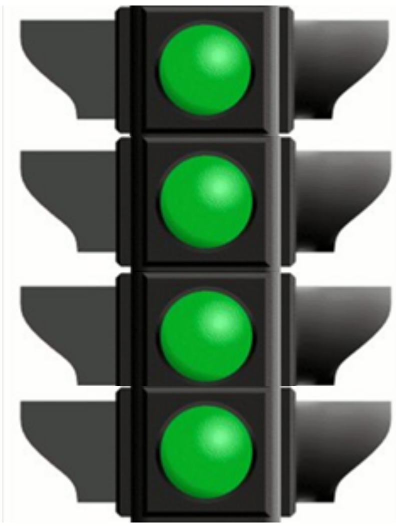
Segment Routing

Berényi Áron

Deployment

- In CY2015, SR will be deployed in all of these markets



WEB

SP Core/Edge

SP Agg/Metro

Large Entreprise

Agenda

- Technology Overview
- Use Cases
- Control and Data Plane
- Traffic Protection

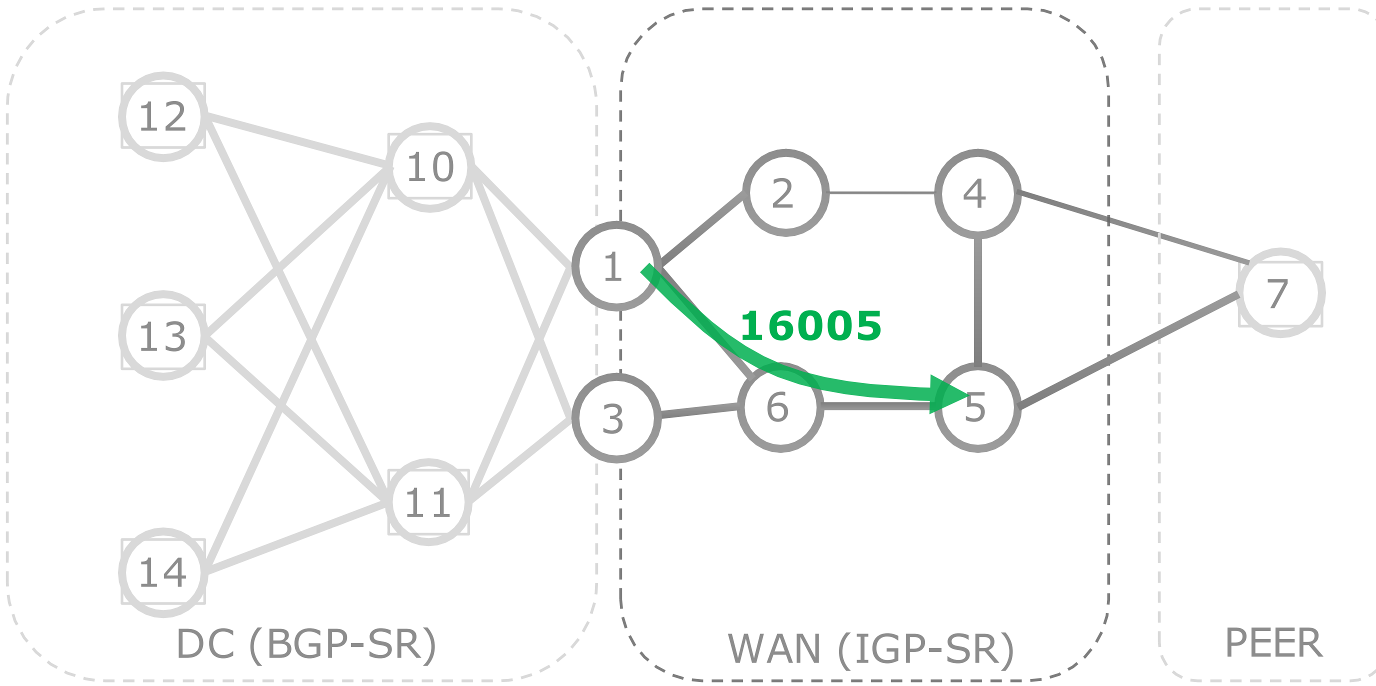
Technology Overview

Segment Routing

- Source Routing
  - the source chooses a path and encodes it in the packet header as an ordered list of segments
  - the rest of the network executes the encoded instructions without any further per-flow state
- Segment: an identifier for any type of instruction
  - forwarding or service

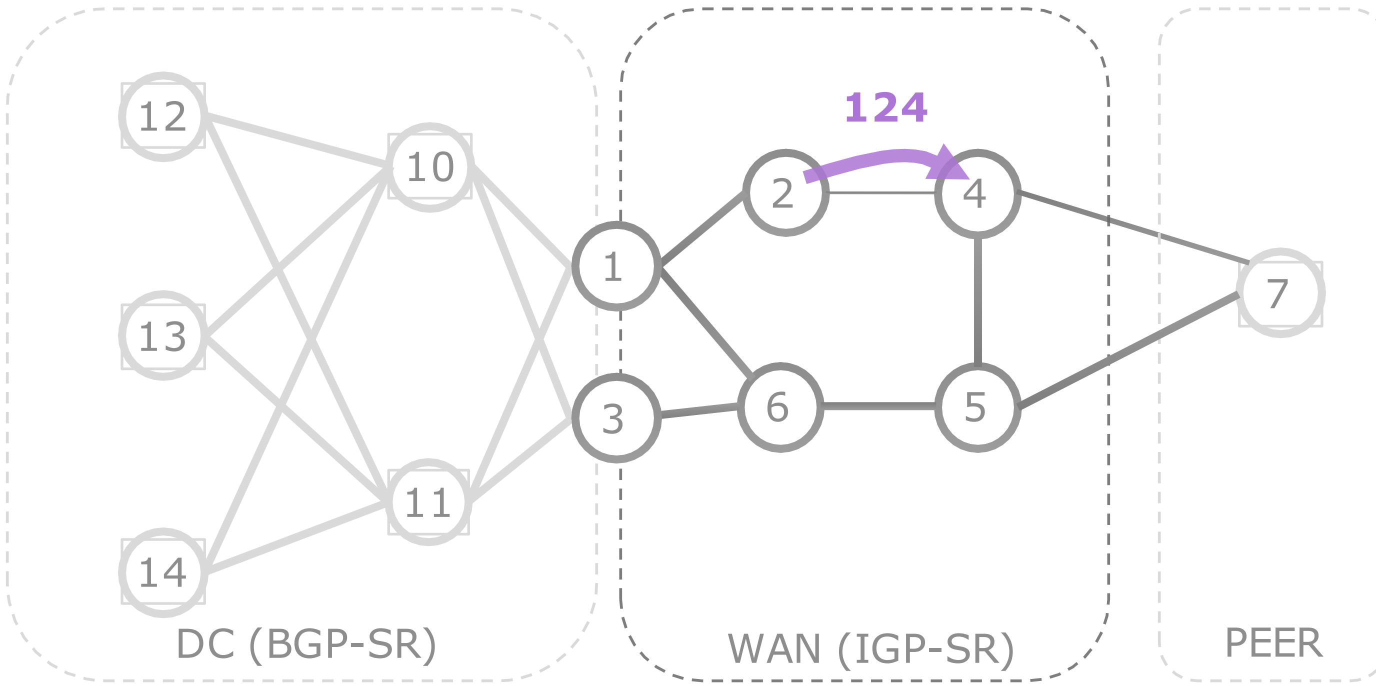
IGP Prefix Segment

- Shortest-path to the IGP prefix
- Global
- 16000 + Index
- Signaled by ISIS/OSPF



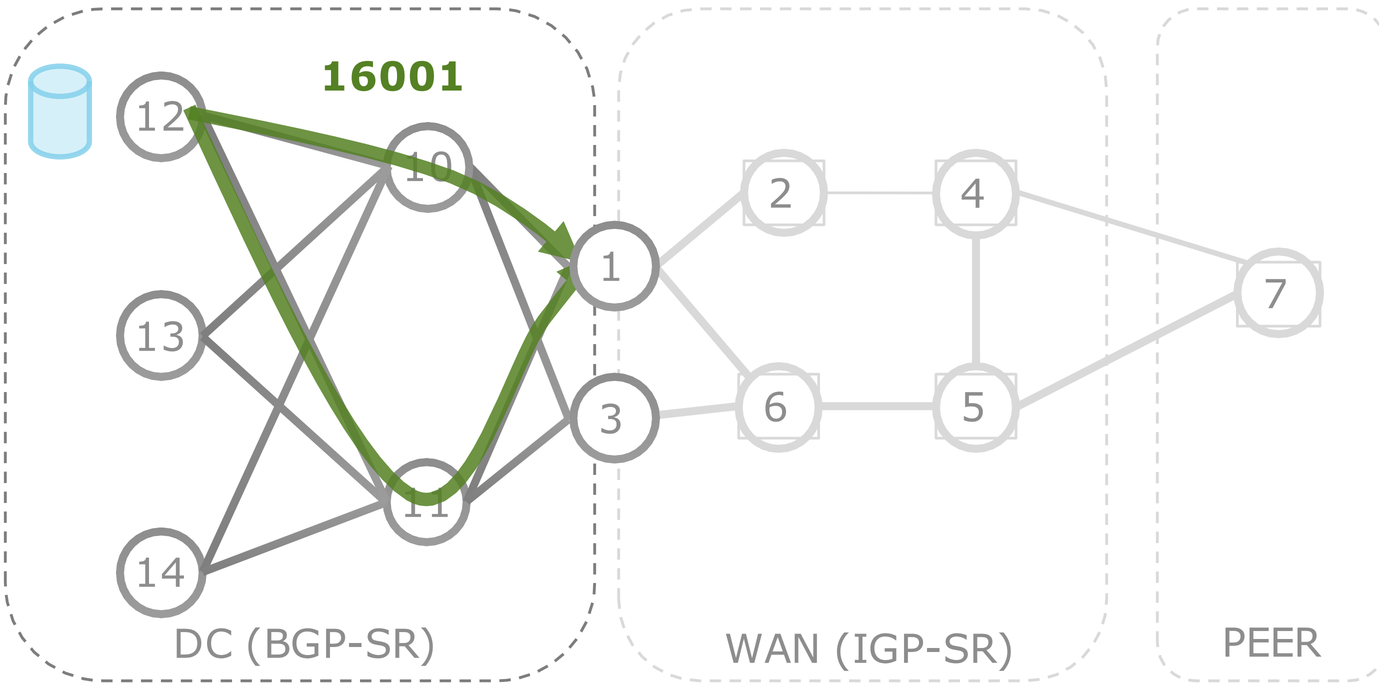
IGP Adjacency Segment

- Forward on the IGP adjacency
- Local
- 1XY
  - X is the "from"
  - Y is the "to"
- Signaled by ISIS/OSPF



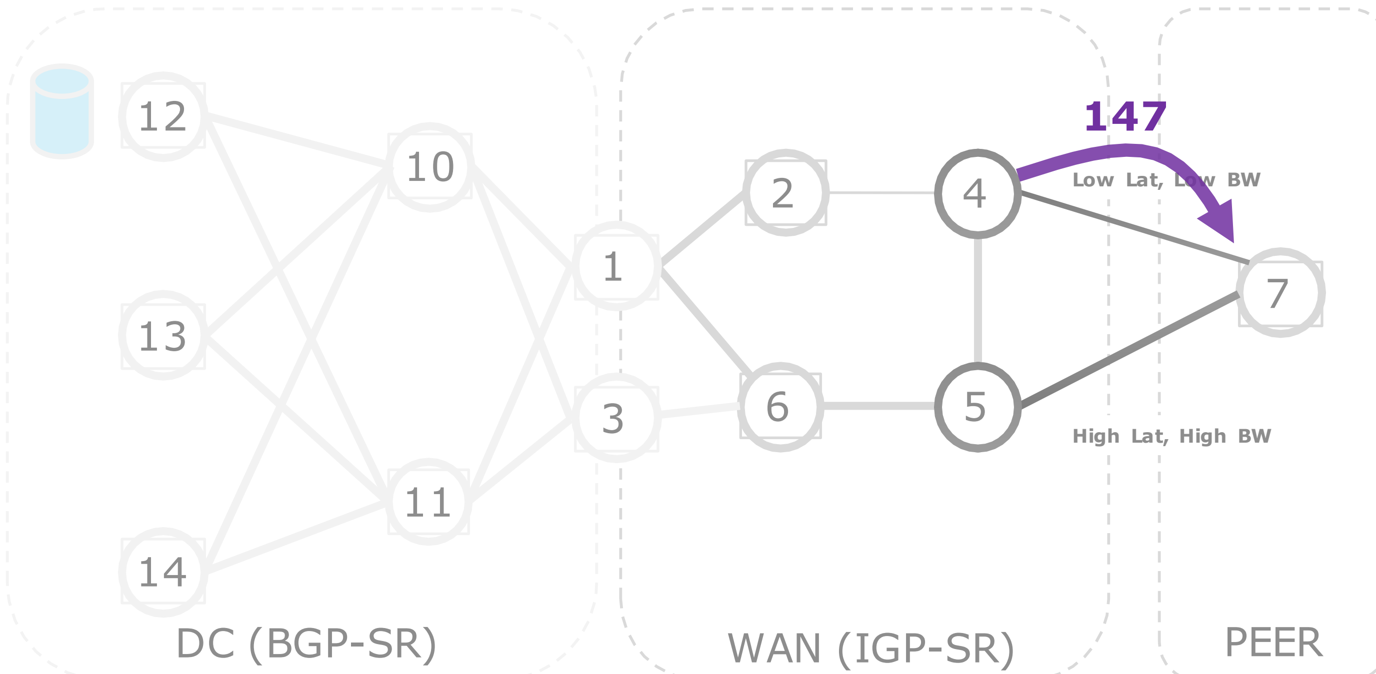
BGP Prefix Segment

- Shortest-path to the BGP prefix
- Global
- 16000 + Index
- Signaled by BGP



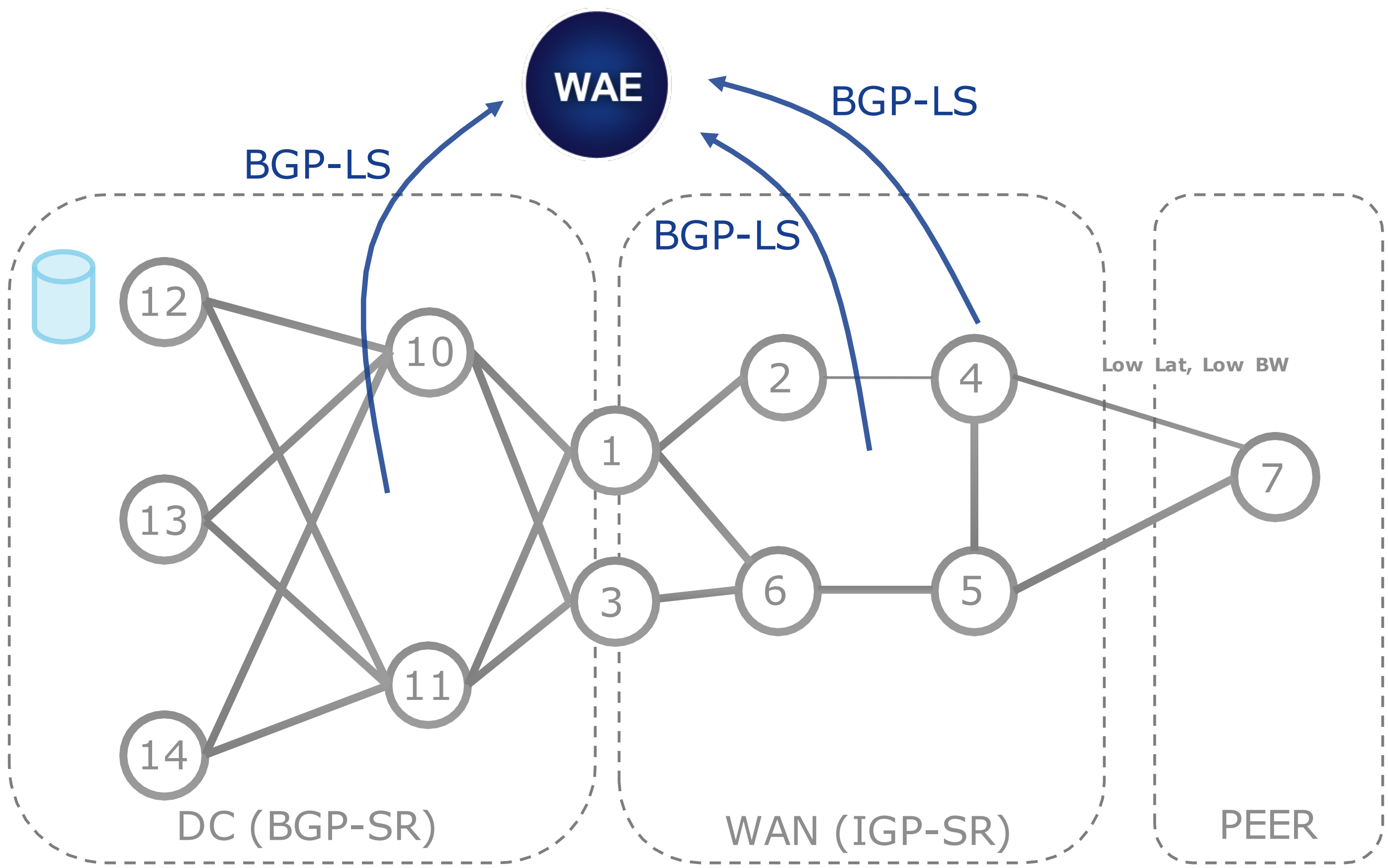
BGP Peering Segment

- Forward to the BGP peer
- Local
- 1XY
  - X is the "from"
  - Y is the "to"
- Signaled by BGP-LS (topology information) to the controller



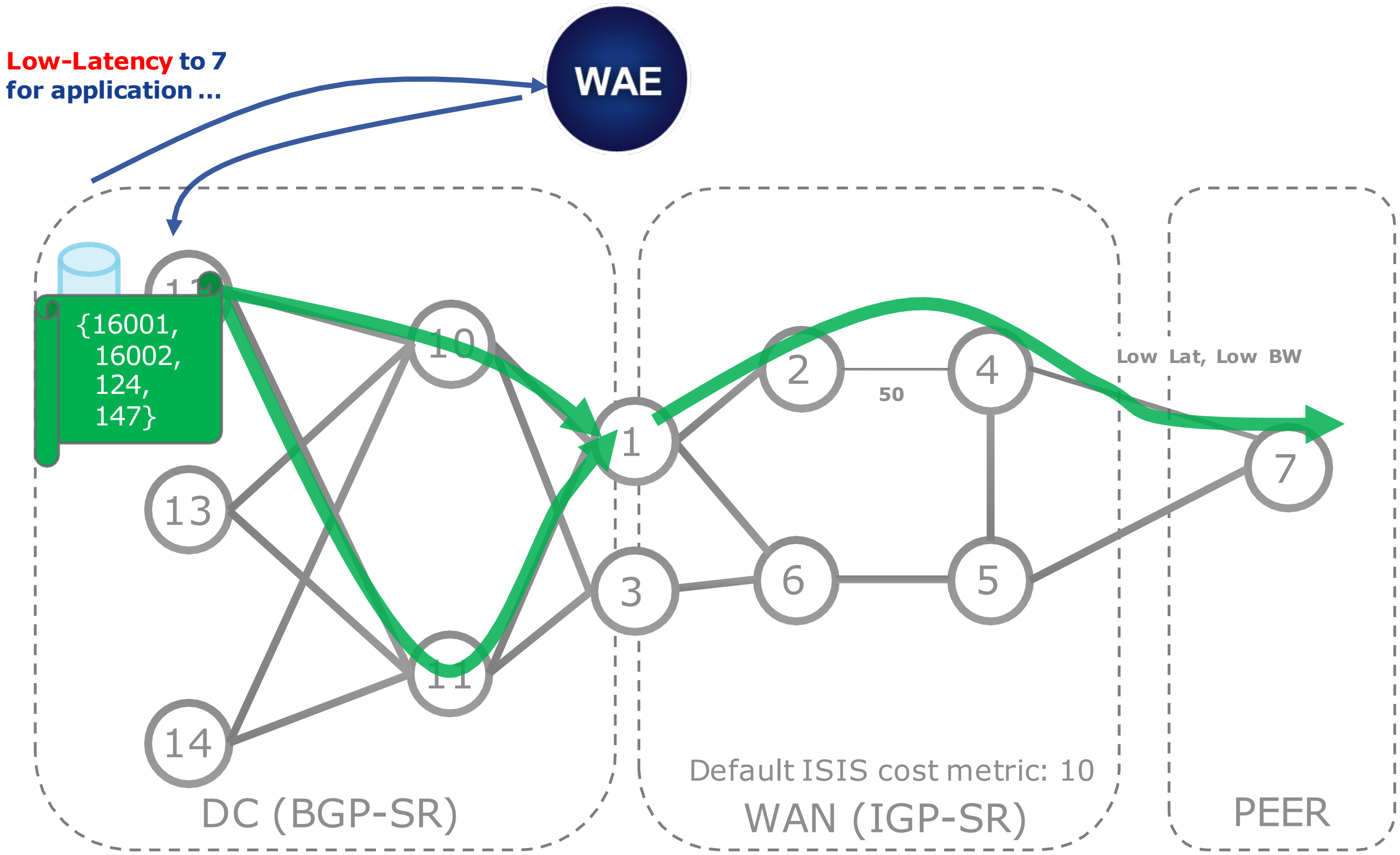
WAN Automation Engine

- WAE collects via BGP-LS
  - IGP segments
  - BGP segments
  - Topology



An end-to-end path as a list of segments

- WAE computes that the green path can be encoded as
  - 16001
  - 16002
  - 124
  - 147
- WAE programs a single per-flow state to create an application-engineered end-to-end policy



Segment Routing Standardization

- IETF standardization in SPRING working group
- Protocol extensions progressing in multiple groups
  - IS-IS
  - OSPF
  - PCE
  - IDR
  - 6MAN
- Broad vendor and customer support

Sample IETF Documents
Segment Routing Architecture (draft-ietf-spring-segment-routing)
Problem Statement and Requirements (draft-ietf-spring-problem-statement)
IPv6 SPRING Use Cases (draft-ietf-spring-ipv6-use-cases)
Segment Routing Use Cases (draft-filsfils-spring-segment-routing-use-cases)
Topology Independent Fast Reroute using Segment Routing (draft-francois-spring-segment-routing-ti-lfa)
IS-IS Extensions for Segment Routing (draft-ietf-isis-segment-routing-extensions)
OSPF Extensions for Segment Routing (draft-ietf-ospf-segment-routing-extensions)
PCEP Extensions for Segment Routing (draft-ietf-pce-segment-routing)

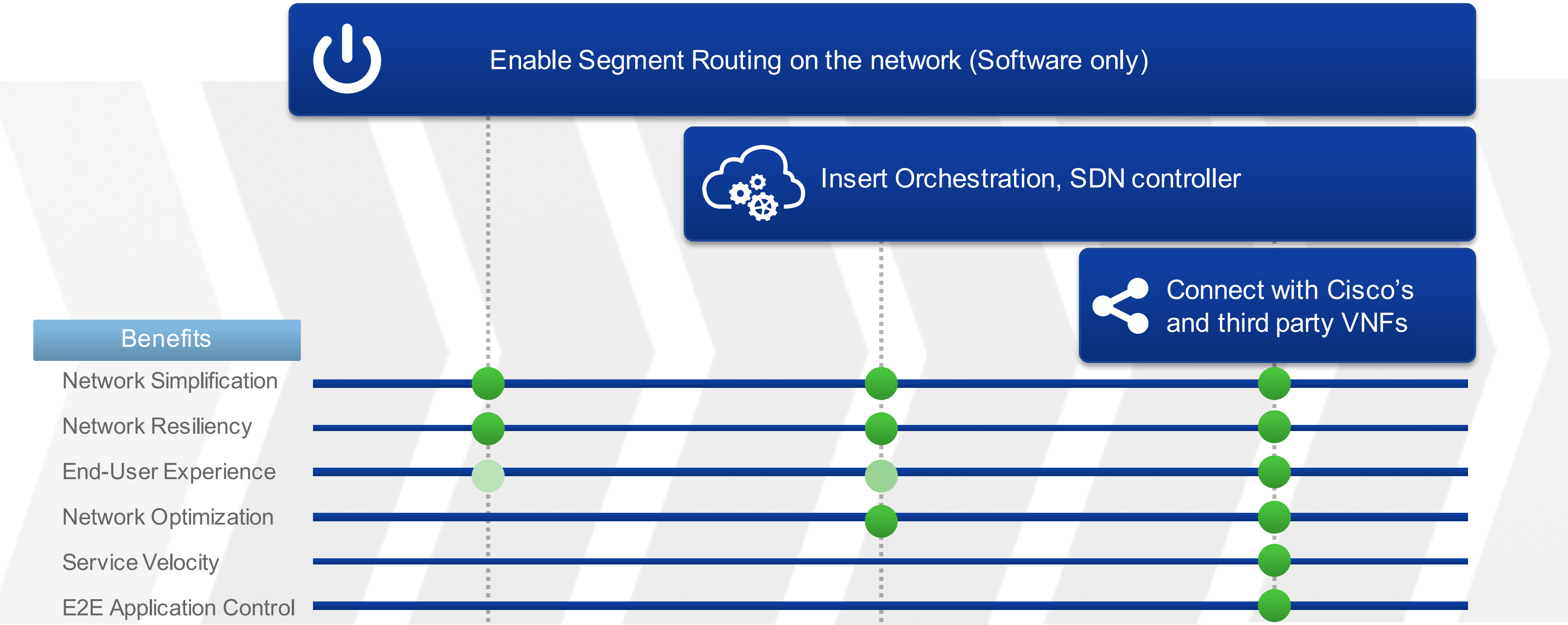
Close to 30 IETF drafts in progress

Segment Routing Product Support

- Platforms: ASR9000, CRS-1/CRS-3, ASR1000, ASR9XX, ISR4XXX
- IS-IS IPv4/IPv6
  - Node/Adjacency SID adertisement
  - LDP interworking (mapping server/client)
  - Traffci protection (Topology Independent LFA link protection)
- OSPFv2
  - Node SID advertisement
  - Traffic Protection (LFA)
- SR Traffic Engineering manual/PCEP
- OAM ping/trace

Application Engineered Routing Journey

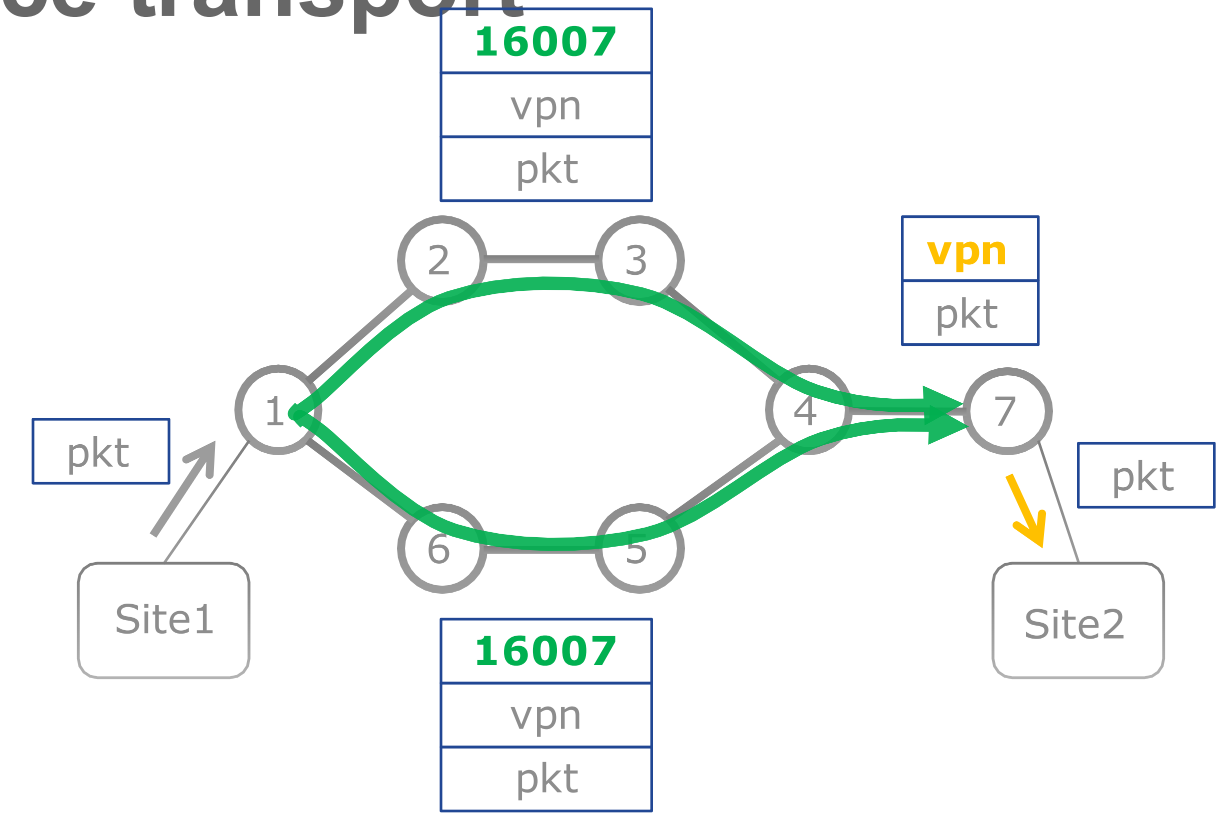
Adding value at your own pace



Use-Cases

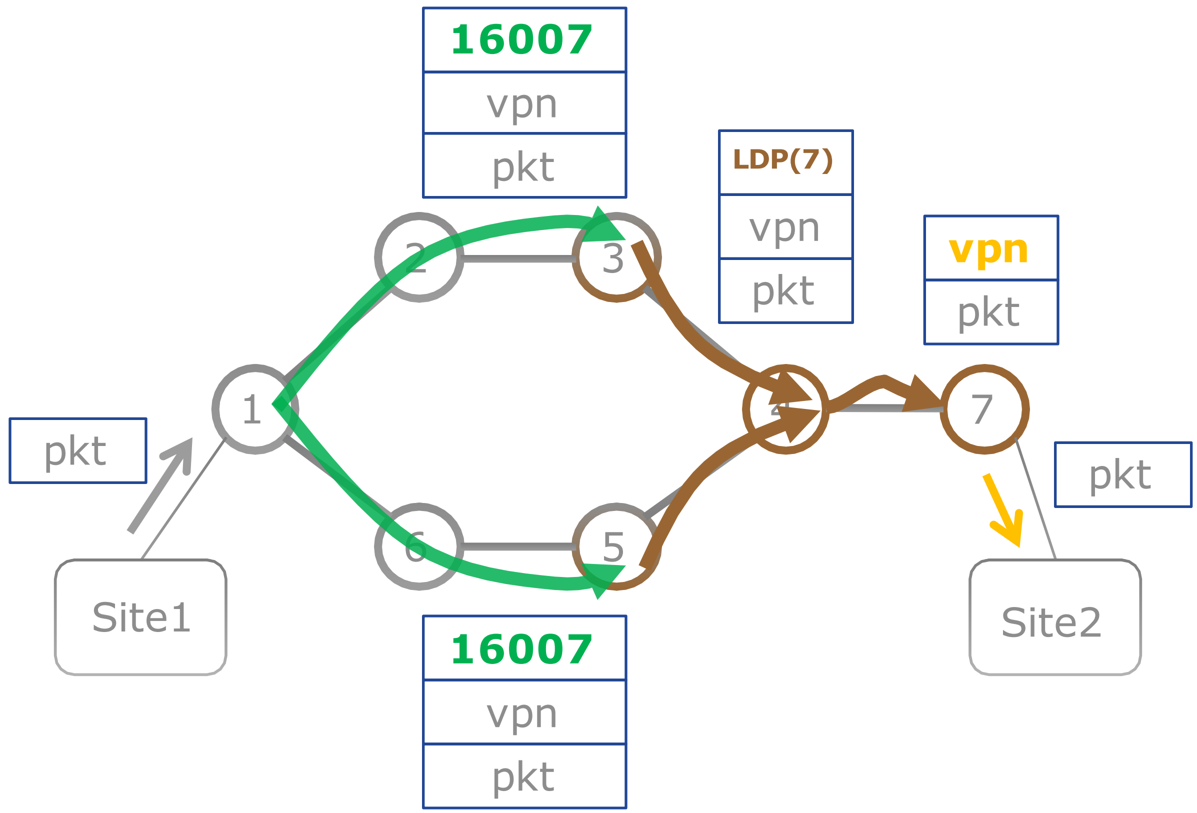
IPv4/6 VPN/Service transport

- IGP only
  - No LDP, no RSVP-TE
- ECMP



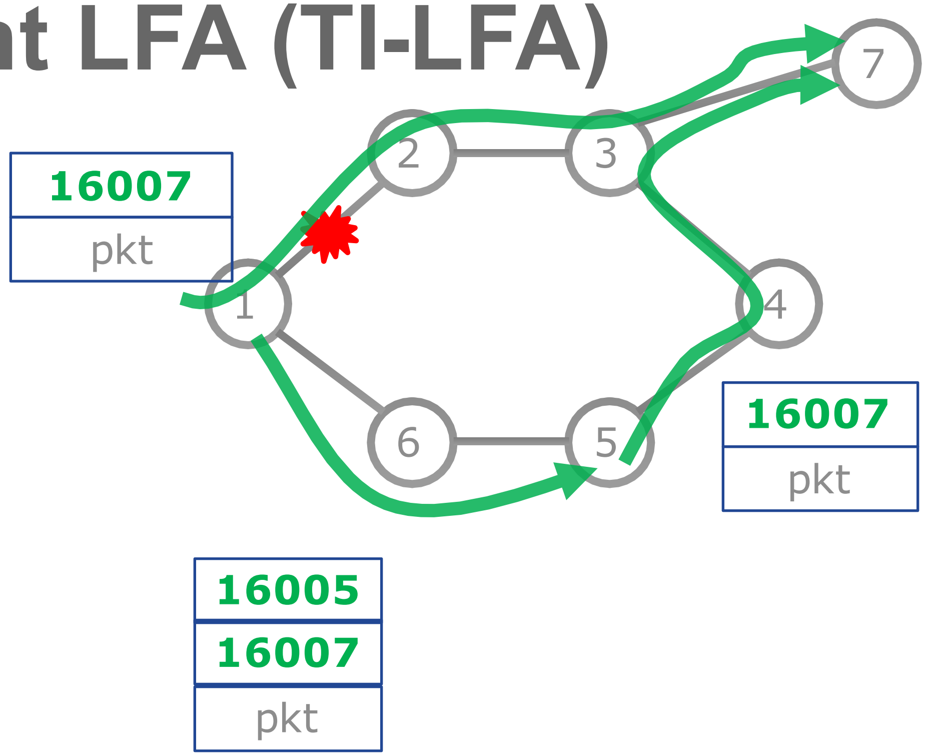
Seamless interworking with LDP

- Seamless deployment



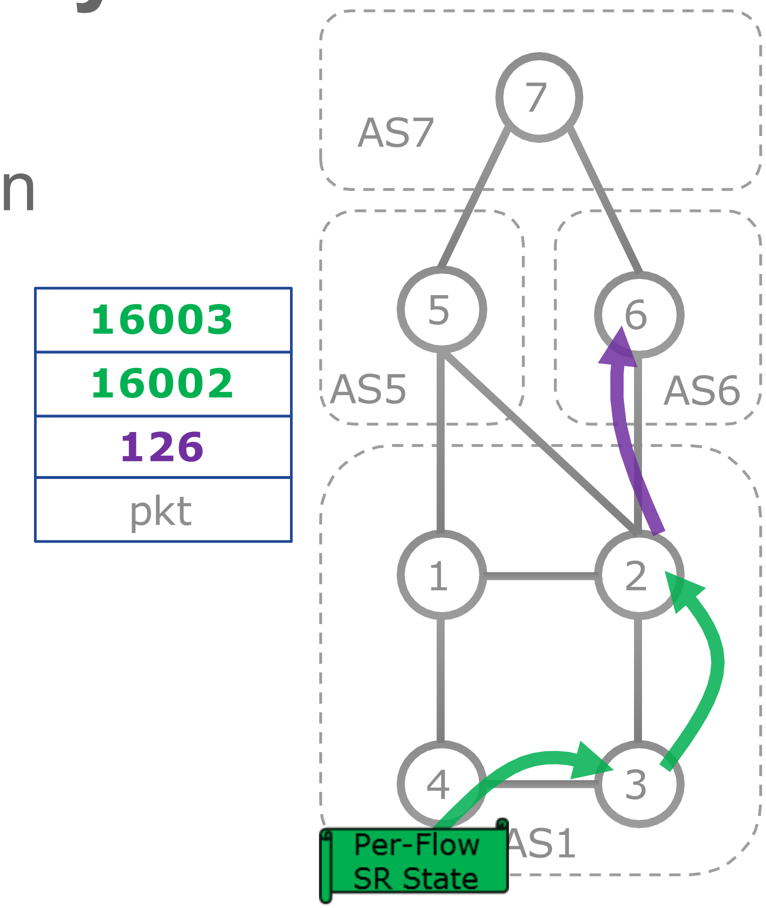
Topology-Independent LFA (TI-LFA)

- 50msec FRR in any topology
- IGP Automated
  - No LDP, no RSVP-TE
- Optimum
  - Post-convergence path
- No midpoint backup state
- Detailed operator report
  - S. Litkowski, B. Decraene, Orange
- Mate Design
  - How many backup segments
  - Capacity analysis



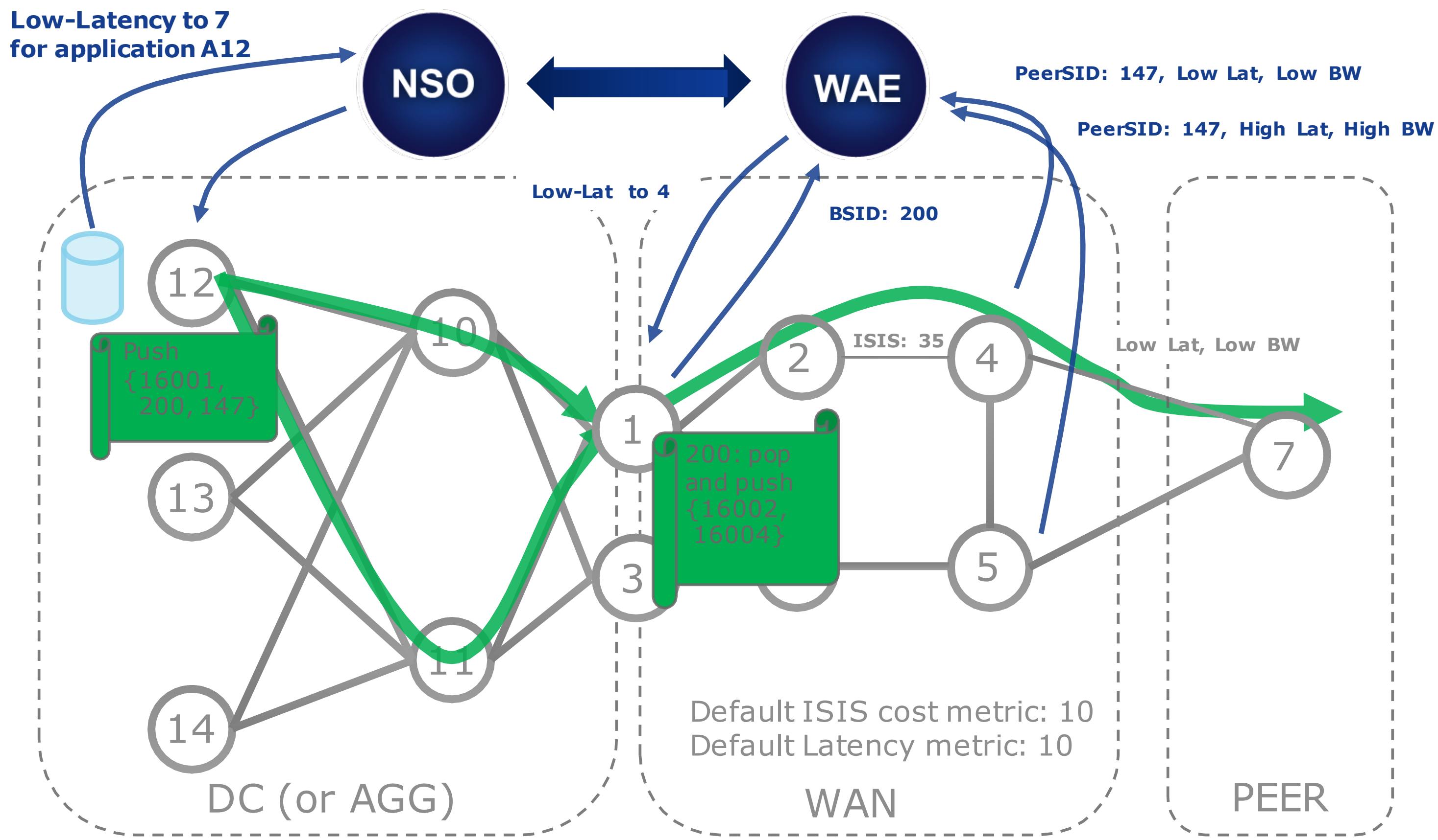
Optimized Content Delivery

- On a per-content, per-user basis, the content delivery application can engineer
  - the path within the AS
  - the selected border router
  - the selected peer
- Also applicable for engineering egress traffic from DC to peer
  - BGP Prefix and Peering Segments



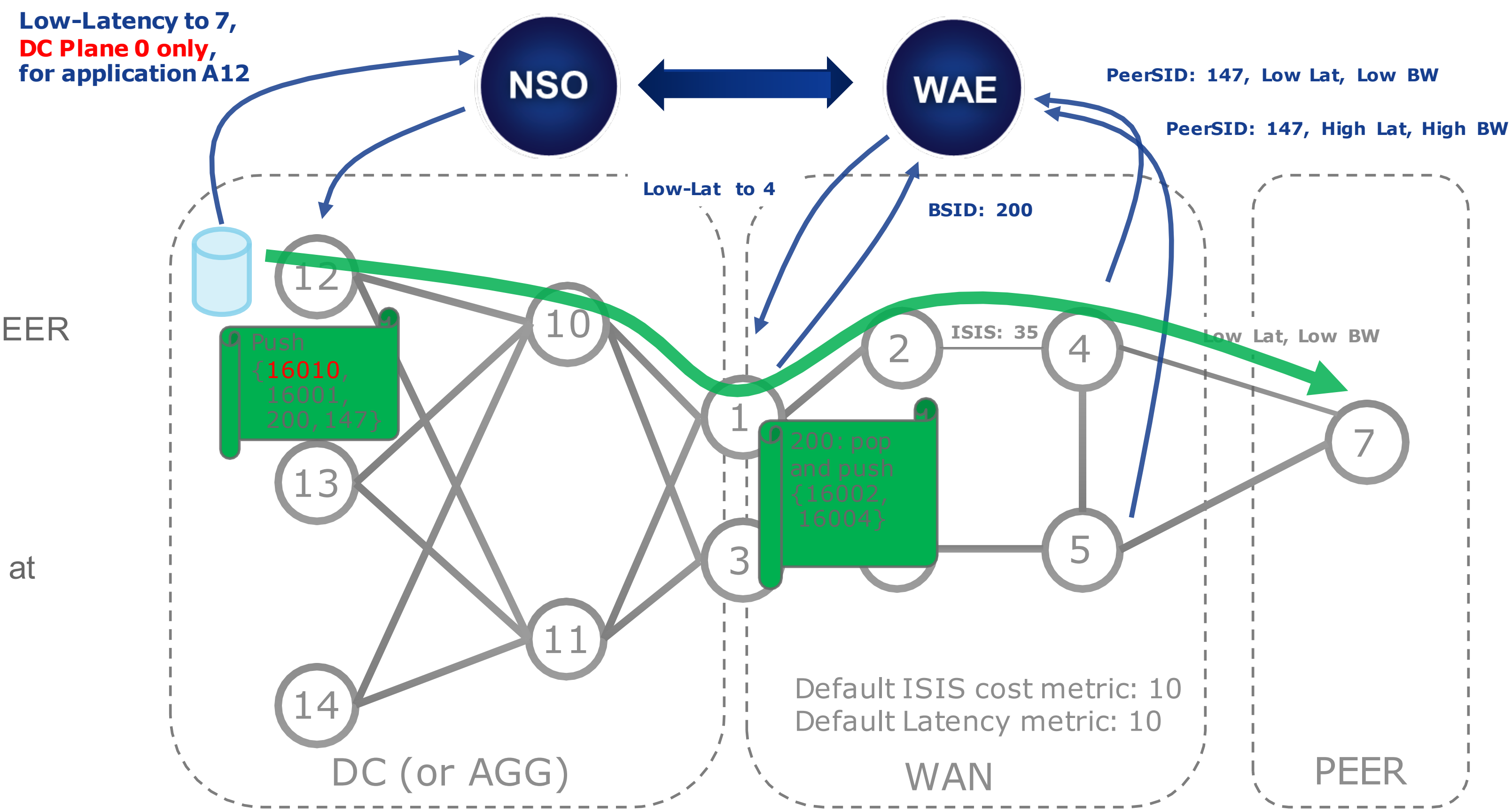
Application Engineered Routing

- Per-application flow engineering
- End-to-End
  - DC, WAN, AGG, PEER
- Millions of flows
  - No signaling
  - No midpoint state
  - No reclassification at boundaries



Application Engineered Routing

- Per-application flow engineering
- End-to-End
  - DC, WAN, AGG, PEER
- Millions of flows
  - No signaling
  - No midpoint state
  - No reclassification at boundaries

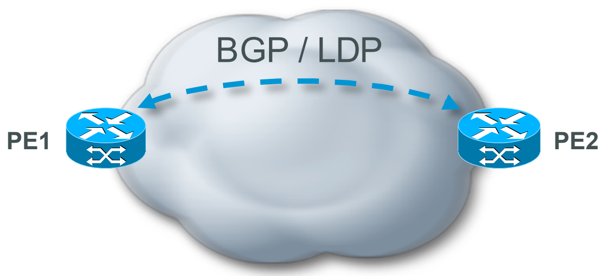


cisco

Control Plane and Data Plane

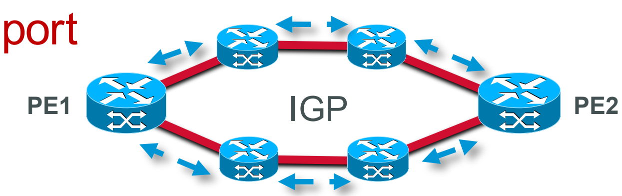
MPLS Control and Forwarding Operation with Segment Routing

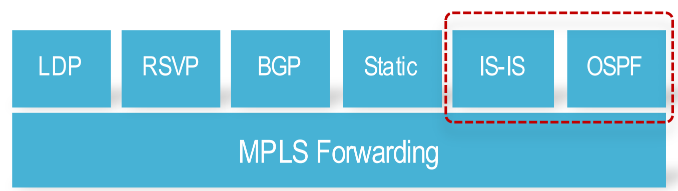
Services



No changes to control or forwarding plane

Packet Transport

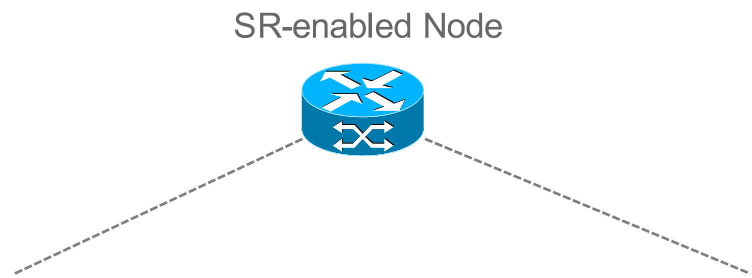




IGP label distribution for IPv4 and IPv6, same forwarding plane

SID Encoding

- Prefix SID
  - SID encoded as an index
  - Index represents an offset from SRGB base
  - Index globally unique
  - SRGB may vary across LSRs
  - SRGB (base and range) advertised with router capabilities
- Adjacency SID
  - SID encoded as absolute (i.e. not indexed) value
  - Locally significant
  - Automatically allocated for each adjacency



SRGB = [16000 - 23999]. Advertised as base = 16,000, range = 7,999
Prefix SID = 16041. Advertised as Prefix SID Index = 41
Adjacency SID = 24000. Advertised as Adjacency SID = 24000

SR IS-IS Control Plane Overview

- Level 1, level 2 and multi-level routing
- Prefix Segment ID (Prefix-SID) for host prefixes on loopback interfaces
- Adjacency SIDs for adjacencies
- Prefix-to-SID mapping advertisements (mapping server)
- MPLS penultimate hop popping (PHP) signaling
- MPLS explicit-null label signaling

IS-IS Configuration

- Required
  - Wide metrics
  - SR enabled under address family IPv4 unicast
- Optional
  - Prefix-SID configured under loopback(s) AF IPv4
- MPLS forwarding enabled automatically on all (non-passive) IS-IS interfaces
- Adjacency-SIDs are automatically allocated for each adjacency

Configuring Segment Routing for IPv4 Using IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 net 49.0001.1720.1625.5001.00
 address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
 !
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid absolute 16041
  !
 !
 interface GigabitEthernet0/0/0/0
  point-to-point
  address-family ipv4 unicast
  !
 !
!
```

Enable Segment Routing for IPv4 with MPLS data plane

Advertise prefix SID 16041 (index 41) for Loopback0

Configuring Segment Routing for IPv6 Using IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 net 49.0001.1720.1625.5001.00
 address-family ipv6 unicast
  metric-style wide
  segment-routing mpls
 !
 interface Loopback0
  passive
  address-family ipv6 unicast
   prefix-sid absolute 16061
  !
 !
 interface GigabitEthernet0/0/0/0
  point-to-point
  address-family ipv6 unicast
  !
 !
!
```

Enable Segment Routing for IPv4 with MPLS data plane

Advertise prefix SID 16041 (index 41) for Loopback0

SR OSPF Control Plane Overview

- IPv4 Prefix Segment ID (Prefix-SID) for host prefixes on loopback interfaces
- MPLS penultimate hop popping (PHP) signaling
- MPLS explicit-null label signaling

OSPF Configuration

- OSPFv2 control plane
- Required
  - Enable segment-routing under instance or area(s)
    - Command has area scope, usual inheritance applies
  - Enable segment-routing forwarding under instance, area(s) or interface(s)
    - Command has interface scope, usual inheritance applies
- Optional
  - Prefix-SID configured under loopback(s)
- MPLS forwarding enabled on all OSPF interfaces with segment-routing forwarding configured

Configuring Segment Routing for IPv4 Using OSPF (Cisco IOS XR)

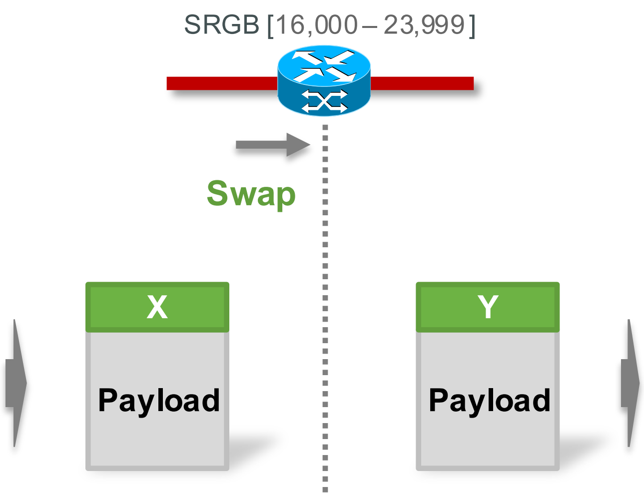
```
router ospf DEFAULT
 router-id 172.16.255.1
 segment-routing mpls
 segment-routing forwarding mpls
 area 0
  interface Loopback0
   passive
   prefix-sid absolute 16041
  !
  interface GigabitEthernet0/0/0/0
   network point-to-point
  !
 !
!
```

Enable Segment Routing with MPLS data plane

Advertise prefix SID 16041 (index 41) for Loopback0

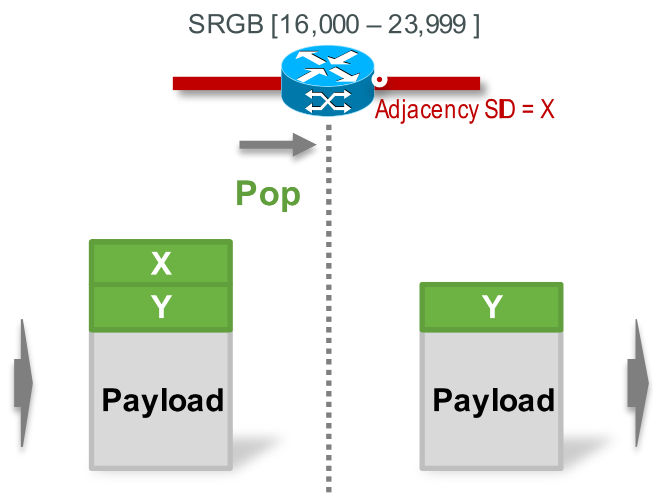
MPLS Data Plane Operation

Prefix SID



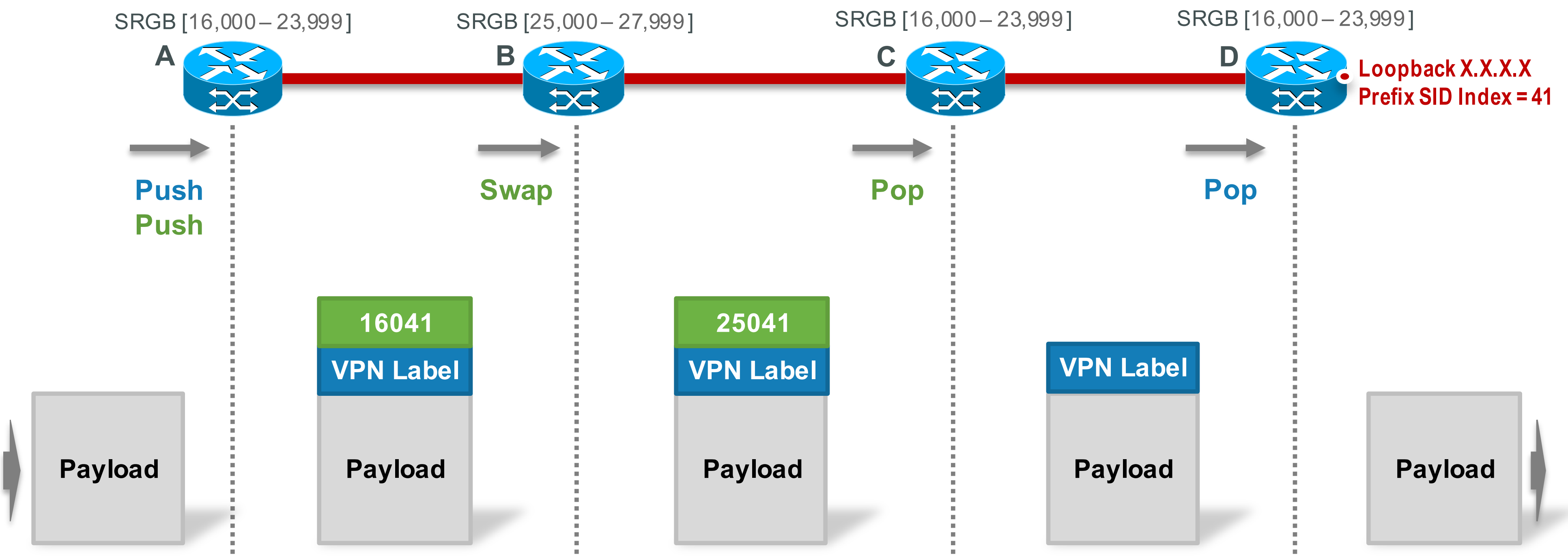
- Packet forwarded along IGP shortest path
- Packet may be leverage ECMP load balancing
- Swap operation performed on input label
- Input label (X) and output label (Y) will have same value when downstream neighbor has same SRGB
- Penultimate hop may perform a pop operation (PHP) if signaled by egress LSR

Adjacency SID

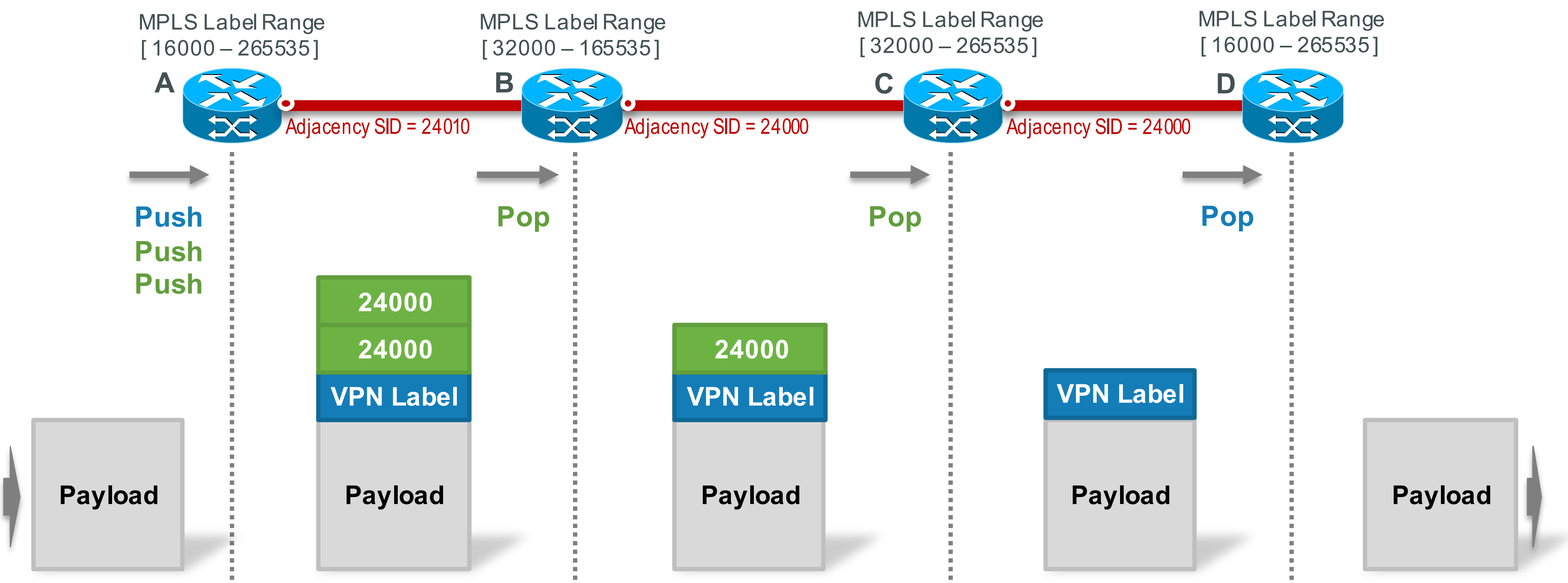


- Packet forwarded along IGP adjacency
- Pop operation performed on input label
- Input topmost label (X) and output label (Y) may or may not have same value
- Penultimate hop always pops last adjacency SID

MPLS Data Plane Operation (Prefix SID)

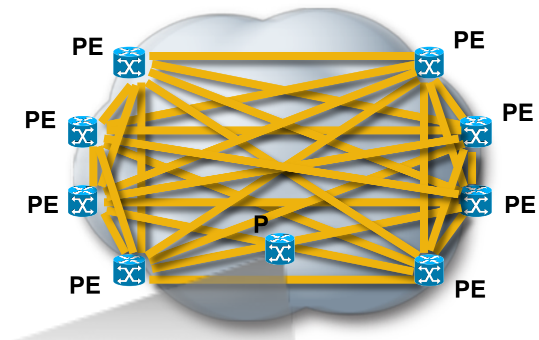


MPLS Data Plane Operation (Adjacency SIDs)



MPLS LFIB with Segment Routing

- LFIB populated by IGP (ISIS / OSPF)
- Forwarding table remains constant (Nodes + Adjacencies) regardless of number of paths
- Other protocols (LDP, RSVP, BGP) can still program LFIB



	In Label	Out Label	Out Interface
Network Node Segment Ids	**L1**	L1	Intf1
	L2	L2	Intf1
	...	...	...
	L8	L8	Intf4
	L9	L9	Intf2
Node Adjacency Segment Ids	**L10**	Pop	Intf2
	...	...	...
	Ln	Pop	Intf5

Forwarding table remains constant

Traffic Protection

Topology Independent LFA (TI-LFA) – Benefits

- 100%-coverage 50-msec link and node protection
- Simple to operate and understand
  - automatically computed by the IGP
- Prevents transient congestion and suboptimal routing
  - leverages the post-convergence path, planned to carry the traffic
- Incremental deployment
  - also protects LDP traffic

Topology Independent LFA – Implementation

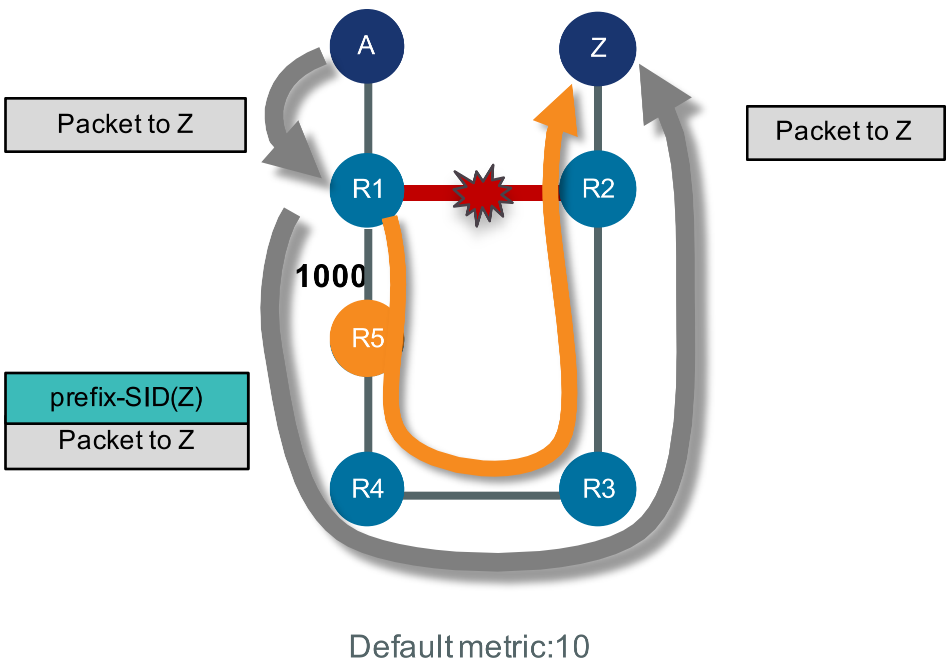
- Leverages existing and proven LFA technology
  - P space: set of nodes reachable from node S (PLR) without using protected link L
  - Q space: set of nodes that can reach destination D without using protected link L
- Enforcing loop-freeness on post-convergence path
  - Where can I release the packet?

    At the intersection between the post-convergence shortest path and the Q space
  - How do I reach the release point?

    By chaining intermediate segments that are assessed to be loop-free

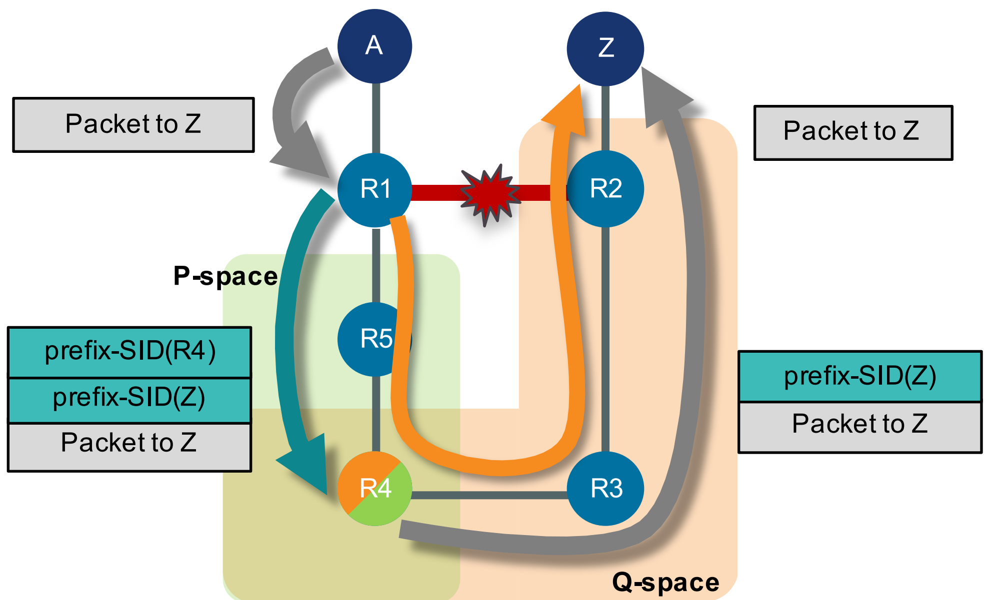
TI-LFA – Zero-Segment Example

- TI-LFA for link R1R2 on R1
- Calculate LFA(s)
- Calculate post-convergence SPT
- Find LFA on post-convergence SPT
- R1 will steer the traffic towards LFA R5



TI-LFA – Single-Segment Example

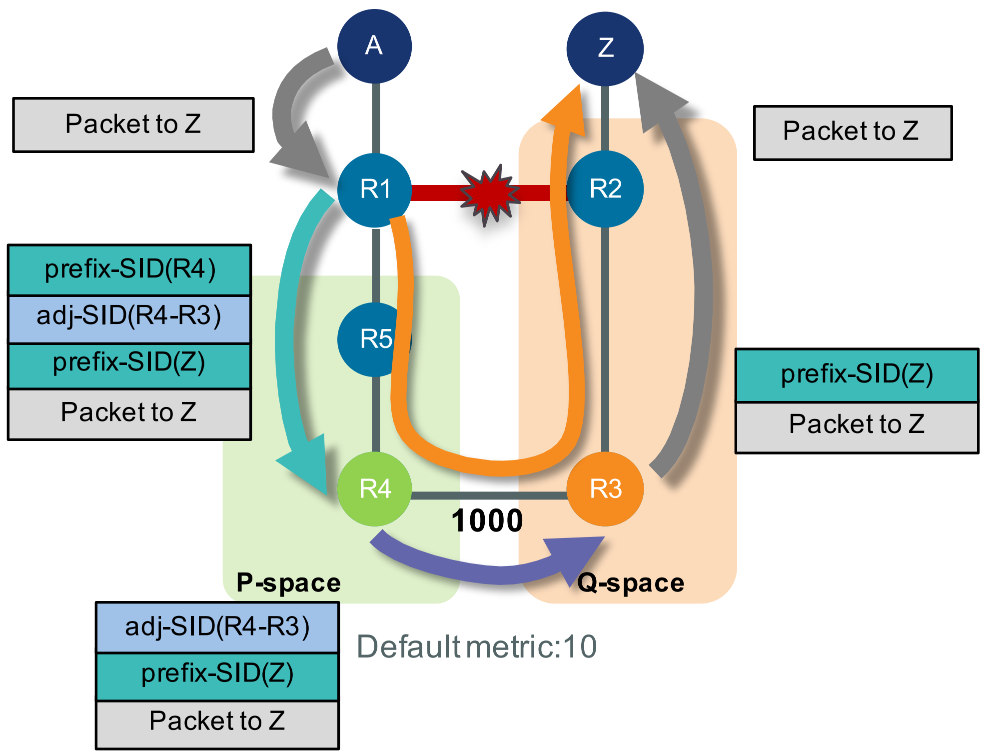
- TI-LFA for link R1R2 on R1
- Calculate P and Q spaces
  - They overlap in this case
- Calculate post-convergence SPT
- Find PQ node on post-convergence SPT
- R1 will push the prefix-SID of R4 on the backup path



Default metric:10

TI-LFA – Double-Segment Example

- TI-LFA for link R1R2 on R1
- Calculate P and Q spaces
- Calculate post-convergence SPT
- Find Q and adjacent P node on post-convergence SPT
- R1 will push the prefix-SID of R4 and the adj-SID of R4-R3 link on the backup path



Configuring Topology Independent Fast Reroute for IPv4 using Segment Routing and IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 net 49.0001.1720.1625.5001.00
 address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
 !
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid absolute 16041
  !
 !
 interface GigabitEthernet0/0/0/0
  address-family ipv4 unicast
   fast-reroute per-prefix
   fast-reroute per-prefix ti-lfa
  !
 !
!
```

Enable TI-LFA for IPv4 prefixes on interface GigabitEthernet0/0/0/0

Configuring Topology Independent Fast Reroute for IPv6 using Segment Routing and IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 net 49.0001.1720.1625.5001.00
 address-family ipv6 unicast
  metric-style wide
  segment-routing mpls
 !
 interface Loopback0
  passive
  address-family ipv6 unicast
   prefix-sid absolute 16061
  !
 !
 interface GigabitEthernet0/0/0/0
  address-family ipv6 unicast
   fast-reroute per-prefix
   fast-reroute per-prefix ti-lfa
  !
 !
!
```

Enable TI-LFA for IPv6 prefixes on interface GigabitEthernet0/0/0/0

Conclusion

- Simple routing extension to enable source routing
- Packet path is determined by prepended segment identifiers (one ore more)
- Dataplane agnostic (MPLS, IPv6)
- Nework Scalability and agility by reducing network state and simplifying control plane
- Traffic protection with 100% coverage with more optimal routing

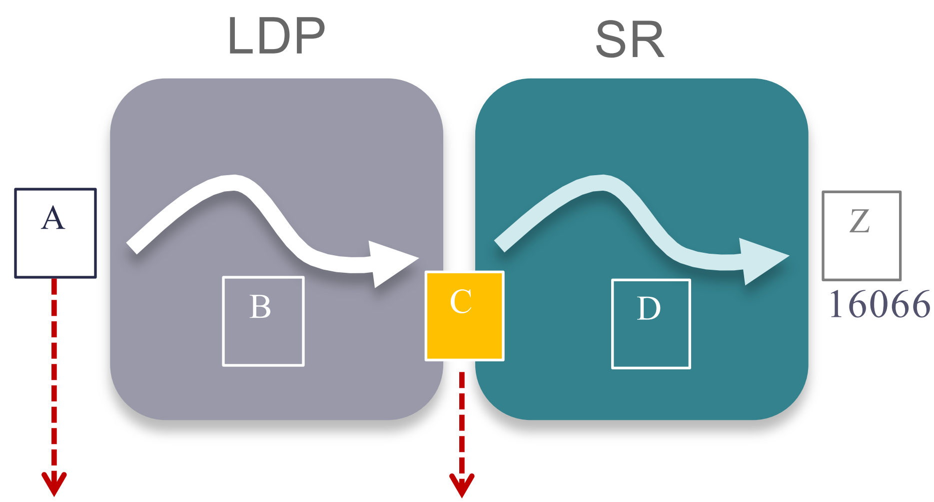
Thank you

TOMORROW starts here.

SR and LDP Interworking

LDP to SR

- When a node is LDP capable but its next-hop along the SPT to the destination is not LDP capable
  - no LDP outgoing label
- In this case, the LDP LSP is connected to the prefix segment
- C installs the following LDP-to-SR FIB entry:
  - incoming label: label bound by LDP for FEC Z
  - outgoing label: prefix segment bound to Z
  - outgoing interface: D
- This entry is derived automatically at the routing layer

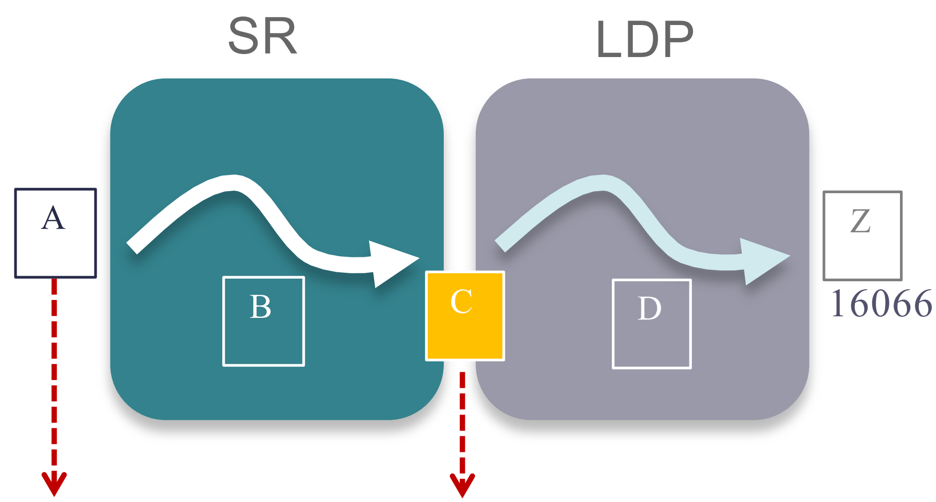


Prefix	Out Label (LDP), Interface
Z	16, 0

Input Label (LDP)	Out Label (SID), Interface
32	16066, 1

SR to LDP

- When a node is SR capable but its next-hop along the SPT to the destination is not SR capable
  - no SR outgoing label available
- In this case, the prefix segment is connected to the LDP LSP
  - Any node on the SR/LDP border installs SR-to-LDP FIB entry(ies)



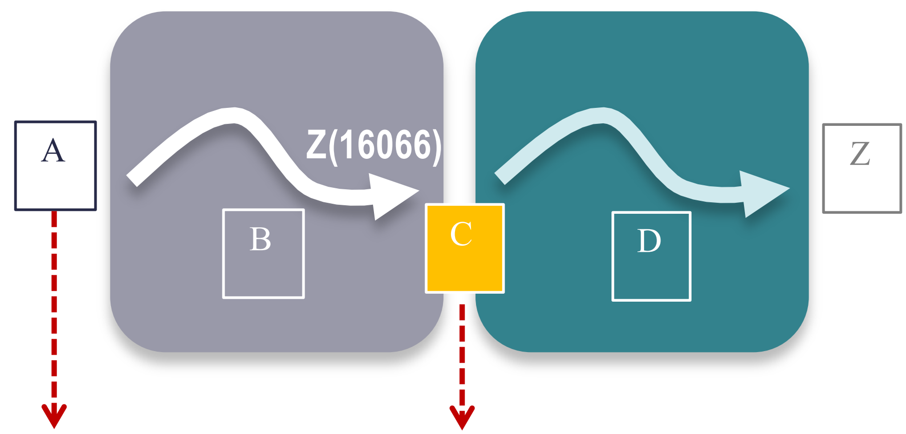
Prefix	Out Label (SID), Interface
Z	?, 0

Input Label (SID)	Out Label (LDP), Interface
?	16, 1

Mapping Server

- A wants to send traffic to Z, but
  - Z is not SR-capable, Z does not advertise any prefix-SID

  → which label does A have to use?
- The Mapping Server advertises the SID mappings for the non-SR routers
  - for example, it advertises that Z is 16068
- A and B install a normal SR prefix segment for 16066
- C realizes that its next hop along the SPT to Z is not SR capable hence C installs an SR-to-LDP FIB entry
  - incoming label: prefix-SID bound to Z (16066)
  - outgoing label: LDP binding from D for FEC Z
- A sends a frame to Z with a single label: 16066



Prefix	Out Label (SID), Interface
Z	?, 0

Input Label (SID)	Out Label (LDP), Interface
?	16, 1

Active Mapping Policy Preferences

- Active SID Mapping policy
  - A set of non-overlapping SID mapping entries derived from locally configured SID mappings and SID mappings received from other nodes
- Backup SID Mapping policy
  - SID mapping entries that overlap with at least one Active SID mapping entry
- When two or more SID mapping entries overlap, which one will be used?
  - Sort all overlapping entries according to preference rules*
    - Locally configured entries are treated the same as remote entries
  - Only the most preferred entry is inserted in the Active SID mapping policy
  - The other SID-entries are inserted in the Backup SID mapping policy

* Highest router-id > smallest prefix numerical value > smallest first SID value > largest range > latest received

Configuring a Mapping Server for SR and LDP Interworking for IPv4 Using IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 net 49.0001.1720.1625.5001.00
 address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
  segment-routing prefix-sid-map receive
  segment-routing prefix-sid-map advertise-local
 !

...

!
segment-routing
 address-family ipv4
  prefix-sid-map
   172.16.255.1/32 4041 range 8
  !
 !
!
```

Construct active mapping policy using remotely learned and locally configured mappings (mapping client)

Advertise local mapping policy (mapping server)

Local prefix-to-SID mapping policy
172.16.255.1/32 – 4041
:
172.16.255.8/32 - 4048

Configuring a Mapping Client for SR and LDP Interworking for IPv4 Using IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 net 49.0001.1720.1625.5001.00
 address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
  segment-routing prefix-sid-map receive
 !
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid absolute 16041
  !
 !
 interface GigabitEthernet0/0/0/0
  point-to-point
  address-family ipv4 unicast
  !
 !
!
```

Construct active mapping policy using remotely learned and locally configured mappings (mapping client)

cisco