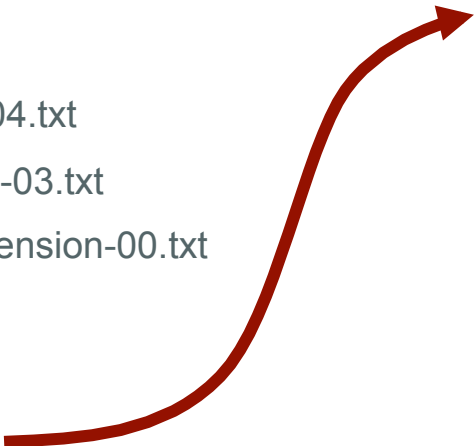


# Segment Routing

- Architecture:
  - draft-filsfils-rtgwg-segment-routing-01.txt
- Use Cases
  - Generic: draft-filsfils-rtgwg-segment-routing-use-cases-02.txt
  - FRR: draft-francois-sr-frr-00.txt
  - OAM: draft-geib-spring-oam-usecase-00.txt
  - To be published: Service Chaining, IPv6
- Segment Routing with MPLS
  - draft-filsfils-spring-segment-routing-mpls-00.txt
  - SR/LDP interop: draft-filsfils-spring-segment-routing-ldp-interop-00.txt
- Protocol Extensions
  - draft-previdi-isis-segment-routing-extensions-04.txt
  - draft-psenak-ospf-segment-routing-extensions-03.txt
  - draft-psenak-ospf-segment-routing-ospfv3-extension-00.txt
  - draft-sivabalan-pce-segment-routing-02.txt
- Many authors of several drafts
  - see each draft for detailed authorship
  - even more contributors...



C. Filsfils  
S. Previdi  
A. Bashandy  
Cisco Systems, Inc.

P. Francois  
IMDEA Networks

B. Decraene  
S. Litkowski  
Orange

M. Horneffer  
R. Geib  
Deutsche Telekom

I. Milojevic  
Telekom Srbija

R. Shakir  
British Telecom

S. Ytti  
TDC Oy

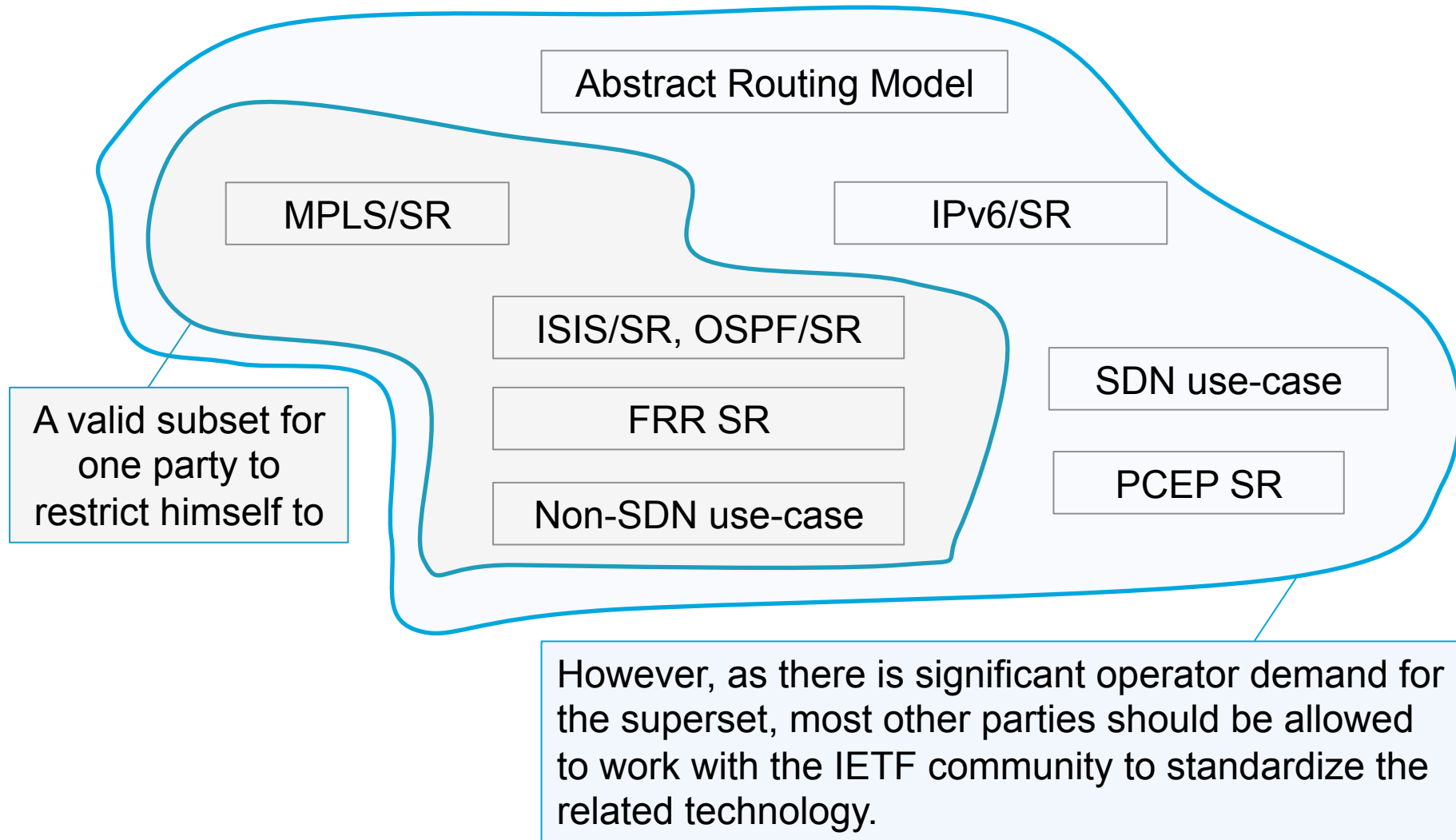
W. Henderickx  
Alcatel-Lucent

J. Tantsura  
S. Kini  
Ericsson

E. Crabbe  
Google, Inc.

H. Gredler  
Juniper Networks

# Visually



# Segment Routing

- **Source Routing**: the source chooses a path and encodes it in the packet header as an ordered list of segments
- **Segment**: an identifier for any type of instruction
  - Service
  - Context
  - Locator
  - IGP-based forwarding construct
  - BGP-based forwarding construct
  - Local value or Global Index
  - Others...
- **Central Orchestration and Optimization**
  - Provide a more responsive and scalable interaction between WAN orchestration, the applications and the network
  - Provide an interesting hybrid approach between centralization and distribution

# Segment Routing Dataplanes

- **MPLS**
  - no change to MPLS dataplane
  - an ordered list of segments is represented as a stack of labels
  - the 12 leftmost bits of the segments are unused
  - a completed segment is popped
- **IPv6**
  - consistent with IPv6 Specification (4.4 of RFC2460)
  - an ordered list of segments is represented as a a new extension header type (TBD)

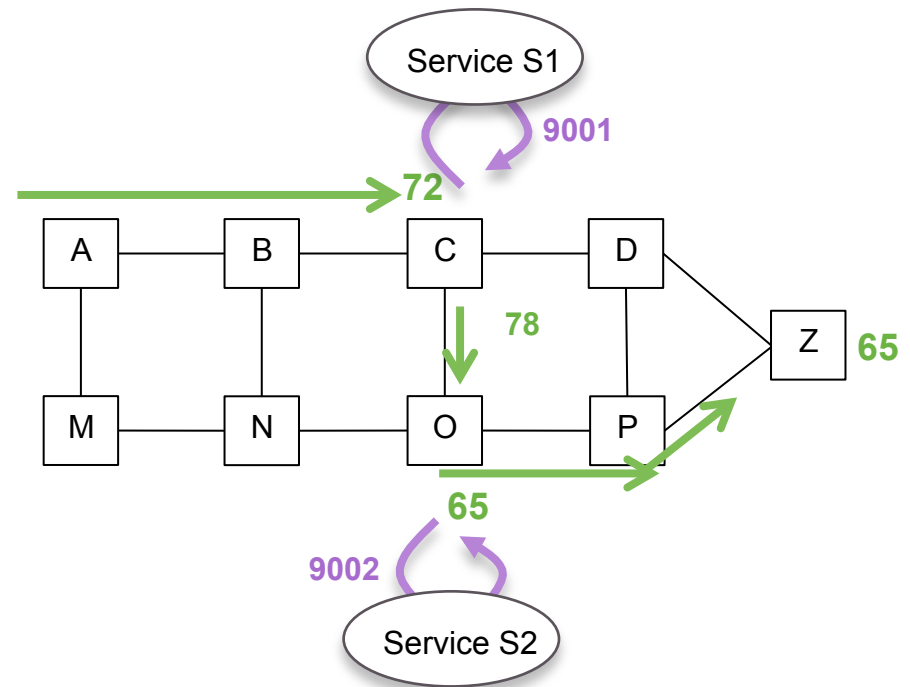
# Real

[www.segment-routing.net](http://www.segment-routing.net)

- Implementation since February 2013
- Commitment to interoperability and standardization
  - Multivendor support , multi-operator support

# Illustration

- 72, 78, 65: global segments representing the shortest-path respectively to C, O and Z
- 9001: local segment to C representing a local service S1
- 9002: local segment to O representing a local service S2
- Ingress node A enforces a source route of forwarding and service instructions on flow F by appending the SR list {72, 9001, 78, 9002, 65} on its packets



# Segment Operation

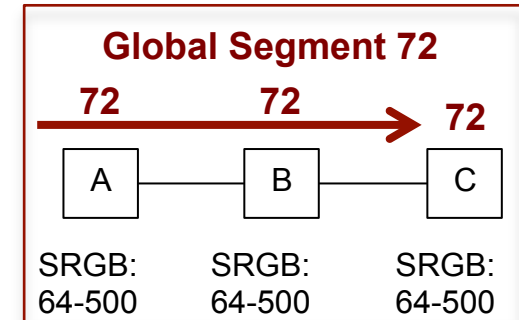
- **Active segment:** the segment defining what current instruction to apply on the packet
  - MPLS: top of the stack
  - IPv6: the segment identified by the pointer
- **PUSH:** an ordered list of segments is inserted
  - MPLS: push at the top of the stack
  - IPv6: inserted at the old pointer position and the pointer is set to the head of the inserted list
- **NEXT:** the active segment is completed, the next segment becomes active
  - MPLS: pop of the top label
  - IPv6: increment the pointer
- **CONTINUE:** the active segment is not completed and stays active
  - MPLS: swap potentially for the same global value (if the nhop has the same SRGB)
  - IPv6: do not increment the pointer

# Local Segment

- The instruction associated with a local segment is only supported by the node originating it.
  - any other node does not install a remote local segment in its FIB
- For example
  - if node N allocates segment 9001 to the local forwarding instruction “complete the segment and forward the packet onto interface I” then it advertises this local instruction with absolute value 9001
  - No other node installs that segment in its SR FIB and hence no conflict can arise.



# Global Segment



- The instruction associated with a global segment is understood by any node in the SR domain. Any node in the domain install the related instruction in its FIB.
- Global segments fall in a sub-space of the segment space called the SRGB: SR Global Block
  - Absolute: all the nodes in the SR domain have the same SRGB
  - Indexed: any node can have a different SRGB
- Most SR operators will allocate the same SRGB at every node in the SR domain. We will use this representation throughout this presentation.
- If all the nodes are configured with the same SRGB [64-500], then the global segment 72 can be installed at any node to perform the global instruction “forward along the shortest-path to C”
- The same segment value is found at each SR node: 72

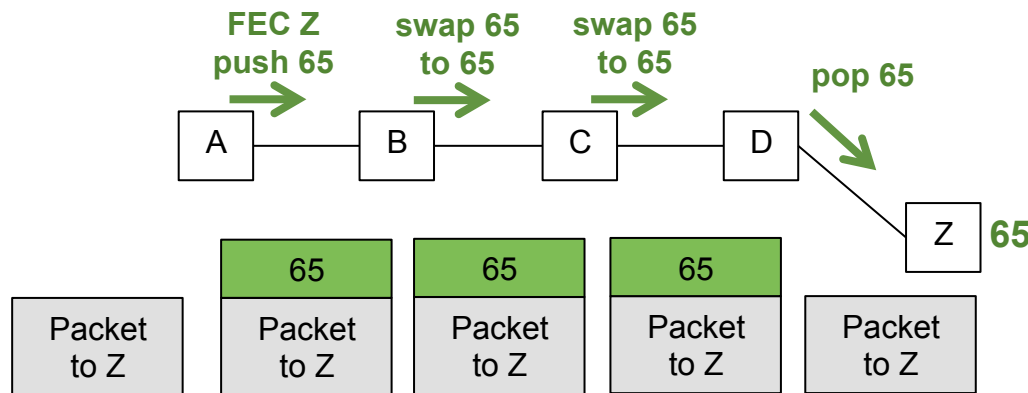
# Segment Routing Key Properties

- **Explicit routing**
  - strict or loose
- **The flow state is in the header, not in the network**
  - Control and Dataplane State Reduction
  - A node holds a state per global segment in the SR domain
    - > typically less than 5000
  - A node holds a state per local segment it originates
    - > Typically less than 500
  - For a flow F, only its ingress node N holds a per-flow state for F. Any other node does not hold any state for F. While they can be millions of flows crossing a midpoint, its SR FIB scale is likely less than 5500

# IGP Segments

- draft-previdi-isis-segment-routing-extensions-04.txt  
draft-psenak-ospf-segment-routing-extensions-03.txt  
draft-psenak-ospf-segment-routing-ospfv3-extension-00.txt
- Prefix Segment
  - a segment associated to a prefix
  - steers traffic along ECMP-aware shortest-path to the prefix
  - global Segment
- Node Segment
  - a segment allocated to a loopback that identifies a specific node
  - a frequent sub-case of prefix segment
- Adjacency Segment
  - steers traffic onto an adjacency or a set of adjacencies
  - local segment

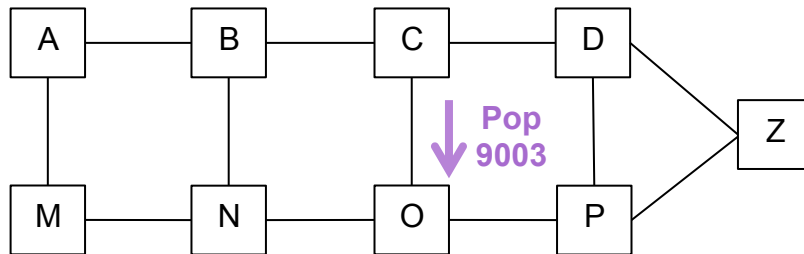
# Node Segment



A packet injected anywhere with top segment 65 will reach Z via shortest-path

- Z advertises a global node segment 65 with its loopback
  - simple ISIS sub-TLV extension
  - we assume the same SRGB at every node
- All remote nodes install in their FIB the node segment 65 to Z

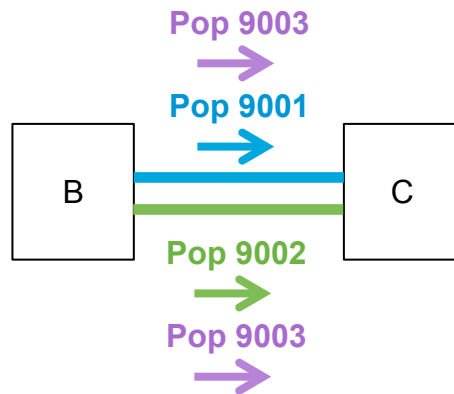
# Adjacency Segment



A packet injected at node C with segment 9003 is forced through datalink CO

- C allocates a local segment 9003 and maps it to the instruction “complete the segment and forward along the interface CO”
- C advertises the adjacency segment in ISIS
  - simple sub-TLV extension
- C is the only node to install the adjacency segment in FIB

# Datalink and Bundle



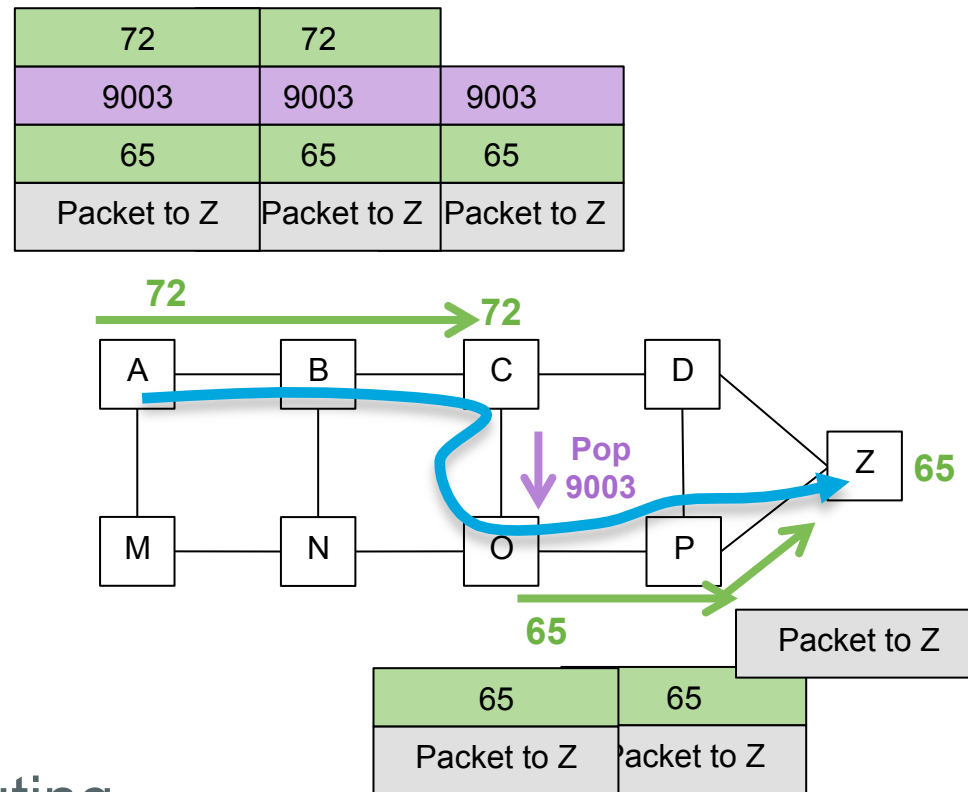
9001 switches on blue member

9002 switches on green member

9003 load-balances on any member of the adj

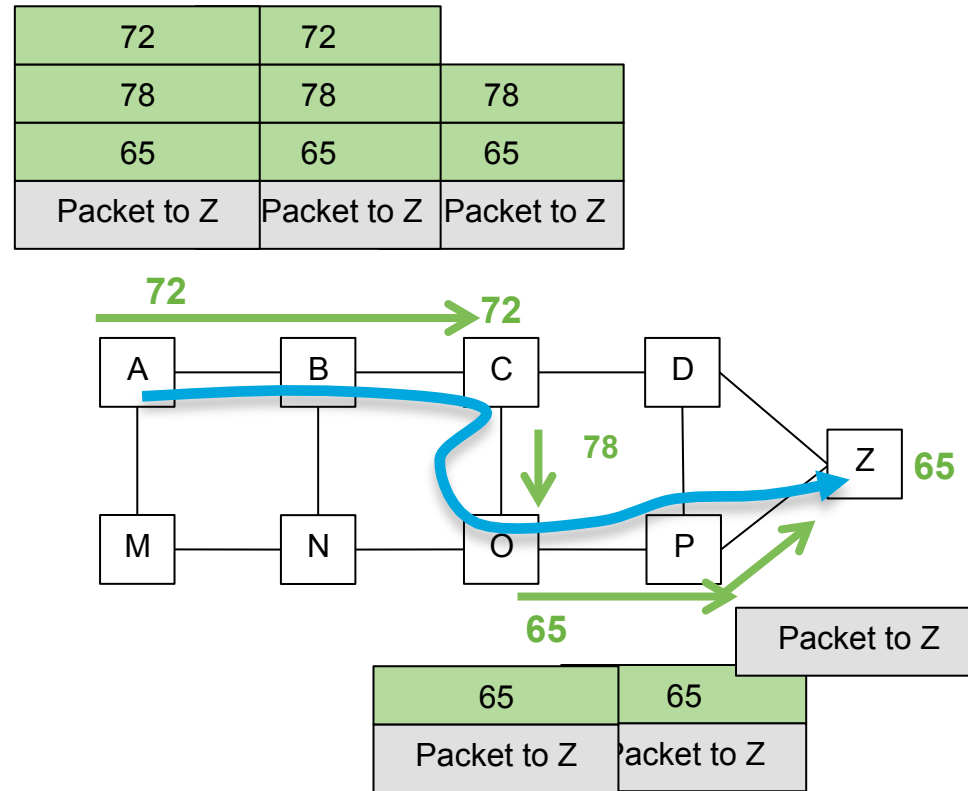
- Adjacency segment represents a specific datalink to an adjacent node
- Adjacency segment represents a set of datalinks to the adjacent node
- Multiple segments can be associated with an ISIS adjacency

# Combining Segments



- Source Routing
- Any explicit path can be expressed: ABCOPZ

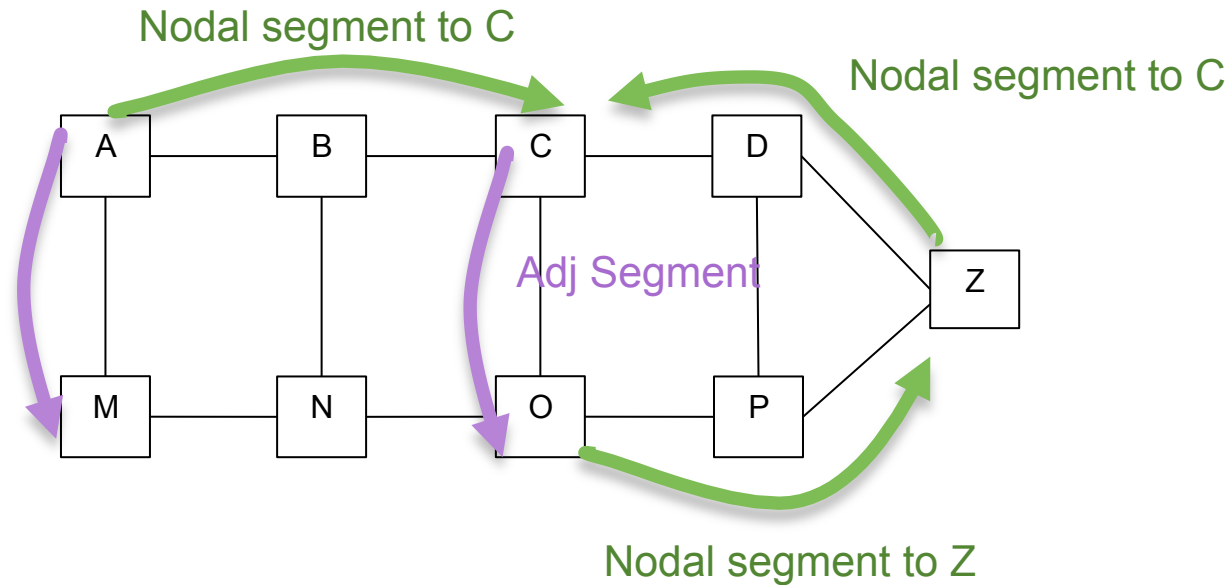
# Combining Segments



- Node Segment is at the heart of the proposal
  - ecmp multi-hop shortest-path
  - in most topologies, any path can be expressed as list of node segments



# ISIS/OSPF automatically install segments



- Simple extension
- Excellent Scale: a node installs  $N+A$  FIB entries
  - $N$  node segments and  $A$  adjacency segments

# Local Service Segment

- 72, 78, 65: global segments representing the shortest-path respectively to C, O and Z
- 9001: local segment to C representing a local service S1
- 9002: local segment to O representing a local service S2
- Ingress node A enforces a source route of forwarding and service instructions on flow F by appending the SR list {72, 9001, 78, 9002, 65} on its packets
- 9001 and 9002 represent local services

