NEXT-GENERATION TRAFFIC ENGINEERING WITH SPRING

INTRODUCTION + LIVE DEMO

- Basically two choices within an autonomous system

- LDP
  - provides any-to-any connectivity
  - less control plane overhead
  - no traffic engineering capabilities, shortest path only

- RSVP-TE
  - explicit signalization of LSPs
  - can provide any-to-any connectivity if configured
  - provides traffic engineering capabilities
  - creates states in the network for each LSP
    - especially problematic on P Routers

3 LSPs   12 LSPs   3 LSPs

3 LSPs   3 LSPs
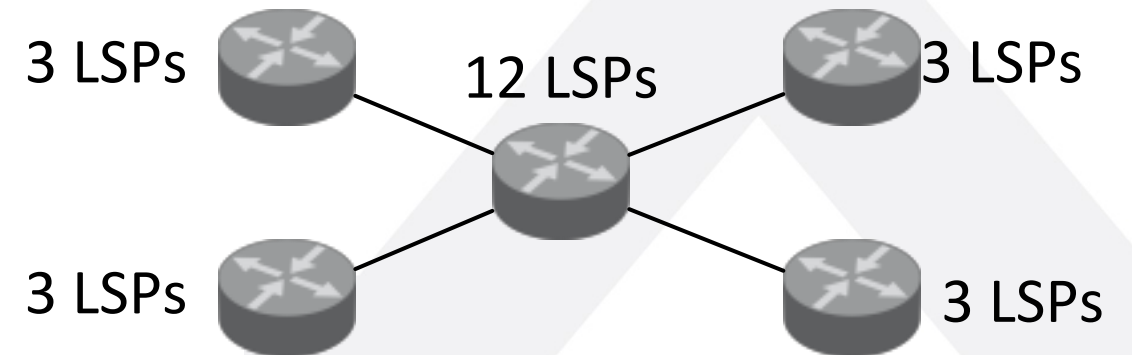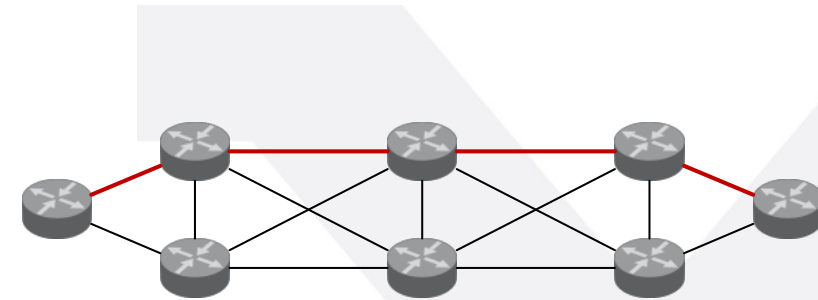
SPRING

- New approach standardized in the IETF
  - **S**ource **P**acket **R**outing **i**n Networkin**G**

- Formerly known as Segement Routing (SR)

- SPRING envisions the network as a collection of topological sub-paths – also called segments.
- Create any kind of desired path by stitching one or more individual segments
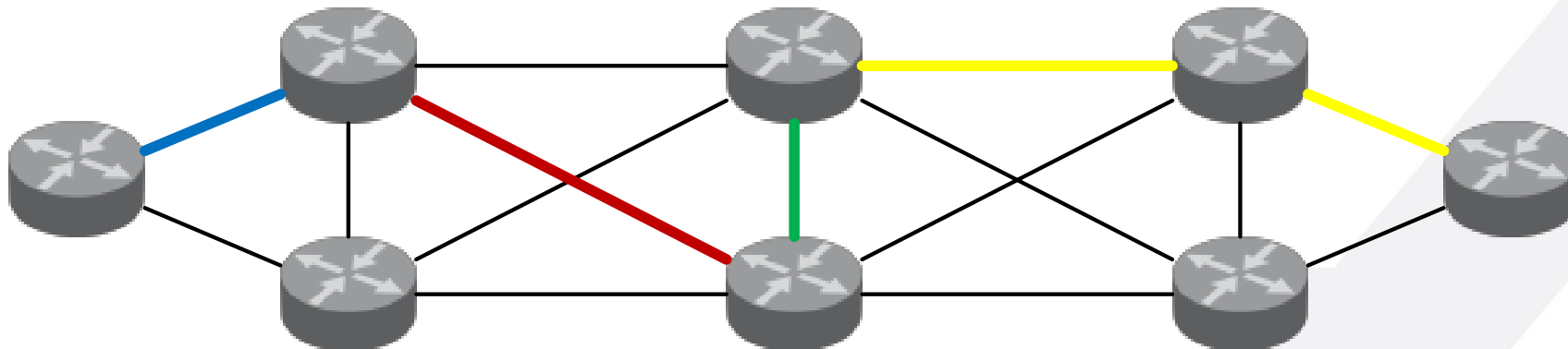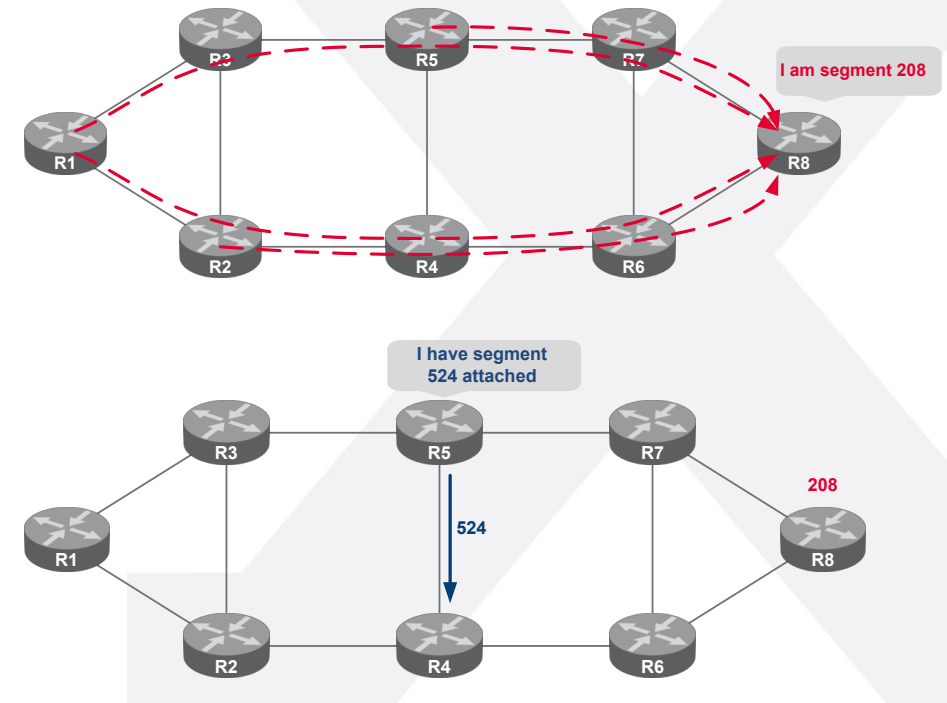
SPRING

- A segment can represent any instruction
  - Service
  - Context
  - IGP-based forwarding construct
- Ordered list of segments represents an ordered chain of topological and service instructions

- Per flow state only at ingress SPRING edge node
  - Ingress edge node pushes the segment list on the packet

- For MPLS Networks : No change required on dataplane (also available for Plain IPv6 forwarding)

SPRING IGP Segment Details

- Prefix Segment
  - Steers traffic along ECMP-aware shortest path to the related IGP Prefix
  - Global segment within the SPRING IGP domain
  - **Node Segment**: a prefix segment allocated to a prefix that identifies a specific node (e.g. the prefix is its loopback)
  - Like LDP based LSP

- Adjacency Segment
  - Steers traffic onto an adjacency or a set of adjacencies
  - Local segment related to a specific SPRING node
  - Brings TE capabilities into SPRING

- Router R8 advertises its node segment **208** with his loopback address
  - simple ISIS sub-TLV extension
  - simple OSPF Opaque sub-TLV extension
- All remote nodes install the node segment to R8 in the SPRING data plane along the IGP shortest path to R8/32
  - Packet injected anywhere with top segment/label **208** will reach router R8 via ECMP-aware shortest path

I am segment 208
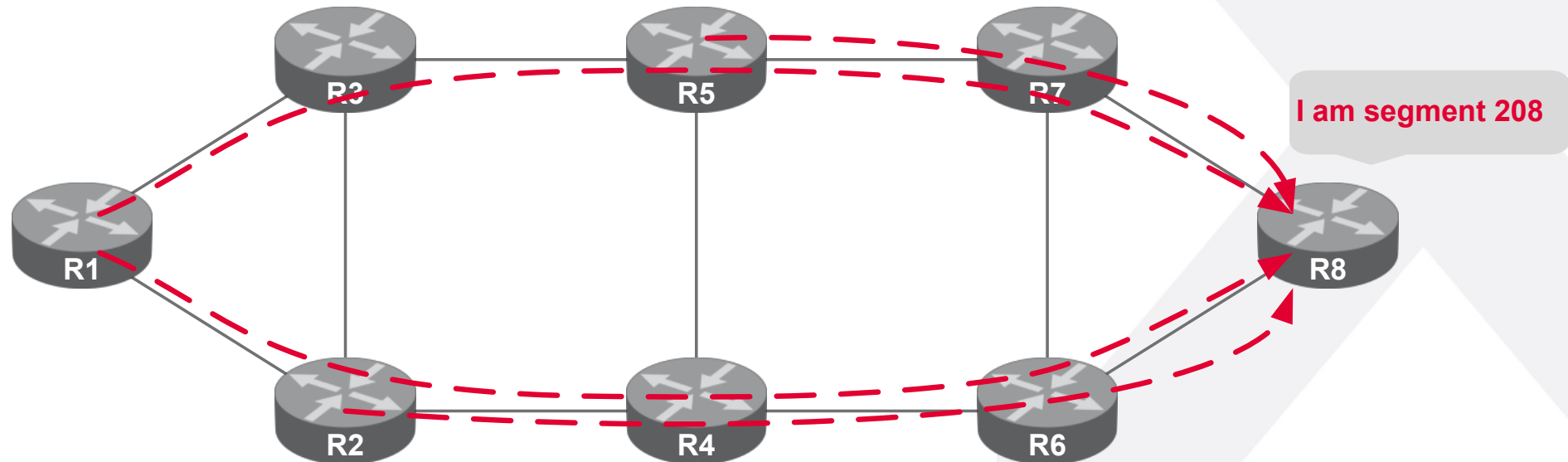
- Router R5 allocates a local segment **524** for its adjacency R5-R4 and advertises the segment in IGP
    - simple ISIS sub-TLV extension
    - simple OSPF Opaque sub-TLV extension
- R5 is the only node to install the adjacency segment in the SR data plane
    - Packet reaching node R5 with segment/label 524 is forwarded through link R5->R4
    - R4 will assign a different label for the same link towards R5
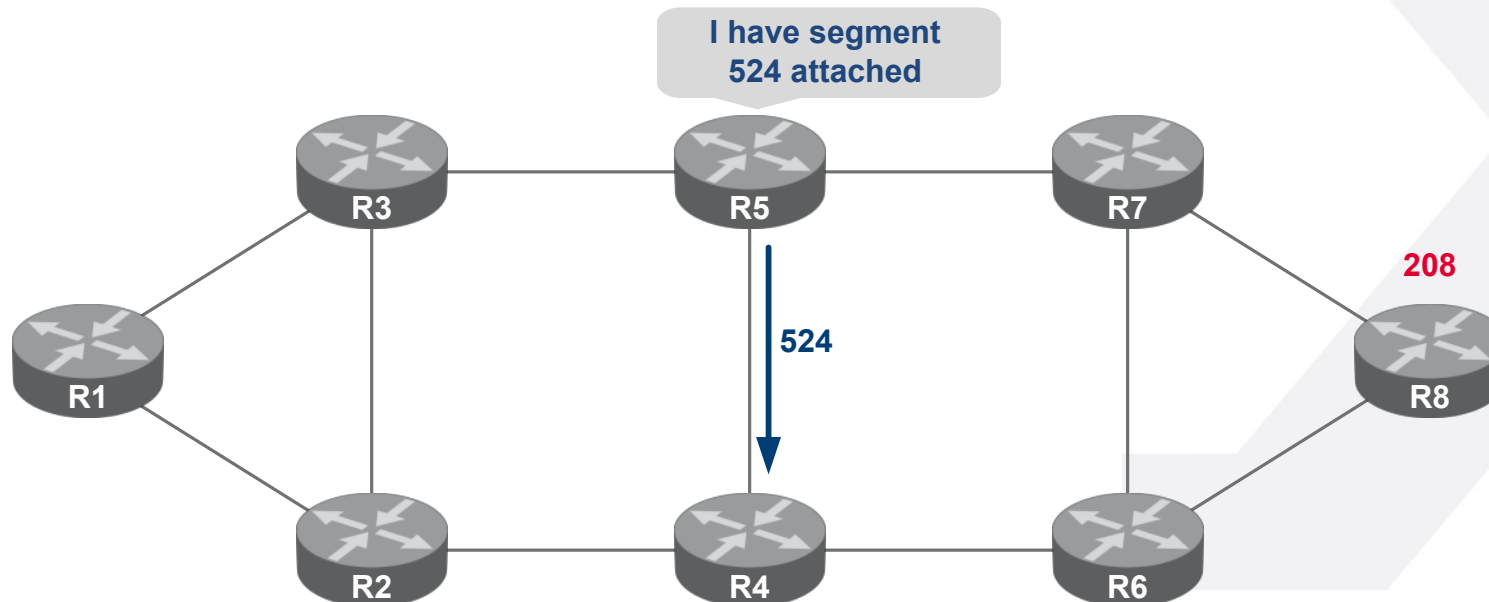
- Any path can be expressed using a combination of IGP prefix (node) segments and adjacency segments at the ingress Router
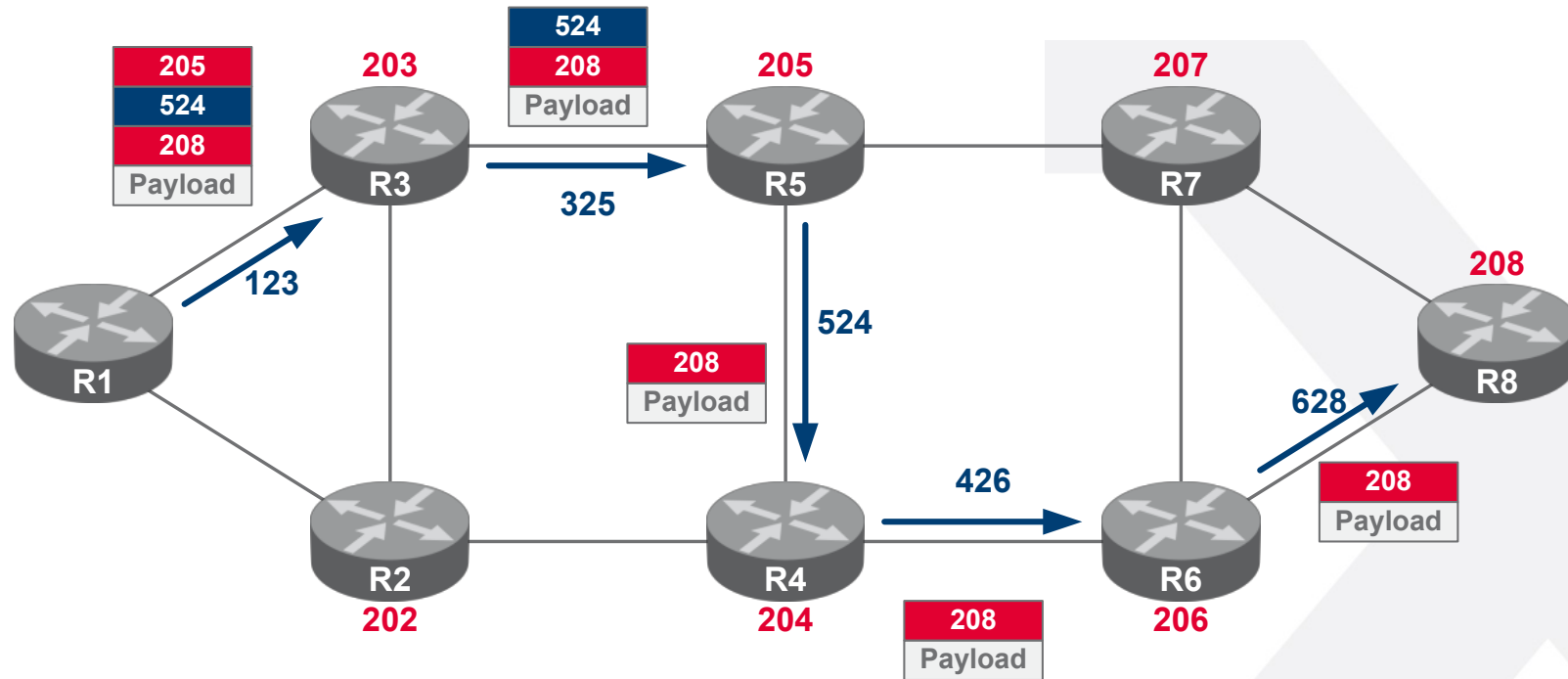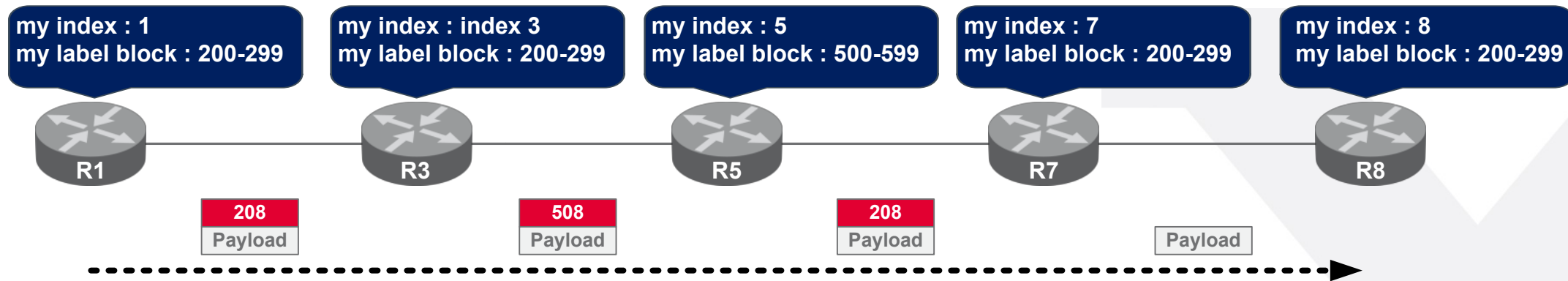

- Penultimate Hop Popping may already pop Labels as shown above

- A segment could even be a pre signaled RSVP LSP (Drafts exist)

Global vs. Local Segment/Label

- Global Labeling "breaks" MPLS Architecture (local significance)

- Workaround 1: Announce local label block + configured unique node segment id
  - R1 node-id=1, R2 node-id=2 …
  - R1 label block 100-199, R2 label block 200-299
  - R1 node segment = 101, R2 node segment = 202

- Workaround 2: Announce global label block + configured unique node segment id
  - R1 node-id1, R2 node-id2 …
  - R1 label block 100-199, R2 label block 100-199
  - R1 node segment = 101, R2 node segment = 102

- Operator manages SRGB (Segment Route Global Block) similar to other resources (e.g. IP address)
  - it ensures unique allocation of a global segment within the SR domain

Global vs. Local Segment/Label - Example

- Workaround 1 results in label swapping as we know it

my index : 1 my label block : 200-299	my index : index 3 my label block : 200-299	my index : 5 my label block : 500-599	my index : 7 my label block : 200-299	my index : 8 my label block : 200-299

R1 — R3 — R5 — R7 — R8

| 208
Payload | 508
Payload | 208
Payload | | Payload |

- Workaround 2 results in a „swap" operation that does not actually change the value

my index : 1 my label block : 200-299	my index : index 3 my label block : 200-299	my index : 5 my label block : 200-299	my index : 7 my label block : 200-299	my index : 8 my label block : 200-299

R1 — R3 — R5 — R7 — R8

| 208
Payload | 208
Payload | 208
Payload | | Payload |
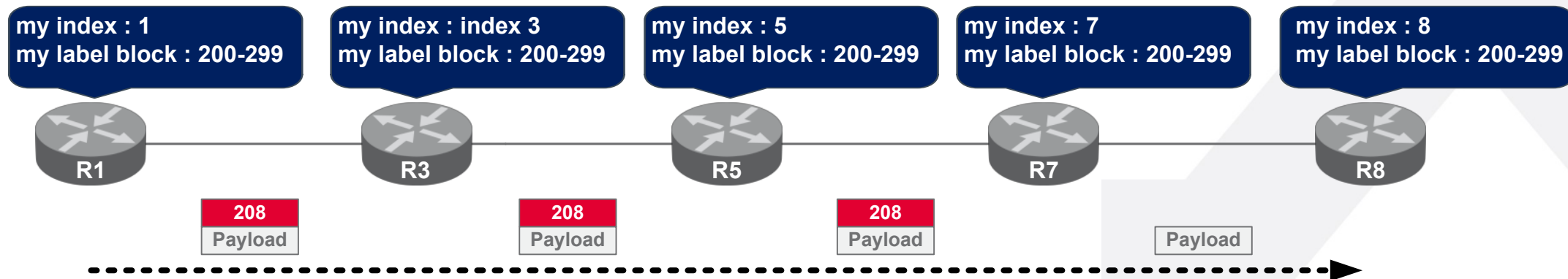
- Policy state is in the packet header, not on the transit routers

- There is no midpoint state (n^2 scale in RSVP-TE)

- There is no extra protocol (RSVP-TE, LDP)

- There is support for Native ECMP

  - up to a certain label stack depth (could be combined with entropy label to overcome)

- Few announced segments are required to achieve TE

- Segment-routing traffic engineering supports distributed or centralized computation

  - bandwidth reservations only centralized

- Labels are known throughout the whole IGP area

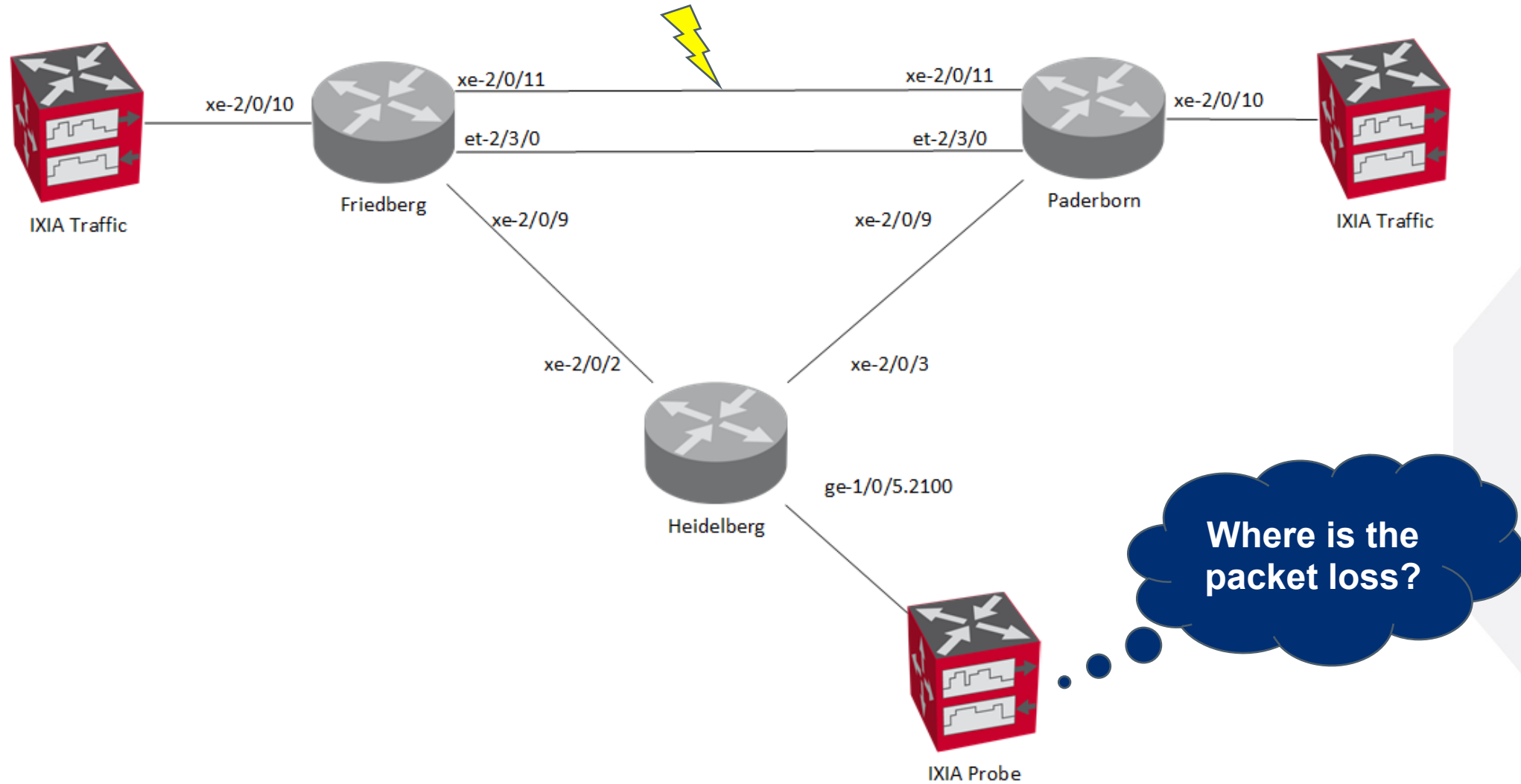 - no need for targeted LDP sessions for Remote-LFA

```
xuser@friedberg> show isis database heidelberg.00-00 extensive
<output omitted>
  TLVs:
    Area address: 49 (1)
    LSP Buffer Size: 1492
    Speaks: IP
    Speaks: IPV6
    IP router id: 192.0.2.1
    IP address: 192.0.2.1
    Hostname: heidelberg
    IS extended neighbor: friedberg.00, Metric: default 10
      IP address: 192.0.2.128
      Neighbor's IP address: 192.0.2.129
      Local interface index: 328, Remote interface index: 337
      P2P IPV4 Adj-SID - Flags:0x30(F:0,B:0,V:1,L:1,S:0), Weight:0, Label: 299792
    IS extended neighbor: paderborn.00, Metric: default 10
      IP address: 192.0.2.145
      Neighbor's IP address: 192.0.2.144
      Local interface index: 330, Remote interface index: 331
      P2P IPV4 Adj-SID - Flags:0x30(F:0,B:0,V:1,L:1,S:0), Weight:0, Label: 299808
    IP extended prefix: 192.0.2.1/32 metric 0 up
      8 bytes of subtlvs
      Node SID, Flags: 0x40(R:0,N:1,P:0,E:0,V:0,L:0), Algo: SPF(0), Value: 1
    IP extended prefix: 192.0.2.128/31 metric 10 up
    IP extended prefix: 192.0.2.144/31 metric 10 up
    Router Capability:  Router ID 192.0.2.1, Flags: 0x01
      SPRING Capability - Flags: 0xc0(I:1,V:1), Range: 4096, SID-Label: 800000
      SPRING Algorithm - Algo: 0
```

Deterministic Network Probing

- draft-lapukhov-segment-routing-large-dc-00

- draft-filsfils-spring-segment-routing-msdc

  - Active probing is a well-known technique for monitoring network elements health, constituting of sending continuous packet streams simulating network traffic to the hosts in the data-center. Segment routing makes possible to prescribe the exact paths that each probe or series of probes would be taking toward their destination. This allows for fast correlation and detection of failed paths, by processing information from multiple actively probing agents.

Further Reading

- SPRING - Problem Statement and Requirements

  - https://tools.ietf.org/html/rfc7855

- Segment Routing with MPLS data plane

  - https://tools.ietf.org/html/draft-ietf-spring-segment-routing-mpls-05

- Segment Routing Architecture

  - https://tools.ietf.org/html/draft-ietf-spring-segment-routing-09

- IS-IS Extensions for Segment Routing

  - https://tools.ietf.org/html/draft-ietf-isis-segment-routing-extensions-09

- OSPF Extensions for Segment Routing

  - https://tools.ietf.org/html/draft-ietf-ospf-segment-routing-extensions-10

- OSPFv3 Extensions for Segment Routing

  - https://tools.ietf.org/html/draft-ietf-ospf-ospfv3-segment-routing-extensions-07

- Orange SPRING Interoperability testing report

  - http://www.slideshare.net/StephaneLitkowski/mpls-sdn-2015-spring-interoperability-testing

Thank you for your Attention!

Questions?!

Sebastian Graf
Solutions Architect

Tel +49 69 2443714 25
Mob +49 151 15154018
sgraf@xantaro.net