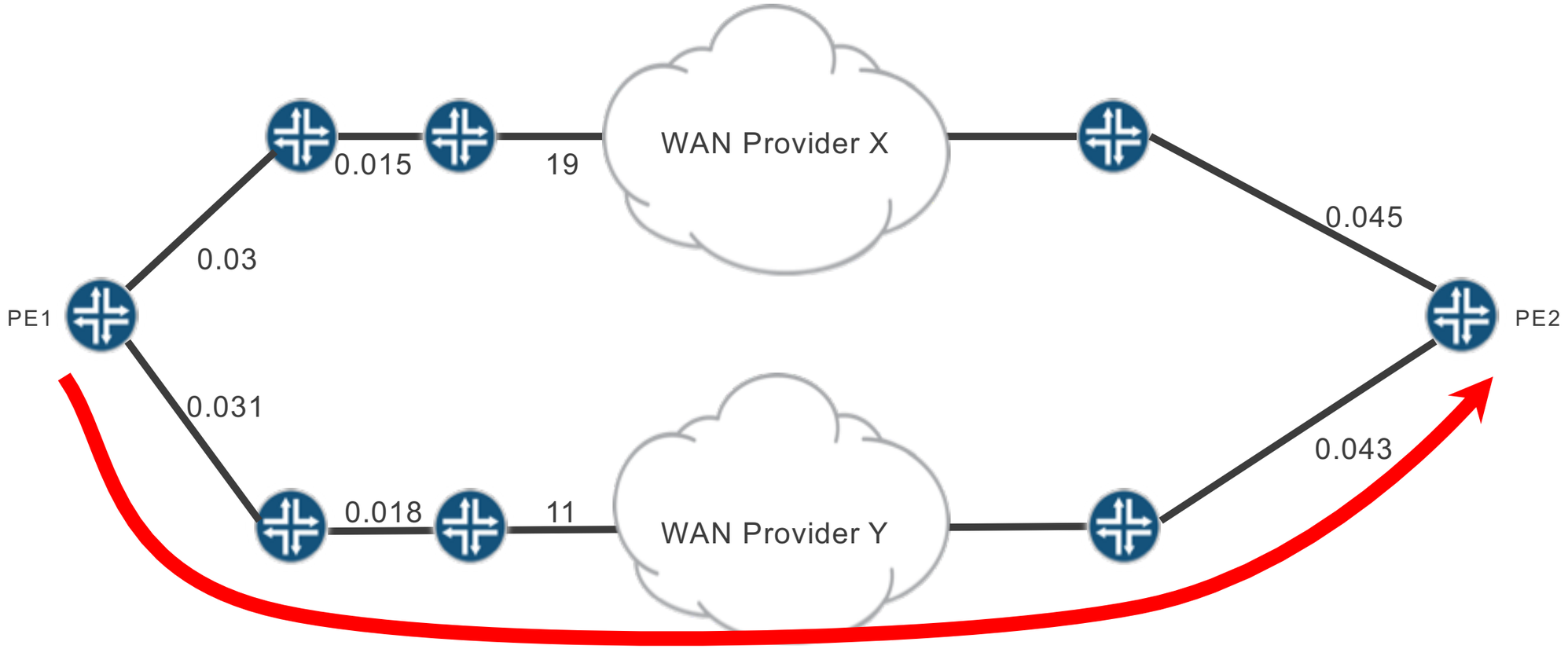
Cost function = lowest latency path that meets the required path constraints (BW etc)



Lowest latency path →

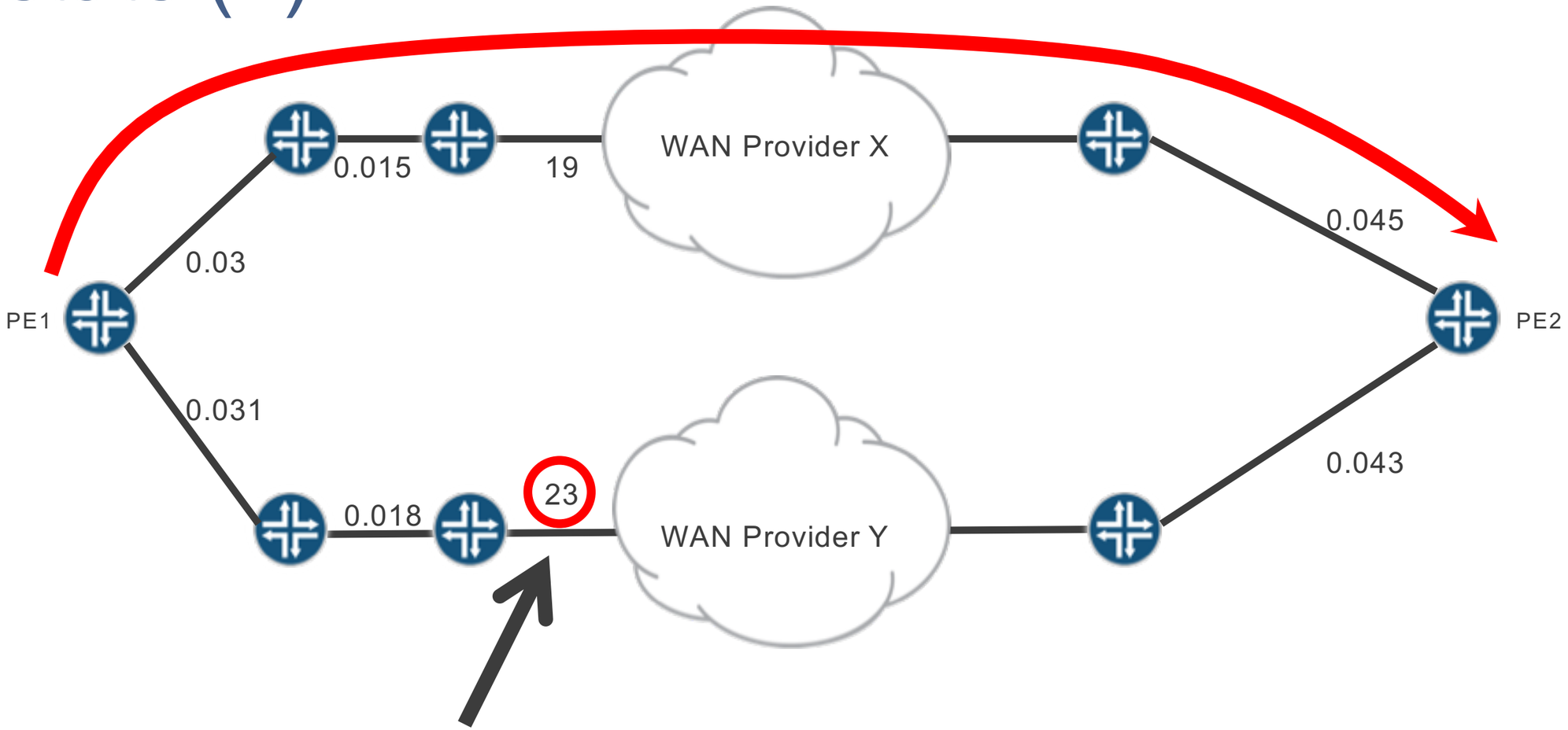
Red numbers show latency (reported via Streaming Telemetry)

Automated actions, based on latency stats (1)



Link delay values are measured by routers and reported to Controller periodically.
Red LSP carries delay-sensitive traffic so Controller uses delay as the cost function when computing the path.

Automated actions, based on based on latency stats (2)



Step change in delay via Provider Y (due to reroute or protection event within Provider Y)
Controller reroutes red LSP via Provider X.

Useful References

- PCEP and BGP-LS deep-dive webinar
  - https://www.ipspace.net/PCEP_and_BGP-LS_Deep_Dive
- IETF PCEP Working Group drafts and RFCs
  - https://datatracker.ietf.org/wg/pce/documents/
- BGP-LS RFC
  - https://tools.ietf.org/html/rfc7752
- NorthStar Controller documentation
  - https://www.juniper.net/documentation/en_US/northstar4.0.0/information-products/pathway-pages/4.0.0/index.html
- See blogs at https://forums.juniper.net/t5/user/viewprofilepage/user-id/24095