STATUS (Stacked Tunnels for Source Routing) BOF

IETF-87, Berlin

July 29, 2013

Note Well

This summary is only meant to point you in the right direction, and doesn't have all the nuances. The IETF's IPR Policy is set forth in BCP 79; please read it carefully.

The brief summary:

- **By participating with the IETF, you agree to follow IETF processes.**
- **If you are aware that a contribution of yours (something you write, say, or discuss in any IETF context) is covered by patents or patent applications, you need to disclose that fact.**
- **You understand that meetings might be recorded, broadcast, and publicly archived.**

For further information, talk to a chair, ask an Area Director, or review the following:

BCP 9 (on the Internet Standards Process)

BCP 25 (on the Working Group processes)

BCP 78 (on the IETF Trust)

BCP 79 (on Intellectual Property Rights in the IETF)

Goal of this BOF, in a nutshell

- Source routing has existed (almost?) as long as networking
  - Historically, we have made little use of it in the IP suite
- Goal of the BOF is to
  - Answer the question, is source routing an idea whose time has come again?
  - If yes, provide input as to how it should be done within the IETF

Goals, in more detail

- What are the use cases driving this work?
- How much data plane variety is needed? MPLS? IPv6? IPv4?
- What are the basic assumptions about the stability of the existing data planes?
- What management/control architecture do we want?
- What does a label represent?

Agenda

- Introduction
- Use cases
- Architectures
- Open Discussion
- Wrap-up

STATUS Use Case:

Performance Engineered LSPs.

Rob Shakir, E2E Network Architect, BT.

STATUS BoF, IETF 87.
July 29, 2013.

Key Requirement.

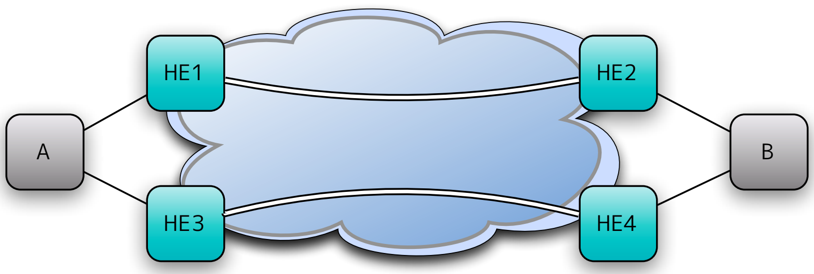
Create LSPs within an IP/MPLS infrastructure which:

- Are routed away from the SPT based on performance constraints (affinity, latency, SRLG etc.).
- Or based on coupling with other LSPs within the network (e.g., for diversity or bidirectionality).
- Provide adequate scale to support per-service or per-flow constraints.
- Are routed according to distributed CSPF or centrally based on service requirements.

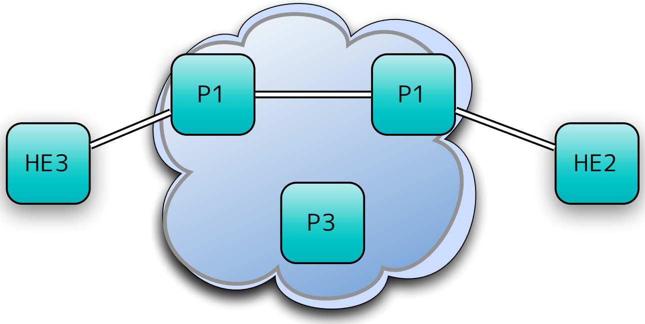
Background – What's the problem we're trying to address?

- In IP/MPLS networks, we have a concept of one "base" topology – which is the SPT.
  - One set of logic applied to choose IGP costs – used to route all services within this topology.

- **Problem for a core network supporting multiple services:** Not all services have the same logic as to the constraints for their routing through the infrastructure.



Co-routing service placement based on consideration of other services within the network.



Pinned paths where services are constrained based on underlying path resources.

- How do we meet the requirement for such constraints?
  - Transport networks have generally provided such constrained paths.
  - More applications requiring performance guarantees.
    - For all traffic (e.g., Broadcast).
    - A subset (e.g., voice within a multi-service VPN).

Problem: Provide means to introduce routing constraints which diverge from the SPT on a per-service or per-flow basis, utilising the existing underlying IP/MPLS network infrastructure.

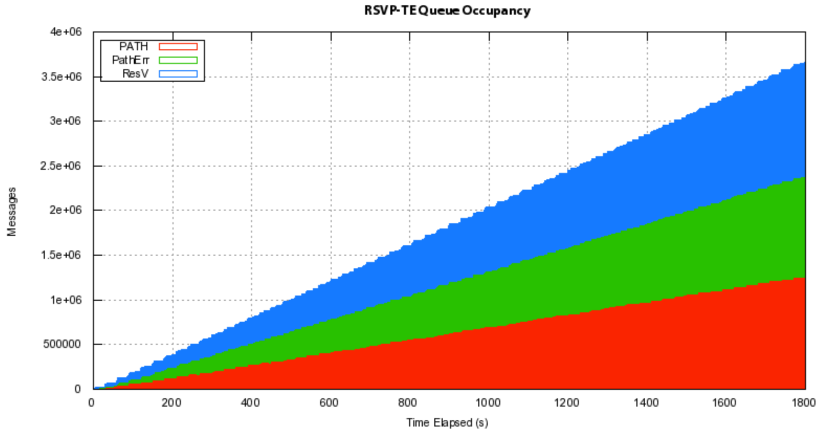
Path Constraints and Technology Options.

- **Requirement for a number of types of constrained service/flow routing:**
  - Co-routing.
    - Considering SRLG/Node/Link diversity or bi-directional paths.
  - Affinity-based routing.
    - Diverging from SPT based on constraining available paths by colour/admin-group.
  - Performance-managed services.
    - Latency, available bandwidth, etc.

- Clearly, a number of these constraints can be delivered by RSVP-TE today.

- **Per-service/flow routing requires a significant increase in the number of RSVP-TE LSPs when compared to current deployments:**
  - Number of LSPs is greater than full mesh (already not recommended).
  - Scale limit of mid-point signalling during large failures.

- **Limited additional functionality is offered by having mid-point state.**
  - Generally only admission control.
  - Required in a subset of path routing scenarios.



Mid-point Overloading – Post-Mortem Model

Unbounded RSVP-TE queue growth based on inability to process PATH messages within LSP retry time – **LSPs never successfully re-signal.**

Suggested Architecture.

- **Use segment routing (STATUS approach) to provide means to instantiate the data-plane paths.**
  - Removes constraint of number of LSPs that can be created by removing mid-point state.
  - Stacked labels indicate the path to be taken through the network.
  - Some consideration of per-label semantics may be required (particularly for reversion).

- **Re-use existing CSPF machinery where it is applicable.**
  - Distributed path calculation based on IGP or TED influences selection (e.g., affinity etc.)
  - Extended IGP attributes can provide increased CSPF-functionality.
    - *draft-previdi-isis-te-metric-extensions-03*
    - *draft-ietf-ospf-te-metric-extensions-04*

- **Re-use existing PCE to provide SID stacks where global visibility is required.**
  - Co-computation of coupled services (bidirectional, diverse with divergent head-ends).
  - Stateful PCE can provide admission control where required.
    - Only maintain state for the subset of services where it is required.

- Use case described in more detail in *draft-shakir-rtgwg-sr-performance-engineered-lsps.*

- Encourage swift progress towards standardised solutions to meet these requirements.
  - Extend IETF technologies to meet an IP/MPLS functionality gap.
  - Existing WGs provide a good forum with experts in each domain.

Thanks!
Questions/Comments?

Segment Routing based OAM use case

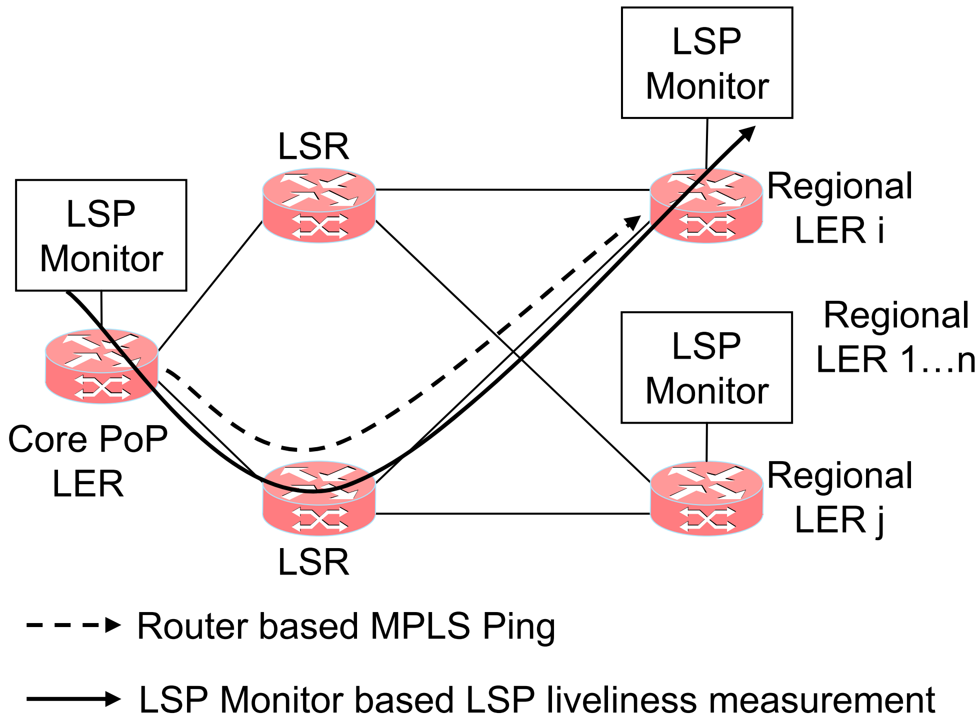
IETF 87, Berlin

Rüdiger Geib, Deutsche Telekom

State of the art MPLS OAM: limited functionality, not scaleable in carrier backbone networks. Despite RFC4379 being a very smart and useful solution.

Monitoring MPLS data plane liveliness

- on the fly by router based MPLS OAM commands. Involves control plane, not useful for permanent LSP monitoring.
- permanent monitoring with LSP data plane traffic only: requires dedicated hardware, e.g. at each LER. Limited to path monitoring between routers with attached LSP monitor. Requires dedicated monitoring HW per PoP, ensuring execution of all LSPs is tricky in the presence of ECMP, requires a monitoring result collection and so on.



Segment Routing based OAM use case

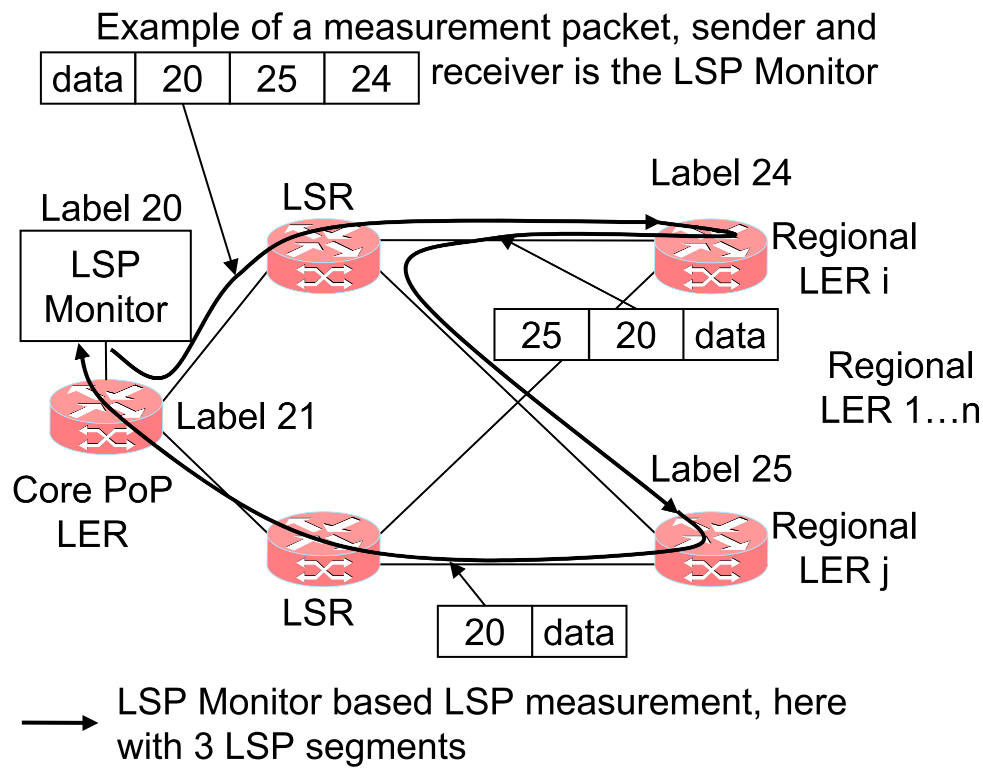
IETF 87, Berlin

Rüdiger Geib, Deutsche Telekom

Segment based Routing allows for scaleable LSP monitoring

Monitoring MPLS data plane liveliness

- source based routing allows execution of arbitrary LSP chains.
- then a ping with data plane loop can be built.
- by ISIS the LSP Monitor is aware of the network topology and its state.
- a single LSP monitor is able to adress all LSPs of a domain. A redundant design is possible if desired.
- Example to the right: the LSP monitor checks data plane liveliness between LER i and LER j. In general, by the method shown all LSPs can be monitored.



Segment Routing based OAM use case

IETF 87, Berlin

Rüdiger Geib, Deutsche Telekom

State and expectation of DT on segment routing based OAM

- Technical Engineering and NOC convinced of the advantages of SR based MPLS backbone monitoring.
- A self developed prototype is connected to our commercial MPLS backbone.
- SR as proposed by the architecture of the draft-filsfils-rtgwg-segment-routing author team should be set on IETF standards track. ASAP.

State of Google on segment routing based OAM

- A pretty much identical approach to that shown on the page before may be found in Google's presentation(s) on „localizing packet loss", e.g. page 15ff of https://ripe65.ripe.net/presentations/828-RIPE65.Talk29.Google_Blackbox_Monitoring.pdf

Segment Routing Use Cases @ DT

Network Complexity, Disjoint Paths, QoS/Service Based Routing.

Reducing Complexity in the Network Architecture

Multi Protocol Label Switching based on RSVP:

- Complex label and path setup protocol (RSVP).
- Physical as well as logical links need to be provisioned, monitored, and maintained.
- Overlay topology.

Multi Protocol Label Switching based on LDP:

- Label distribution protocol (LDP) in addition to IGP.
- LDP needs to be synchronized with IGP.

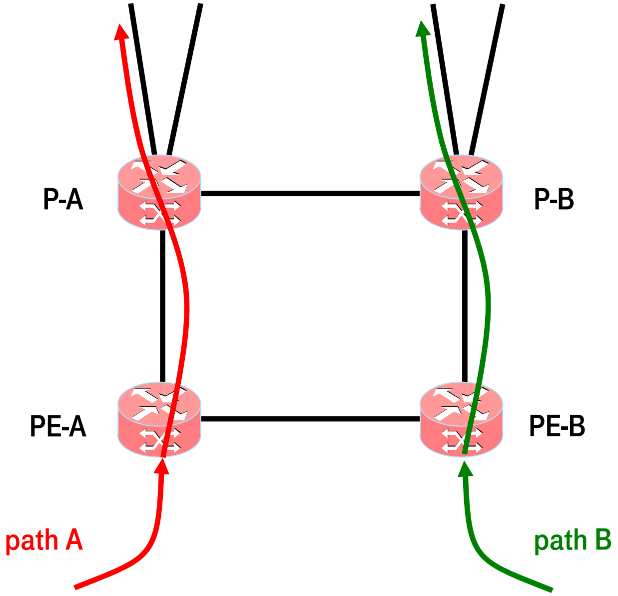
Segment Routing (with MPLS labels):

- Label TLV in IGP.
- Additional complexity only where needed for additional functionality.

Disjoint Paths (1/2)
Traditional Solutions

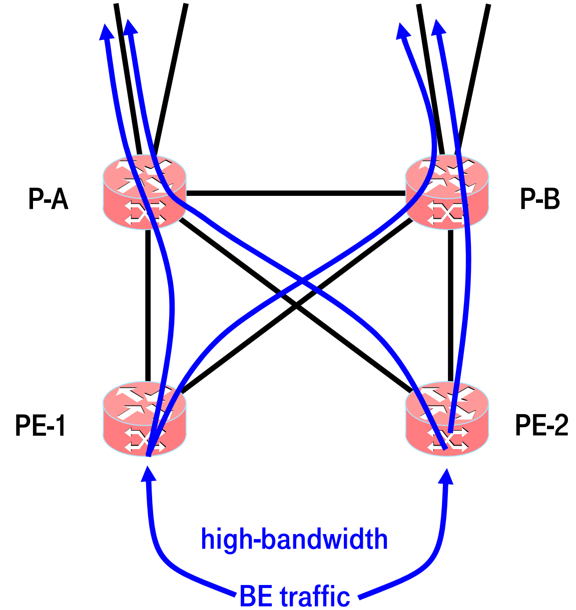
Mobile network:

- Sigtran traffic requires disjoint paths.
- Topology tailored to provide disjoint paths.
- RSVP based MPLS FRR provides fast re-route.



Fixed network:

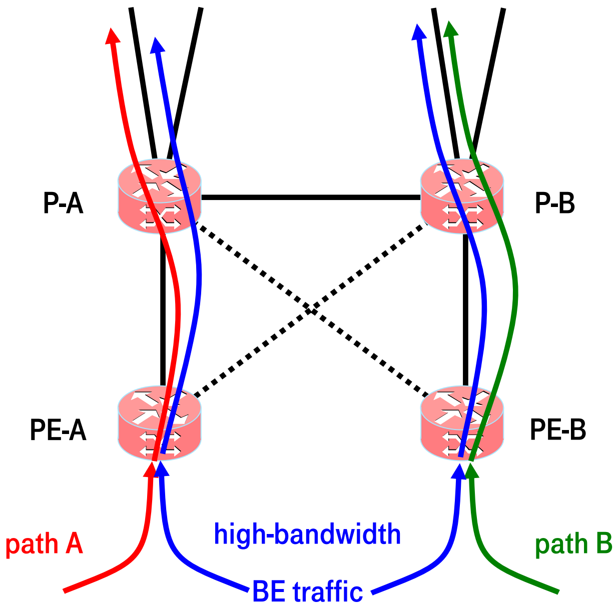
- No traditional requirement for disjoint paths.
- Topology optimized for high bandwith demand and efficiency.
- IP FRR where needed.



Disjoint Paths (2/2)
Current and Future Solution (?)

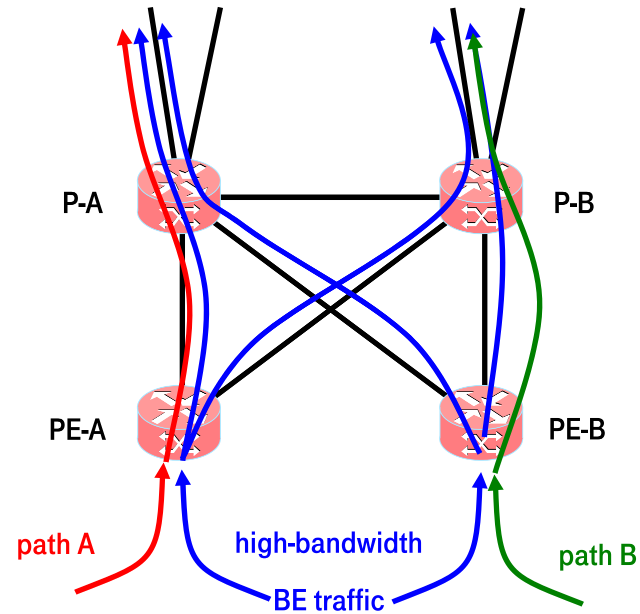
Merged network:

- Topology tailored for both disjoint paths and IP-FRR.
- Limited efficiency.



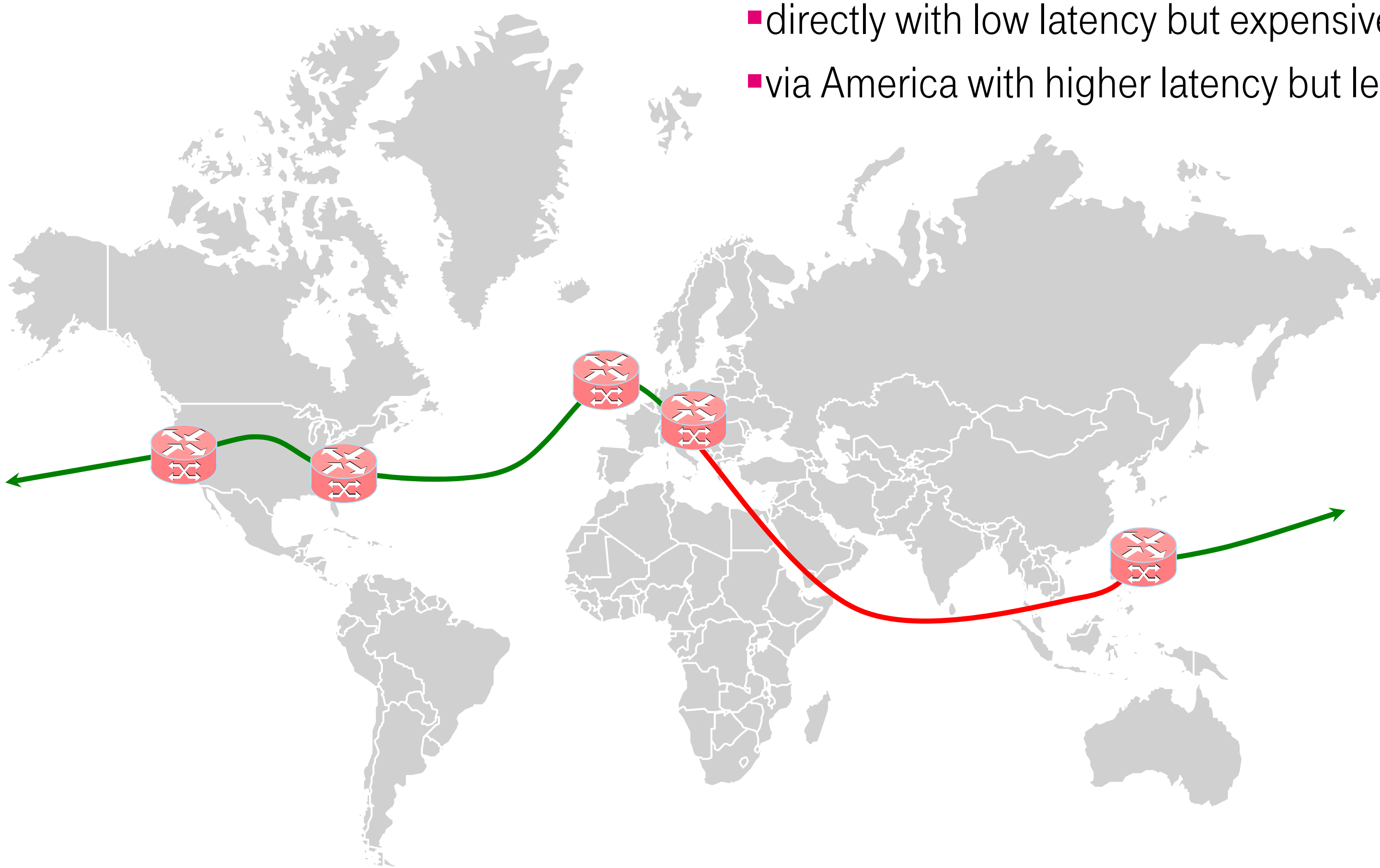
Optimized future network with SR:

- Basic topology optimized for IP-FRR and efficiency.
- Sigtran traffic constrained with A/B anycast segment to provide disjoint paths.



QoS/Service Based Routing (1/2)
Routing of Asia-Europe traffic

- directly with low latency but expensive
- via America with higher latency but less cost



QoS/Service Based Routing (2/2)
Routing of Asia-Europe traffic

Traditional approach:

- Set up full mesh of RSVP tunnels.
- Optimize RSVP for latency.
- Optimize IGP/LDP for capacity (alternatively a second full mesh of RSVP tunnels).
- Route delay-sensitive traffic on RSVP, other on IGP (or second set of tunnels).

➔ Adds all complexity and operational efforts (configuration, monitoring, maintenance) of a full mesh of RSVP tunnels and overlay topology.

With Segment Routing:

- Optimize IGP for capacity and cost-efficiency.
- Set up anycast segment for direct links between Asia and Europe.
- Add special segment to delay-sensitive Asia/Europe traffic only (QoS or service based).

➔ Little extra efforts once segment routing is rolled out.

References

Simplicity, MPLS label in IGP:

- draft-filsfils-rtgwg-segment-routing-use-cases-01, section 2

Disjoint Paths:

- draft-filsfils-rtgwg-segment-routing-use-cases-01, section 4.1.1

QoS Based Routing:

- draft-filsfils-rtgwg-segment-routing-use-cases-01, section 4.1.2

Segment Routing
Fast ReRoute use case

Bruno Decraene Orange
Stéphane Litkowski Orange

Fast ReRoute

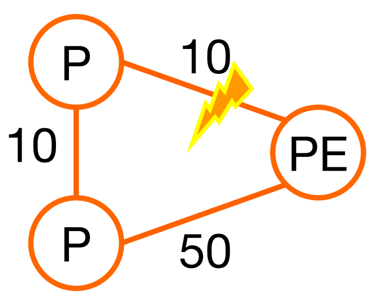
- IGP Convergence performance sometimes not enough for apps
- Fast ReRoute (FRR) required to fill the gap
- We are looking for a Simple, Scalable, Full coverage FRR solution.
- Current State for LDP networks
  - Adding RSVP-TE is sometimes too complex / not scalable
  - We successfully deployed LFA thanks to its simplicity…
  - More is needed to increase the coverage

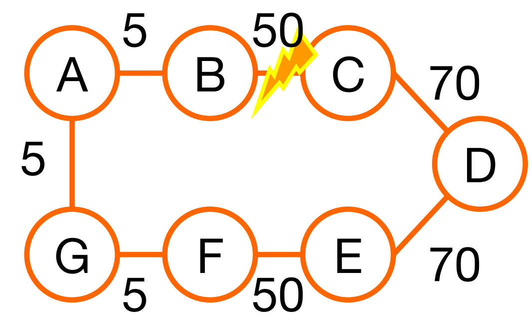
Remote LFA

- draft-ietf-rtgwg-remote-lfa
  - Extends LFA coverage through non-connected protecting nodes.
  - Requires the stacking of tunnels to “*source route*” packets via a specific node
- Completely matches STATUS « Stacked Tunnels for Source Routing »
- Currently requires dynamically established T-LDP sessions
  - Overhead and Inter-op issues
  - Troubleshooting is complicated due to transient session establishment
- IGP-learned labels would be simpler

Directed LFA

- Remote LFA provides 100% coverage in most real live networks, but does not cover some specific topologies:





- Directed LFA (DLFA) is proven to give 100% coverage
  - for link protection provided links have symmetric costs
  - draft-francois-sr-frr
- DLFA requires the advertisement of a label for each IGP adjacency.
  - to enforce a directed forwarding over the adjacency.

FRR Egress node protection / Service protection

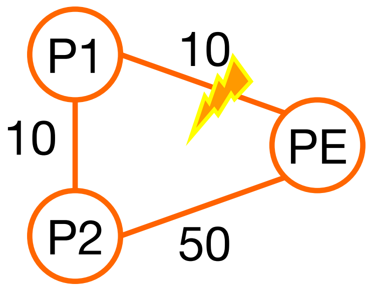
- Egress node protection addresses the failure of the egress node
  - draft-ietf-mpls-seamless-mpls-01#section-5.1.8.4
  - draft-minto-2547-egress-node-fast-protection
  - draft-filsfils-rtgwg-segment-routing-use-cases-01#section-3.3
- The protector node advertises its ability to mirror/protect the BGP routes advertised by the primary ASBR
- Requires a context label for de-multiplexing at the protector node
  - to distinguish the specific failure.

Protocol extensions for FRR purposes

- Simplifies RLFA (no T-LDP sessions)
- Required for DLFA and Egress node protection
- Having a single generic tool for all applications is beneficial
  - Avoids N specific extensions for N features
  - Allows other features to benefit from it
    - Niche use cases tend to not justify extensions
    - All use cases benefit from easier interoperability
    - Incremental deployments are more likely feasible
- Segment Routing covers all those FRR uses cases
  - draft-previdi-isis-segment-routing-extensions
  - draft-psenak-ospf-segment-routing-extensions
  - draft-filsfils-rtgwg-segment-routing-use-cases

Incremental deployment in a LDP network

- As first step, Segment Routing use may be restricted to FRR backup path.
  - Keeping LDP for nominal traffic, like the way it currently is.
- If nodes are already SR capable, SR FRR can be deployed incrementally on a per PLR basis. (with incremental benefit).
  - i.e. enabling SR FRR on P1



- In the absence of SR capable node in the network, SR FRR can be deployed incrementally on a per PLR + (last) P + (first) Q basis.
  - i.e. enabling SR FRR on P1 (PLR) & SR on P2 (P) and PE (Q)
  - Note that on the Q, SR may be replaced by a T-LDP session (which is natively the case in the above example)
- More details in draft-filsfils-rtgwg-segment-routing-use-cases-01#section-6.4

Thank you

STATUS BoF (working title: Segment Routing)

Use Cases: Converged Multi-network Operation

Prepared by: Victor Kuarsingh
Network Strategy and Technology Development

July 29, 2013 (Ver 03)

ROGERS™

Problem Statement

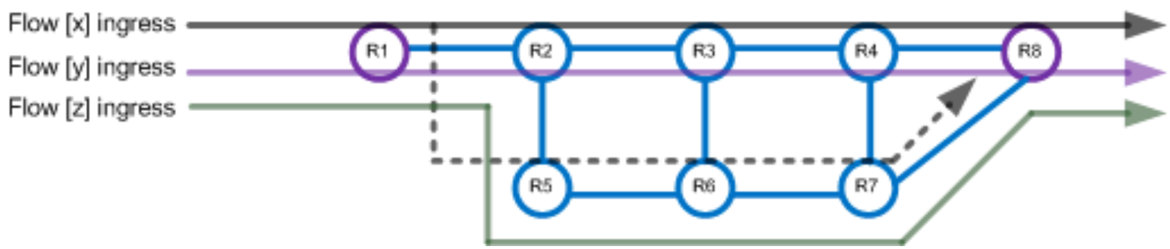
- Multiple networks and service domains
- Multiple services on each network, services span networks
- Networks built for primary function(s) (size, redundancy, etc)
- Services require differing capabilities from the network (path, redundancy, bandwidth, QoS)
- Significant amount of traffic only requires ECMP, lowest cost path routing
- Some traffic flows require explicit path
- Current options require many protocols, significant work, states in network and complexity (operationally challenging)

Requirements for Solution

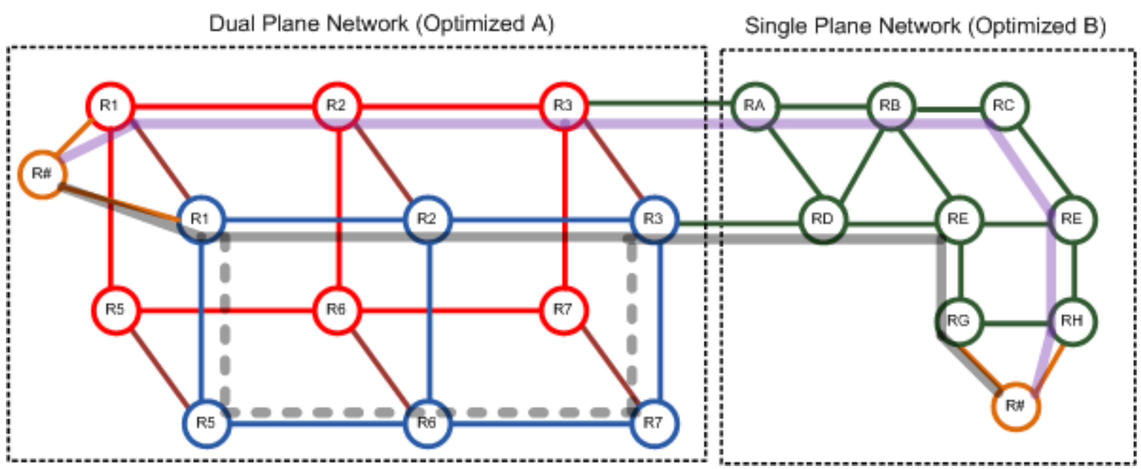
- Single solution and protocol suite across different networks able to support different services
- Reduced/contained network state
- Simple and Programmable - SDN compliant
- Support for explicit path, dual plane, restricted path and FRR
- Controlled/Limited traffic engineering logic
- Capable to stack services, allowing them to use network differently based on service requirements

Examples: Stack Services, Domains and Path

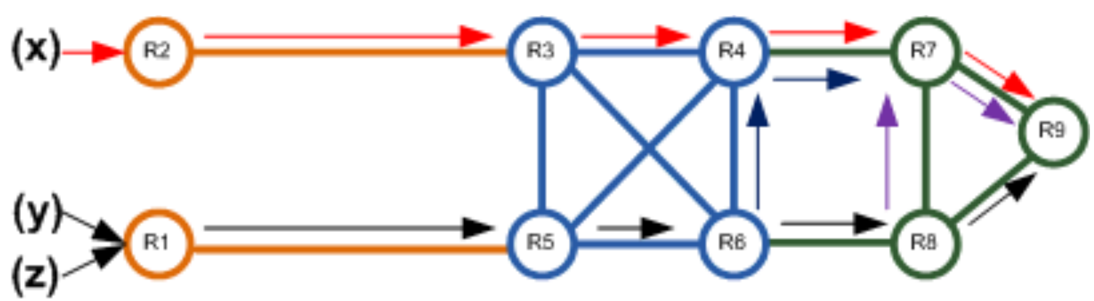
- Services stacked onto network, each requiring different forwarding behaviors



- Services across multiple domains
- Traffic engineering in dissimilar networks

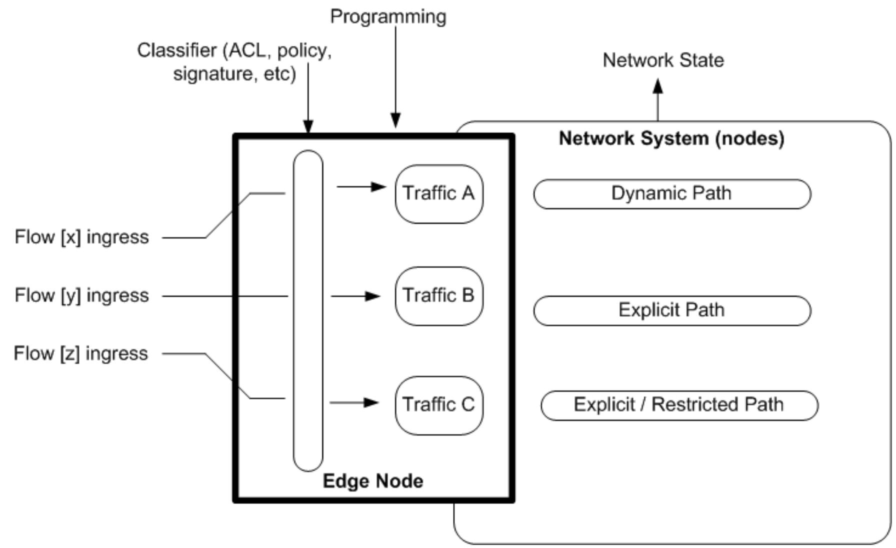


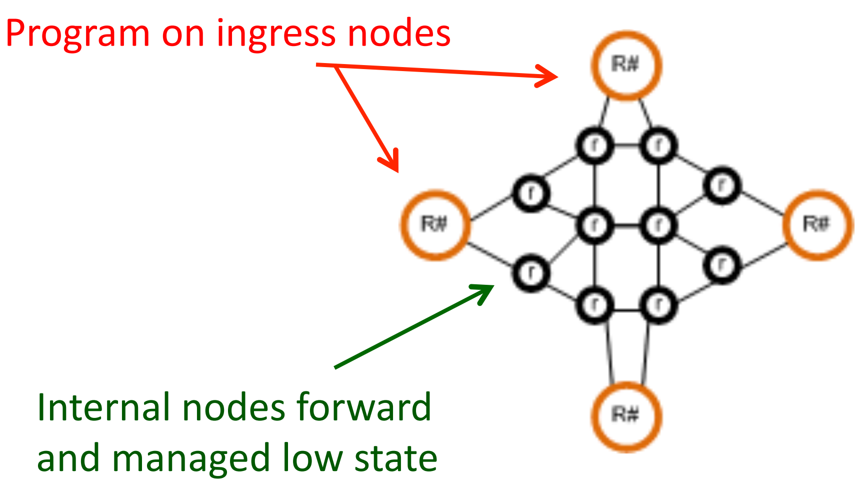
- Traffic engineering for best (lowest cost) path, explicit path, and restricted path
- Different per service



Programming and SDN Interaction

- Automation of the network is essential for future operation
  - Current operational modes not scalable
- SDN (path programming within this document's context) is desired, with per-flow/service network treatment
- Minimize the number of elements where programming must occur, and simplify configuration required





Questions?

Stacked Tunnels, OSPF/ISIS as Label Distribution Protocols, and MPLS Architecture

RFC 3031 ("Multiprotocol Label Switching Architecture") provides all the concepts and terminology needed to describe (a) source routed tunnels, including the stacked-label variation, and (b) OSPF/ISIS as a label distribution protocol.

This presentation places stacked-label source routed tunneling, and OSPF/ISIS as a label distribution protocol in the context of the MPLS Architecture, using the terminology defined there.

MPLS Architecture and Explicit Routing

- *Explicitly Routed Tunnels*
  - "If a Tunneled Packet travels from Ru to Rd over a path other than the Hop-by-hop path, we say that it is in an "Explicitly Routed Tunnel"" (section 3.27.2 of RFC3031)
- Explicitly routed LSPs as a way to realize explicitly routed tunnels
  - "An "Explicitly Routed LSP Tunnel" is a LSP Tunnel that is also an Explicitly Routed LSP" (section 3.27.3 of RFC3031)
    - RSVP-TE (or CR-LDP) as the label distribution protocol for explicitly routed LSPs

MPLS Architecture and LSP Hierarchy

- *LSP Hierarchy*: LSP Tunnels within LSPs (see section 3.27.4 of RFC3031)
- “The label stack mechanism allows LSP tunneling to nest to any depth” (section 3.27.4 of RFC3031)

Stacked LSPs for Explicitly Routed Tunnels

- Instead of an explicitly routed LSP, one can use LSP hierarchy (stack of LSPs) to realize explicitly routed tunnels
- All LSPs in the stack have the same "LSP ingress"
- "LSP egress" of a given LSP in the stack is an intermediate point of the next LSP in the stack
- LSP egress of a given LSP in the stack could be either single or multi-hop away from the LSP egress of the next LSP in the stack
- Such stack of LSPs provides the functionality to forward a packet through a sequence of "LSP egress" of the LSPs on the stack
  - the sequence of "LSP egress" represents the explicit route
- The label stack mechanism allows stack of LSPs to nest to any depth

Incoming and Outgoing label in Label Swap

- From RFC3031:
  - label swap: the basic forwarding operation consisting of looking up an incoming label to determine the outgoing label, encapsulation, port, and other data handling information
- It is legal in the MPLS Architecture to preserve the label as a packet transits an LSR
- This means the incoming and outgoing label values just happen to be the same, and so the swap produces the same output as its input
- If for a given LSP all the LSRs traversed by the LSP have the same incoming and outgoing label, then the value of the label stays the same from the LSP ingress all the way to the LSP egress

MPLS Architecture and Label Distribution Protocol (1)

- "*THE ARCHITECTURE DOES NOT ASSUME THAT THERE IS ONLY A SINGLE LABEL DISTRIBUTION PROTOCOL*. In fact, a number of different label distribution protocols are being standardized." (section 3.6 of RFC3031)

MPLS Architecture and Label Distribution Protocols (2)

- *Local/remote label distribution peers*
  - When two LSRs are IGP neighbors, we will refer to them as "local label distribution peers". When two LSRs may be label distribution peers, but are not IGP neighbors, we will refer to them as "remote label distribution peers" (section 3.27.5 of RFC3031)

OSPF/ISIS as a Label Distribution Protocol (1)

- OSPF/ISIS advertisements by a router carry label bindings for LSPs that transit through the router
- The router can be either single or multi-hop away from routers that receive these bindings
  - Local label distribution peers are the IGP neighbors of the router
  - Remote label distribution peers are other routers in the same IGP domain
- The router could be either single or multi-hop away from the egress of these LSPs
  - Existing label distribution protocols (LDP, RSVP-TE, etc.) can be used to establish multi-hop LSP fragments
  - See also slide 5

OSPF/ISIS as a Label Distribution Protocol (2)

- Each OSPF/ISIS router passes Link State Advertisements originated by other routers unmodified
  - Including the label binding information
  - Similar to how BGP Route Reflectors handle routes/labels
- Provides scalable support for remote label distribution peering
  - Label distribution protocol messages (Link State Advertisements) are exchanged only between IGP neighbors
    - Control plane peering only between IGP neighbors
  - No control plane peering between a router and each of its remote label distribution peers

Example

Assume that R1 wants to forward certain set of packets along the explicitly routed tunnel (R1, R2, R3, R4, R5)

- (R1, R2), (R2, R3), (R3, R4), and (R4, R5) are local label distribution peers (as they are IGP neighbors); all other pairs are remote label distribution peers

Explicitly routed tunnel (R1, R2, R3, R4, R5) is realized via the following LSP Stack:

- LSP1: (R1, R4, R5)
- LSP2: (R1, R3, R4)
- LSP3: (R1, R2, R3) - top of the stack

To get from the first hop of LSP1 to the second hop of LSP1, the packet has to go through an LSP tunnel, LSP2.To get from the first hop of LSP2 to the second hop of LSP2, a packet has to go through an LSP tunnel, LSP3.

R1 receives (via OSPF/ISIS) from its local label distribution peer R2 label bindings originated by R4, R3, and R2.

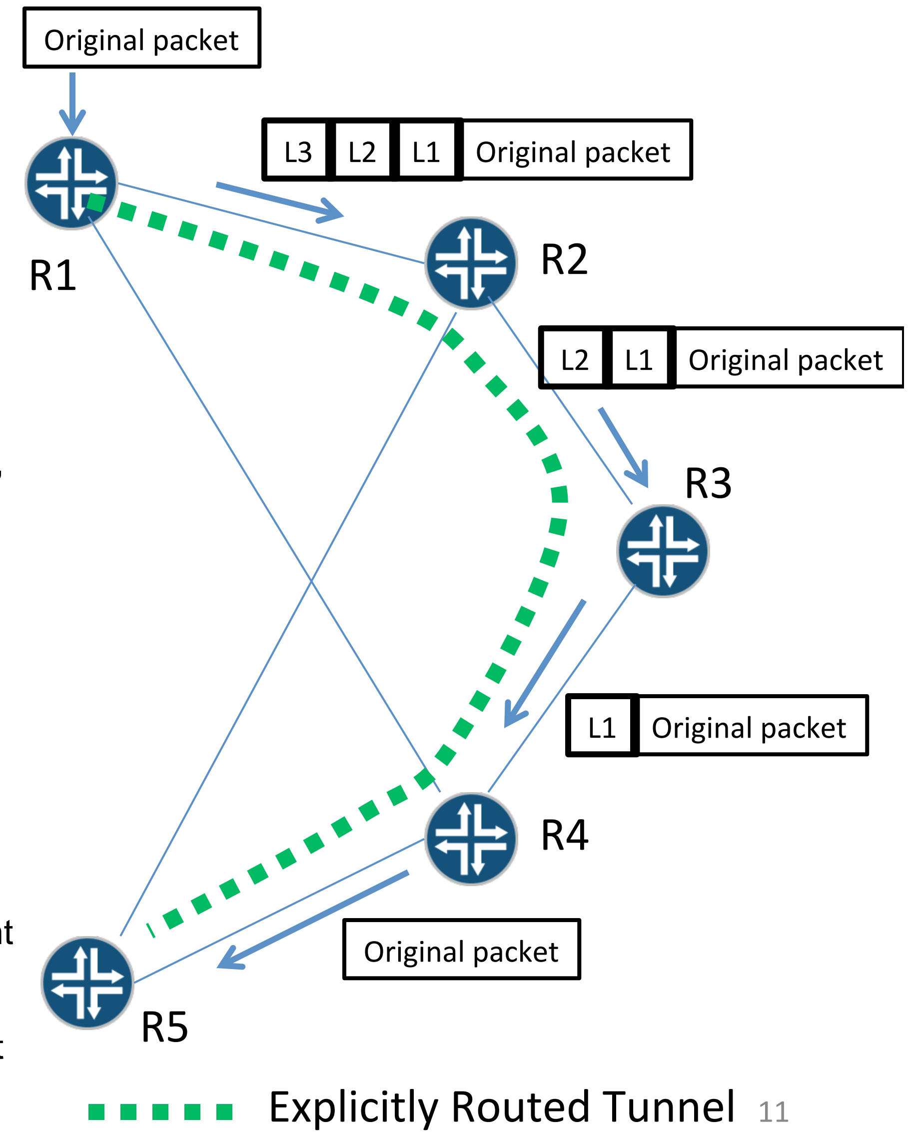
Label Stack construction at R1:

<u>Step 1</u>: R1 uses label binding L1 originated by R4 for LSP1 (R1, R4, R5). So, R1 starts building the label stack by pushing L1 onto the label stack.

<u>Step 2</u>: R1 uses label binding L2 originated by R3 for LSP2 (R1, R3, R4). So, R1 pushes L2 into the stack. At this point the stack contains (L2, L1).

<u>Step 3</u>: R1 uses label binding L3 originated by R2 for LSP3 (R1, R2, R3). So, R1 pushes L3 into the stack. At this point the stack contains (L3, L2, L1).

<u>Step 4</u>: Since R1 and R2 are local label distribution peers, label stack construction is completed



Conclusion

- Stacked LSPs for explicit/source routed tunnels and use of ISIS/OSPF as a label distribution protocol, both fit comfortably within the MPLS architecture.
- As the MPLS architecture defines the necessary concepts and terminology, no new architecture, concepts or terms are required.

Advertising MPLS LSPs in the IGP

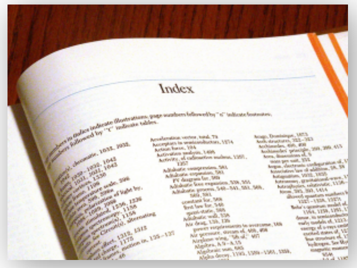
`draft-gredler-ospf-label-advertisement`

`draft-gredler-isis-label-advertisement`

Hannes Gredler
hannes@juniper.net
IETF87, Berlin

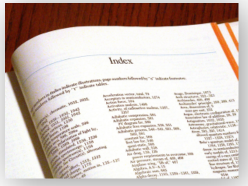
Two Types of Advertisement

Label Range,
Index to Label Range
(=sink-tree paths)



Multi-Hop Segments
(= explicit paths)

Automatic Meshing



Desire for ***automatic meshing*** of transport paths (= LDP alike sink trees)

Transport path tree is computed according to an ***algorithm –*** e.g.

- SPT(Metric)
- SPT(Delay)
- MRT(Metric)

Problem ***may*** be solved using domain-wide labels

… Really ?

History tends to repeat itself … On Domain-wide-labels Quote from IETF60 (San Diego, August 2004) minutes …

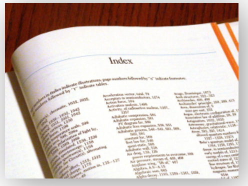
Source Routed MPLS LSP using Domain Wide Label

draft-tian-mpls-lsp-source-route-01, Albert Tian

This draft introduces global labels as a means of source routing or loose source routing a packet in an MPLS network. If each destination of interest (say all of the loopback addresses used as router-IDs) had a label which was both globally known and globally used by all routers, then one could source route a packet by stacking up labels for each of the routers in the source route. This technique could be applied to fast reroute.

In the ensuing discussion it was pointed out that the idea of global labels had been discussed in the early days of MPLS*. At that time it was decided that labels would only have* ***local significance*** *(within the forwarding plane). Many workgroup members* ***expressed opposition to such a fundamental change*** *to the MPLS architecture. Alex Zinin stated, "****The workgroup is done with architecture. Unless it can be proved that architecture is not sufficient then this doesn't fit in charter****."*

Index to Label Range



No requirement for domain-wide label

Problem has been solved before:

RFC4761, Kompella & Rekhter, VPLS BGP Signaling

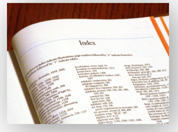
```
3.2.1.  Label Blocks

   To accomplish this, we introduce the notion of "label blocks".  A
   label block, defined by a label base LB and a VE block size VBS, is a
   contiguous set of labels {LB, LB+1, ..., LB+VBS-1}.  Here's how label
   blocks work.  All PEs within a given VPLS are assigned unique VE IDs
   as part of their configuration.  A PE X wishing to send a VPLS update
   sends the same label block information to all other PEs.  Each
   receiving PE infers the label intended for PE X by adding its
   (unique) VE ID to the label base.  In this manner, each receiving PE
   gets a unique demultiplexor for PE X for that VPLS.
```

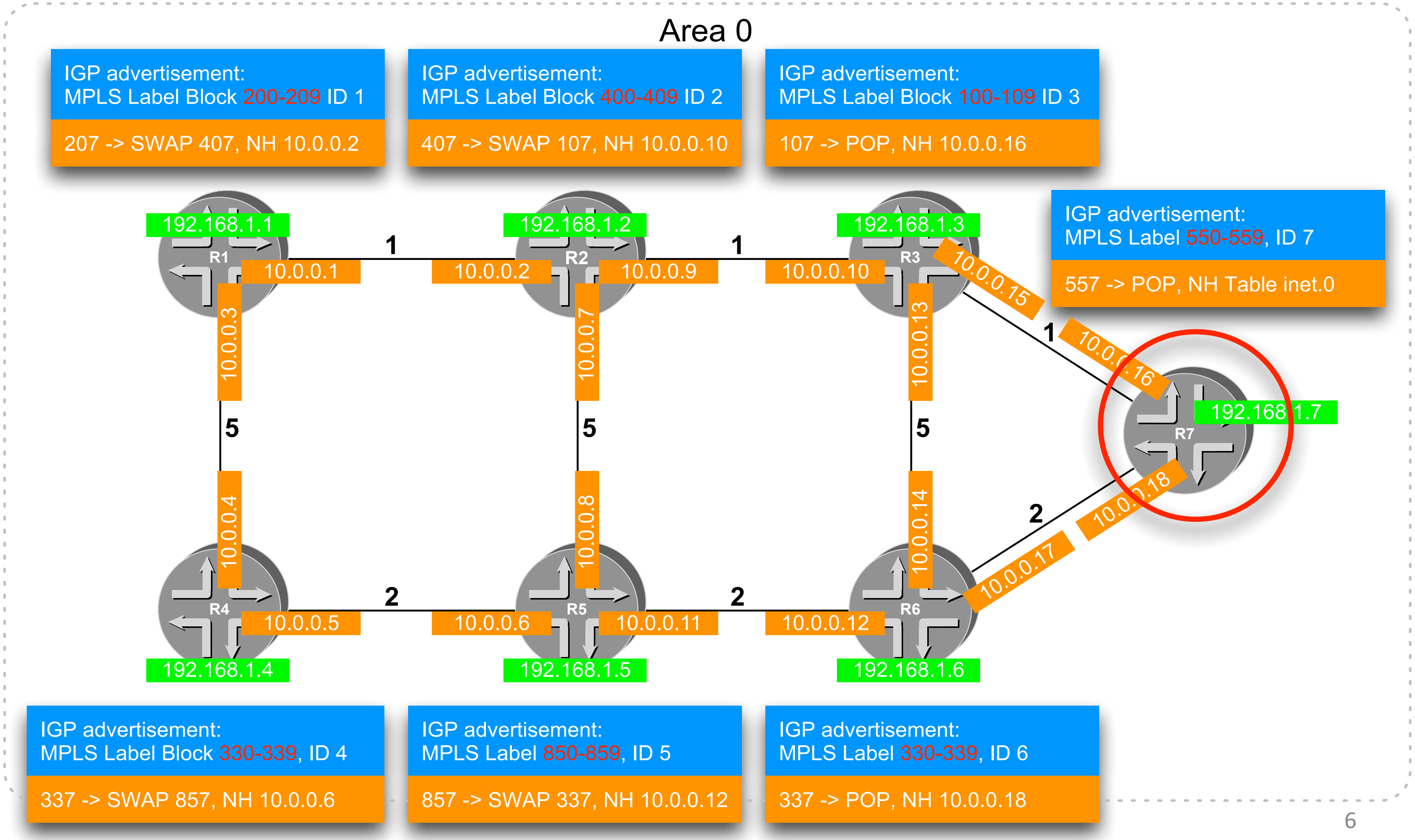
Mirror Label Block ***source semantics*** from VPLS into ***destination semantics***

Assign each router a domain-wide unique ID

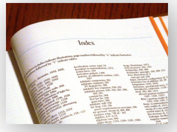
- ID is an index to locate the actual label value, inside label block
- Each LSR allocates and advertises a block of ***locally significant*** labels
- The block should be large enough to accommodate the range of assigned IDs



Example:
LFIB construction for R7, Algorithm: SPT(Metric)



Index to Label RANGE



Index advertises { egress router context, Algorithm } tuples
- Currently support for two contexts
  - IPv4 “Internet” forwarding table
  - IPv6 “Internet” forwarding table

Extensible for ***additional*** algorithms
- (SPT-delay, SPT-TE, MRT, ARC …)

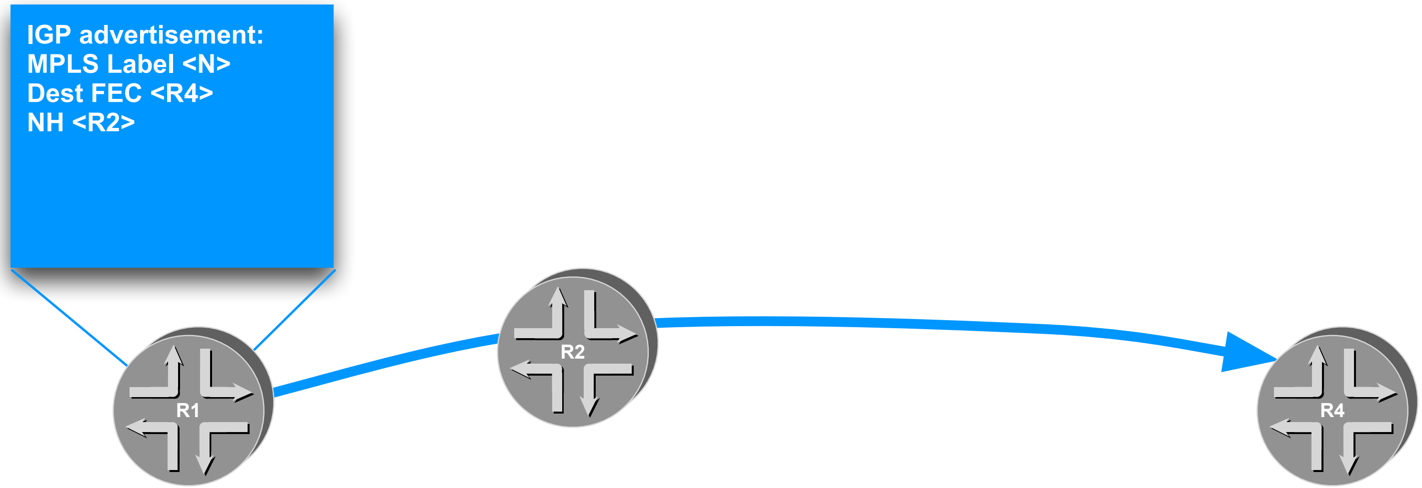
What if I want to have ***identical*** labels in my domain ?
- Most devices have ***configurable*** label-ranges per protocol
- Routers who do not have support for this still can ***interoperate***
  - Protocol operations ***do not assume*** identical label-range for downstream routers

Protocol semantics of MPLS LSPs (1)

RFC 3031

LSP in its most generic form:

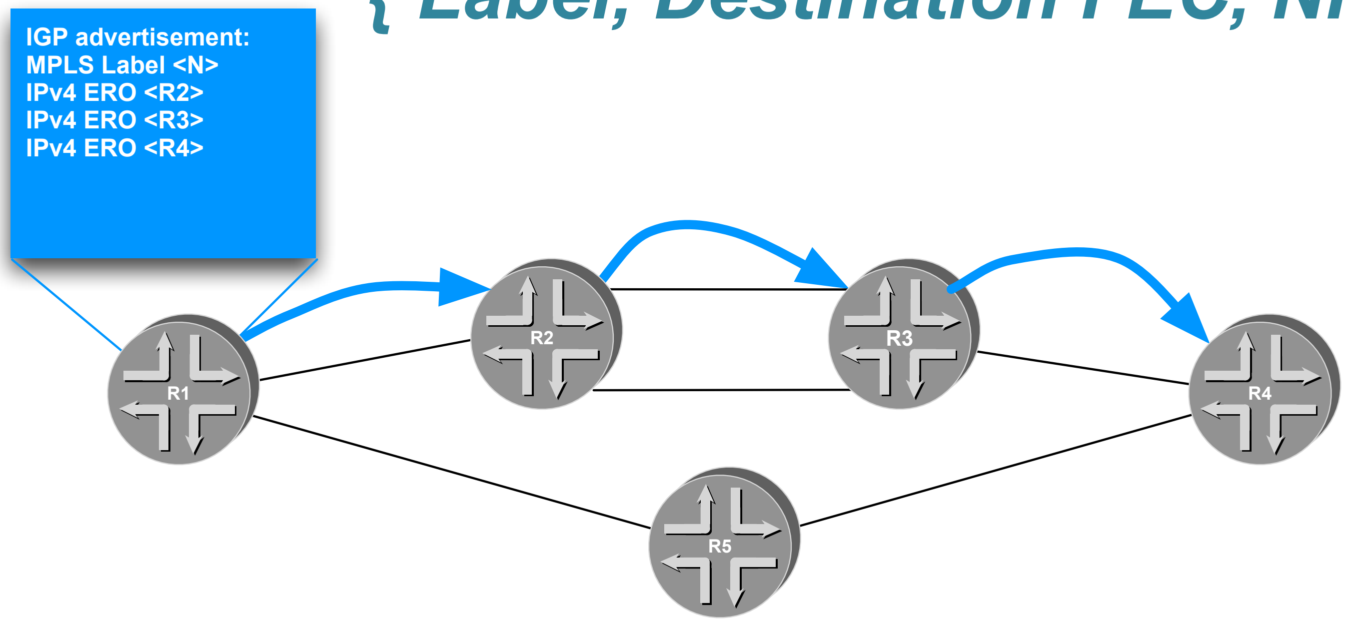
{ Label, Destination FEC, NH }



Protocol semantics of MPLS LSPs (2)

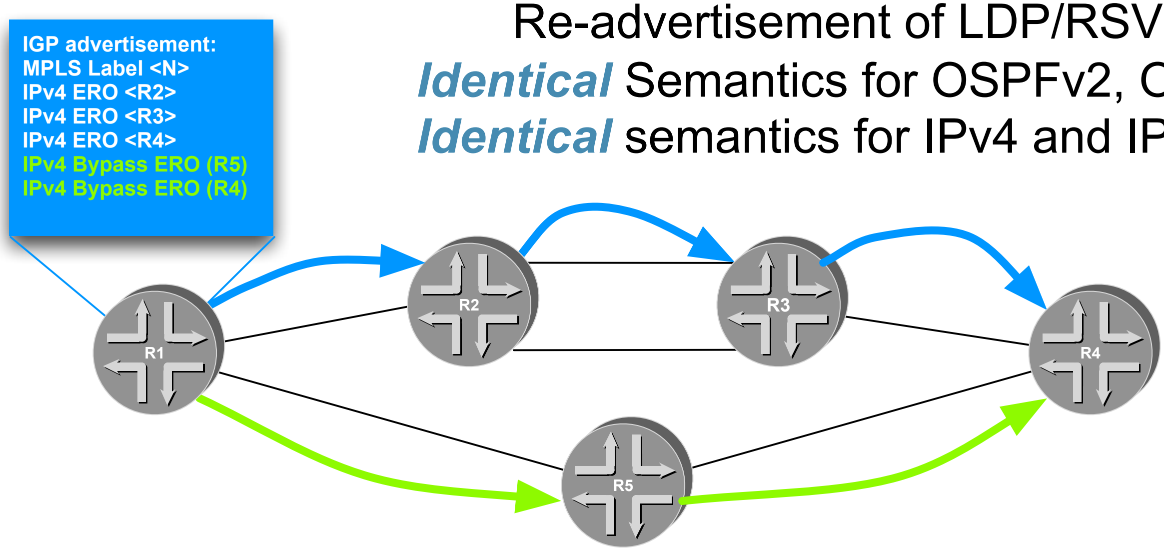
LSP supporting list of NHs = ERO

{ Label, Destination FEC, NH[ ] }



Protocol semantics of MPLS LSPs (3)

Segments and Labels are ***Node*** properties
Nexthop notion gives the Segment/Label a direction
Support for ***more than one*** Nexthop = Path ERO
FEC/Prefix describes what is at the tail-end of the path
(= Support for stitching to LDP/RSVP/LBGP)
Support for ***more than one*** path (Primary, Bypass) for a given destination
Identical semantics for
Re-advertisement of existing IGP label stack
Re-advertisement of LDP/RSVP/LBGP label
Identical Semantics for OSPFv2, OSPFv3 and IS-IS
Identical semantics for IPv4 and IPv6 FECs



Merge draft-gredler-isis-label-advertisement-03 into draft-previdi-isis-segment-routing-extensions-01

[Docs] [txt|pdf] [Tracker] [Email] [Nits]

Versions: 00 01

```
IS-IS for IP Internets                                  S. Previdi, Ed.
Internet-Draft                                              C. Filsfils
Intended status: Standards Track                            A. Bashandy
Expires: December 30, 2013                           Cisco Systems, Inc.
                                                            B. Decraene
                                                           S. Litkowski
                                                                 Orange
                                                                R. Geib
                                                       Deutsche Telekom
                                                           I. Milojevic
                                                          Telekom Srbija
                                                               R. Shakir
                                                        British Telecom
                                                                   S. Y
                                                                    TDC
                                                             W. Henderi
                                                            Alcatel-Luc
                                                               J. Tants
                                                               Ericsson
                                                          June 28, 2013
```

IS-IS Extensions for Segment Routing
draft-previdi-isis-segment-routing-extensions-00

[Docs] [txt|pdf|xml|html] [Tracker] [Email] [Diff1] [Diff2] [Nits]

Versions: 00 01 02 03

```
IS-IS for IP Internets                                  H. Gredler, Ed.
Internet-Draft                                    Juniper Networks, Inc.
Intended status: Standards Track                              S. Amante
Expires: November 22, 2013              Level 3 Communications, Inc.
                                                               T. Scholl
                                                                  Amazon
                                                                L. Jalil
                                                                 Verizon
                                                            May 21, 2013
```

Advertising MPLS labels in IS-IS
draft-gredler-isis-label-advertisement-03

[Docs] [txt|pdf|xml|html] [Tracker] [Email] [Diff1] [Diff2] [Nits]

Versions: 00 01

```
IS-IS for IP Internets                                  S. Previdi, Ed.
Internet-Draft                                              C. Filsfils
Intended status: Standards Track                            A. Bashandy
Expires: January 2, 2014                             Cisco Systems, Inc.
                                                              H. Gredler
                                                  Juniper Networks, Inc.
                                                            B. Decraene
                                                           S. Litkowski
                                                                 Orange
                                                                R. Geib
                                                       Deutsche Telekom
                                                           I. Milojevic
                                                          Telekom Srbija
                                                               R. Shakir
                                                        British Telecom
                                                                 S. Ytti
                                                                  TDC Oy
                                                           W. Henderickx
                                                          Alcatel-Lucent
                                                             J. Tantsura
                                                                Ericsson
                                                           July 1, 2013
```

IS-IS Extensions for Segment Routing
draft-previdi-isis-segment-routing-extensions-01

Merge draft-gredler-ospf-label-advertisement-03 into draft-psenak-ospf-segment-routing-extensions-01

```
[Docs] [txt|pdf] [Tracker] [Email] [Nits]

Versions: 00 01

Network Working Group                                    P. Psenak, Ed.
Internet-Draft                                           S. Previdi, Ed.
Intended status: Standards Track                             C. Filsfils
Expires: December 30, 2013                           Cisco Systems, Inc.
                                                           June 28, 2013


                    OSPF Extensions for Segment Routing
             draft-psenak-ospf-segment-routing-extensions-00
```

```
[Docs] [txt|pdf|xml|html] [Tracker] [Email] [Diff1] [Diff2] [Nits]

Versions: 00 01

Open Shortest Path First IGP                             P. Psenak, Ed.
Internet-Draft                                           S. Previdi, Ed.
Intended status: Standards Track                             C. Filsfils
Expires: January 2, 2014                             Cisco Systems, Inc.
                                                              H. Gredler
                                                  Juniper Networks, Inc.
                                                               R. Shakir
                                                         British Telecom
                                                            July 1, 2013


                    OSPF Extensions for Segment Routing
             draft-psenak-ospf-segment-routing-extensions-01
```

```
[Docs] [txt|pdf|xml|html] [Tracker] [Email] [Diff1] [Diff2] [Nits]

Versions: 00 01 02 03

Open Shortest Path First IGP                             H. Gredler, Ed.
Internet-Draft                                    Juniper Networks, Inc.
Intended status: Standards Track                               S. Amante
Expires: November 22, 2013                Level 3 Communications, Inc.
                                                               T. Scholl
                                                                  Amazon
                                                                L. Jalil
                                                                 Verizon
                                                            May 21, 2013


                     Advertising MPLS labels in OSPF
                 draft-gredler-ospf-label-advertisement-03
```

Draft-Merge Summary

- Common
  - Protocol Semantics
- Differences
  - draft-gredler-* builds on top of ***RFC3031*** and ***does not require*** a new architecture.
  - Draft-gredler-* proposes to re-use only ***existing data plane*** for source-routing (=MPLS)

ISSUE: SR-V6 dataplane is a *significant* change of IPv6 dataplane

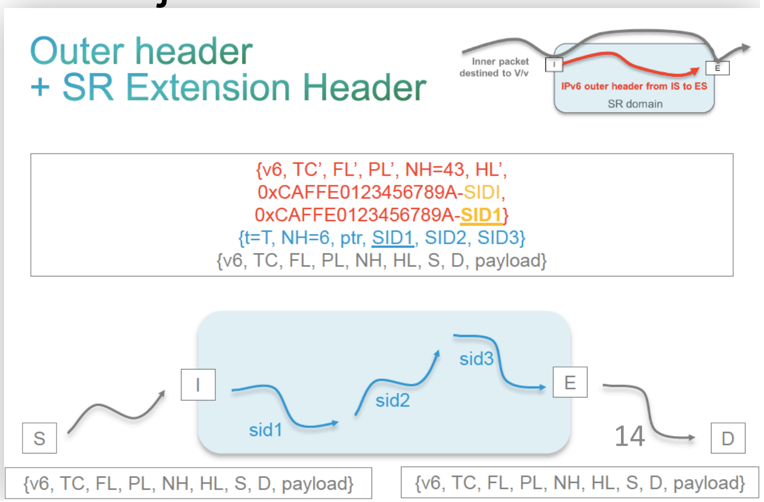
Proposal to create a ***source routing extension header***
- Issue: requires ***new Hardware*** to overcome first N-bytes lookup buffer limits, of deployed hardware

Data planes in routers (IPv4, IPv6, MPLS)
- Why need for data plane #4 ?

Use cases ?

MPLS ***is*** the ***'routing header'*** of IPv4 and IPv6 today

SR-v6 dataplane is ***more complex*** than MPLS Label operation
- Requires IPv6 address rewrite capabilities to adjust SID segment pointer



Segment Routing

IETF 87

Clarence Filsfils – cf@cisco.com

C. Filsfils, S. Previdi, A. Bashandy, B. Decraene, S. Litkowski, M. Horneffer,
I. Milojevic, R. Shakir, S. Ytti, W. Henderickx, J. Tantsura, Ericsson, E. Crabbe,
H. Gredler, and a few other contributors…

Technology

- Generality of a segment
  - Intra and inter domain
  - Forwarding construct
  - Service construct
  - virtualization
  - Abstract Routing Model (draft-filsfils-rtgwg-segment-routing), see nanog video
  - SR is not a label-in-IGP solution. Label-in-IGP is a subset of SR !
- Agnostic Control Plane
- Instantiation in two dataplanes
  - MPLS
  - IPv6

Productization

- Wide and rapid industry adoption
- Committed deployments received from operators within 5 months of first review
  - MPLS
  - IPv6

The last 9 months

- Oct: first SR presentation to operators
  - Lead Operator group formed, see co-authors, weekly meeting since then
  - Commitment to velocity, transparency and multi-vendor agreement
- Feb: initial implementation released as per commitment
- Mar: MPLS WC and IPv6 conference
- Mar: first draft submitted to IETF-86
  - Multi-vendor technology agreement and interoperability plans (cisco, Alcatel and Ericsson)
  - draft-gredler-... co-authors want more details as draft. We commit to detailed drafts by end of May
- May
  - Team shares 6 detailed drafts with draft-gredler-... co-authors and seek merge agreement
- June
  - Merge agreement on a subset of SR: MPLS/SR instantiation, use-cases, FRR, ISIS and OSPF: great collaborative work
- The point:
  - Detailed and thoughtful work
  - Velocity
  - Collaborative
  - Commitments met

Next few months - IETF

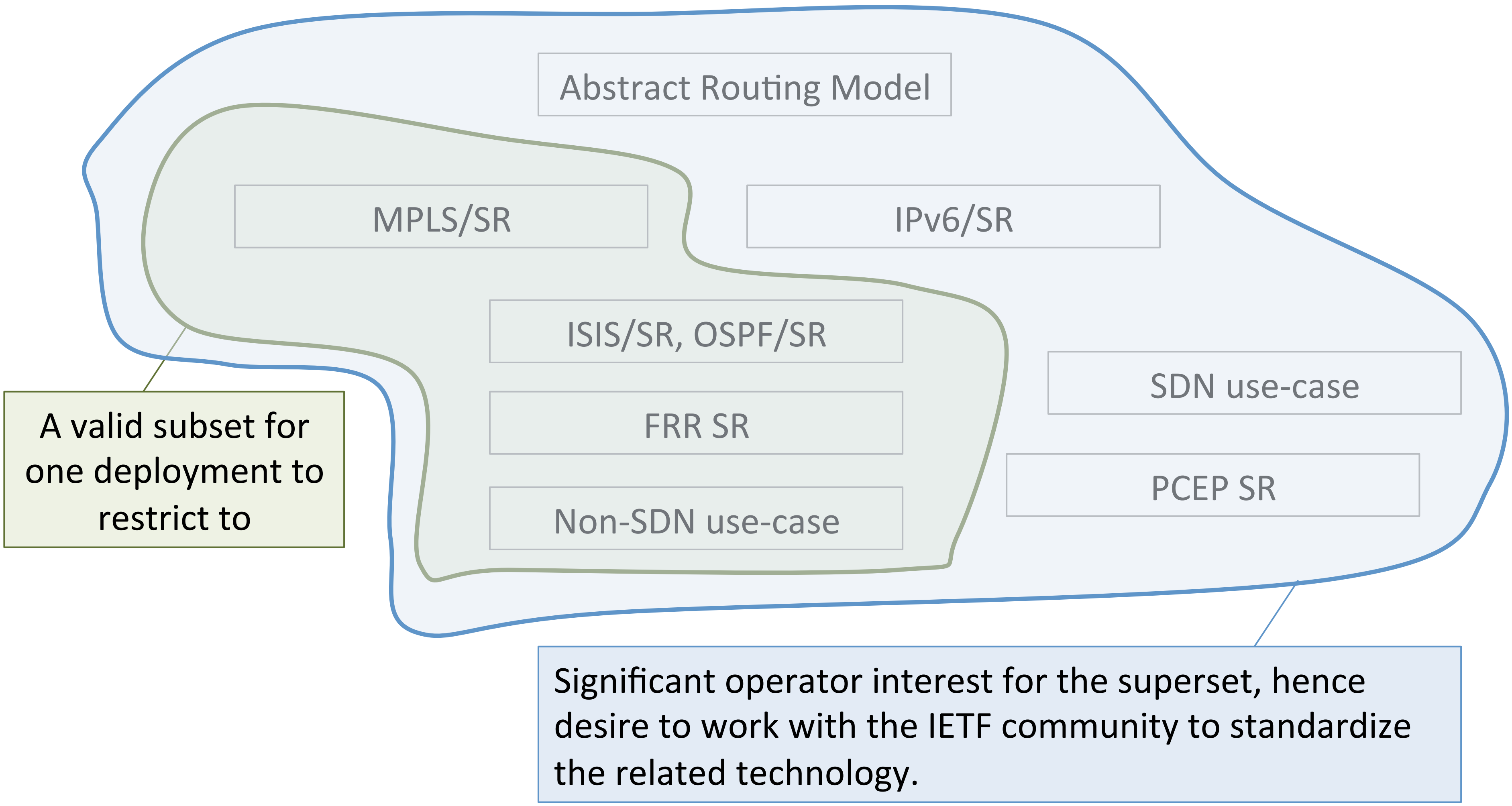
Topic	IETF Reference	WG
Abstract Routing Model	draft-filsfils-rtgwg-segment-routing	RTGWG
MPLS Instantiation	New draft to be submitted (based on section 5 of draft-filsfils-rtgwg-segment-routing)	MPLS
IPv6 Instantiation	New draft to be submitted	IPv6
Use Cases	draft-filsfils-rtgwg-segment-routing-use-cases	RTGWG
Perf Eng. LSP with SR	draft-shakir-rtgwg-sr-performance-engineered-lsps	RTGWG
ISIS SR Extensions	draft-previdi-isis-segment-routing-extensions	ISIS
OSPF SR Extensions	draft-psenak-ospf-segment-routing-extensions	OSPF
FRR SR	draft-francois-sr-frr	RTGWG
PCEP SR Extensions	draft-sivabalan-pce-segment-routing	PCEP

- Positive collaboration and consensus

Progress

- Agreement between draft-sr and draft-gredler co-authors on a subset of SR:
  - ISIS, OSPF: merge already submitted
  - MPLS/SR, FRR and use-case: work in progress
- Disagreement on IPv6/SR, should not be an issue
  - We thus accepted to organize the documents such that those that only want to support the MPLS instantiation can do so
  - We believe that there is a clear demand (e.g. as confirmed by feedback in the room) , we are committed to a positive collaboration process
  - For IPv6, together with operators (Comcast, Rogers…), Martha Steenstrup and academia, we will submit the IPv6/SR draft proposal for the next IETF and will work with the community to improve it as required

Visually


Abstract Routing Model
MPLS/SR
IPv6/SR
ISIS/SR, OSPF/SR
FRR SR
Non-SDN use-case
SDN use-case
PCEP SR
A valid subset for one deployment to restrict to
Significant operator interest for the superset, hence desire to work with the IETF community to standardize the related technology.

Next IETF

- While a new WG might be formed, we would like to be able to present and review our proposal in their home WG's
  - ISIS
  - OSPF
  - MPLS
  - 6man
  - PCEP
  - RTGWG
- Significant operator support and vendor consensus

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

Sections	Presentations today
2. IGP-based MPLS Tunneling	Martin, Victor
3. FRR	Bruno
4.1.1 Disjointness in dual-plane networks	Martin
4.1.2. CoS-based Traffic Engineering	Martin
4.4. Deterministic non-ECMP Path	Rob
5.2. SDN /SR use-case	Victor
6.4. Leveraging SR benefits for LDP-based traffic	Bruno
7. OAM	Rudiger

There is a lot of requirements, and SR meets all of them, despite their variety, with few extensions to core protocols. No new protocol is added.

Conclusion

- Multi-vendor/operator constructive collaboration
- Many requirements/use-cases supported by small extensions to well-established core protocols
  - ISIS, OSPF
  - LFA
  - PCEP
  - MPLS
  - IPv6
- Significant industry interest and contribution to SR
- Your feedback and contribution are welcome!

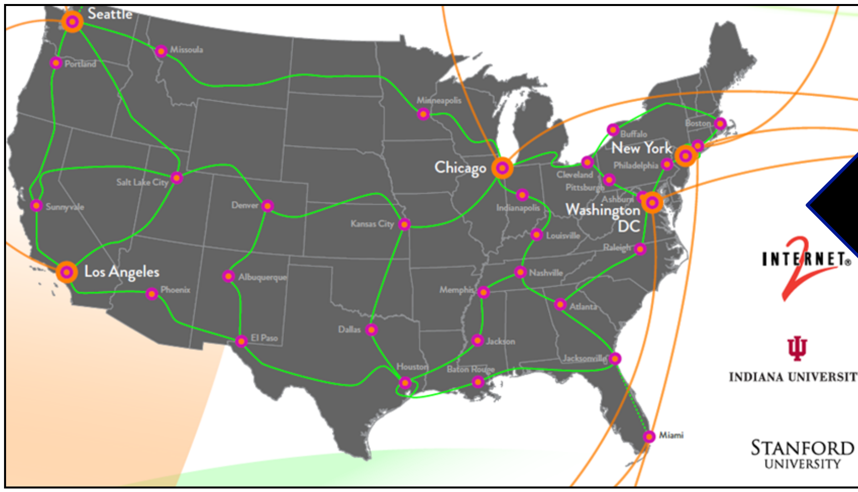
draft-ashwood-sdnrg-state-reduction-00

Presentation to STATUS BOF

Peter Ashwood-Smith
peter.ashwoodsmith@huawei.com

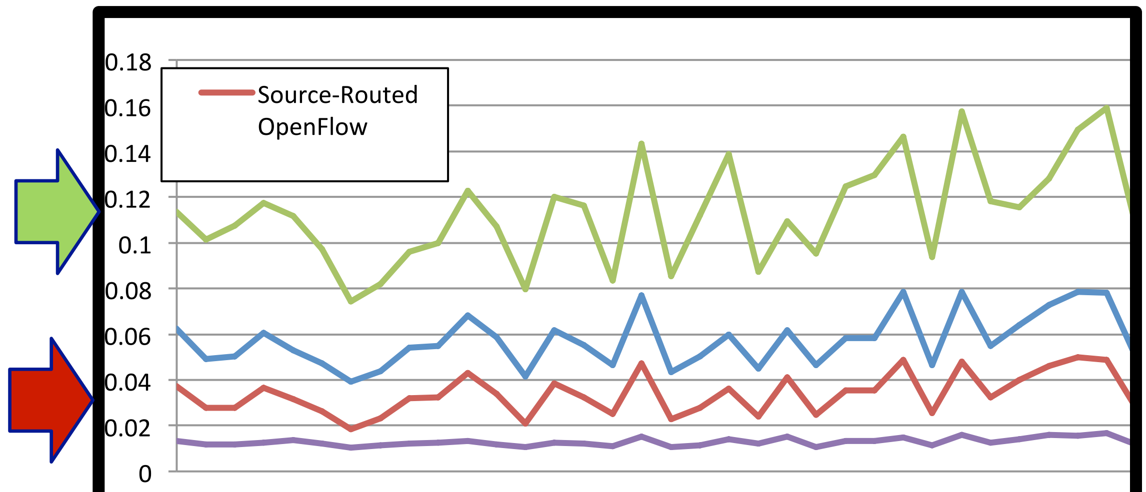
Mourad Soliman
MouradSoliman@cmail.carleton.ca

Source/Segment Routed SDN Simulation



We simulated a network of 36 nodes (OS³E), various flows/paths and compared source/segment routed v.s traditional forwarding. Simulator implemented in NS-3.

Observed 3 x Improvement

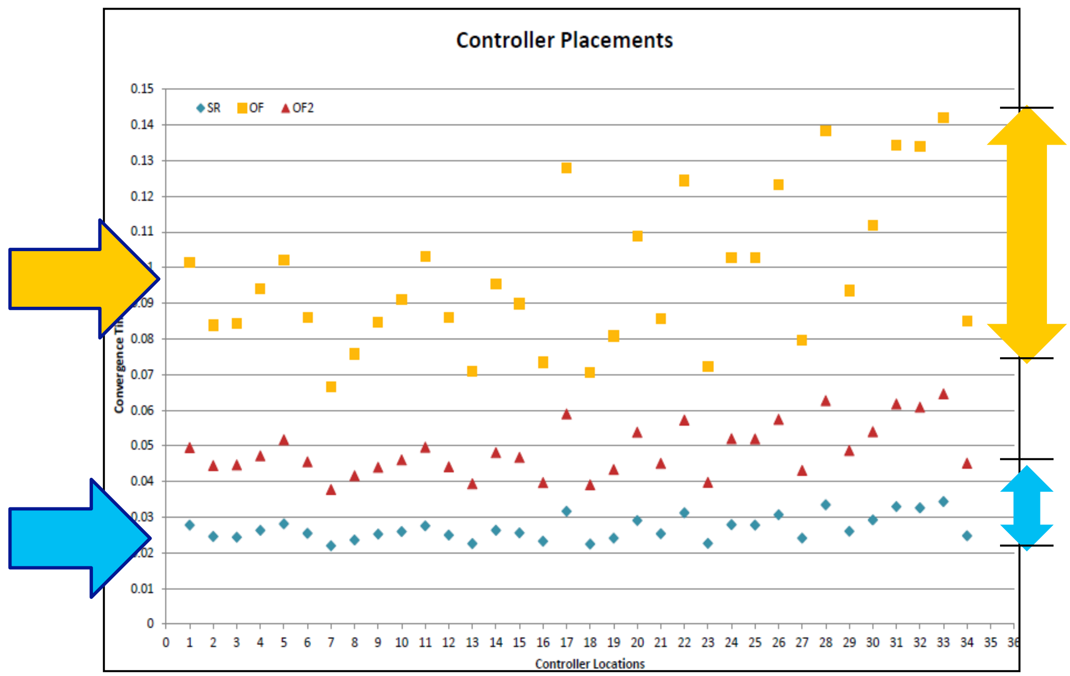


Convergence time for each source

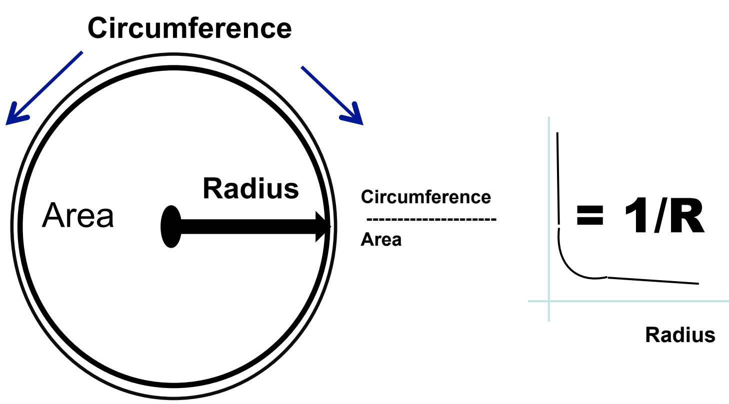
Observed 86 % reduced standard deviation in convergence for all controller placements with source / segment routed v.s. hop-by-hop forwarded SDN

SDN hop-by-hop is very sensitive to controller position.

Source/Segment Routing reduced sensitivity by 86%



Relative Source/Segment Routing performance as function of increasing diameter

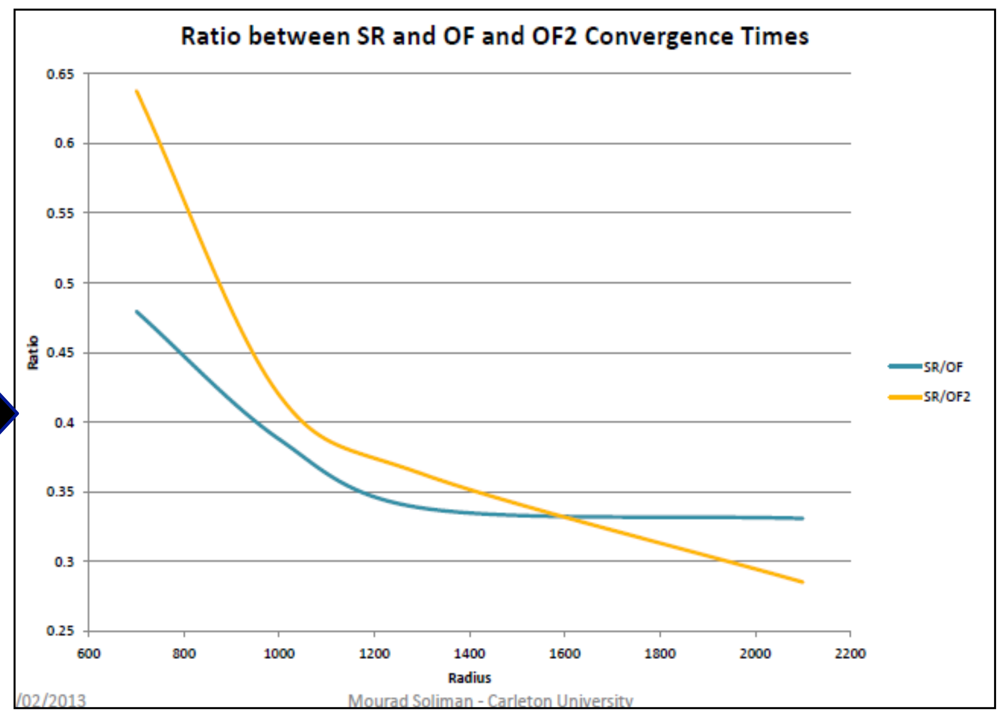


Source/Segment Route Burden is:
O(Circumference)

Hop-by-Hop Burden is:
O(Area)

Plotted Source/Segment Route convergence divided by hop-by-hop convergence for different network sizes.

We observed 1/R trend… ie linearly better with bigger R.



Conclusions

Definite advantages to source routing/segment routing in SDN.

Advantages grow linearly with network diameter so bigger => more impressive advantages.

Simulations give strong support for segment routing source routing/SDN use case claim.

In SDN with POF can implement smaller overhead.

Thank-You

STATUS (Stacked Tunnels for Source Routing) BOF

IETF-87, Berlin

July 29, 2013

Note Well

This summary is only meant to point you in the right direction, and doesn't have all the nuances. The IETF's IPR Policy is set forth in BCP 79; please read it carefully.

The brief summary:

- **By participating with the IETF, you agree to follow IETF processes.**
- **If you are aware that a contribution of yours (something you write, say, or discuss in any IETF context) is covered by patents or patent applications, you need to disclose that fact.**
- **You understand that meetings might be recorded, broadcast, and publicly archived.**

For further information, talk to a chair, ask an Area Director, or review the following:

BCP 9 (on the Internet Standards Process)

BCP 25 (on the Working Group processes)

BCP 78 (on the IETF Trust)

BCP 79 (on Intellectual Property Rights in the IETF)

Goal of this BOF, in a nutshell

- Source routing has existed (almost?) as long as networking
  - Historically, we have made little use of it in the IP suite
- Goal of the BOF is to
  - Answer the question, is source routing an idea whose time has come again?
  - If yes, provide input as to how it should be done within the IETF

Goals, in more detail

- What are the use cases driving this work?
- How much data plane variety is needed? MPLS? IPv6? IPv4?
- What are the basic assumptions about the stability of the existing data planes?
- What management/control architecture do we want?
- What does a label represent?

Agenda

- Introduction
- Use cases
- Architectures
- Open Discussion
- Wrap-up