Segment Routing

Clarence Filsfils – cf@cisco.com
Distinguished Engineer
Christian Martin – martincj@cisco.com
Sr. Director, Engineering

Agenda

- Introduction
- Technology
- Use Cases
- Conclusion

Introduction

Goals and Requirements

- Make things easier for operators
  - Improve scale, simplify operations
  - Minimize introduction complexity/disruption
- Enhance service offering potential through programmability
- Leverage the efficient MPLS dataplane that we have today
  - Push, swap, pop
  - Maintain existing label structure
- Leverage all the services supported over MPLS
  - Explicit routing, FRR, VPNv4/6, VPLS, L2VPN, etc
- IPv6 dataplane a must, and should share parity with MPLS

Operators Ask For Drastic LDP/RSVP Improvement

- Simplicity
  - less protocols to operate
  - less protocol interactions to troubleshoot
  - avoid directed LDP sessions between core routers
  - deliver automated FRR for any topology
- Scale
  - avoid millions of labels in LDP database
  - avoid millions of TE LSP's in the network
  - avoid millions of tunnels to configure

Operators Ask For A Network Model Optimized For Application Interaction

- Applications must be able to interact with the network
  - cloud based delivery
  - internet of everything
- Programmatic interfaces and Orchestration
  - Necessary but not sufficient
- The network must respond to application interaction
  - Rapidly-changing application requirements
  - Virtualization
  - Guaranteed SLA and Network Efficiency

Segment Routing

- Simple to deploy and operate
  - Leverage MPLS services & hardware
  - straightforward ISIS/OSPF extension to distribute labels
  - LDP/RSVP not required
- Provide for optimum scalability, resiliency and virtualization
- SDN enabled
  - simple network, highly programmable
  - highly responsive

IETF

- Simple ISIS/OSPF extension
- Welcoming contribution

S. Previdi, Ed.
C. Filsfils, Ed.
A. Bashandy
Cisco Systems, Inc.
M. Horneffer
Deutsche Telekom
B. Decraene
S. Litkowski
Orange
I. Milojevic
Telekom Srbija
R. Shakir
British Telecom
S. Ytti
TDC Oy
W. Henderickx
Alcatel-Lucent
J. Tantsura
Ericsson
March 20, 2013

Segment Routing with IS-IS Routing Protocol
draft-previdi-filsfils-isis-segment-routing-02

Abstract

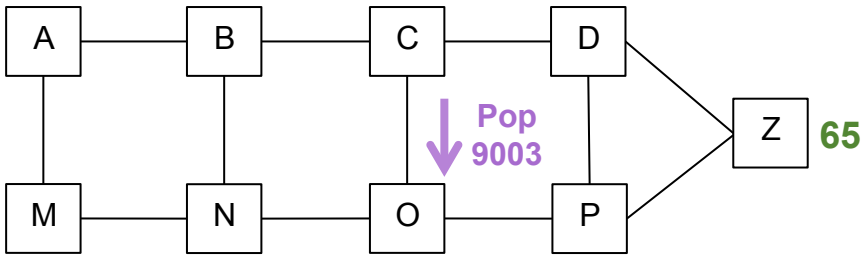
Segment Routing (SR) enables any node to select any path (explicit or derived from IGPs SPT computations) for each of its traffic classes. The path does not depend on a hop-by-hop signaling technique (neither LDP nor RSVP). It only depends on a set of "segments" that are advertised by the IS-IS routing protocol. These segments act as topological sub-paths that can be combined together to form the desired path.

Segment Routing

Segment Routing

- Forwarding state (segment) is established by IGP
  - LDP and RSVP-TE are not required
  - Agnostic to forwarding dataplane: IPv6 or MPLS
- MPLS Dataplane is leveraged without any modification
  - push, swap and pop: all that we need
  - segment = label
- Source Routing
  - source encodes path as a label or stack of segments
  - two segments: node or adjacency

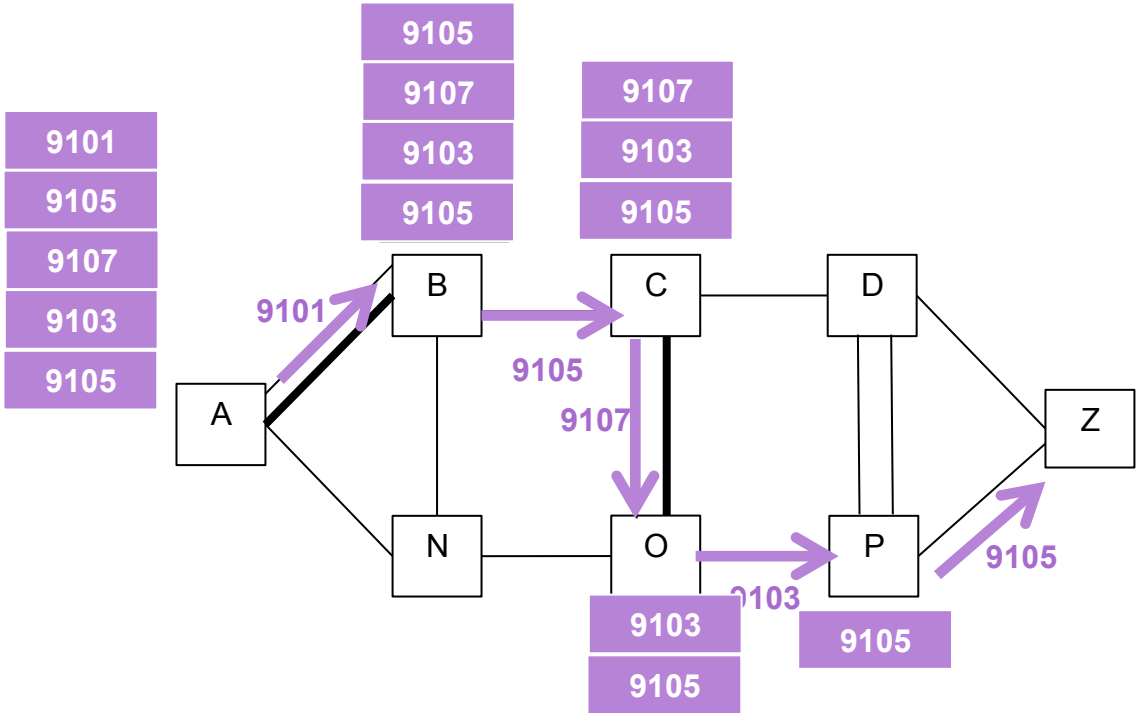
Adjacency Segment



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

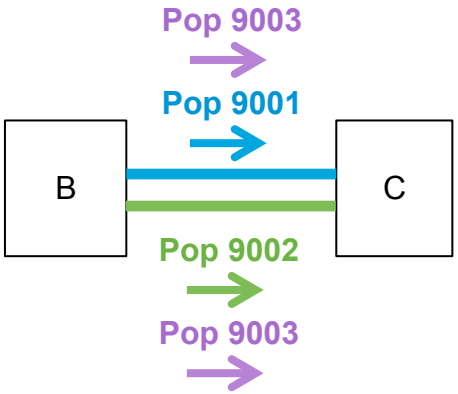
- C allocates a local label
- C advertises the adjacency label in ISIS
  - simple sub-TLV extension
- C is the only node to install the adjacency segment in MPLS dataplane

A path with Adjacency Segments



- Source routing along any explicit path
  - stack of adjacency labels
- SR provides for entire path control

Datalink and Bundle



9001 switches on blue member

9002 switches on green member

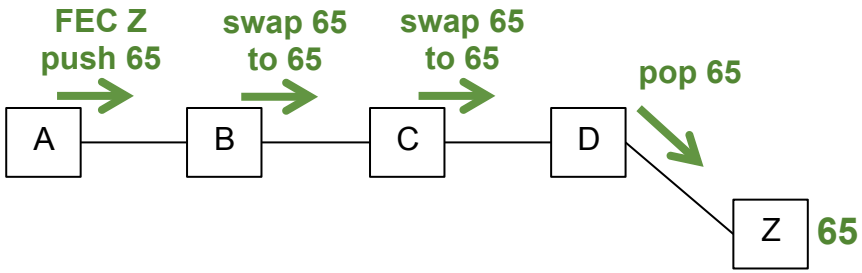
9003 load-balances on any member of the adj

- Adjacency segment represents a specific datalink to an adjacent node
- Adjacency segment represents a set of datalinks to the adjacent node

Node SR Range

- SR requires only 1 label per node in the IGP domain
  - insignificant: < 1% of label space
- Node SR Range
  - a range of labels allocated to the SR control-plane
  - e.g. [64, 5000]
- Each node gets one unique label from SR Range
  - Node Z gets label 65

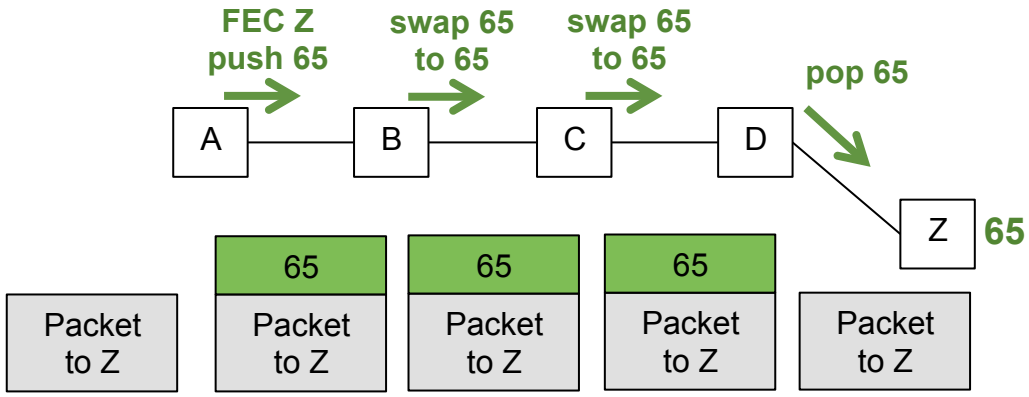
Node Segment



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

- Z advertises its node segment
  - simple ISIS sub-TLV extension
- All remote nodes install the node segment to Z in the MPLS dataplane

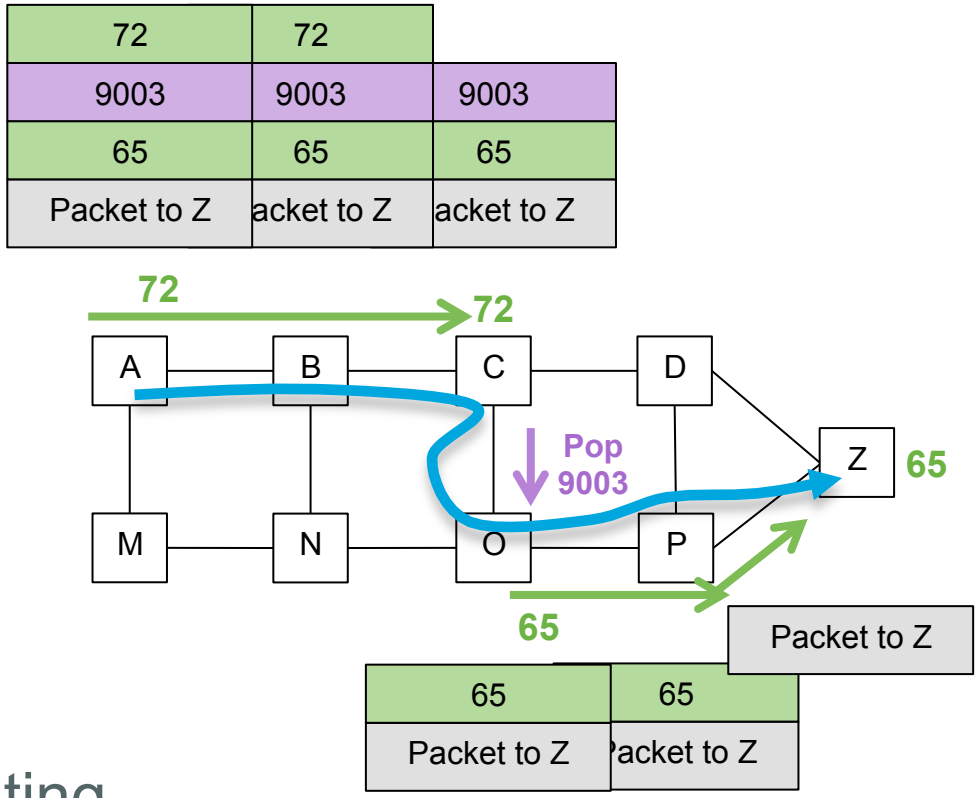
Node Segment



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

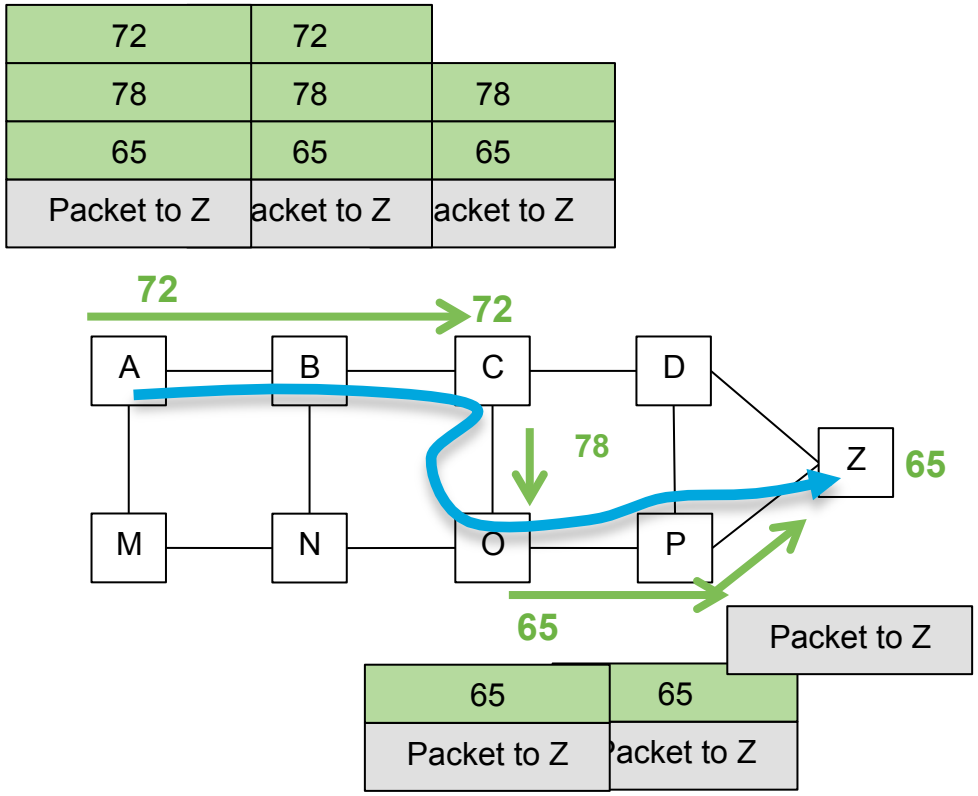
- Z advertises its node segment
  - simple ISIS sub-TLV extension
- All remote nodes install the node segment to Z in the MPLS dataplane

Combining Segments



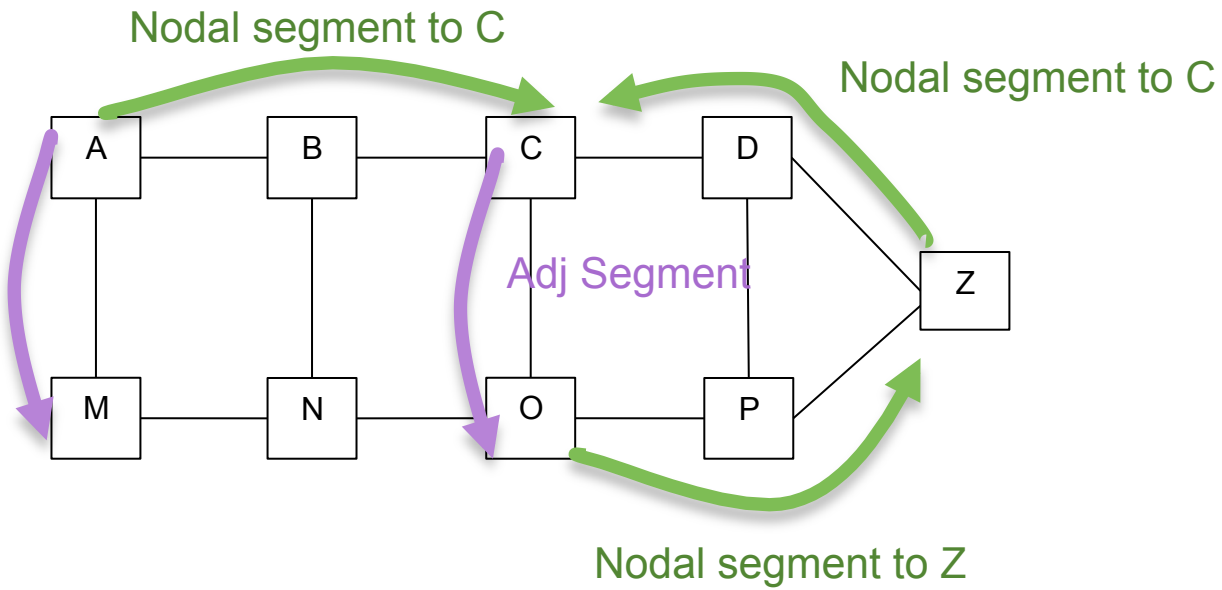
- Source Routing
- Any explicit path can be expressed: ABCOPZ

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

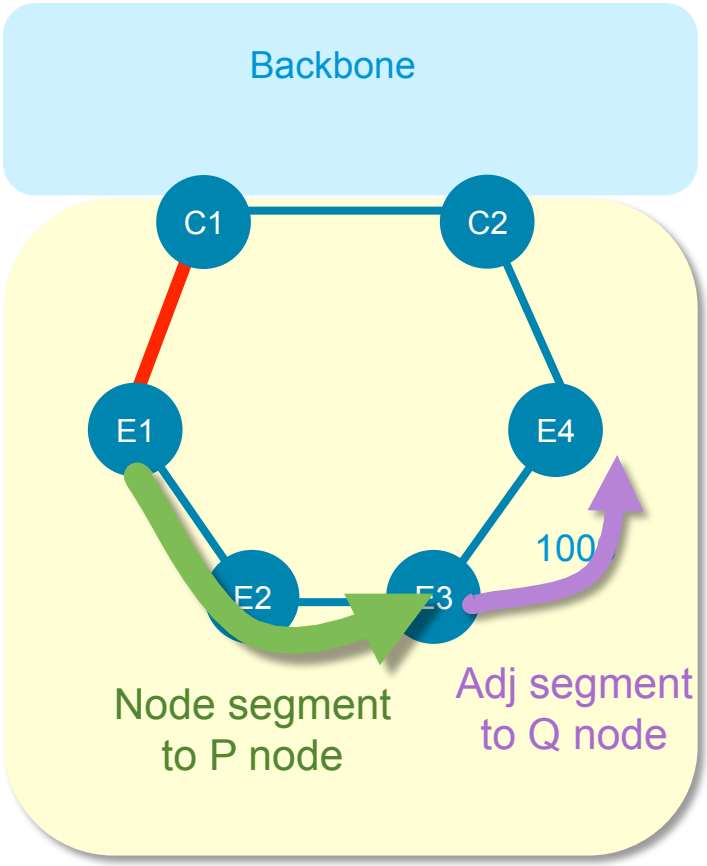
ISIS automatically installs segments



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

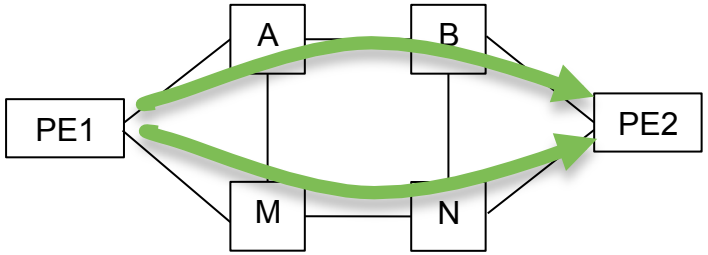
Automated & Guaranteed FRR

- IP-based FRR is guaranted in any topology
  - 2002, LFA FRR project at Cisco
  - draft-bryant-ipfrr-tunnels-03.txt
- Directed LFA (DLFA) is guaranteed when metrics are symetric
- No extra computation (RLFA)
- Simple repair stack
  - node segment to P node
  - adjacency segment from P to Q



Use Cases

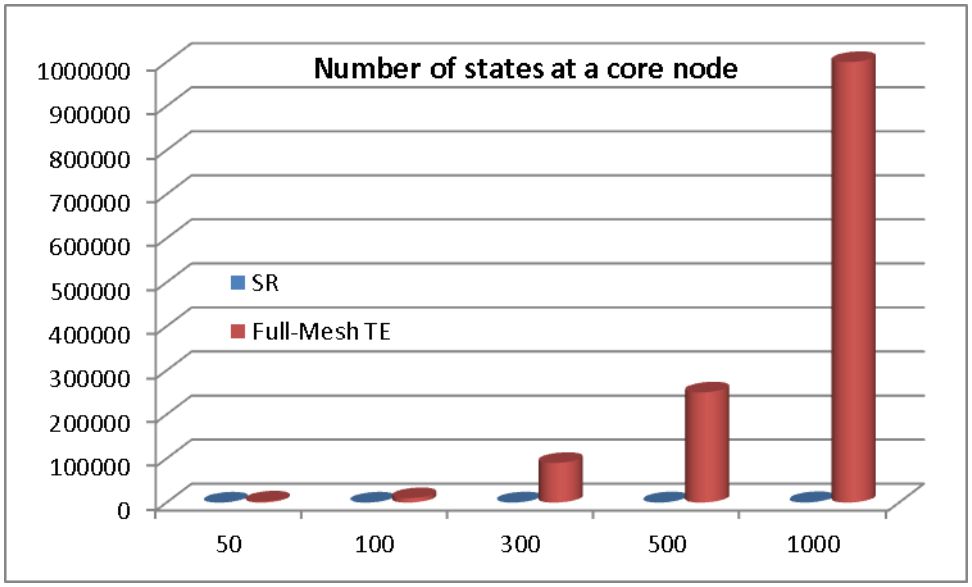
Simple and Efficient Transport of MPLS services



All VPN services ride on the node segment
to PE2

- Efficient packet networks leverage ecmp-aware shortest-path!
  - node segment!
- Simplicity
  - no complex LDP/ISIS synchronization to troubleshoot
  - one less protocol to operate

Scalable TE



- An SR core router scales much than with RSVP-TE
  - The state is not in the router but in the packet
  - N+A vs N^2

N: # of nodes in the network
A: # of adjacencies per node

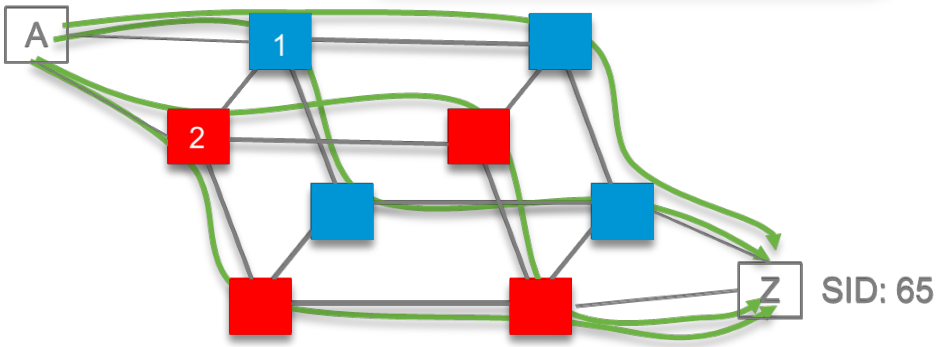
Simple Disjointness

SR avoids state in the core

SR avoids enumerating RSVP-TE tunnels for each ECMP paths

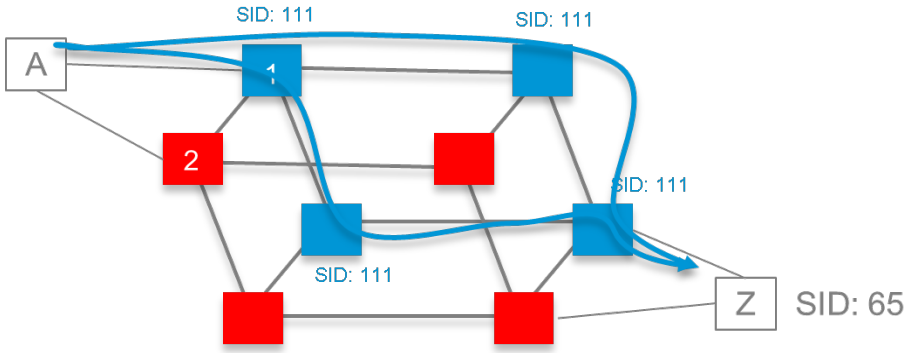
- A sends traffic with [65]

  Classic ECMP “a la IP”



- A sends traffic with [111, 65]

  Packet gets attracted in blue plane and then uses classic ecmp “a la IP”



ECMP-awareness!

CoS-based TE

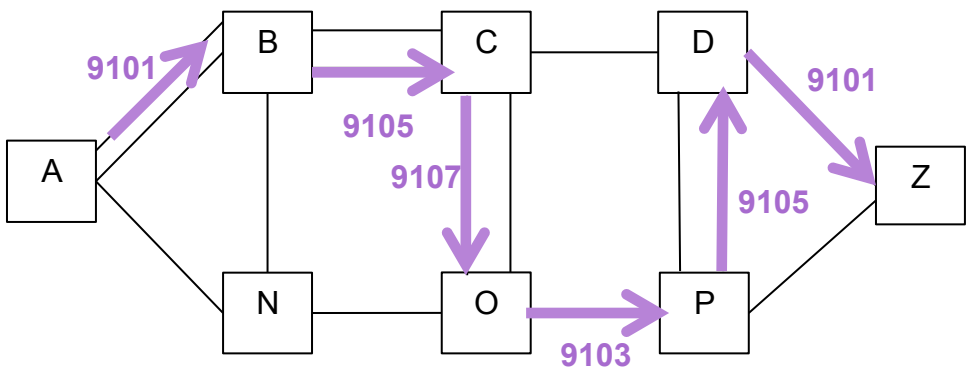
- Tokyo to Brussels
  - data: via US: cheap capacity
  - voip: via russia: low latency
- CoS-based TE with SR
  - IGP metric set such as
    - > Tokyo to Russia: via Russia
    - > Tokyo to Brussels: via US
    - > Russia to Brussels: via Europe
  - Anycast segment “Russia” advertised by Russia core routers
- Tokyo CoS-based policy
  - Data and Brussels: push the node segment to Brussels
    - ➔ ECMP-aware shortest-path to Brussels
  - VoIP and Brussels: push the anycast node to Russia, push Brussels
    - ➔ ECMP-aware shortest-path to Russia, followed by ECMP-aware shortest-path to Brussels

Node segment to Brussels

Node segment to Russia

No TE tunnel enumeration,
no TE state in the core

Full control and OAM



- For Traffic Engineering
- or for OAM

Google

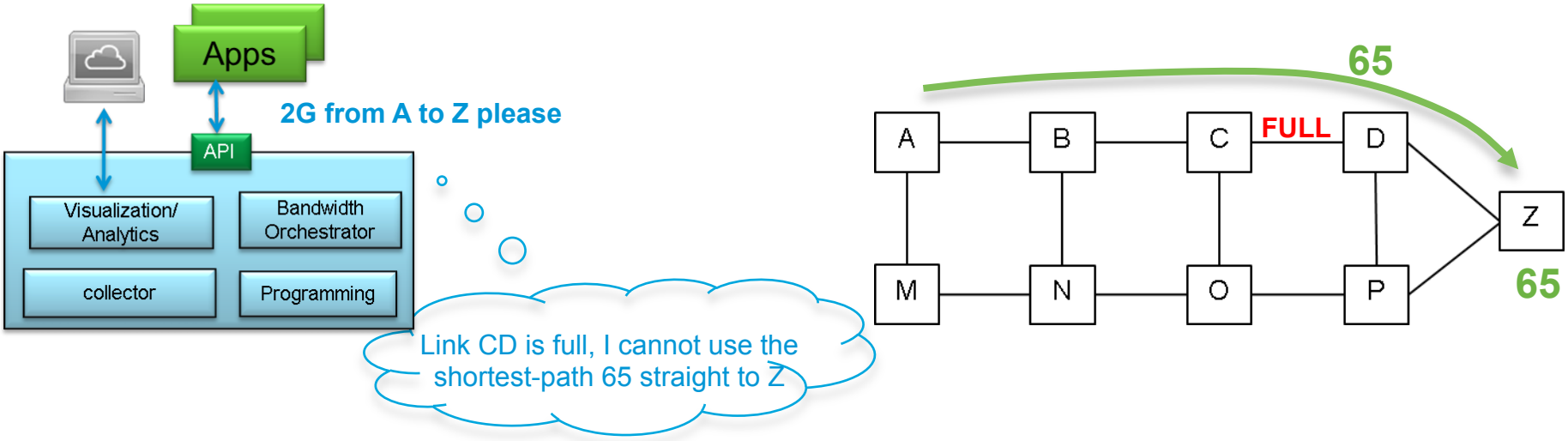
Localizing packet loss

In a large complex network

Nicolas Guilbaud nguilbaud@google.com
Ross Cartlidge rossc@google.com

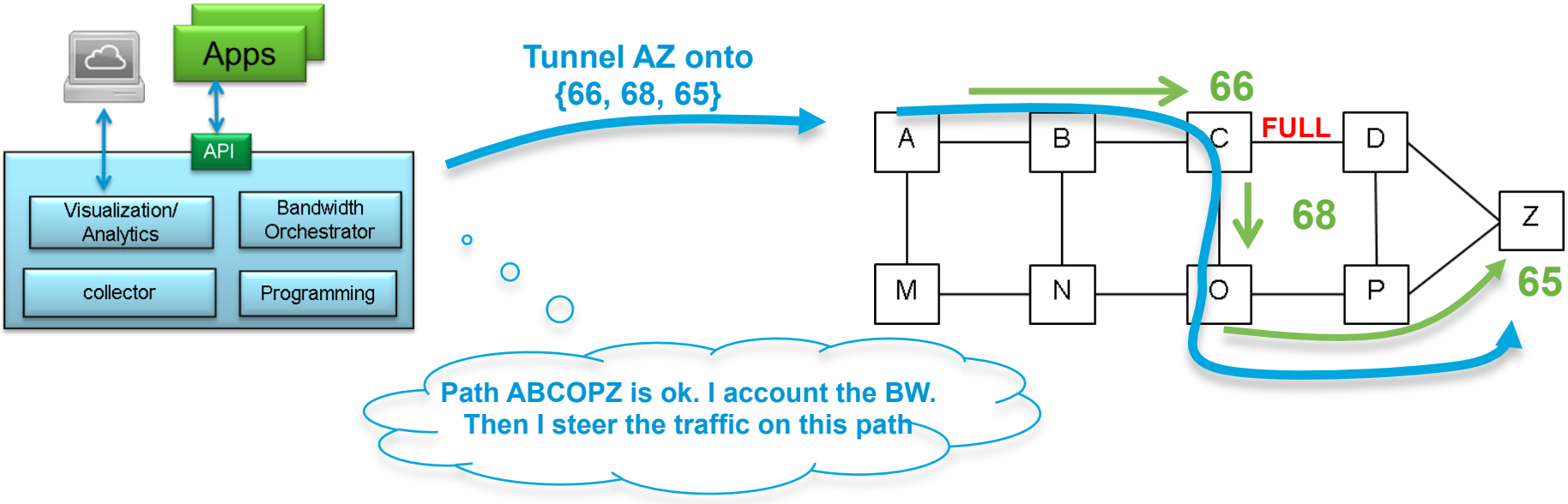
Nanog57, Feb 2013

Application controls – network delivers



- The network is simple, highly programmable and responsive to rapid changes
  - The controller abstracts the network topology and traffic matrix
  - Perfect support for centralized optimization efficiency, if required

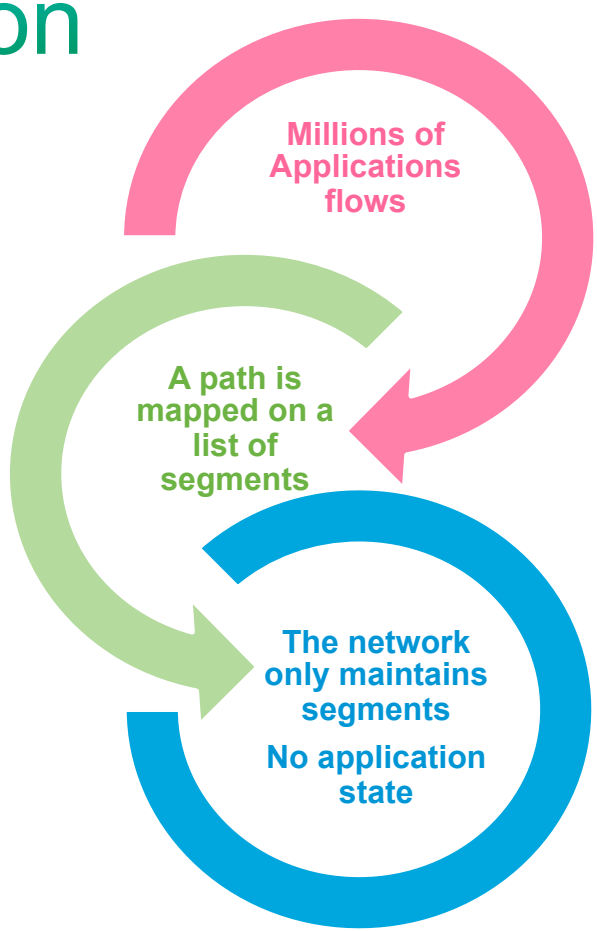
Application controls – network delivers



- The network is simple, highly programmable and responsive to rapid changes

Scalability and Virtualization

- Each engineered application flow is mapped on a path
  - millions of paths
  - maintained in the orchestrator, scaled horizontally
- A path is expressed as an ordered list of segments
- The network maintains segments
  - thousands of segments
  - completely independent of application size/frequency



Conclusion

Segment Routing

- Simple to deploy and operate
  - Leverage MPLS services & hardware
  - straightforward ISIS/OSPF extension
- Provide for optimum scalability, resiliency and virtualization
- Perfect integration with application
- EFT and IETF available – test and contribute