Fast Reroute Techniques in MPLS Networks

George Swallow
swallow@cisco.com

Agenda

- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

What are your Objectives?

- Limit traffic loss
  - for most traffic?
  - for all traffic?
- Quality of Service
  - Statistical guarantees?
  - Hard guarantees?
  - For some of the traffic?
  - For all of the traffic?
- Immunity only to single failures?
- What about cable cuts and power outages that take multiple links and nodes?

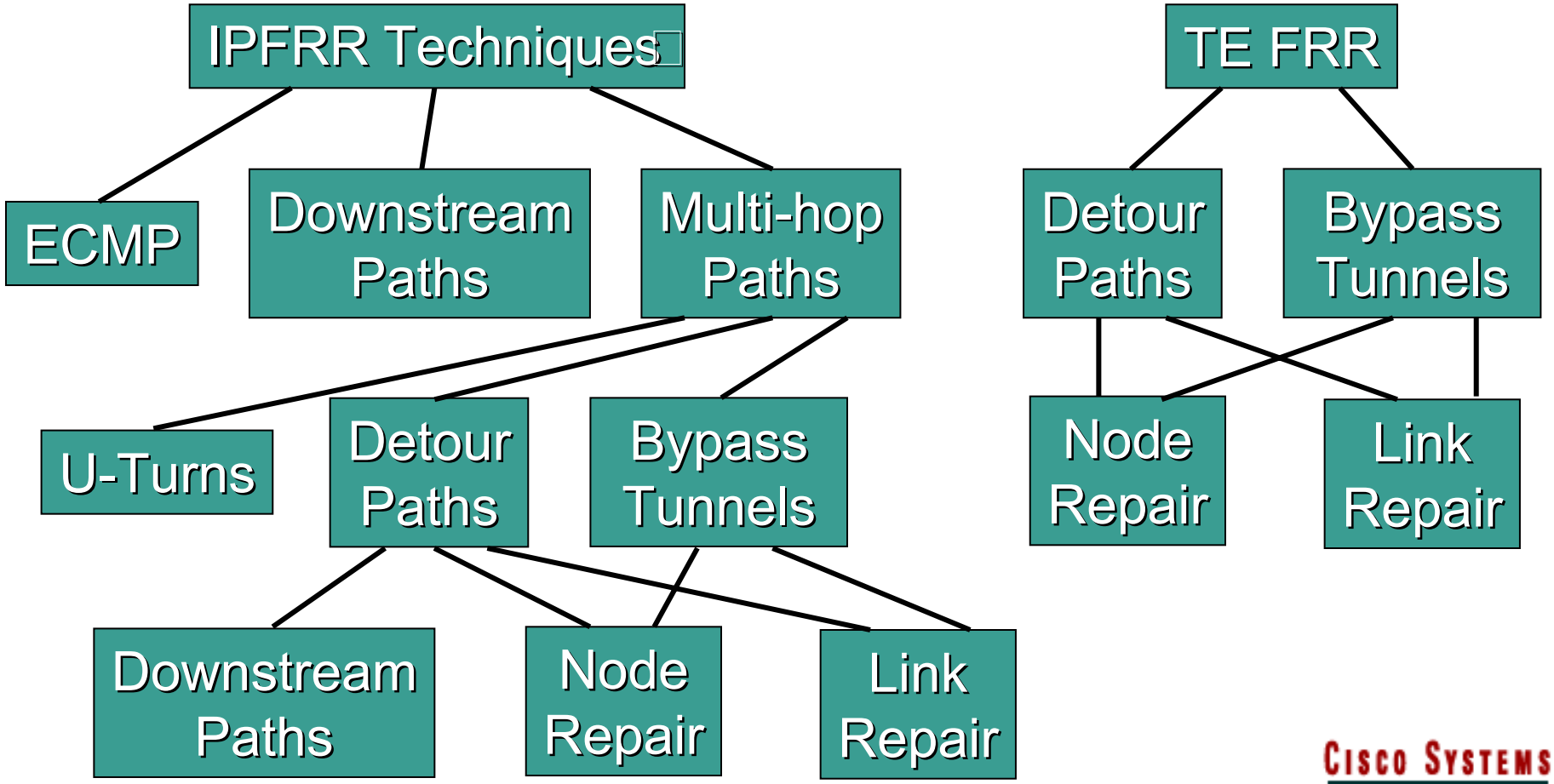
Agenda

- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

The Problem Space

- Detection time
- Distinguishing between Link & Node failures
- Repair techniques
- Repair time
  - On the fly
  - Pre-planned
  - Pre-installed
- Re-convergence issues, e.g. Micro-Loops
  - How you transition for the recovery state to the new normal
- Label distribution

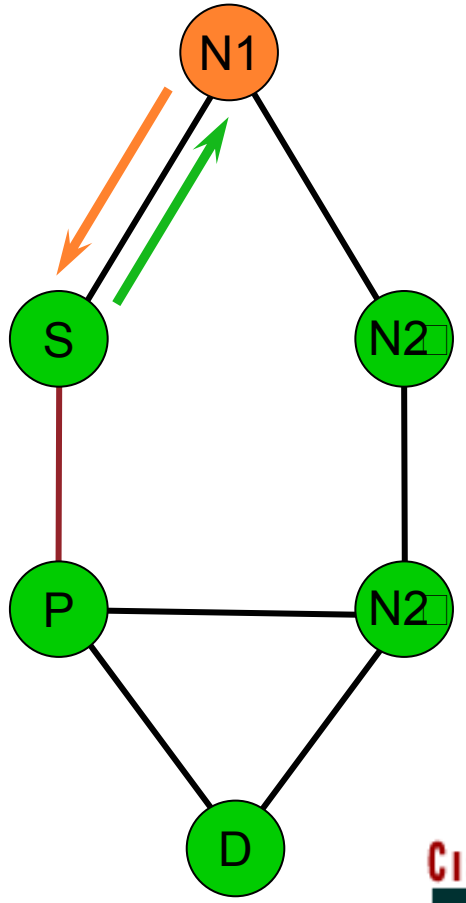
A (partial) Taxonomy



Cisco Systems

Micro-Loops

- Link-state protocols distribute updates to all nodes in an area.
- Nodes make the same routing decision about the same time.
- For a brief period, nodes can be out of sync.
- Inconsistent routing databases lead to µ-Loops



S converges before N1

New Route to D

Old Route to D

Applying IP FRR to MPLS

- Labels are local
- Tunneling techniques require label from tunnel tail
- Label exchange techniques
  - Next Hop Labels
    - Used in RSVP
    - Can be applied to LDP
  - Locally significant globally distributed
  - "Push label"
  - Directed LDP

A Broad Spectrum

We now focus on three points along the spectrum:

- U-turn
  - defines one end of the spectrum
- Traffic Engineering
  - defines the other
- TE FRR for LDP
  - one example of multihop tunnels

Agenda

- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

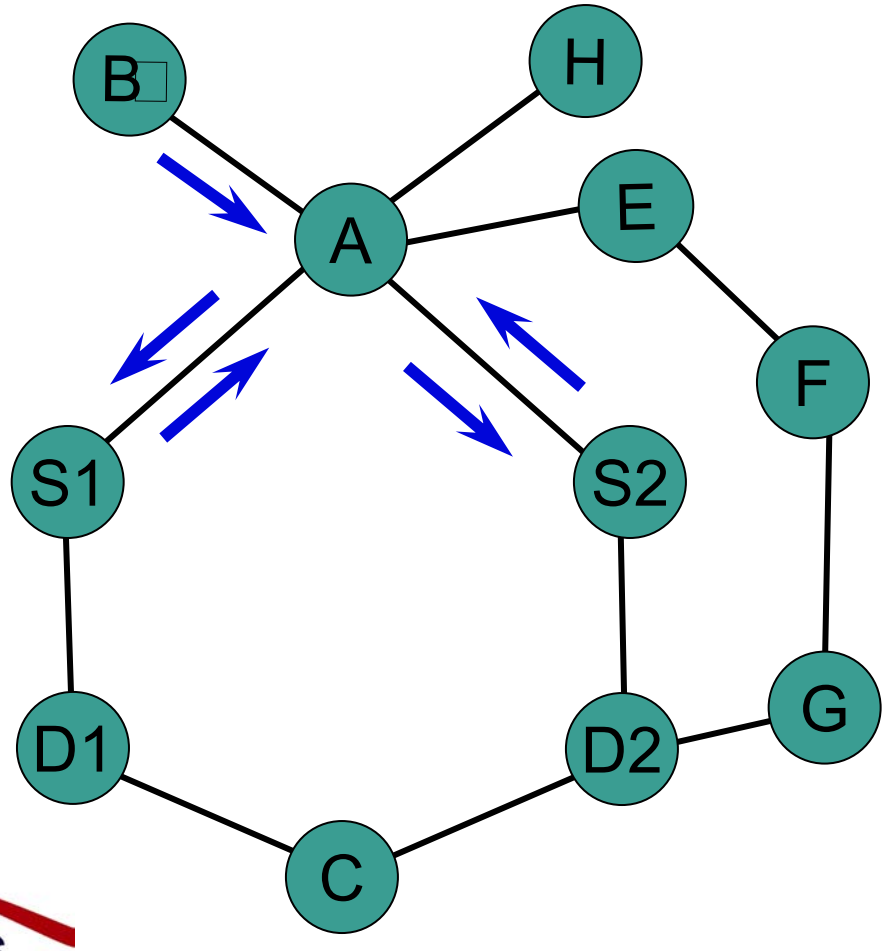
Some comments on U-Turn

- Interesting idea
- Doesn't require extra label exchanges
- For TE a second set of labels is required
- Doesn't cover all traffic
  - Claims high coverage
  - What about multiple failures?
- Changes data plane characteristics

Traffic Engineering & U-Turn

- Traffic from a TE tunnel can arrive on any interface, including interfaces that are downstream routes for some of the traffic
- If the normal LDP label is used for traffic on the tunnel, it will incorrectly trigger a U-turn
- Can be solved with a second set of labels
- But then you never reroute that traffic until it is received a second time

A multiple failure case



Fiber cut takes out links S1-D1, S2-D2

S1 and S2 U-turn traffic to C

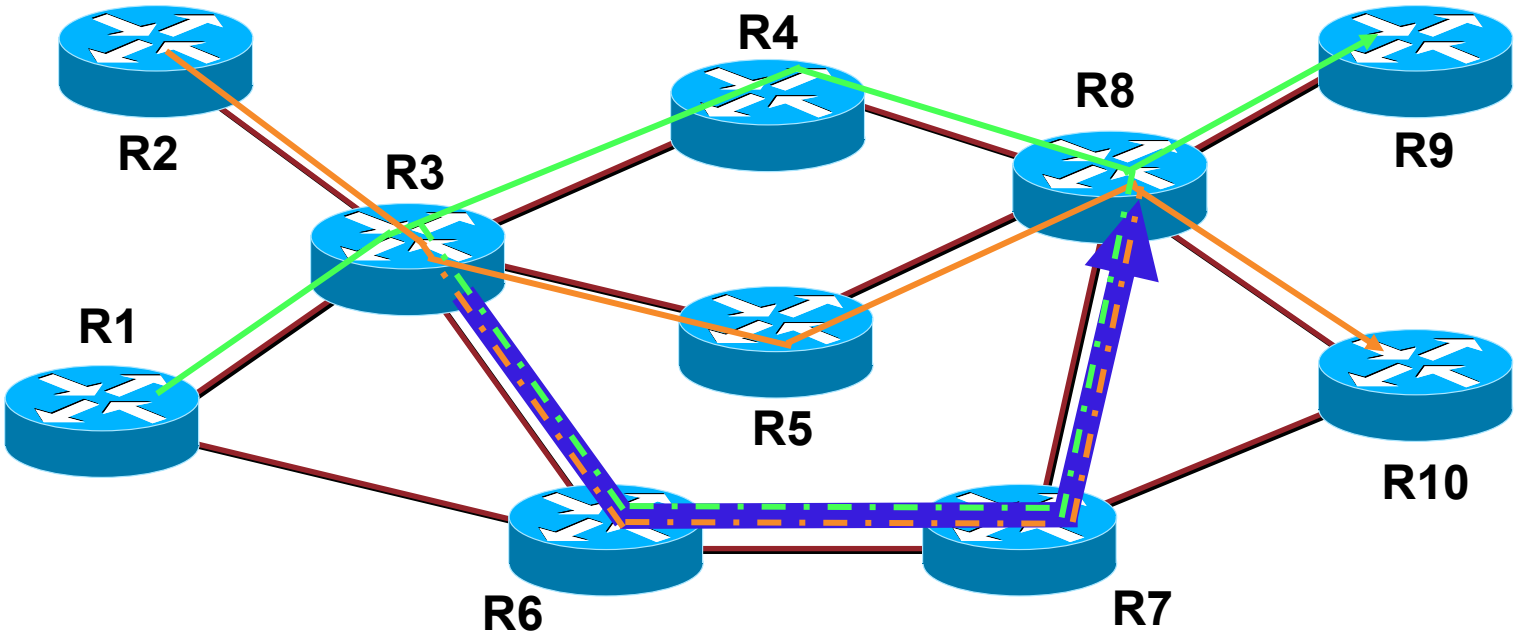
A has two equal cost paths

Packets to C Ping-Pong until TTL expires

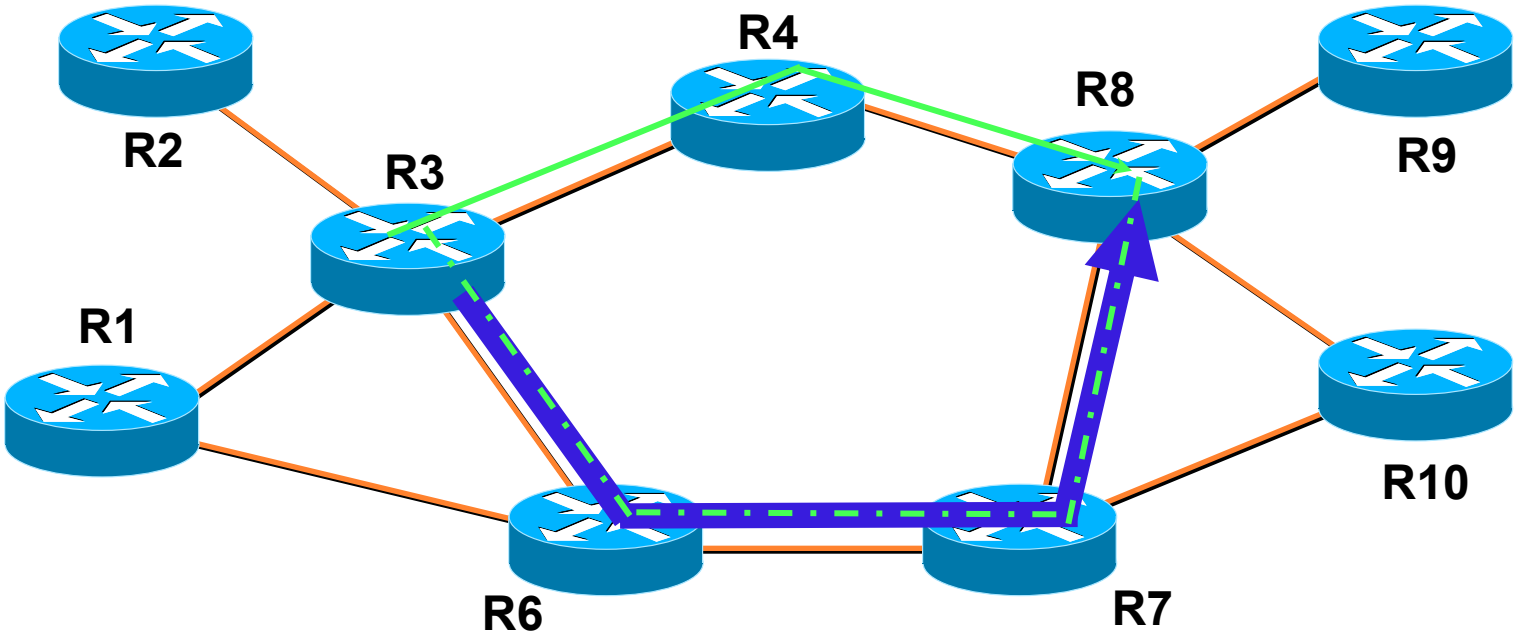
Agenda

- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

TE Fast Restoration



TE FRR for LDP



Bypass Tunnel

Routed Path

Backup Path

Bypass tunnel protects failure of R4 and link to R4

R3 needs R8's labels

For link protection no extra labels are needed

Agenda

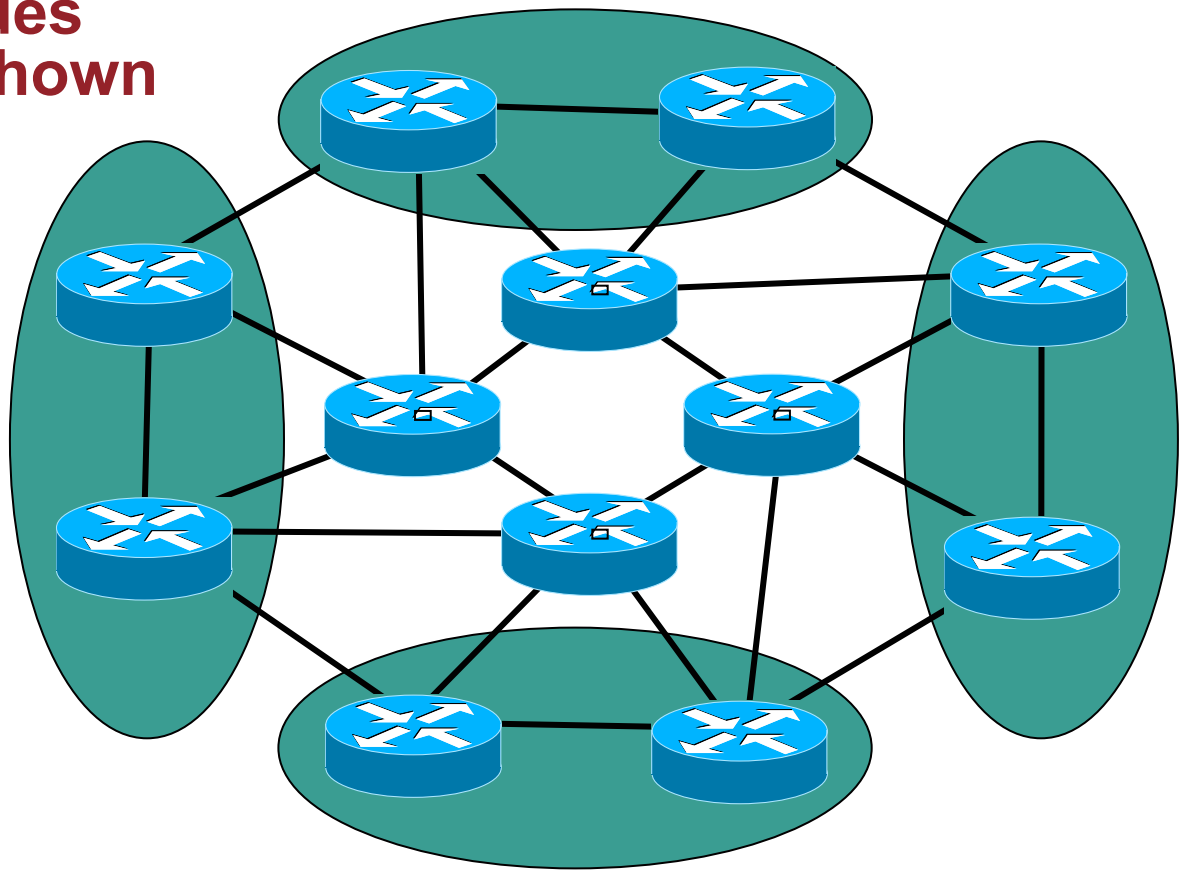
- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

Traffic Engineering

- TE Fast Reroute
- TE FRR for LDP
  - Multi-hop bypass tunnel for IP FRR
- Characteristics of TE FRR
  - 100% Coverage
  - Immunity to µ-Loops
  - Handles SRG Diversity
  - Can offer bandwidth guarantees
  - Doesn't need to be difficult to configure

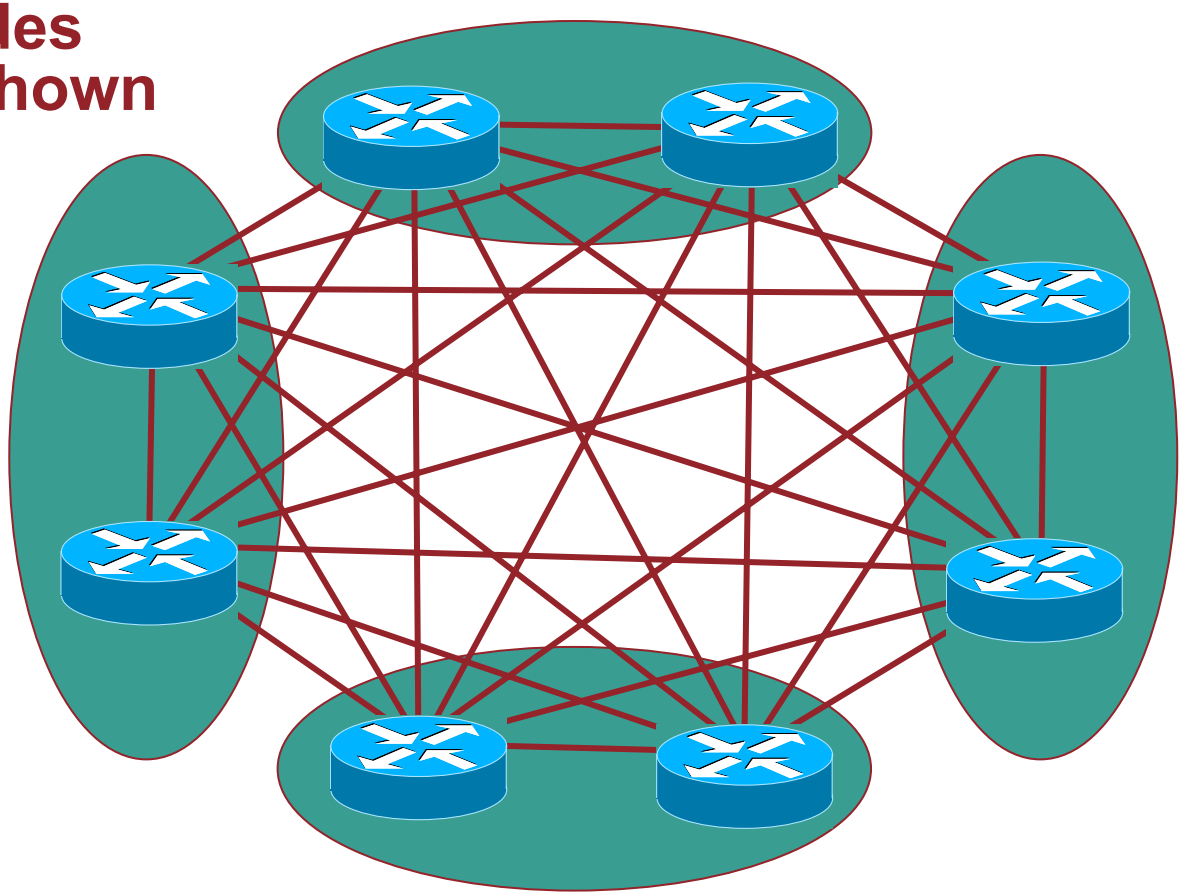
Pop to Pop Physical Topology

Note: Access nodes not shown



PoP to PoP Mesh

Note: Access nodes not shown



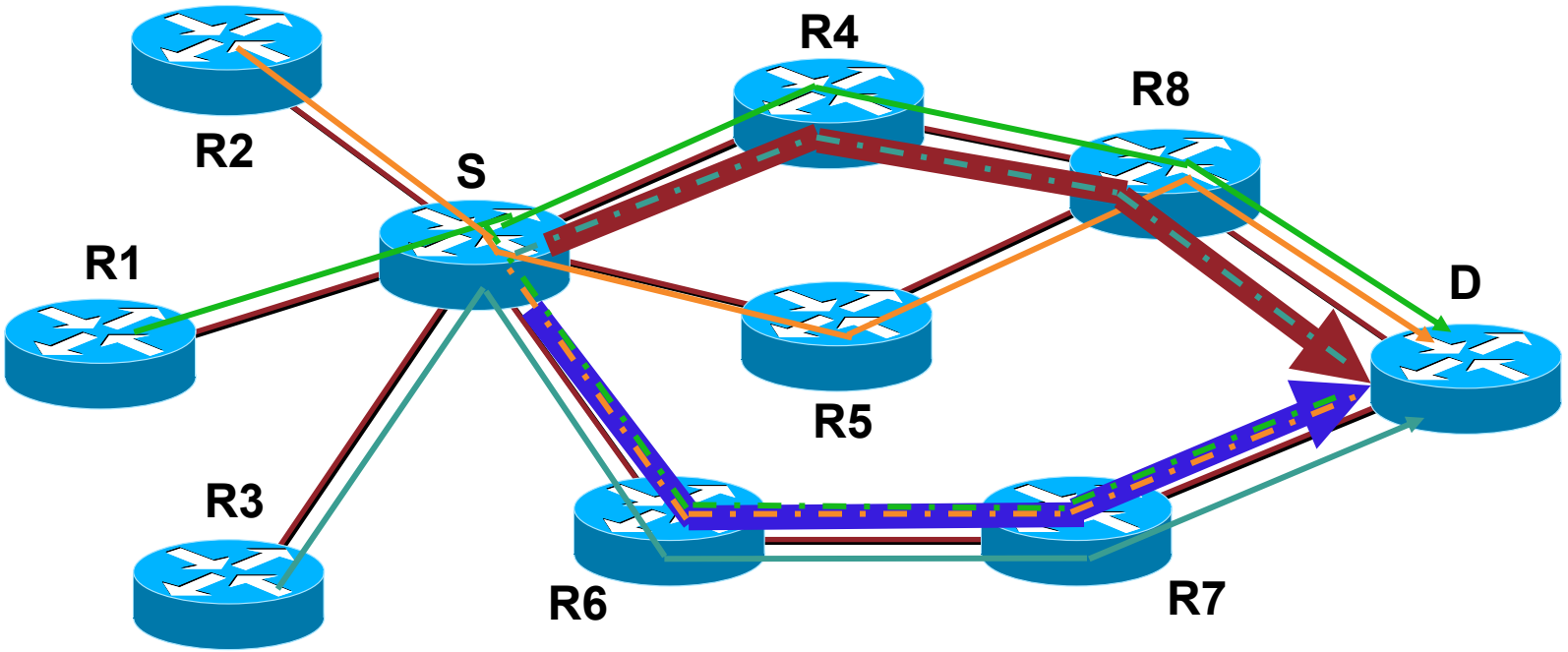
Immunity to Micro Loops

- A node will use a TE Tunnel if the Tail is on the shortest path to the to the destination
- If multiple such tunnels exist it will use the one that is closest to the destination
- Tunnel selection is affected only by topology changes close to the Tail
- If the node has many TE Tunnels then the traffic is likely to still be handled by that node

Automatic Deterministic Connectivity

- Auto-mesh establishes TE tunnels between all nodes in a given set
  - Including core nodes
  - Using only nodes in the mesh
- Extend Auto-mesh to establish two tunnels along SRG independent routes to every other node
- Load share traffic across the two tunnels
- These same tunnels are used for FRR
  - Failed tunnels are not repaired;
  - Traffic is simply spliced into a tunnel that is SRG diverse of the failed link or node

SRG Diverse FRR



SRG diverse Tunnels
S to D

Protected Tunnels

Repair Paths

Agenda

- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

Other things to look for

- Much of the current discussion focuses on the transition from the failure to a temporary recovery state
- Some of the schemes result in prolonging the period between normal states
- Some delay informing all nodes of outage
- Coupling of data plane to control plane
  - Are there valid reasons to come in on the "wrong" interface

Further caveats

- Examine cases where algorithms break down
- Some depend on symmetric metrics
  - In examining many networks I've found asymmetric metrics - usually because of misconfigurations
- Data plane debugging can be complicated - gets worse if you introduce more exceptions that only get invoked some of the time
- Many of the coverage claims don't deal with SRG

Agenda

- What are your requirements?
- The solution space
- U-turns
- Traffic Engineering for LDP
- Traffic Engineering
- Some Observations
- Summary

Summary

- There's a broad spectrum of options, some of which will come to market
- U-turn and Traffic Engineering / FRR mark the ends of that spectrum
- Know what what you're trying to achieve for backup
- Think about real world failures
- Make sure µ-Loops are addressed
- Traffic Engineering is here today