Segment Routing

Atle Heimnes

Cisco SE - Service Provider

Agenda

- Executive summary
- Technology Overview
- Use Cases

Segment Routing Executive Summary

Bakgrunn/historikk

- MPLS eller Tag switching ble oppfunnet på slutten av 1990 taller
  - Motivasjon : kontrollplan på rutere var begrenset(og dyrt) og tanken var at om at rutere kunne forwarde pakker basert på en enkel label informasjon istedet for ip header så ville det være langt mer effektivt.
- IGP førstevalg til å utveksle label informasjon,men
  - pga begrnesede ressurser for kontrollplan og vanskelig å endre IGP protokollene så ble det bestemt at de ville lage en ny protokoll
- LDP ble oppfunnet rundt 1999
  - Oppgave: Genere labler og utveksle label mellom rutere.(Hop by hop)
  - Kjører sammen med IGP og bruker IGP for å definere beste vei gjennom nettet.

Bakgrunn forts.

- RSVP-TE (ca 2003)
  - Alternativ protokoll for generering og distribuering av labler
  - Kan reserverer stier gjennom nettet basert på flere input enn IGP metrics (Båndbredde, spesifik Path, mm)
  - Krever at alle ruterne I en Path holder rede på informasjon om stien.
  - Blir oppfattet som kompleks og resulterer I at RSVP-TE er veldig lite utbredt.
- Når RSVP-TE blir implemetert, så blir det oftest brukt taktisk (der det er brukt for det) og LDP og RSVP-TE kjører ofte samtidig.
- Segment Routing er tatt frem som en erstatning for LDP og RSVP-TE

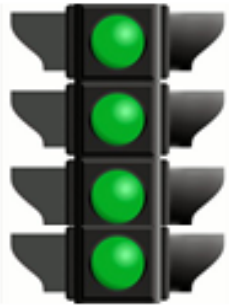
Segment Routing

- 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
- Segment: an identifier for any type of instruction
  - forwarding or service
- Forwarding Plane:
  - MPLS: an ordered list of segments is represented as a stack of labels
  - IPv6: an ordered list of segments is encoded in a routing extension header
- Multi-Vendor solution

Deployed !

Deployment

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



WEB

SP Core/Edge

SP Agg/Metro

Large Entreprise

- First deployments in 2015 – just 15 months after FCS !!

IETF

- Strong commitment for standardization and multi-vendor support
- SPRING Working-Group (started Nov 2013)
- All key documents are WG-status
- Over 25 drafts maintained by SR team
  - Over 50% are WG status
  - Over 75% have a Cisco implementation
- Several interop reports are available
- First RFC document - RFC 7855 (May 2016)

www.segment-routing.net
tools.ietf.org/wg/spring/

Technology and Problem Statement

- Architecture (draft)
- Problem Statement
  - Generic (draft)
  - Resiliency (draft)
  - IPv6 (draft)
  - OAM (draft)
- Applicability
  - SR illustration to problem statement (draft)
  - Centralized Egress Peer Engineering (draft)

Protocol Extension

- ISIS extension for SR (draft)
- OSPF extension for SR (draft)
- OSPFv3 extension for SR (draft)
- BGP-LS extension for SR (draft)
- BGP-LS extension for SR EPE use-case (draft)
- PCEP extension for SR (SR ext, setup method)

FRR

- Topology-Independent LFA FRR with SR (draft)

MPLS instantiation of Segment Routing

- MPLS support for SR (draft)
- SR/LDP interaction and interworking (draft)

IPv6 instantiation of Segment Routing

- IPv6 SR routing extension header (draft)
- IPv6 use-cases (draft)

OAM

- SR/LSP Ping (draft)
- OAM (draft)

Technology Overview

Segment Routing

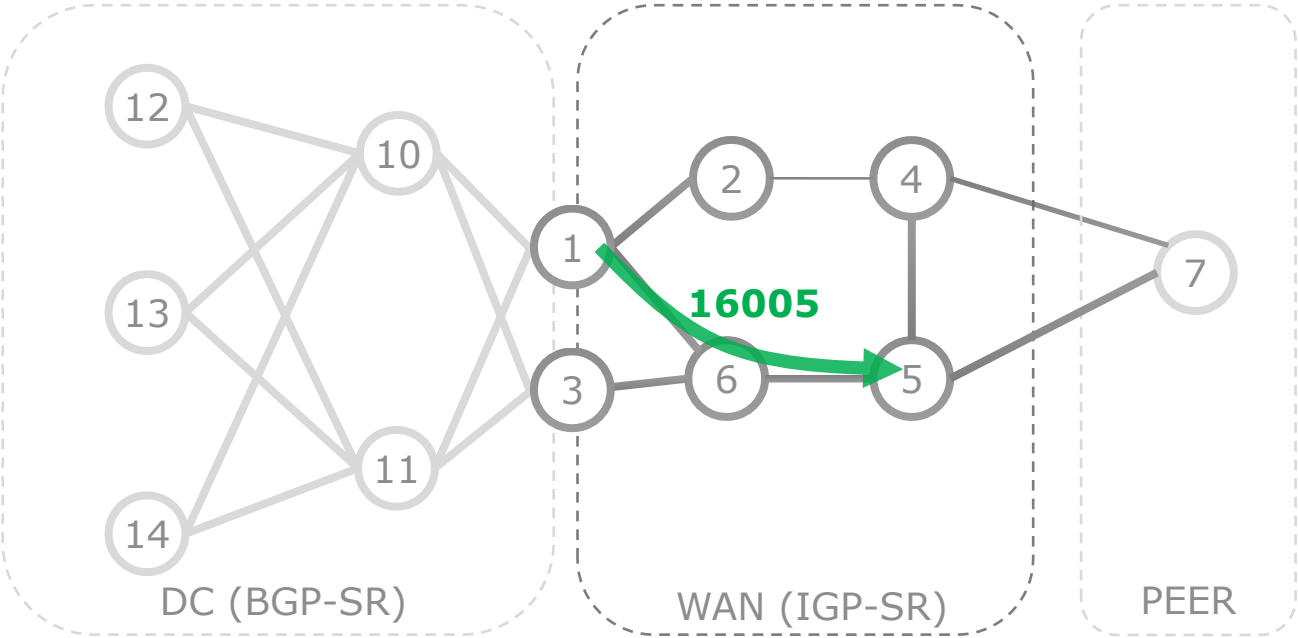
- **Source Routing**
  - the source chooses a path and encodes it in the packet header as an ordered list of segments
  - the rest of the network executes the encoded instructions
- **Segment**: an identifier for any type of instruction
  - forwarding or service

Segment Routing – Forwarding Plane

- **MPLS**: an ordered list of segments is represented as a stack of labels
- **IPv6**: an ordered list of segments is encoded in a routing extension header
- This presentation: **MPLS data plane**

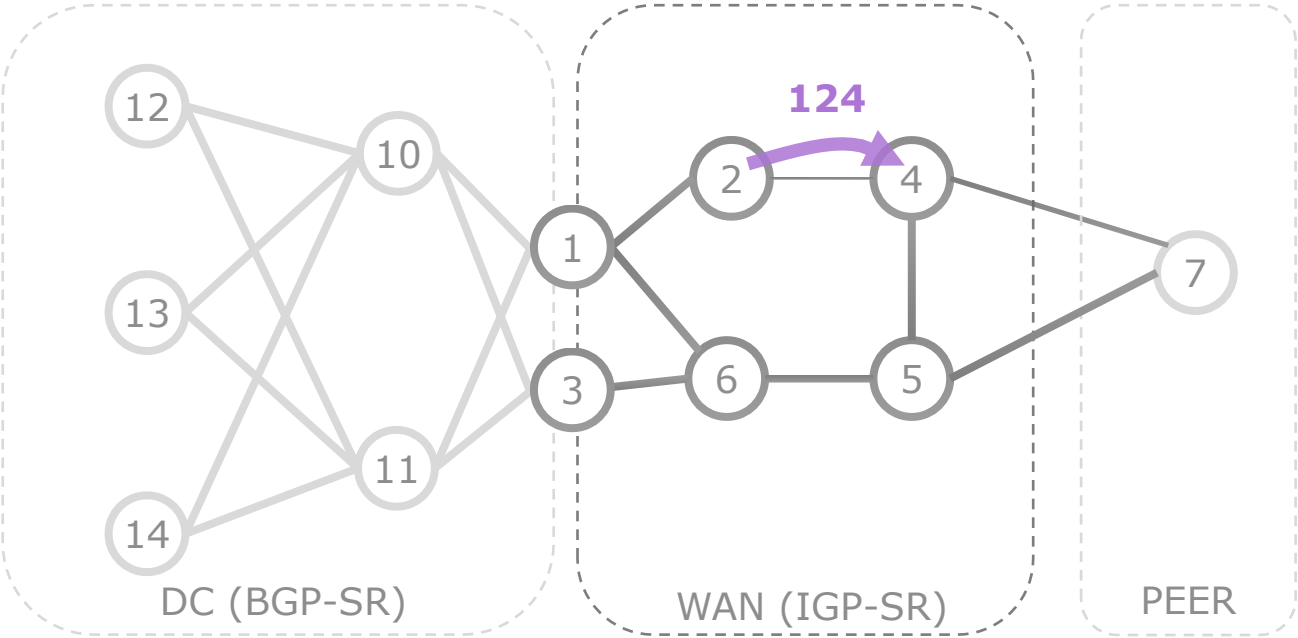
IGP Prefix Segment

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



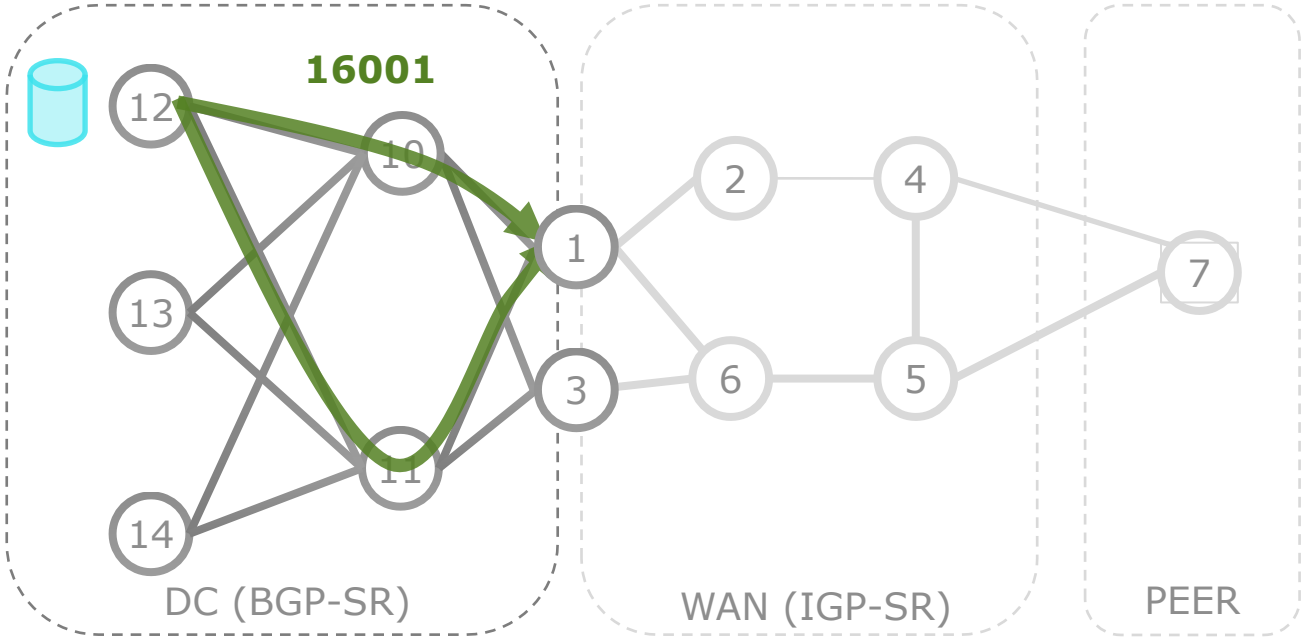
IGP Adjacency Segment

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



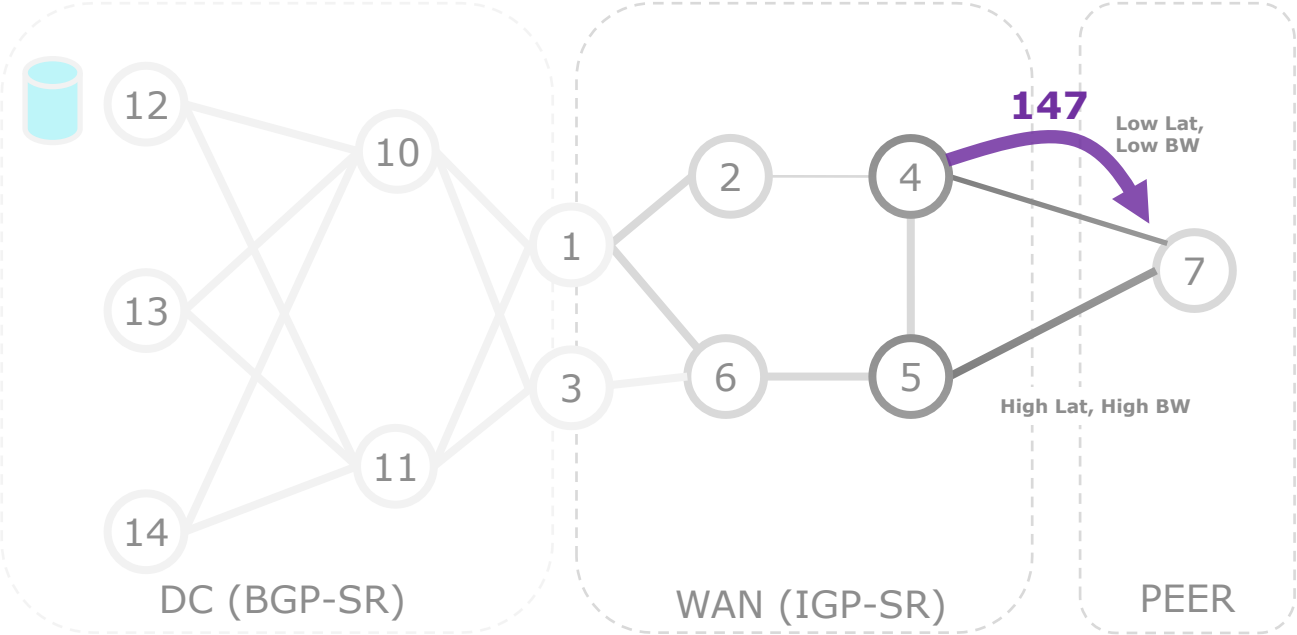
BGP Prefix Segment

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



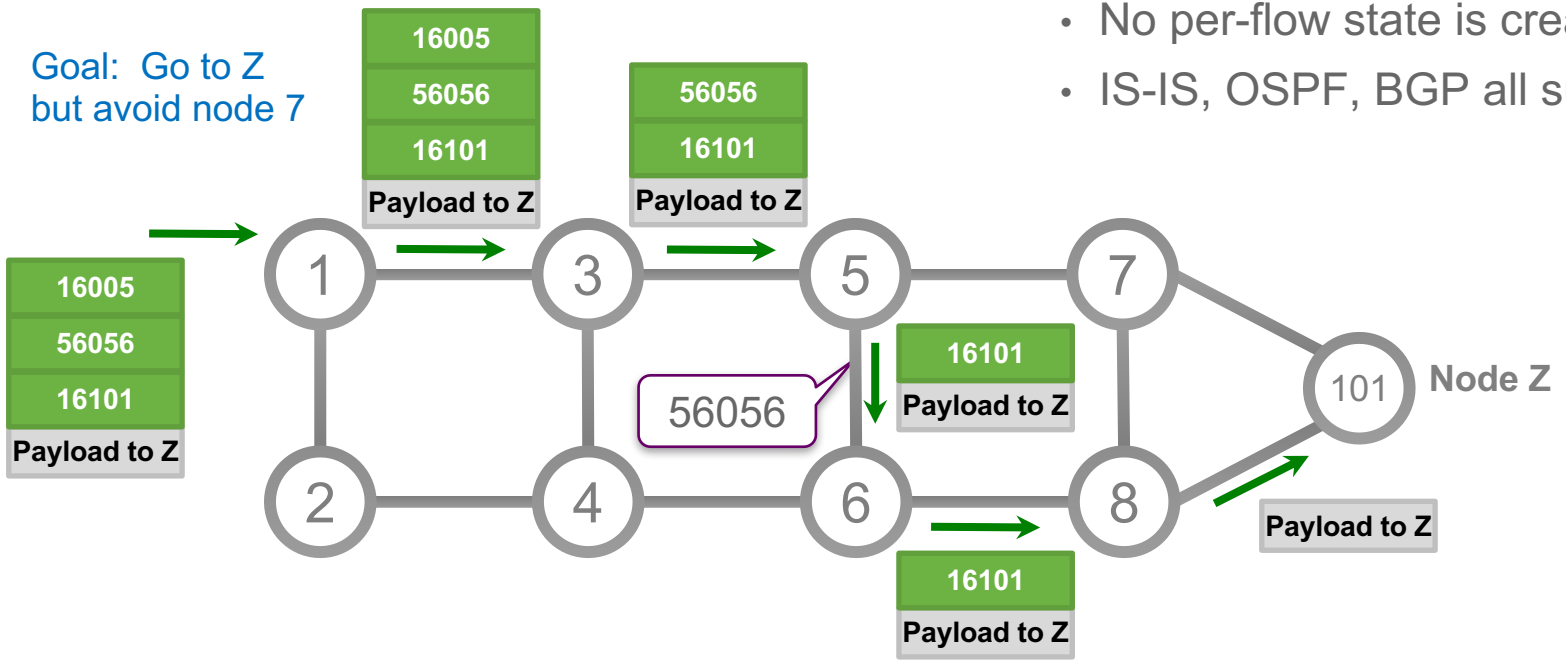
BGP Peering Segment

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



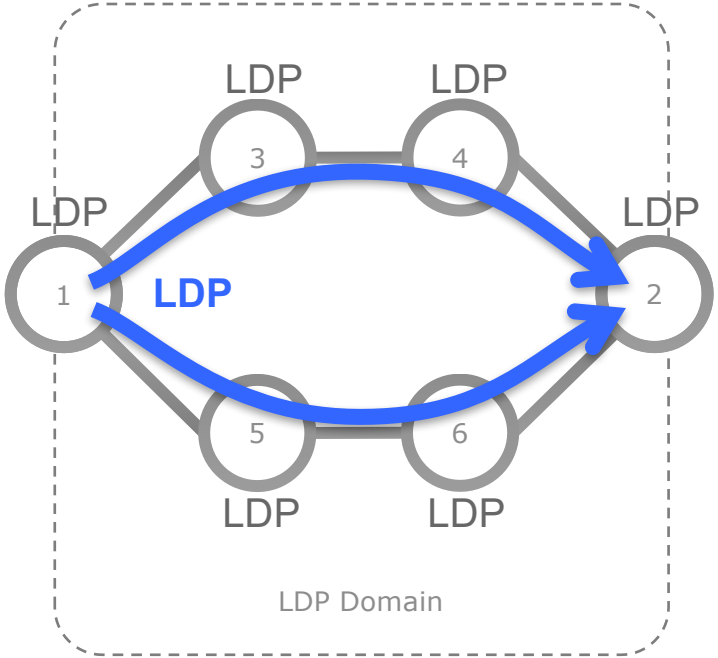
SR operations illustration

- Steer traffic on any path through the network
- Path is specified by list of segments in packet header, a stack of labels
- No path is signaled
- No per-flow state is created
- IS-IS, OSPF, BGP all supported



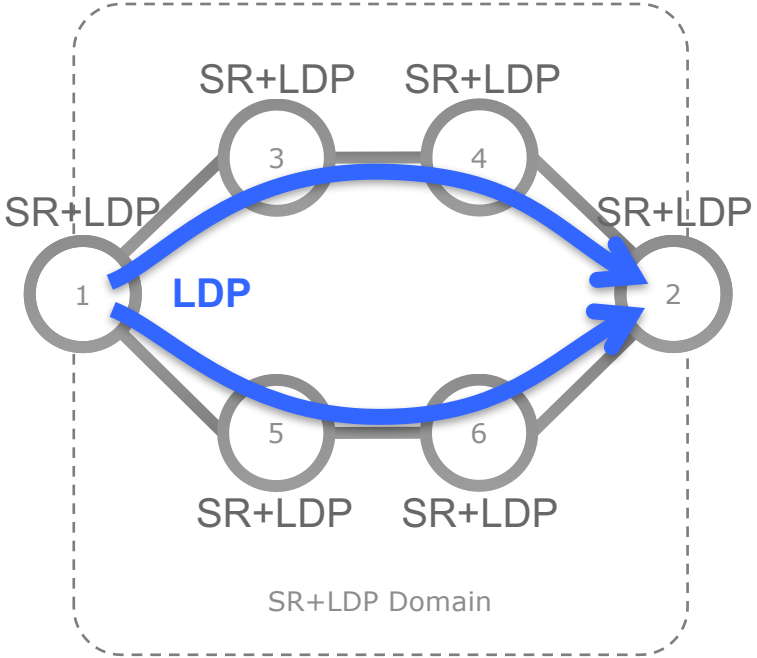
Simplest migration: LDP to SR

- Initial state: All nodes run LDP, not SR



Simplest migration: LDP to SR

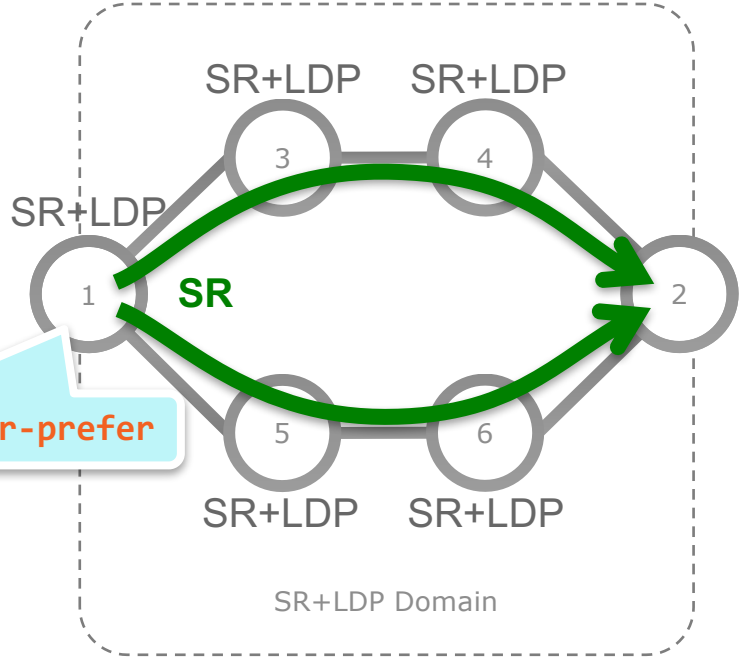
- Initial state: All nodes run LDP, not SR
- Step1: All nodes are upgraded to SR
  - In no particular order
  - Default label imposition preference = LDP



Simplest migration: LDP to SR

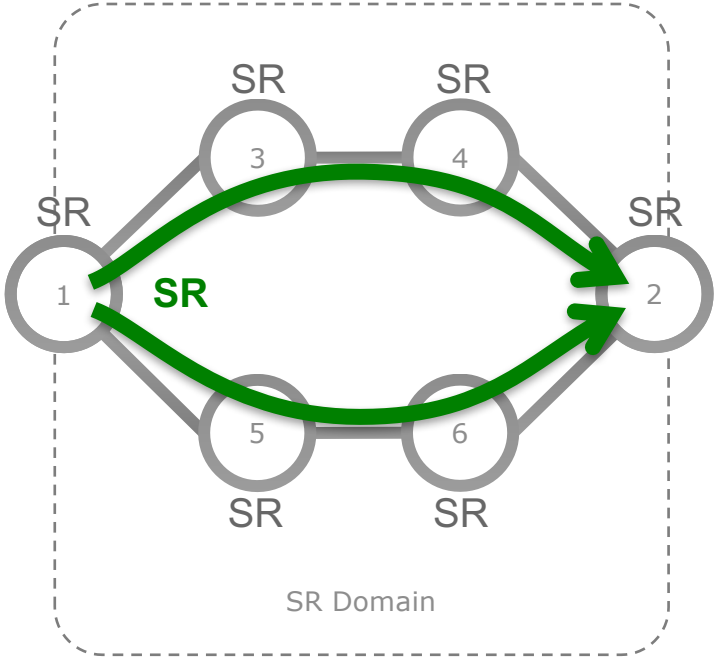
- Initial state: All nodes run LDP, not SR
- Step1: All nodes are upgraded to SR
  - In no particular order
  - leave default LDP label imposition preference
- Step2: All PEs are configured to prefer SR label imposition
  - In no particular order

```
segment-routing mpls sr-prefer
```



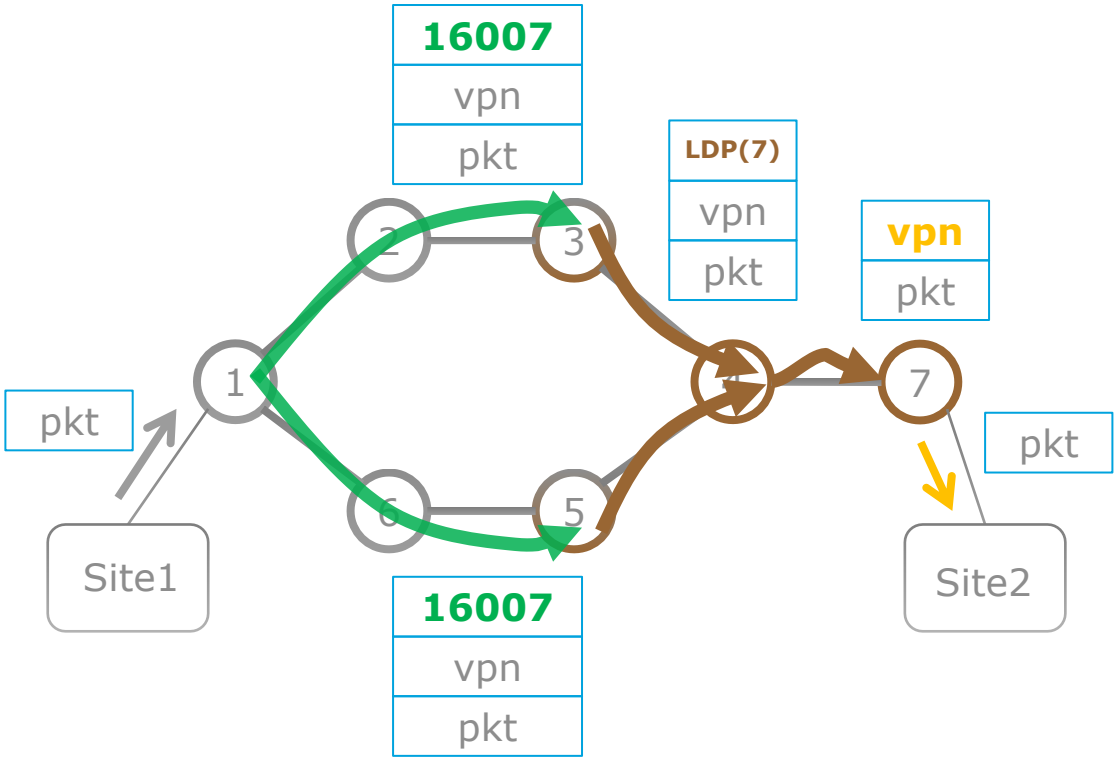
Simplest migration LDP to SR

- Initial state: All nodes run LDP, not SR
- Step1: All nodes are upgraded to SR
  - In no particular order
  - leave default LDP label imposition preference
- Step2: All PEs are configured to prefer SR label imposition
  - In no particular order
- Step3: LDP is removed from the nodes in the network
  - In no particular order
- Final state: All nodes run SR, not LDP



Seamless Interworking with LDP

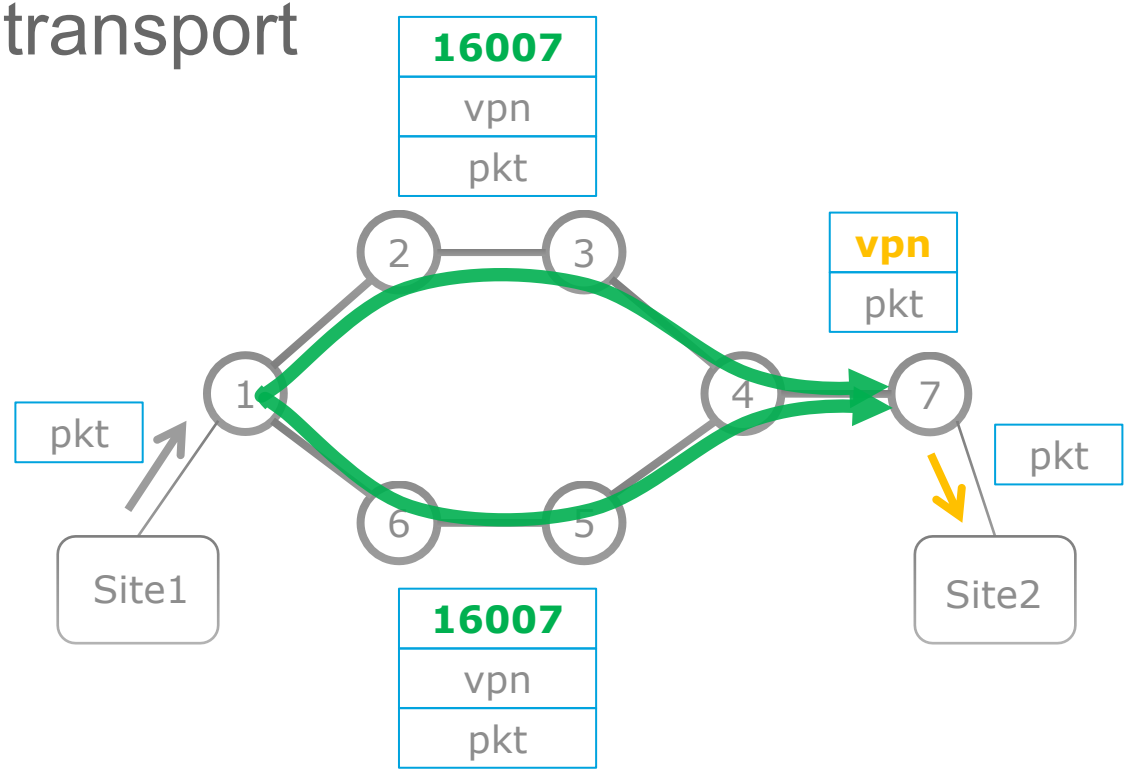
- Seamless Deployment



Use Cases

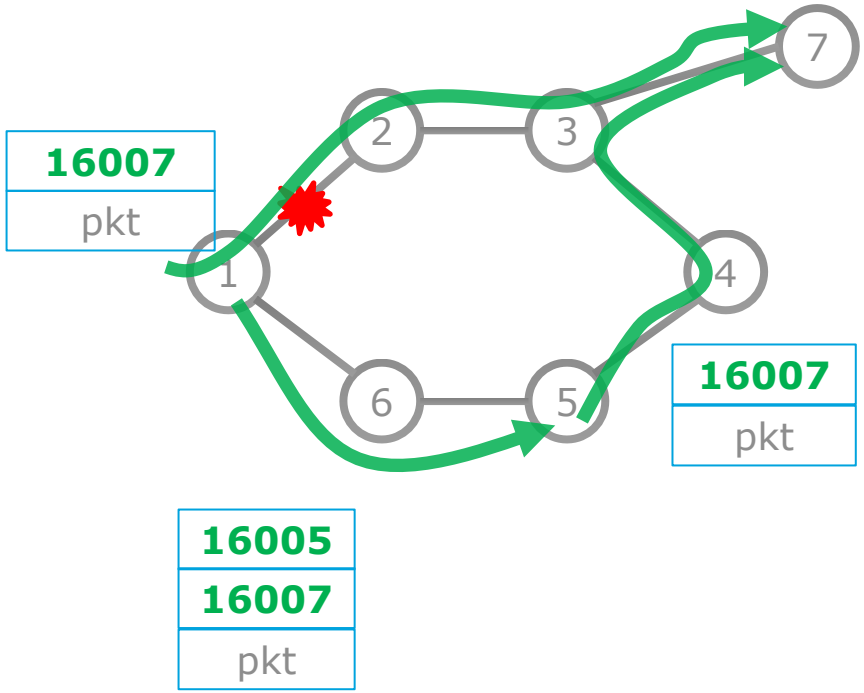
IPv4/6 VPN/Service transport

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



Topology-Independent LFA (TI-LFA FRR)

- 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



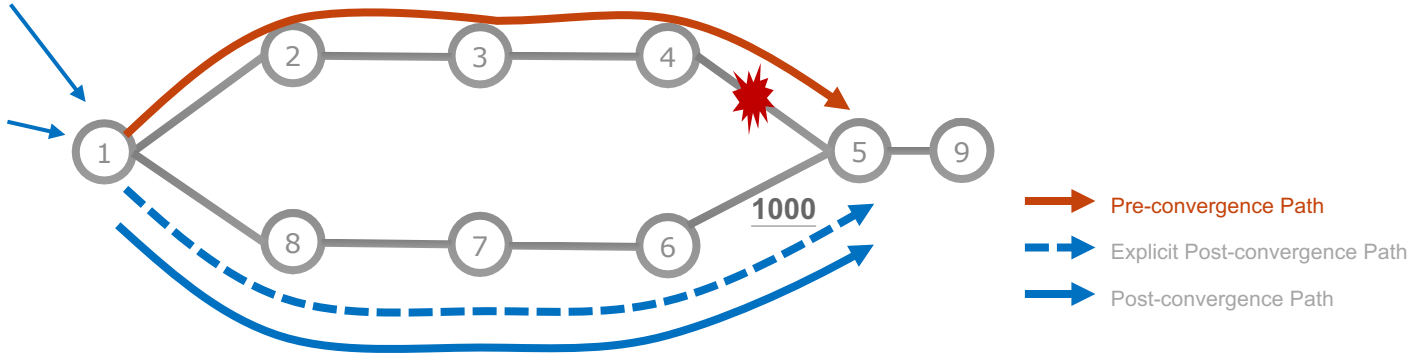
SR Microloop Avoidance

Demo

- Prevent any microloop upon isolated convergence due to
  - link up/down event & metric increase/decrease event
- 2-stage convergence
  - Stage 1: non-looping SID lists to implement the post-convergence path
  - Stage 2: post-convergence path
- If multiple back-to-back convergences, fall back to native IP convergence

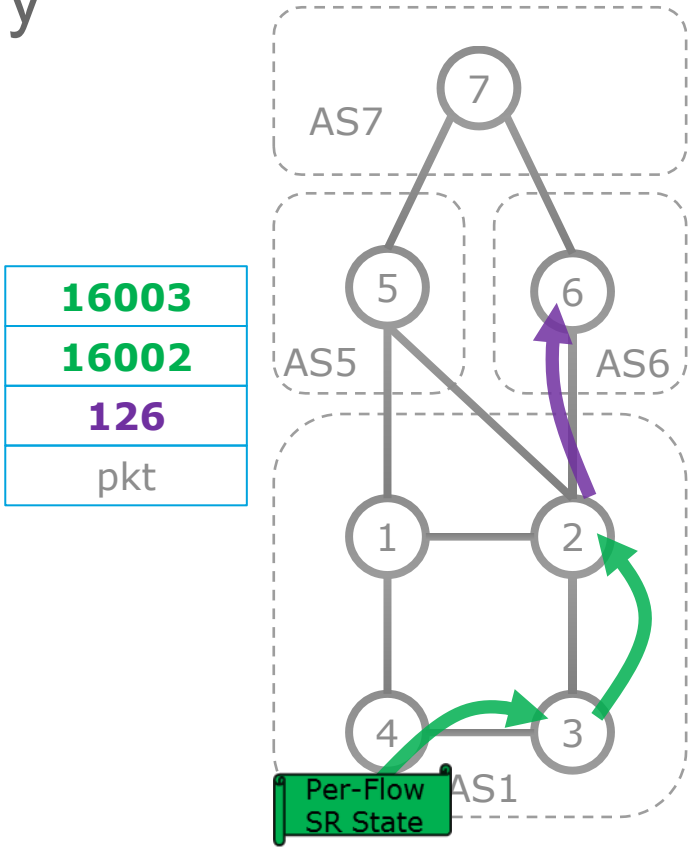
```
microloop avoidance segment-routing
```

FIB @ 1 for Destination 9
Initially: {16009} OIF 2
Stage1: {16006, 24065, 16009}
Stage2: {16009} OIF 8



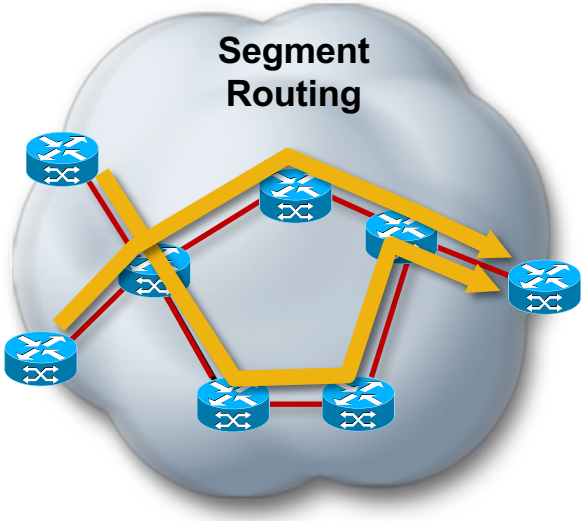
Optimized Content Delivery

- 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



Traffic Engineering with Segment Routing

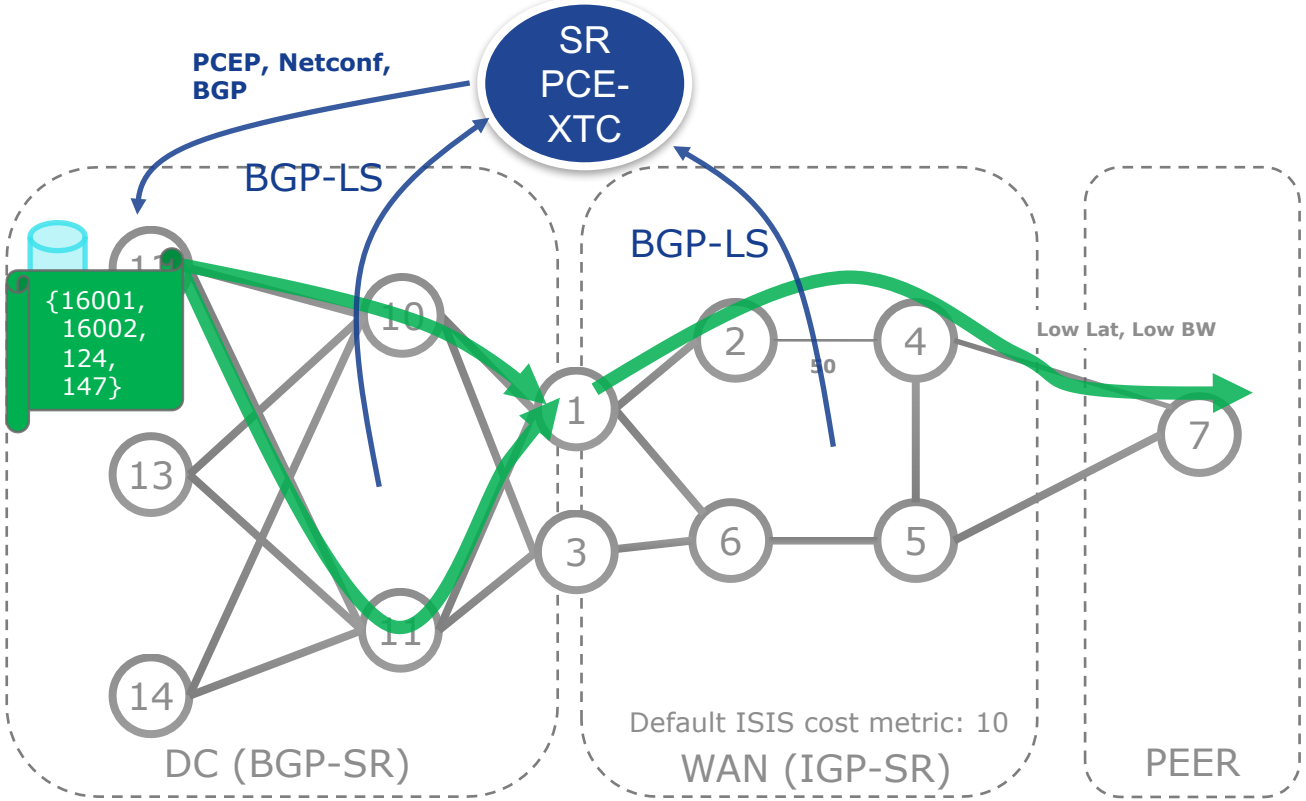
- Provides explicit routing
- Supports constraint-based routing
- Supports centralized admission control
- No RSVP-TE to establish LSPs
- Uses existing ISIS / OSPF extensions to advertise link attributes
- Supports ECMP



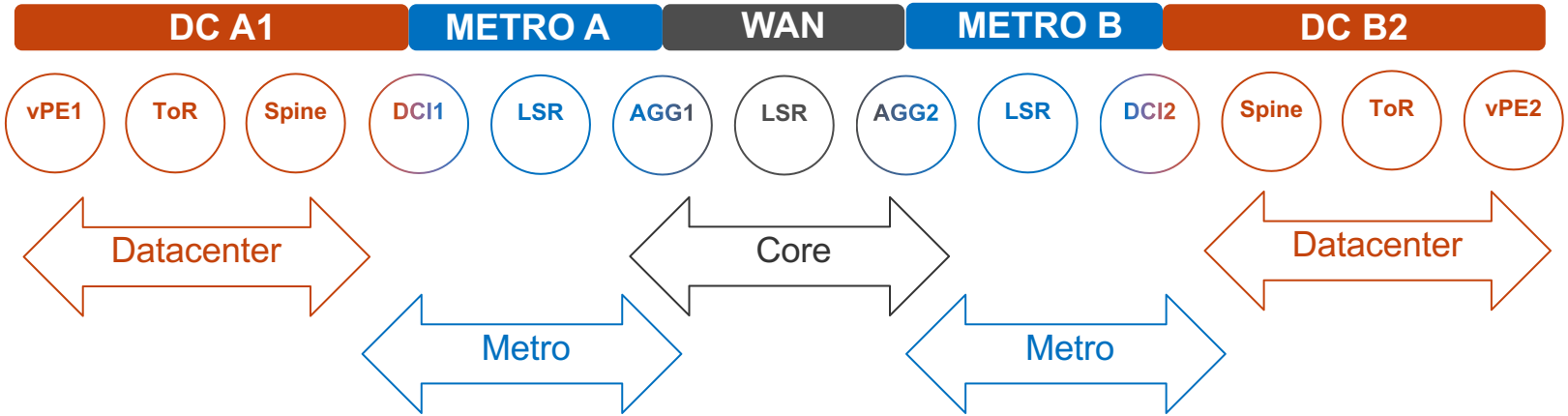
TE LSP

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

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

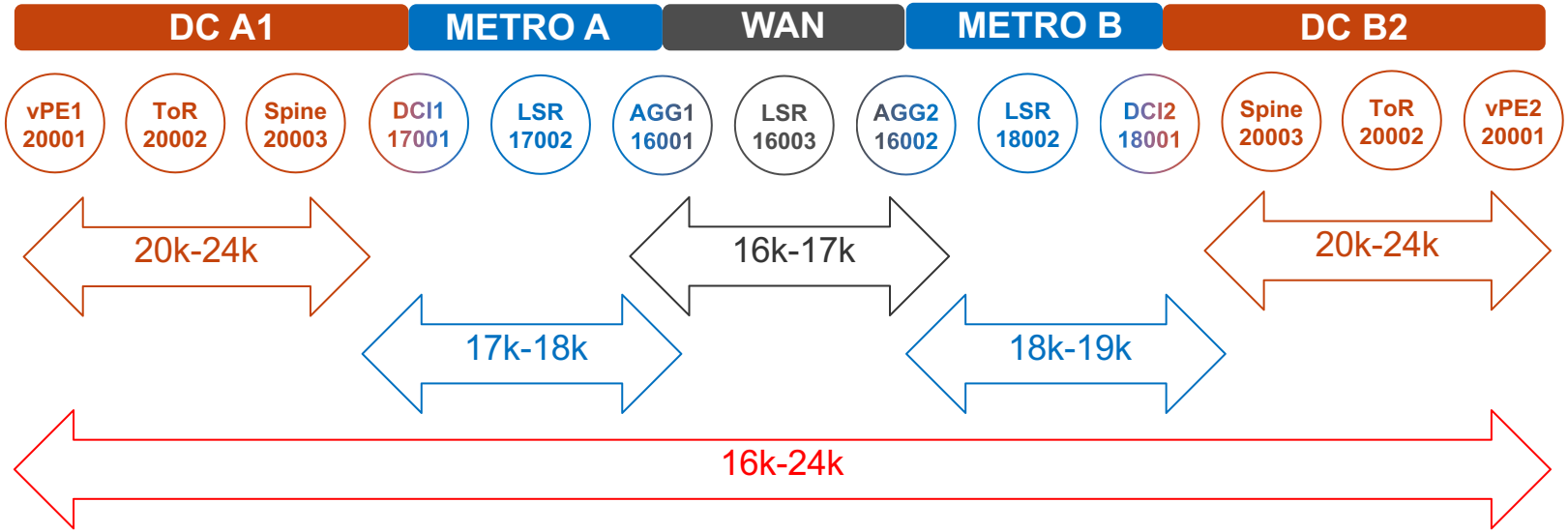


Inter-Domain Policy at Scale



- Segment Routing use-case aiming to scale the network to support hundreds of thousands of network nodes, and tens of millions of physical underlay endpoints
- Applicable to the interconnection of massive-scale DC's and/or large aggregation networks
  - Principles are equally applicable to a network of any size

SRGB and SID allocation



- Homogenous end-to-end SRGB for simplicity
- Globally Unique Prefix SIDs for devices WAN and Metro domains
- Locally Unique Prefix SIDs for Datacenters

Use cases summary

- Simplify transport by removing control protocols (built in)
- Traffic protection with TI-LFA (built in) and anycast SIDs
- Microloop avvoidance (built in)
- Traffic steering
- Easily build out to SDN routing(controller based)
  - => connecting/unifying MPLS domains
- Full SR traffic engineering (Also Multidomain)
  - with ECMP for PATHS and without keeping flow states in the network routers

Conclusion

- Simple routing extensions to implement source routing
- Packet path determined by prepended segment identifiers (one or more)
- Data plane agnostic (MPLS, IPv6)
- Network scalability and agility by reducing network state and simplifying control plane
- Traffic protection with 100% coverage with more optimal routing

Resources

- Stay Informed - Tutorials, Conferences, IETF, Open-source SW
  - http://www.segment-routing.net/
  - Join us – Segment Routing @ LinkedIN
- Get in Touch
  - ask-segment-routing@cisco.com
- “Latest” SR Demonstrations
  - On-demand Next-Hop and SR PCE
  - TI-LFA Node protection
  - Microloop Avoidance
  - SRv6 “Spray” use-case