Segment Routing

Clarence Filsfils – cf@cisco.com

Agenda

- Objective
- Abstract Routing Model
- Use-cases
- Conclusion
- Q&A

SR Objectives

- Tackling issues reported by operators for years
  - IGP-based FRR for any topology
  - Simpler to operate, more scalable explicit routing
- Supporting “SDN”-based services
  - Provide a more responsive and scalable interaction between WAN orchestration, the applications and the network
- Evolution, no revolution
  - Must be simple to operate
  - Must support incremental deployment

Objective for this Nanog talk

- Informative
- Trigger your interest
  - A wealth of details in the upcoming drafts
- Seek your involvement
- Brief
  - We could speak for a full-day as we have much research and use-cases to share and discuss

Details

- draft-filsfils-rtgwg-segment-routing-00
- draft-filsfils-rtgwg-segment-routing-use-cases-00
- draft-previdi-isis-segment-routing-extensions-00
- draft-psenak-ospf-segment-routing-extensions-00
- draft-msiva-pce-pcep-segment-routing-extensions-00

Real

- Excellent endorsement and **leadership** from SP and Entreprise community
- Multi-vendor consensus and collaboration
- By mid June, we will submit detailed IETF drafts
  - Architecture
  - Use-cases
  - ISIS extensions
  - OSPF extensions
  - PCEP extensions
  - FRR
- SR EFT is available since Feb 28
  - 12k, ASR9k, CRS1, CRS3

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

Abstract Routing Model

draft-filsfils-rtgwg-segment-routing-00

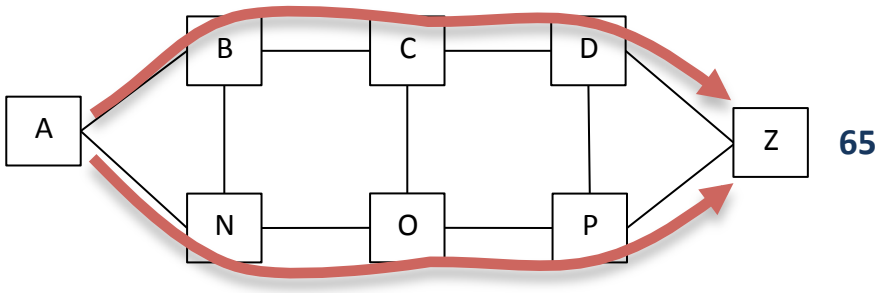
Segment Routing

- A 32-bit segment can represent any instruction
  - Service
  - Context
  - IGP-based forwarding construct
  - Locator
- Ordered list of segments
  - An ordered chain of topological and service instructions
- Per-flow state only at ingress SR edge node
  - Ingress edge node pushes the segment list on the packet

IGP Segments

- Prefix Segment
  - Steers traffic along ECMP-aware shortest-path to the related IGP Prefix
  - Global segment within the SR IGP domain
  - Node Segment: a prefix segment allocated to a prefix that identifies a specific node (e.g. the prefix is its loopback)
- Adjacency Segment
  - Steers traffic onto an adjacency or a set of adjacencies
  - Local segment related to a specific SR node
- SR Global Block
  - A subset of the Segment space
  - All the global segments must be allocated from SRGB
  - Operator manages SRGB like an IP address block: it ensures unique allocation of a global segment within the SR domain

IGP Prefix Segment

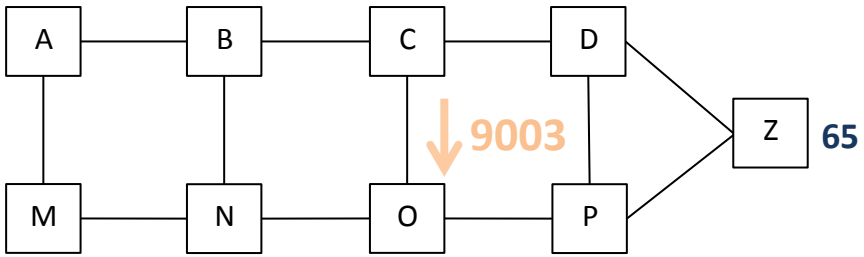


A packet injected anywhere with active segment 65 will reach Z via ecmp-aware shortest-path

- Z advertises its global prefix segment 65 with his loopback address Z/32
  - simple ISIS sub-TLV extension
  - simple OSPF Opaque sub-TLV extension
- All remote nodes install the prefix segment to Z in the SR dataplane along the shortest path to Z/32
- IPv4 and IPv6

draft-previdi-isis-segment-routing-extensions-00
draft-psenak-ospf-segment-routing-extensions-00

IGP Adjacency Segment

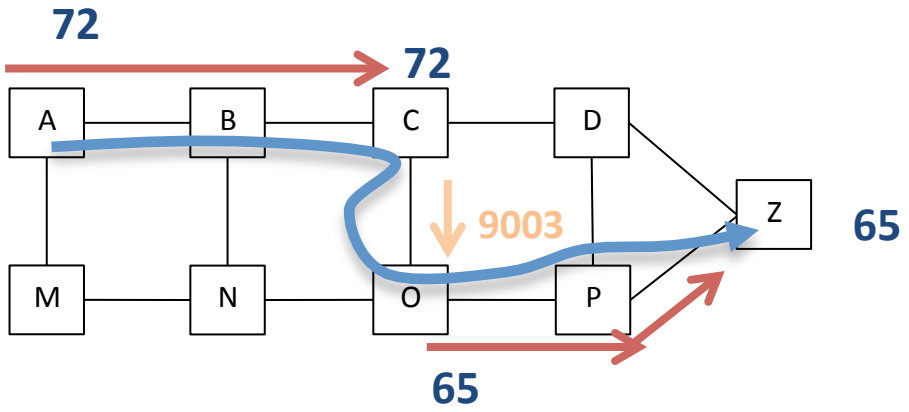


A packet injected at node C with active segment 9003 is forced through datalink CO

- C allocates a local segment 9003 for its adjacency CO
- C advertises the adjacency segment in the IGP
  - Simple ISIS sub-TLV extension
  - simple OSPF Opaque sub-TLV extension
- C is the only node to install the adjacency segment in SR dataplane
- IPv4 and IPv6

draft-previdi-isis-segment-routing-extensions-00
draft-psenak-ospf-segment-routing-extensions-00

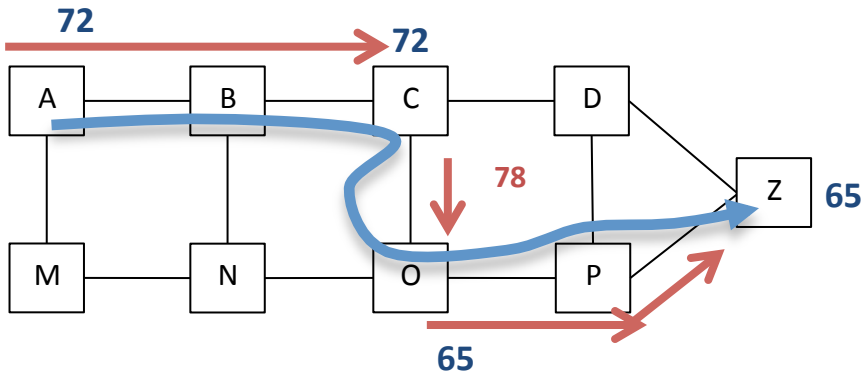
Combining Segments



- Source Routing
- ABCOPZ is expressed as {**72**, **9003**, **65**}

Combining Segments

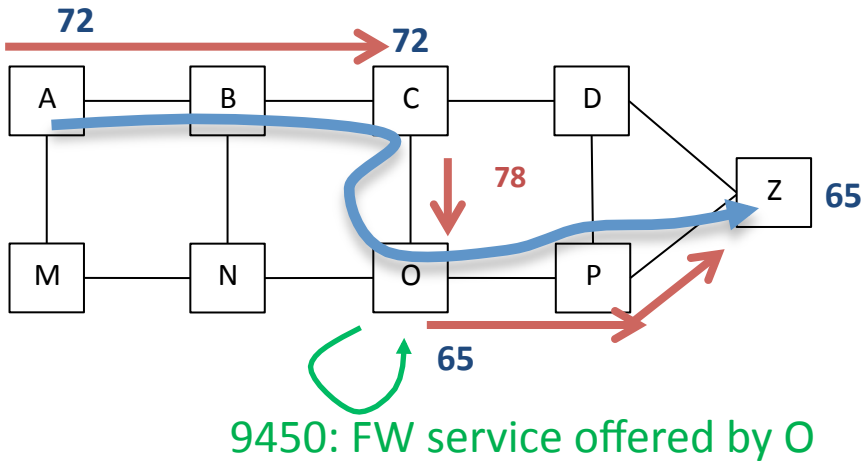
{72, 78, 65}



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

Combining Segments

{72, 78, 9450, 65}



- Service Segments can be part of the source route

SR Control-Plane

- Lightweight extension to ISIS/OSPF
- IPv4 and IPv6
- Agnostic to the dataplane
  - works with any dataplane that supports the encoding of a list of segments on the packet

MPLS dataplane

- The 20 right-most bits of the segment are encoded as a label
- A list of segments is represented as a stack of labels
- The active segment is the top label
- The IGP Prefix segment stays on the top of the stack thanks to a SWAP operation where the ingress and egress label values are the same
- Transports IPv4 and IPv6
- No changes in the operations of the MPLS dataplane
- SR can co-exist and interwork with other MPLS control-plane protocols (LDP, RSVP)

IPv6 dataplane

(without any MPLS dataplane)

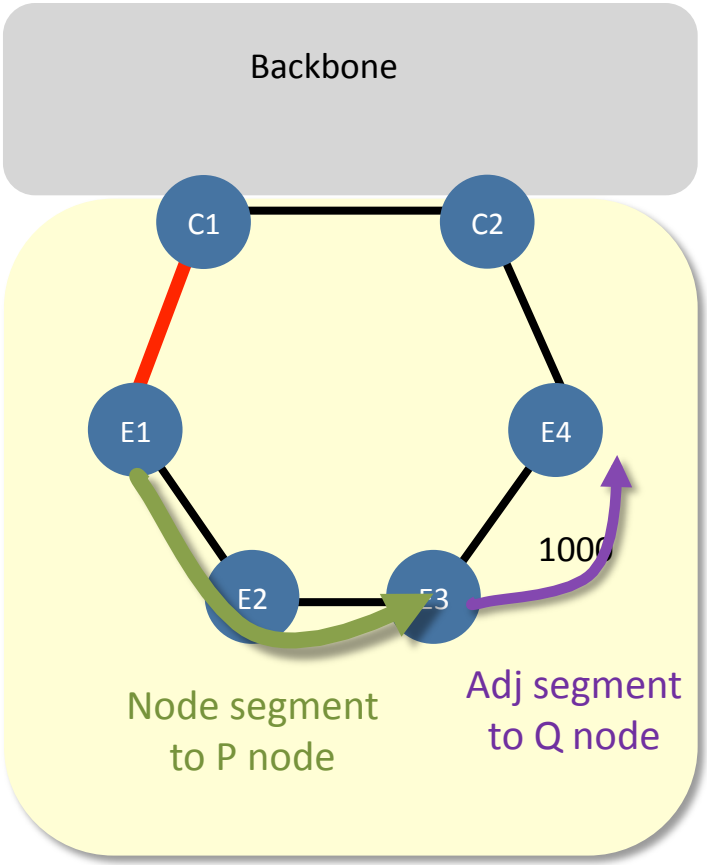
- All the SR ISIS/OSPF Control Plane is dataplane agnostic and hence applies directly to IPv6
- Remaining work: detailing the IPv6 tunneling and new Routing Extension type header
  - High-level description provided at March IPv6 Conference
  - Detailed Draft should be available soon
    - We are working on this in close collaboration with Comcast and other SP/Entreprise operators and academia
    - Any contribution is welcome

Use-Cases

draft-filsfils-rtgwg-segment-routing-use-cases-00

Automated & Guaranteed FRR

- Directed LFA FRR is guaranted in any symmetric topology
  - 2002, LFA FRR project at Cisco
  - draft-bryant-ipfrr-tunnels
- No extra computation (RLFA)
- Simple repair stack
  - node segment to P node
  - adjacency segment from P to Q



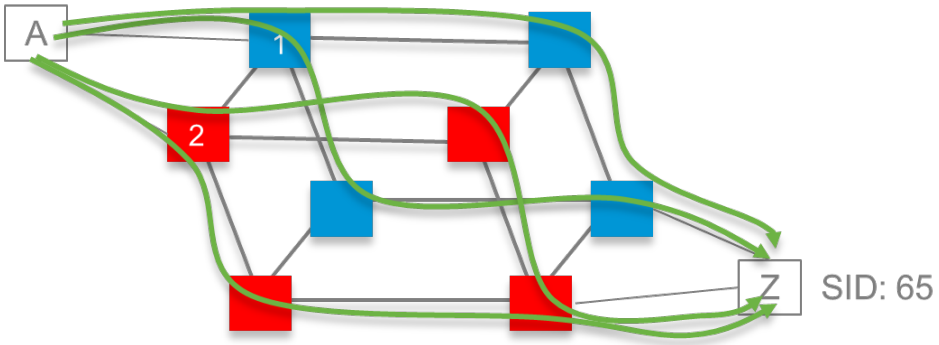
Default metric: 10

Disjointess in Dual-Plane

Anycast SID illustration

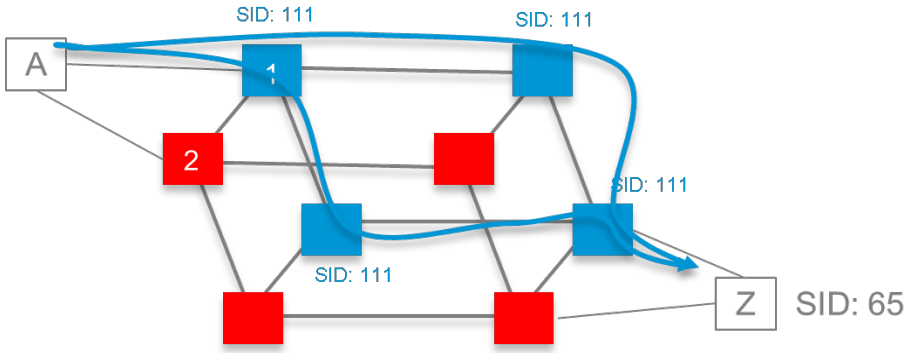
A sends traffic with [65]

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”



CoS-based TE

Anycast SID illustration

- 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

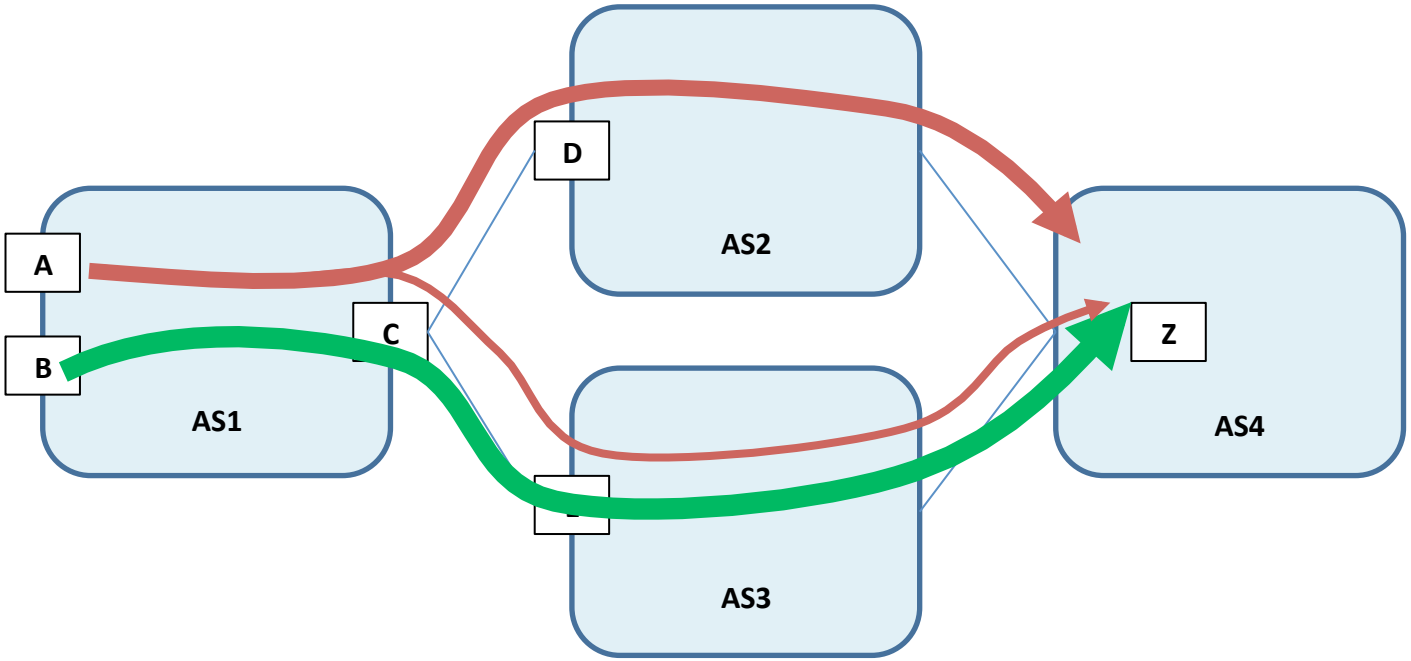
    ➔ ECMP-aware shortest-path to Brussels
  - VoIP and Brussels: push the anycast node to Russia, push Brussels

    ➔ ECMP-aware shortest-path to Russia, followed by
    ECMP-aware shortest-path to Brussels

Node segment to Brussels

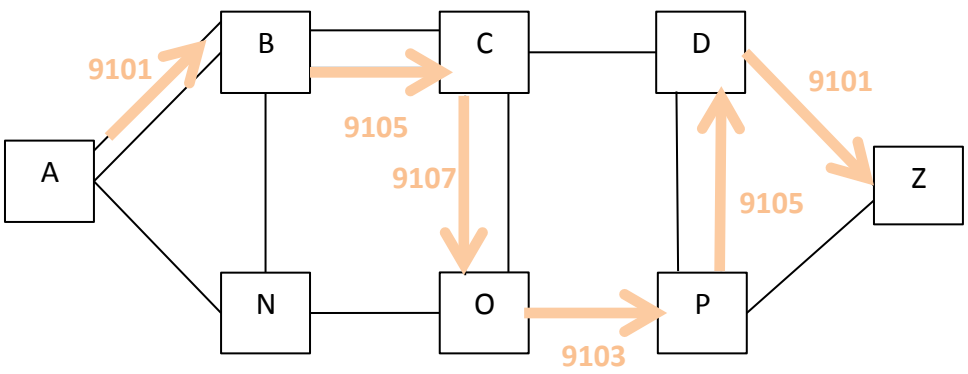
Node segment to Russia

Engineer traffic towards egress peers



- Ingress border routers control how their traffic is balanced between peers
  - Overriding BGP decision at egress border

Full control and OAM



9101
9105
9107
9103
9105
9101

- For Traffic Engineering
- or for OAM

Google

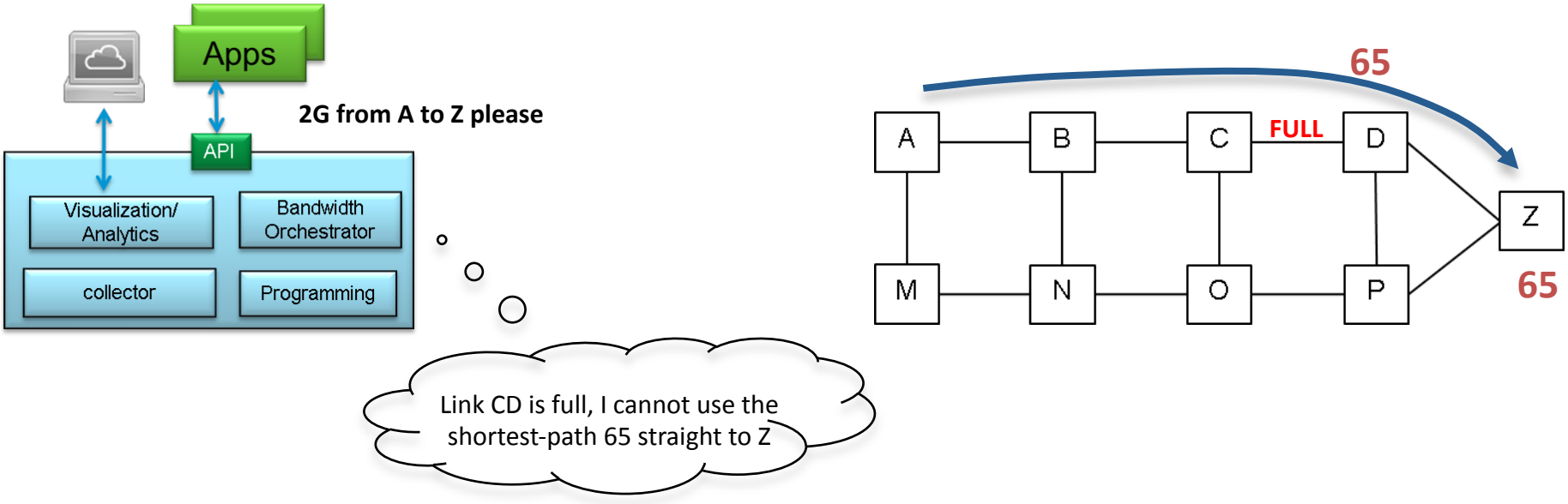
Localizing packet loss

In a large complex network

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

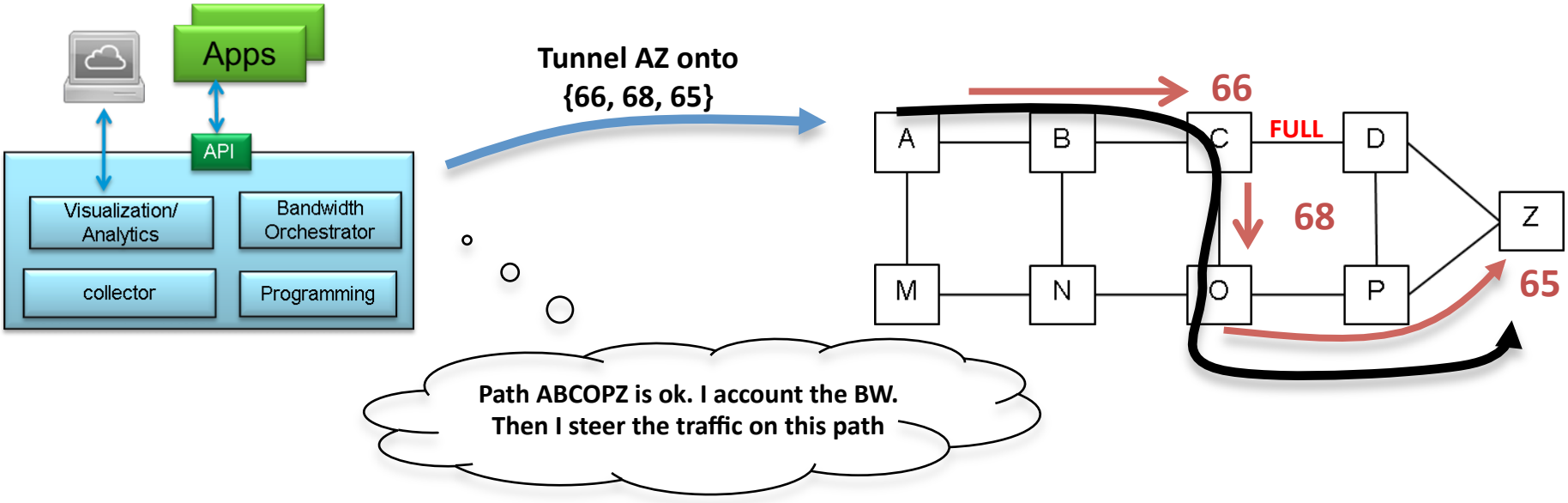
Nanog57, Feb 2013

SDN Orchestration



- The network is simple, highly programmable and responsive to changes instructed by stateful PCE

SDN Orchestration



- The network is simple, highly programmable and responsive to changes instructed by stateful PCE

Conclusion

- Technology is simple
  - Lightweight ISIS/OSPF extensions
  - Immediate applicability to MPLS dataplane
    - IPv4 and IPv6
  - A new type of Routing Extension header for IPv6 pure dataplane
- Numerous use cases
- Significant industry interest
- Multi-vendor/operator constructive collaboration
- Your feedback and contribution is welcome!