

# 6PE fails and other short stories...

Sandy Breeze

# 6PE

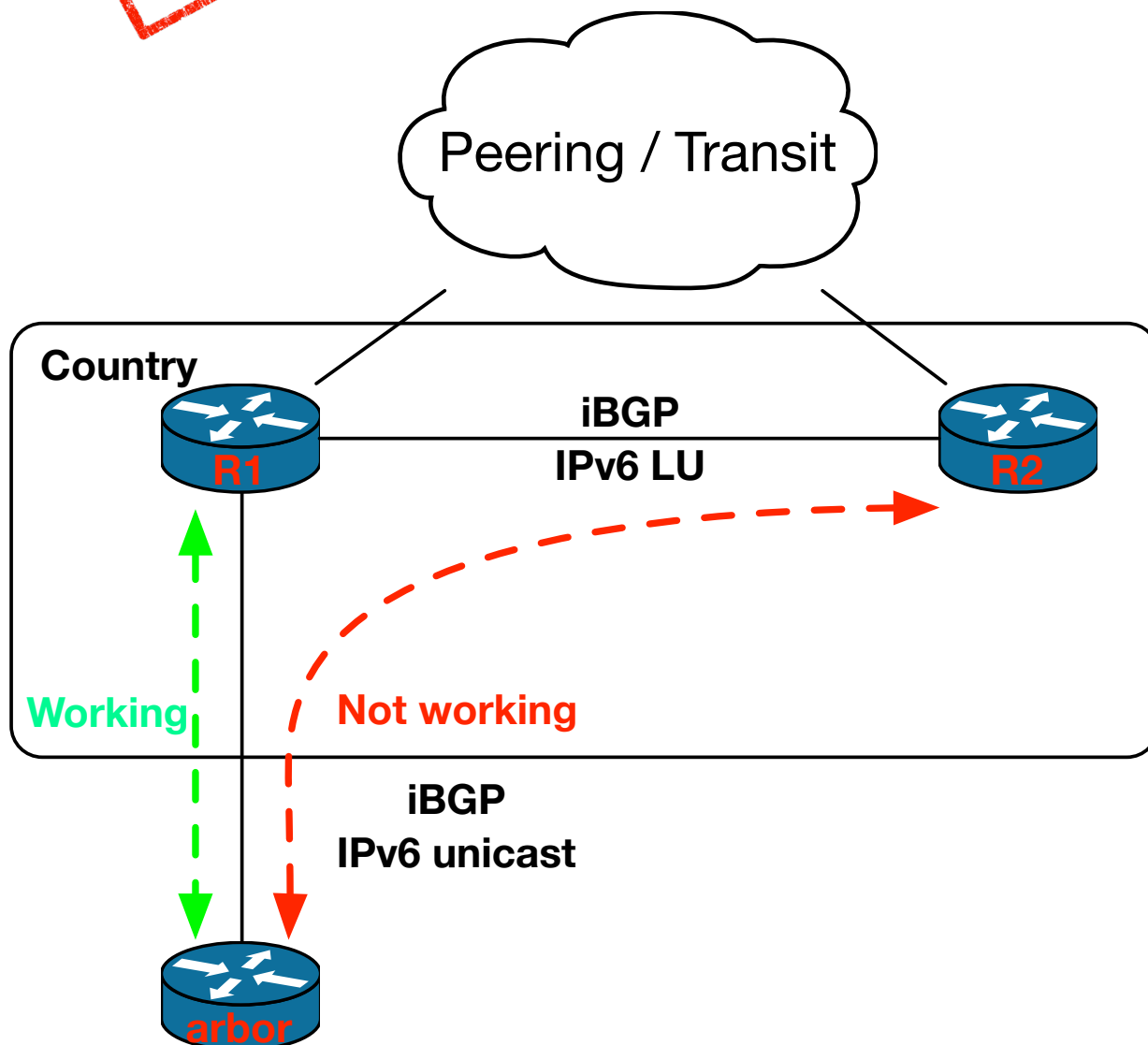
Why did we even do 6PE?

- Been around for ages, implementations should be mature
- LDP6 not going anywhere, never really happened
- Enables VPNv6

# 6PE



# DDoS detection



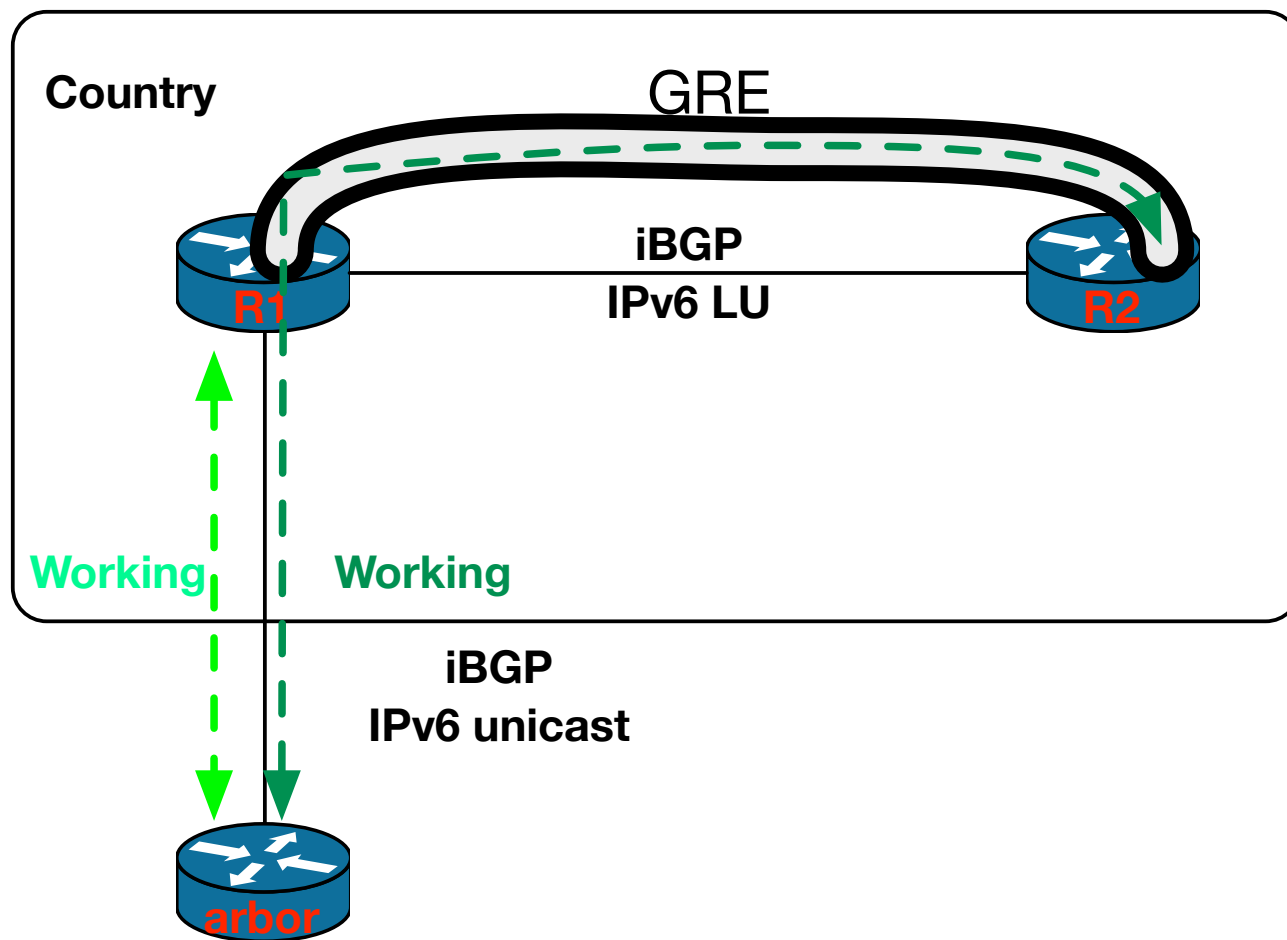
- No BGP session in AFI/SAFI 2/1 if NH in 2/4

IOS-XR

# 6PE



# DDoS detection



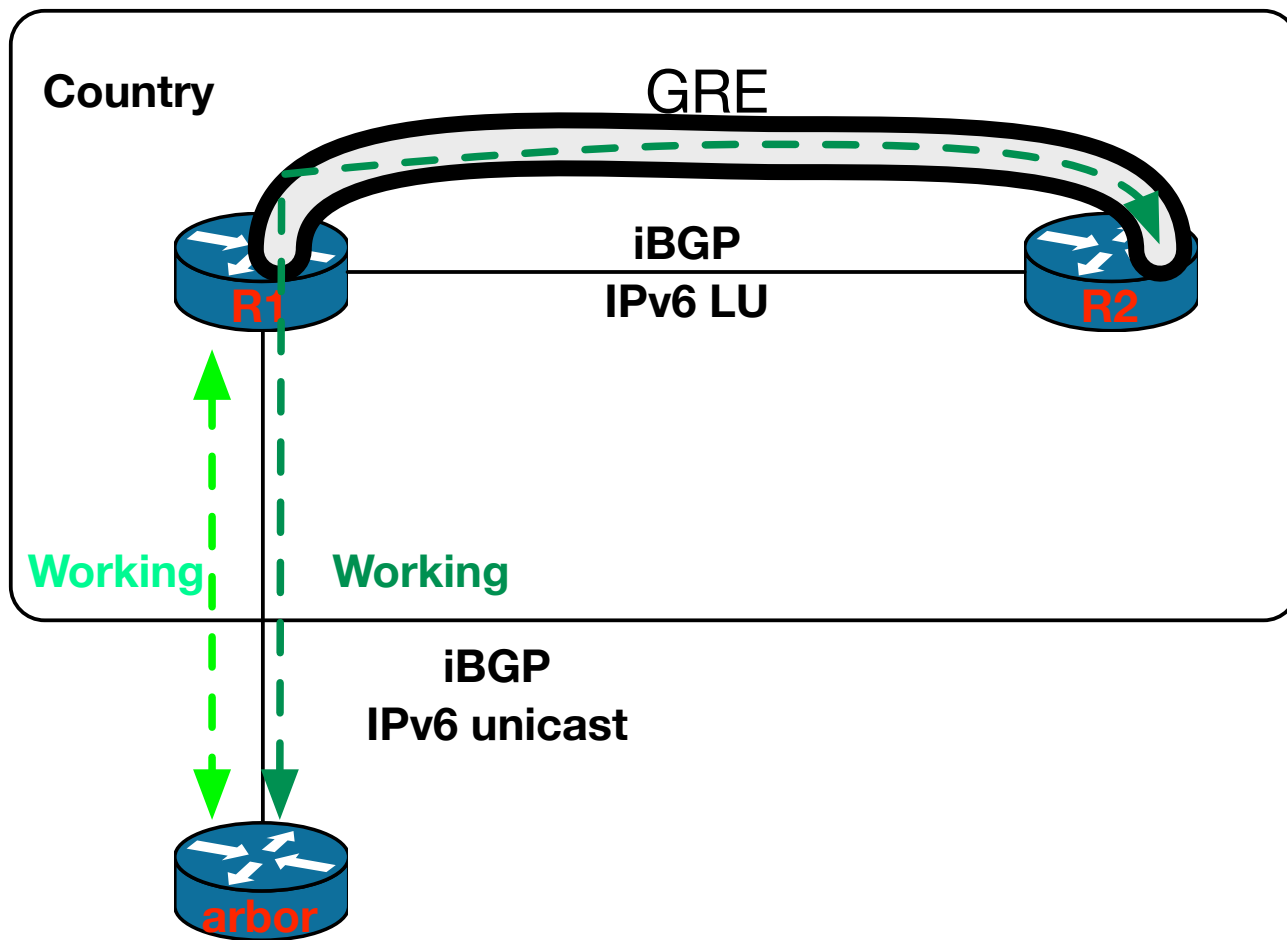
- GRE to the rescue!

IOS-XR

# 6PE



# DDoS detection



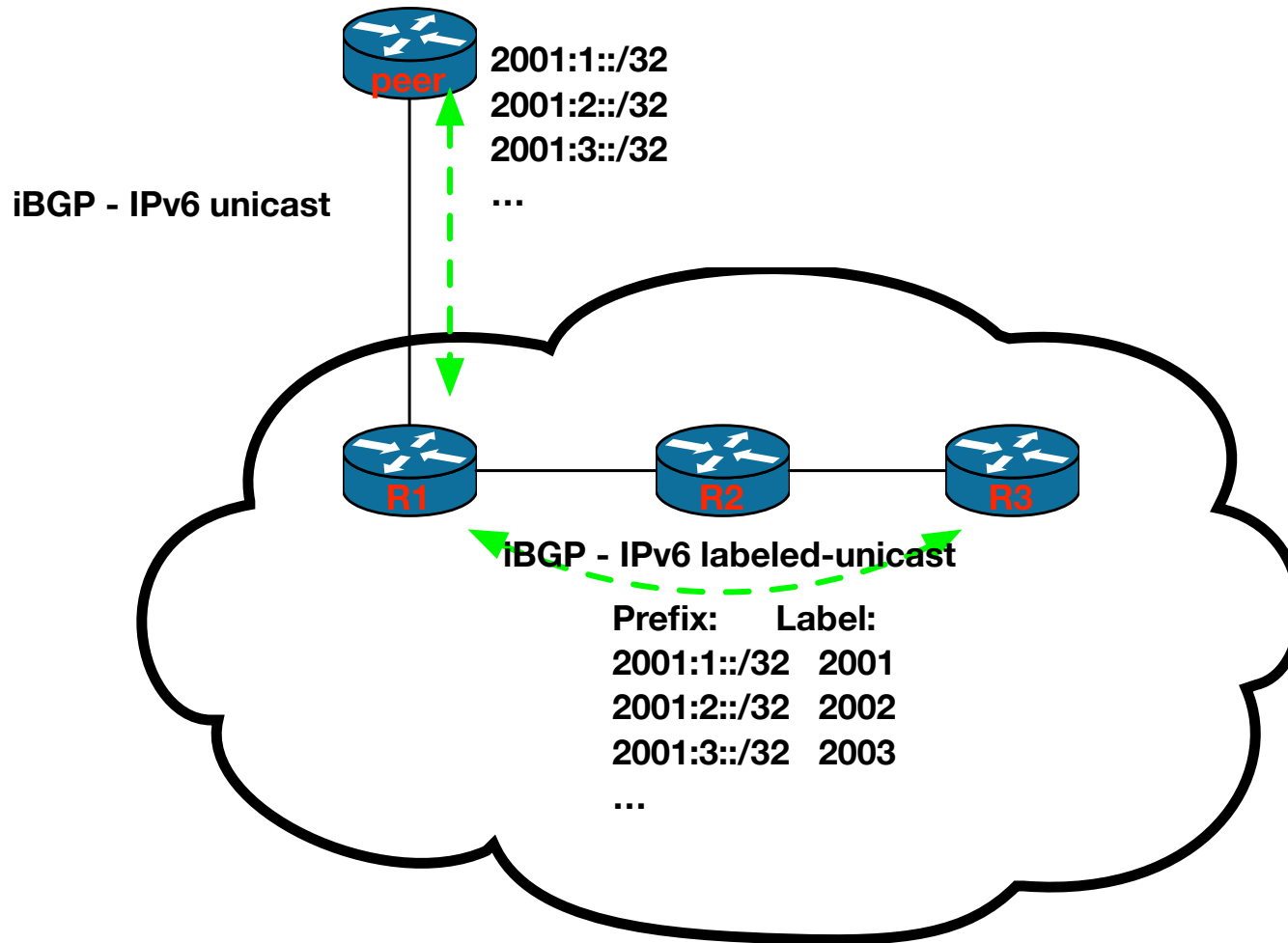
- FRR royally breaks GRE

IOS-XR

# 6PE



## Label Allocation



- Allocate-all does not scale

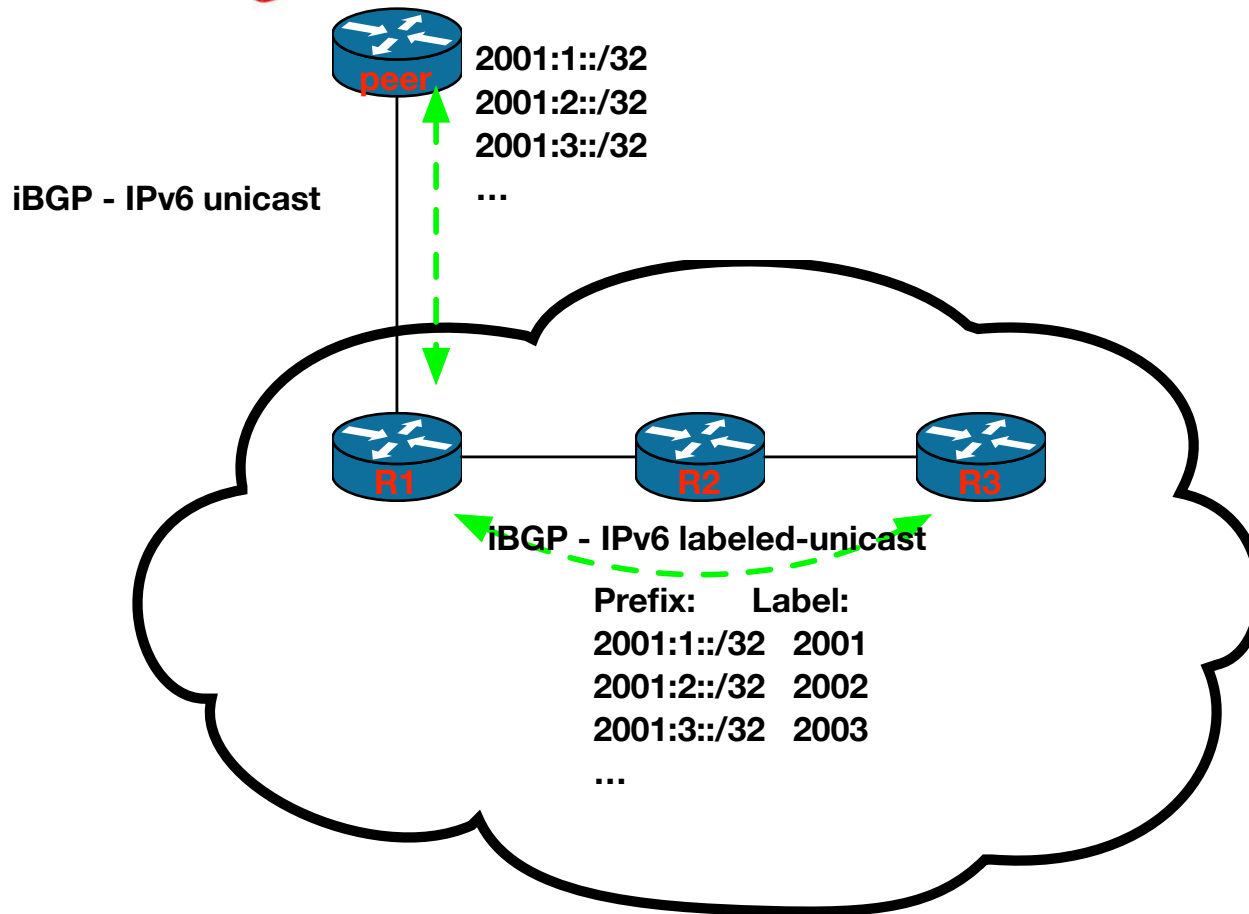
IOS

# 6PE



# Label Allocation

ARISTA



- Default behavior to IPv6 Exp-Null for all
- Configurable behavior to IPv6 Exp-Null

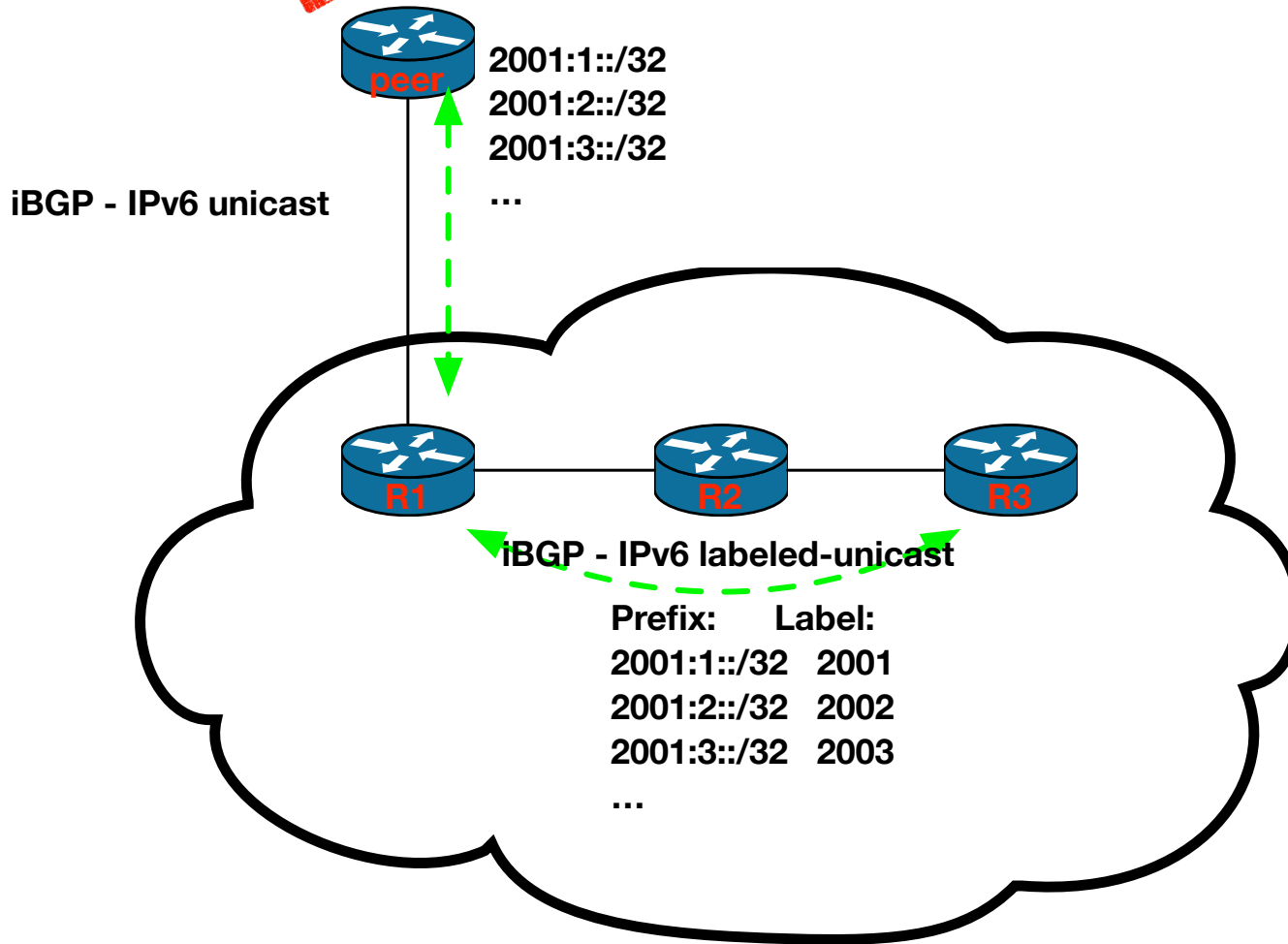
EOS

IOS-XR

# 6PE



# TCAM exhaustion



- Both 72 and 144 space used (If P+T edge)

IOS



# 6PE



## Bugs in 2019

- XRv fails to process 6LU withdraw, loc-RIB grows indefinitely until crash. Status: Fixed.

IOS-XR



- ASR1k continually sends full BGP RIB to IPv6-Unicast peers. Status: Fixed.

IOS-XE



- Shared code-path for all labelled NH's.

EOS



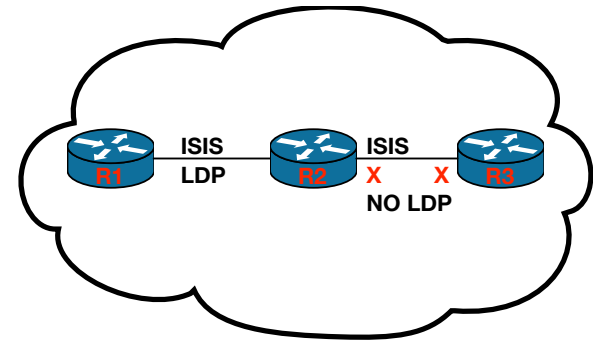
Eg; VPNv4 and 6PE, A withdraw in one will cause blackholing in other. State: Fixed (very quickly!)

# LDP



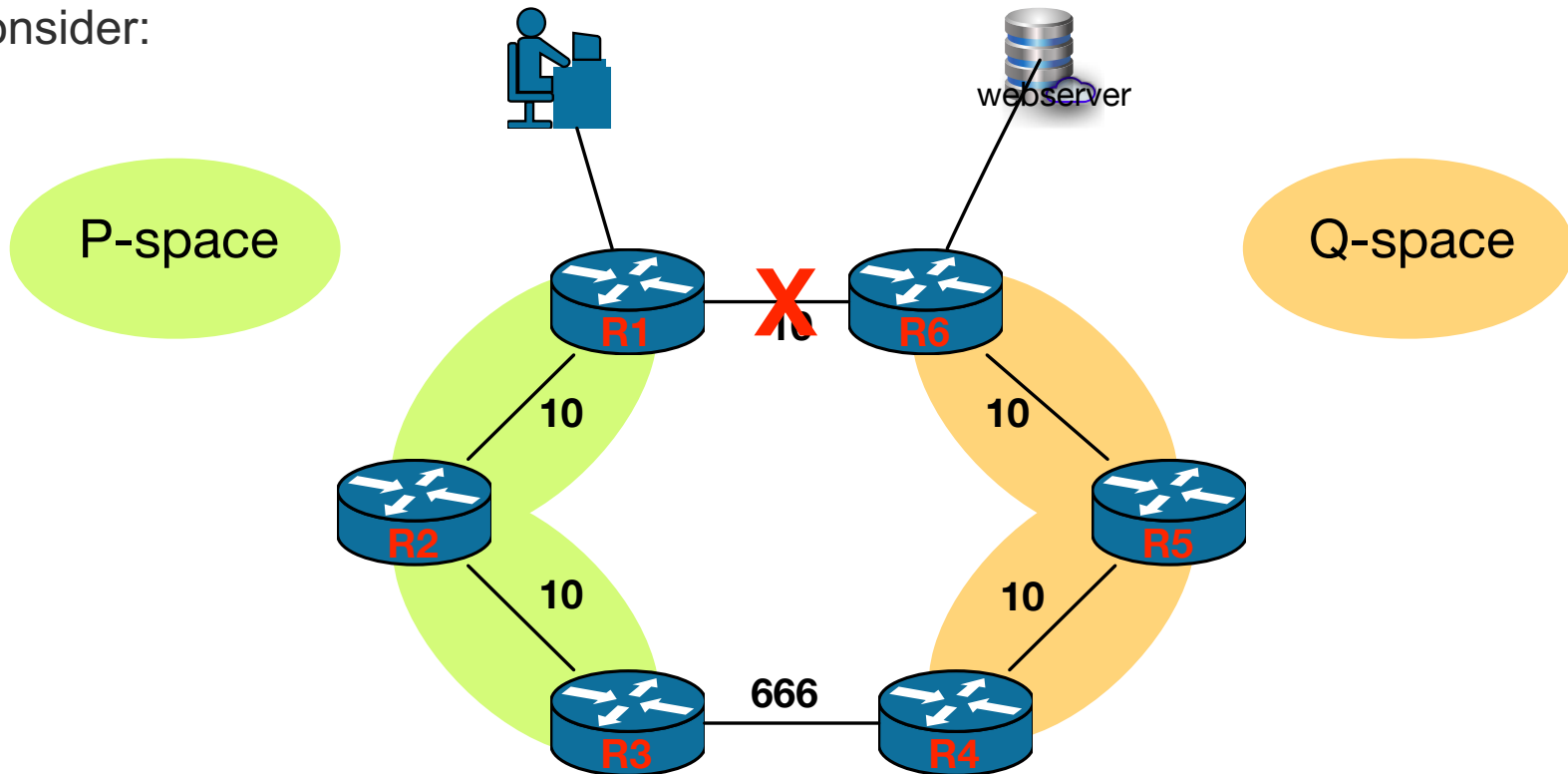
Deployed in 2003 for MPLS L3VPN

- Unnecessarily independent of IGP (in the core)
- Original spec had too many dependencies on IPv4
- RFC7552 LDP6 was too late to the table (June 2015)
- LDP6 not really implemented, and where it is, still no L2/L3VPN support



# Other nuisances: FRR LFA

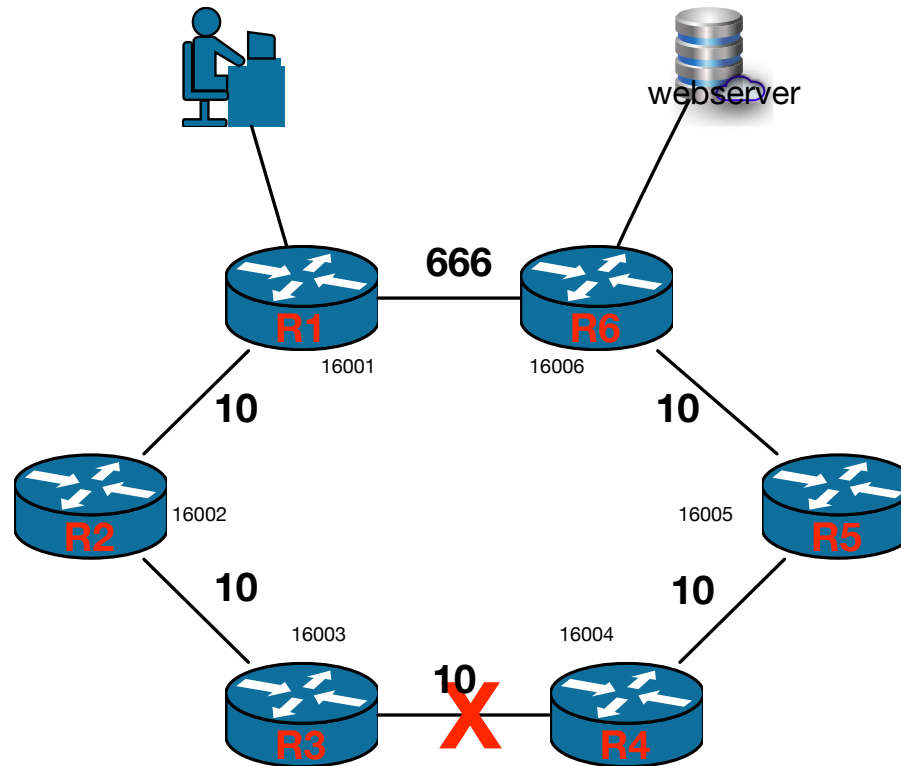
Consider:



- No shared PQ space, no FRR LFA

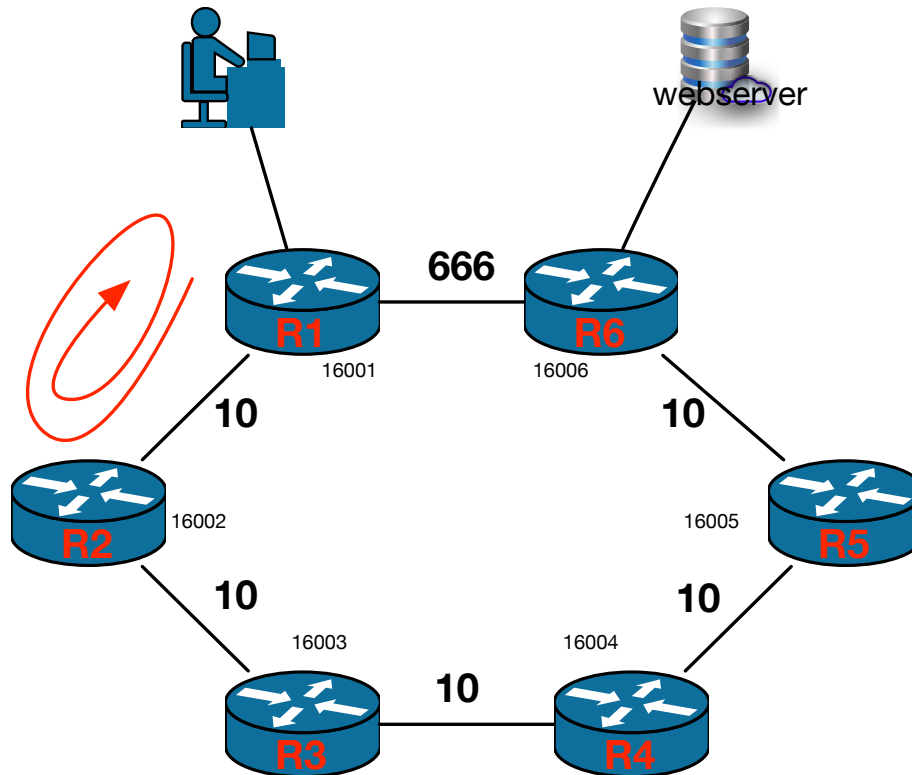
# Other nuisances: $\mu$ -loops

Consider:



# Other nuisances: $\mu$ -loops

Consider:

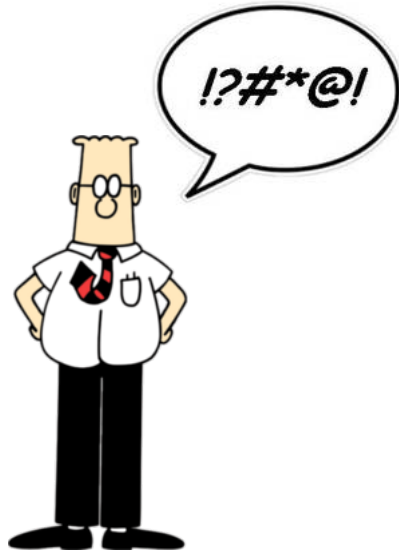


- Loop duration function of convergence time
- Existed since day 1 in IGP

# Problem recap

Problem	Why
6PE	BGP label allocation (can) chew through labels Vendors are still implementing it badly Relies on LDP
LDP	No (implemented) native IPv6 support Not closely coupled with IGP, independent
FRR LFA	Coverage can be bad where topology has no overlapping SPF from source / destination (PQ router space)
u-Loops	Bringing links into service can cause $\mu$ -loops which are based on surrounding speed devices converge

# Summary

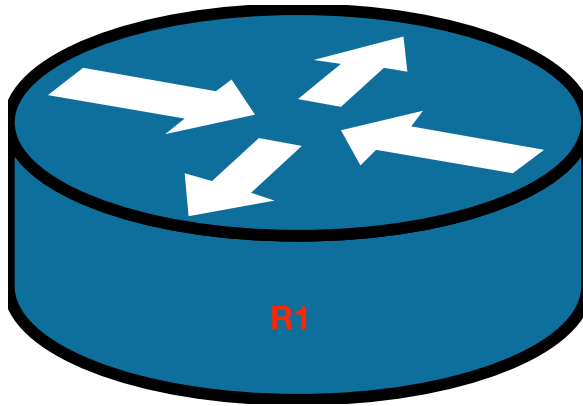


...we had a bad time with 6PE

...another tool on the block



# How do routers allocate labels?



- There is a label manager (LSD)
- Protocols who can allocate / distribute labels are clients of the label manager

## Label Manager: (aka LSD - Label Switching DB)

Application	Count
-----	
LSD (A)	4
LDP (A)	308
BGP-VPNv4 (A) :bgp-default	150494
ISIS (A) :CLUK	100
-----	
TOTAL	150905

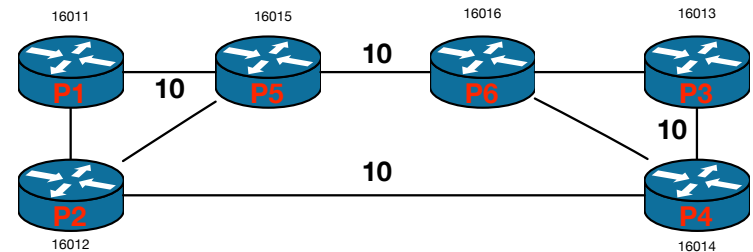
# If an IGP did labels...

...it'd (***probably***) be the best at label allocation and distribution in the world [sic]

# Segment Routing 101? (2 slides)

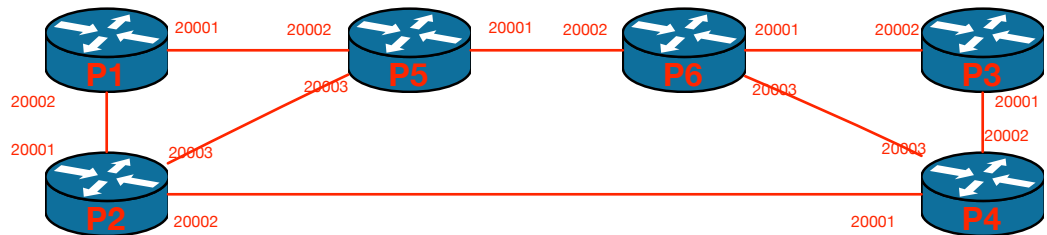
## Node SID

- Global instruction in IGP, which any node in the SR domain can execute
  - Forward to node x via shortest path



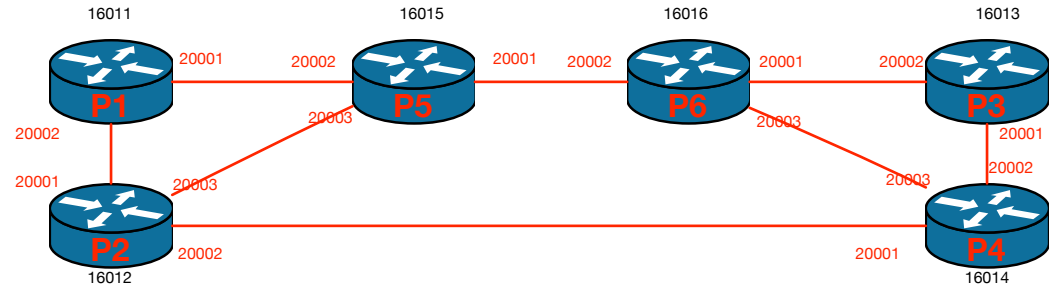
## Adjacency SID

- Instruction which only the node who originated the instruction can execute, eg:
  - Send out interface y



# Visually

Topology:

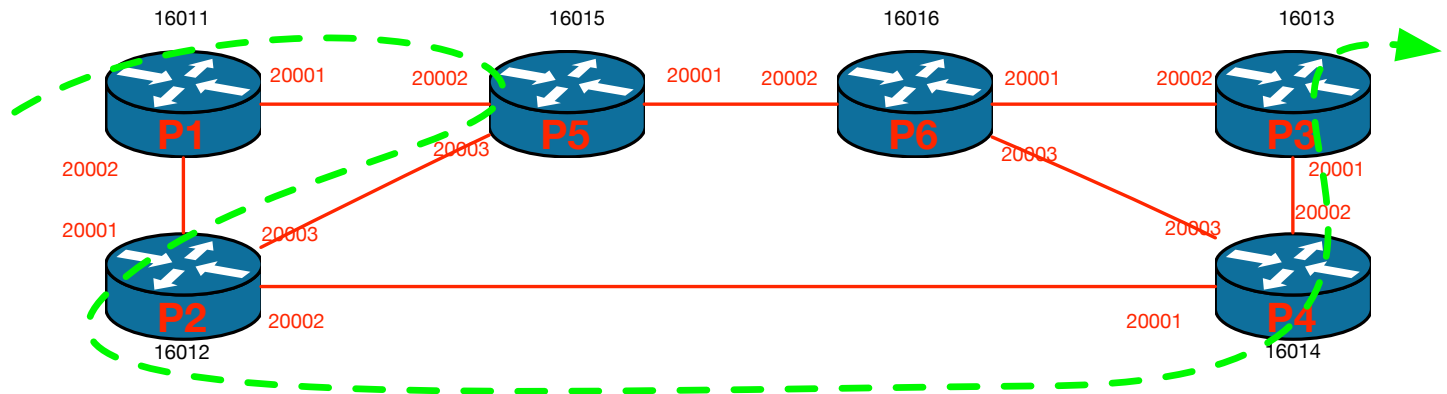


+

Apply SR header at P1:

{ 16015, 20003, 16013 }

=



# Problem recap

Problem	Why
<b>6PE</b>	BGP label allocation chews through labels Vendors are always implementing it badly Relies on LDP
<b>LDP</b>	No (implemented) native IPv6 support Not closely coupled with IGP, independent
<b>FRR LFA</b>	Coverage can be bad where topology has no overlapping SPF from source / destination (PQ router space)
<b>u-Loops</b>	Bringing links into service can cause $\mu$ -loops which are based on surrounding speed devices converge

# Problem recap

Problem	Why
LDP	No (implemented) native IPv6 support Not closely coupled with IGP, independent

- IS-IS will allocate and distribute them

# Problem recap

Problem	Why
6PE	BGP label allocation chews through labels Vendors are always implementing it badly Relies on LDP

- LDP will be gone
- Dual-stack all links in IS-IS
- Move to IPv6 unicast (keep LU for VPNv6 only)

# Problem recap

Problem	Why
---------	-----

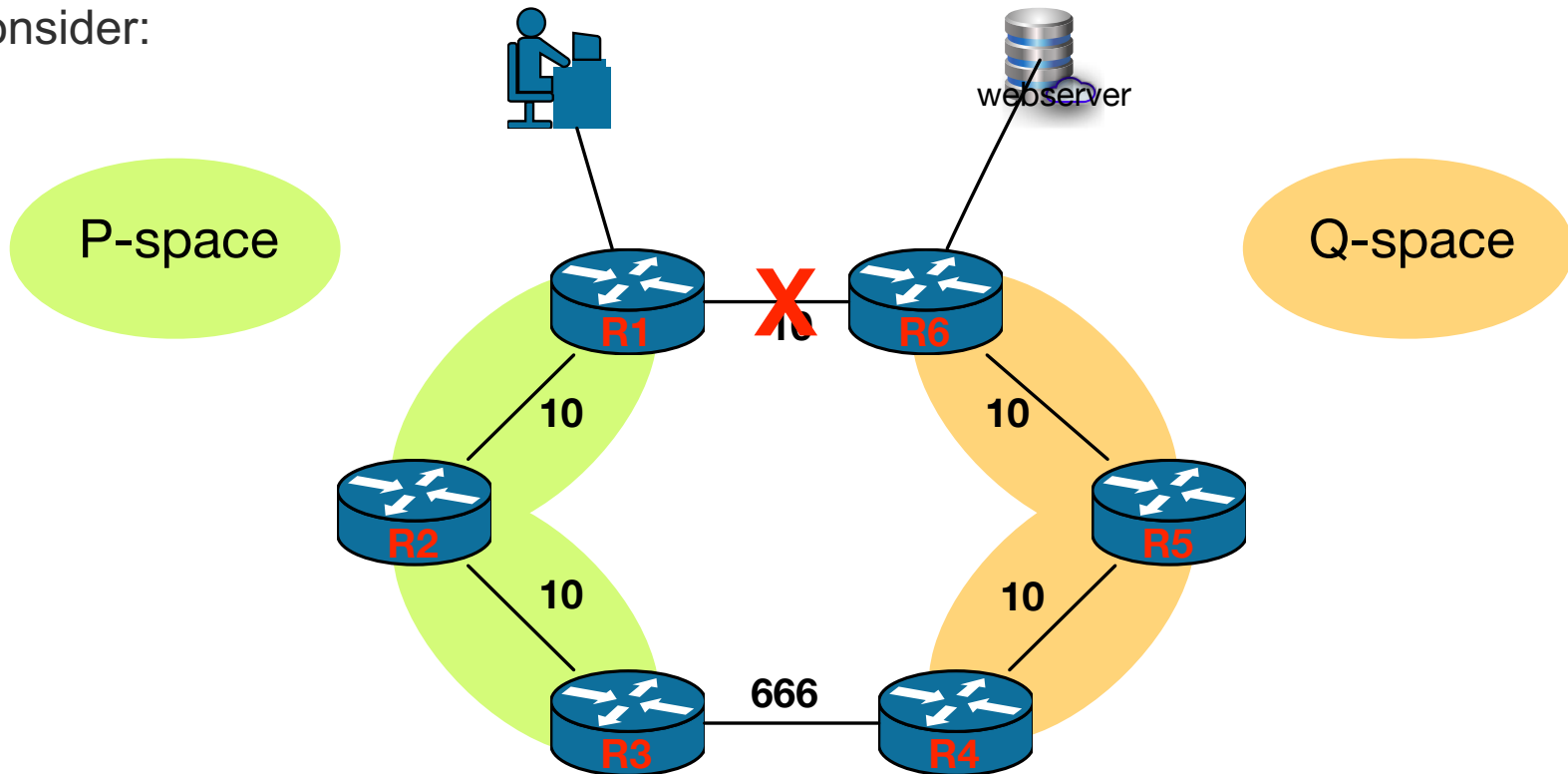
<b>FRR LFA</b>	Coverage can be bad where topology has no overlapping SPF from source / destination (PQ router space)
----------------	---

- SR introduces TI-LFA



# Nuisance: FRR LFA

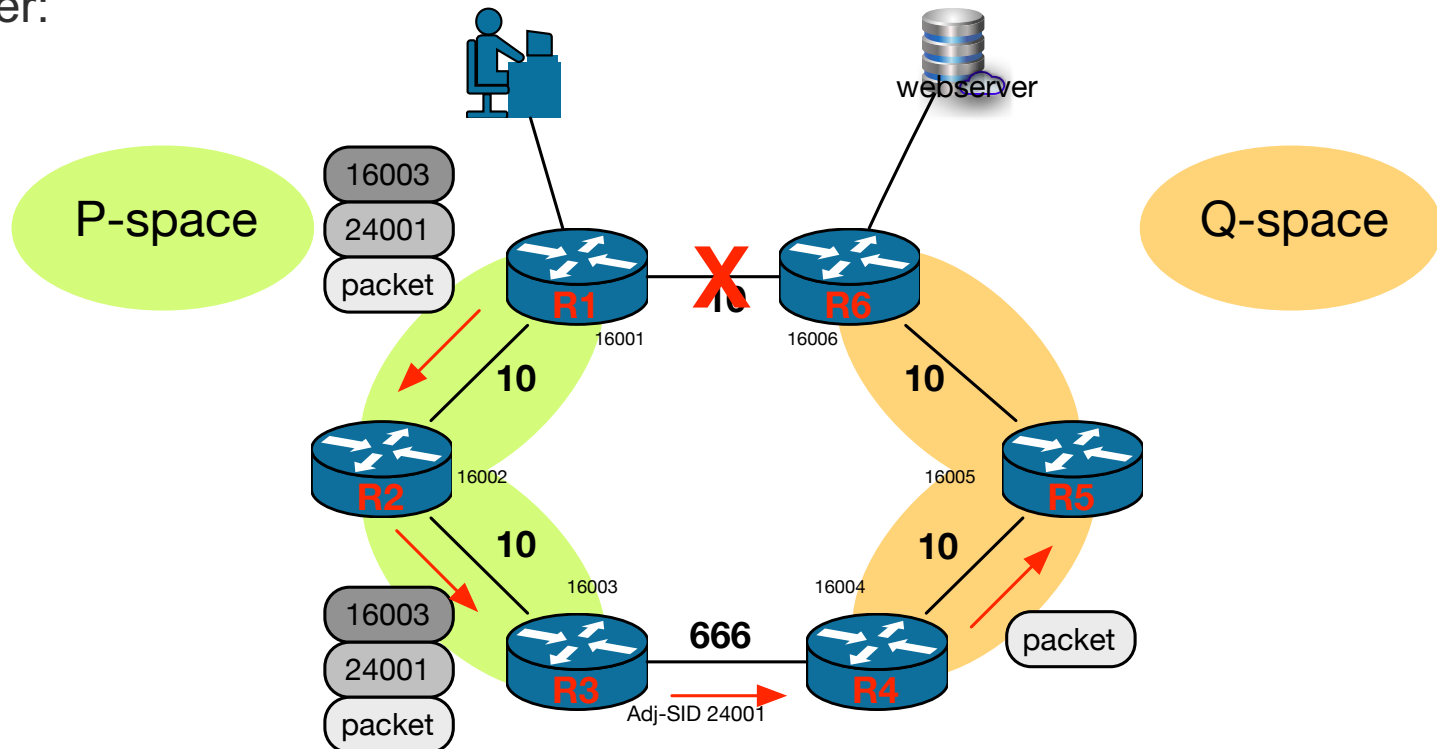
Consider:



- No shared PQ space, no FRR LFA

# TI-LFA

Consider:



- No shared PQ space? No problem. List of adj-sids provide missing bits where no PQ overlap

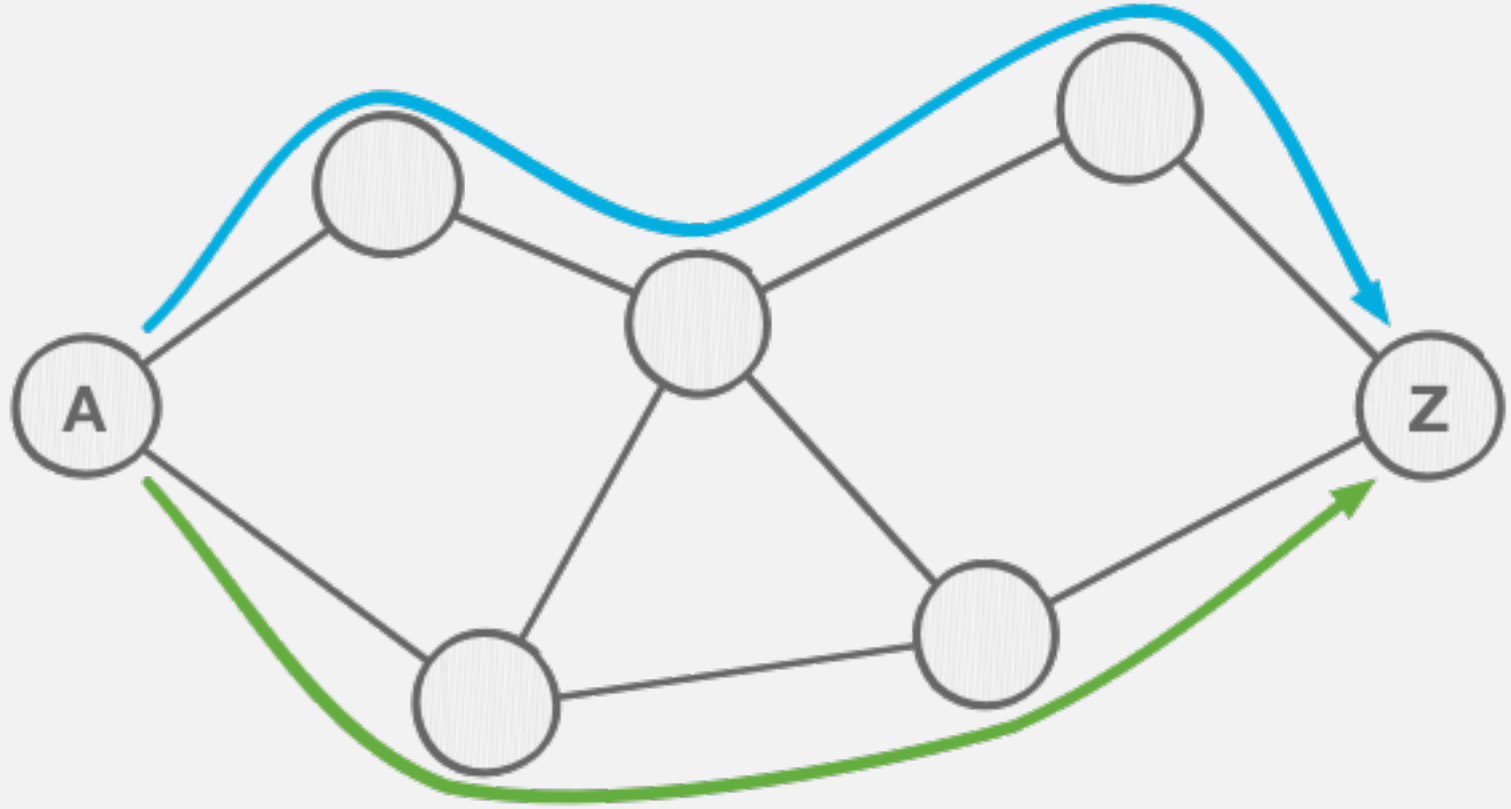
# Problem recap

Problem	Why
---------	-----

## u-Loops

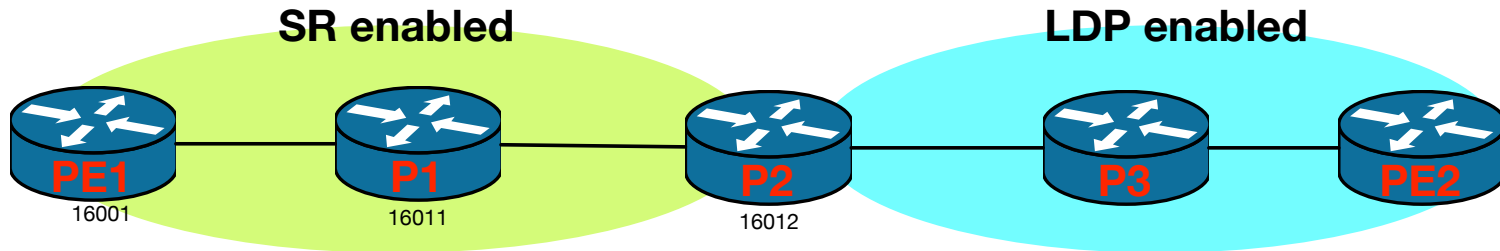
Bringing links into service can cause  $\mu$ -loops which are based on surrounding speed devices converge

- SR introduces micro-loop avoidance, with timer



How to we get there?

# Coexistence



- LDP -> SR
  - PE2 -> P2 as per standard LDP
  - P2 has no LDP binding for PE1 but sees PE1 via SR, so swaps label for SR segment
- SR -> LDP
  - SRMS programs 'remote-binding SID's' for LDP only routers
  - PE1 sees PE2 node SID just as if PE2 had sent it to him
  - P2 has no SR to PE2 so swaps for LDP label to reach PE2

# Feature to OS/Hardware map

	<b>IOS-XR (ASR9k, XRv)</b>	<b>IOS-XE (ASR1k, CSR1kv)</b>	<b>Arista EOS</b>	<b>IOS (6500/7200/GSR)</b>
<b>SR IPv4 Node-SID</b>	Pre: 5.3	From: 3.16S	Pre: 4.18	No support
<b>SR TI-LFA</b>	Pre: 5.3	From: 3.18S	Roadmap: 2019	No support
<b>SR Microloop Avoidance</b>	Pre: 5.3	From: 16.6.1	TBC	No support
<b>SR OAM</b>	Pre: 5.3	From: 3.17S	TBC	No support
<b>SR-DPM</b>	No Support	No Support	TBC	No Support
<b>PW prefer SR</b>	6.4.2	TBC	TBC	No Support
<b>SRMS</b>	Pre: 5.3	From: 3.18S (domain-wide flooding)	TBC	No Support

**Note:** do not actually **use** IOS-XR pre 6.4.2 for any SR-MPLS

# Implementation

- 1) Deploy SR mapping server (SRMS) configuration for all LDP only prefixes in the IGP.
- 2) Rollout IS-IS SR to all SR capable routers, leaving the default behaviour of preferring LDP over SR
- 3) Move all SR routers to prefer SR
- 4) Remove LDP from all SR-to-SR adjacent routers.
- 5) Remove LDP from all SR routers not attached to LDP only routers (watchout for LDP GR)

# Any questions?

