

Tutorial on MPLS + TE + VPN

**Based on MPLS tutorial from
Tim Griffin
and
MPLS/VPN tutorial from
Chris Chase**

1

What's all this talk about MPLS?

- **“MPLS is going to solve all of our problems”**
- **“MPLS is a solution in search of a problem”**
- **“MPLS is all about traffic engineering”**
- **“MPLS is what I wish on all of my competitors”**
- **“MPLS is all about virtual private networks”**
- **“MPLS solves network operations problems”**
- **“MPLS creates network operations problems”**
- **“MPLS is all about lowering operational costs”**
- **“MPLS is going to cost more than its worth”**
- **“MPLS is the natural next step in Internet evolution”**
- **“MPLS is too complicated to survive in the Internet”**

But what is MPLS anyway?

2

Goals of this Tutorial

- **To understand MPLS from a purely technical point of view**
 - avoid the hype
 - avoid the cynicism
- **To understand the broad technical issues without getting lost in the vast number of details**
 - the **gains**
 - the **costs**
 - the **tradeoffs**

3

Keep in Mind

- **MPLS is an emerging technology**
- **Many technical issues have not yet been resolved**
- **Interest and enthusiasm is not universal, but primarily found in large providers (and their vendors)**
- **Standards are rapidly evolving**
- **Implementations are rapidly evolving**
- **Operational experience and expertise still very scarce**

Expect interoperability problems and feature availability problems for the next few years

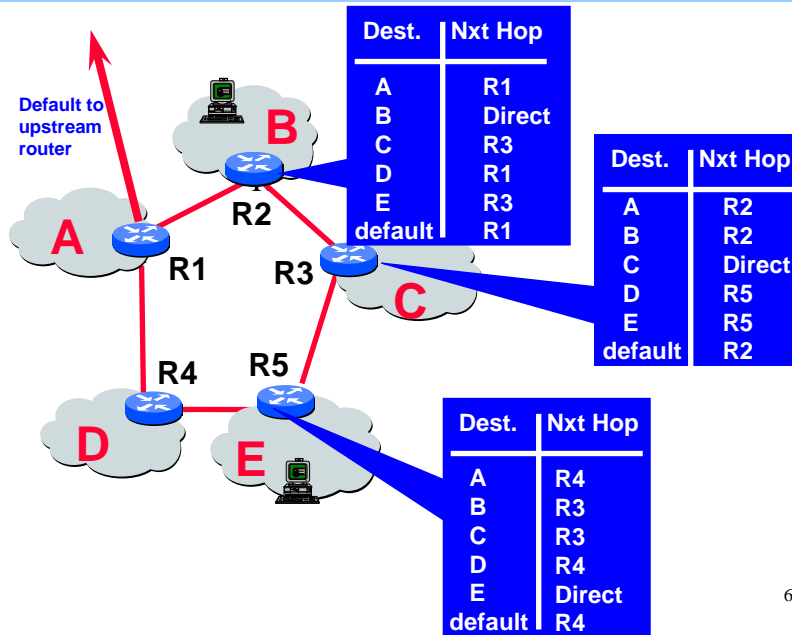
4

Outline

- **Why MPLS?**
 - Problems with current IP routing and forwarding
 - Complexity of overlay model
- **What is MPLS?**
 - Label swapping
 - Label distribution
 - Constraint based routing
- **What applications could exploit MPLS?**
 - Traffic Engineering
 - Virtual Private Networks
 - Both Layer 2 and Layer 3 VPNs

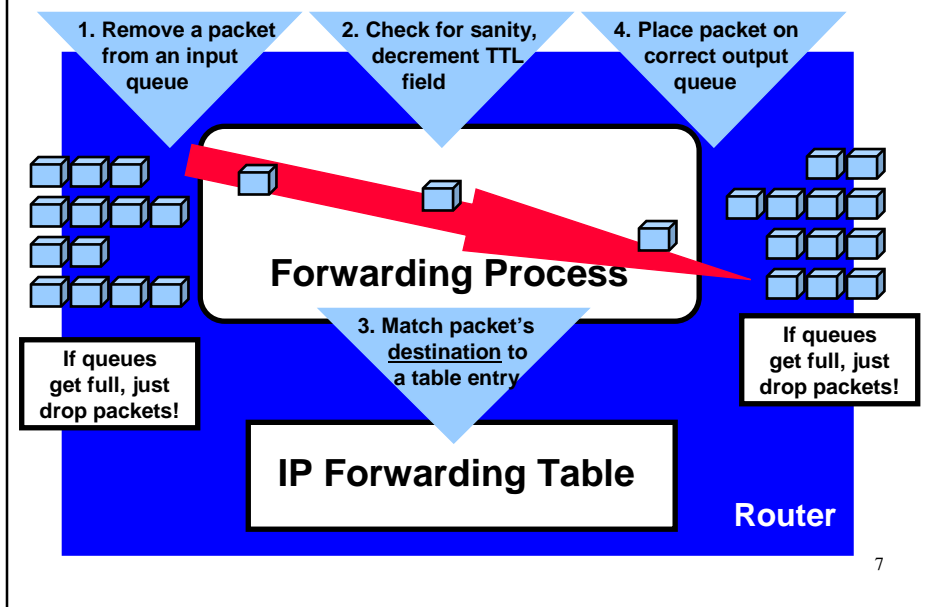
5

IP forwarding paths are implemented with destination-based next hop tables

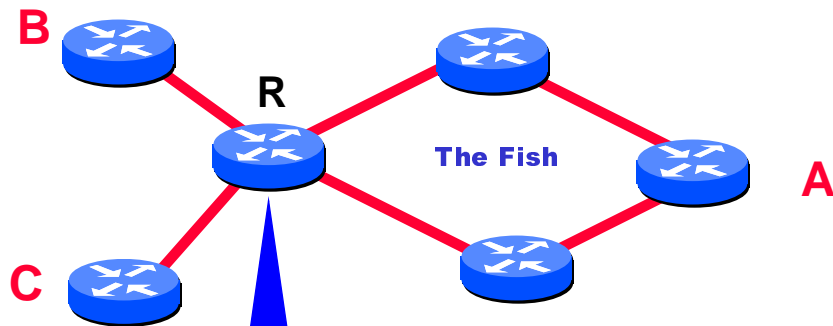


6

IP Forwarding Process

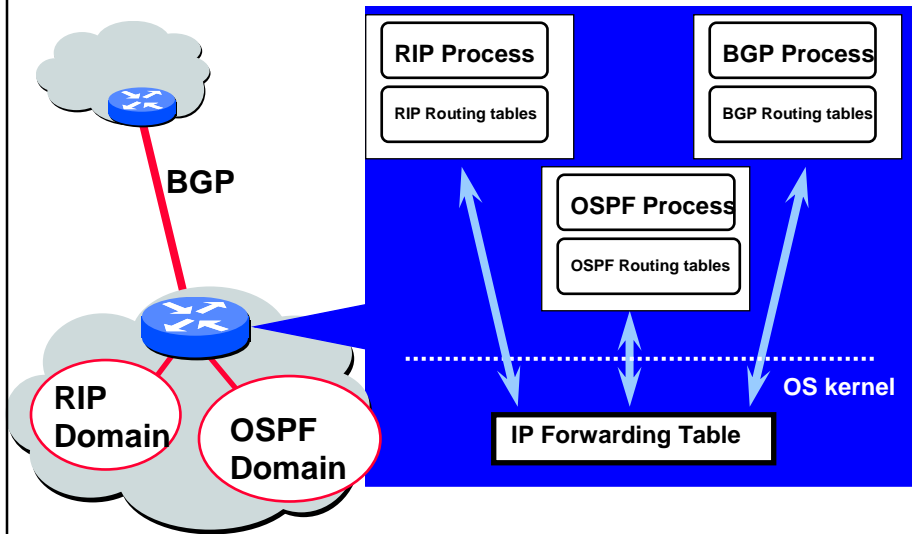


IP routing protocols assume all forwarding is destination-based



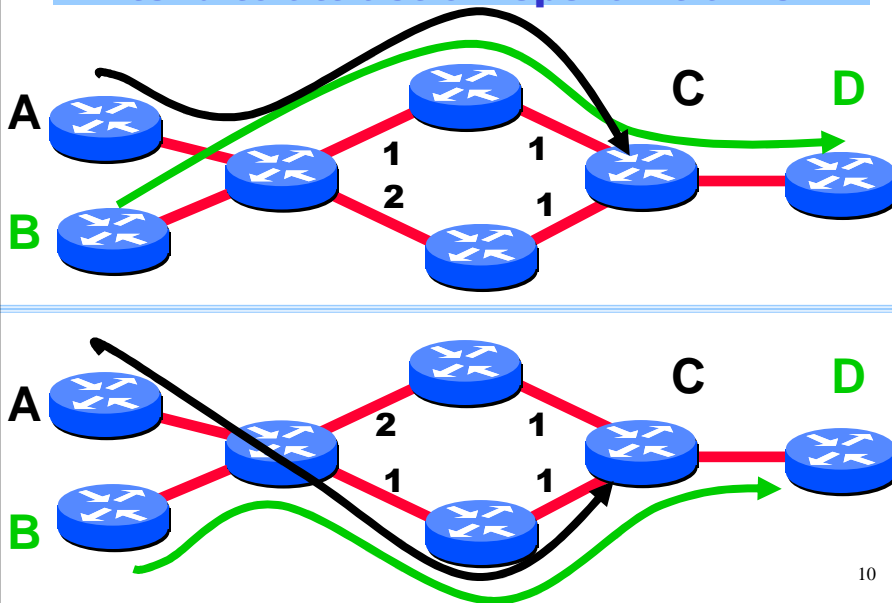
The next-hop forwarding paradigm does not allow router R to choose a route to A based on who originated the traffic, B or C.

IP forwarding tables are maintained by dynamic routing protocols



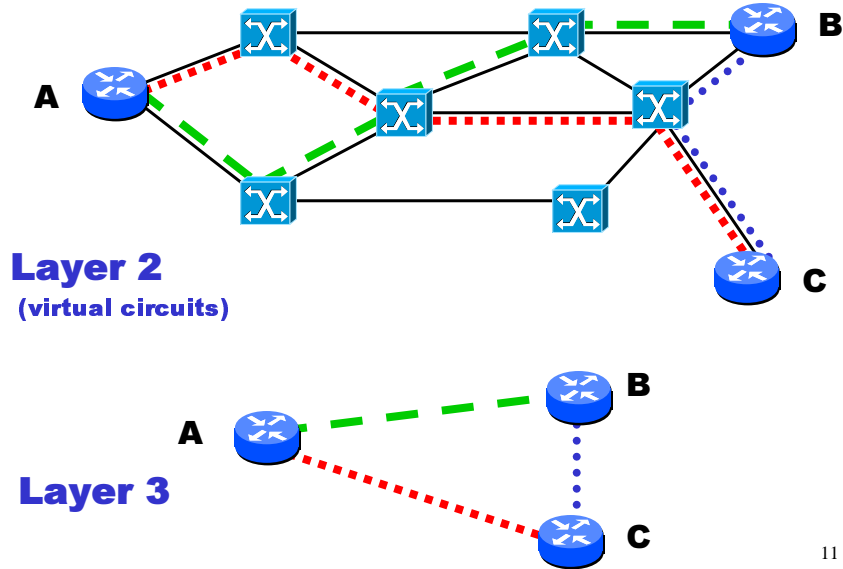
9

Shortest Path Routing: Link weights tend to attract or repel all traffic



10

Overlay Networks



11

Advantages of Overlay Networks

- **ATM and Frame Relay switches offer high reliability and low cost**
- **Virtual circuits can be reengineered without changing the layer 3 network**
- **Large degree of control over traffic**
- **Detailed per-circuit statistics**
- **Isolates layer 2 network management from the details of higher layer services**

12

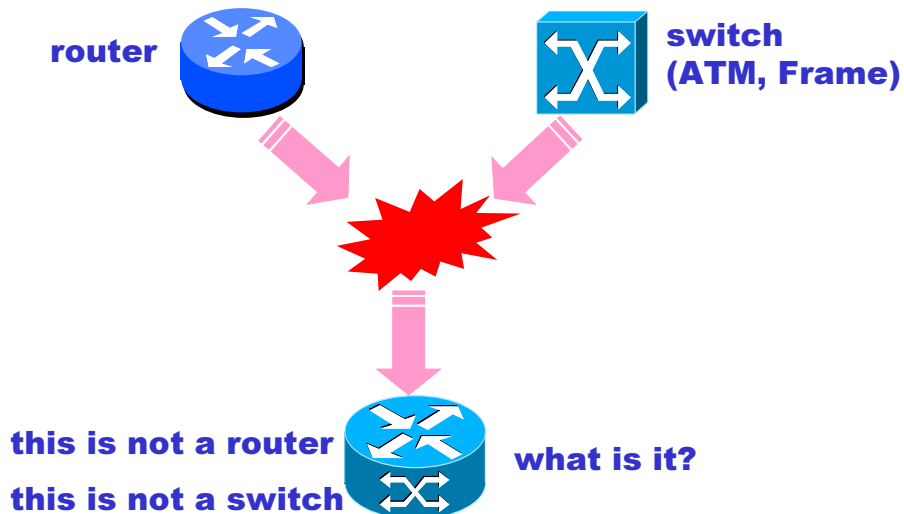
Problems with Overlay Networks

- Often use proprietary protocols and management tools
- Often requires “full meshing” of statically provisioned virtual circuits
- ATM cell tax ---- about 20% of bandwidth
- If layer 3 is all IP, then the overlay model seems overly complicated and costly
- Advances in optical networking cast some doubt on the entire approach

Overlay model is just fine when layer 2 network provides diverse non IP services (e.g., IPv6, AppleTalk, IPX, ...)

13

Blur Layer 2 and 3?



14

Sanity Check?

- **The problems with IP forwarding and routing do not require technologies like MPLS**
 - Many can be addressed with simple solutions. Like the design of simple networks!
 - The problems are not “show stoppers”
 - The MPLS cure will have side effects
 - For many applications, TCP/IP handles congestion very well
- **Technologies like MPLS may be very valuable if they can enable new services and generate new revenue**

15

Sign of the Times: New Sub-IP area in the IETF

- **Multiprotocol Label Switching (mpls)**
- **Common Control and Management Protocols (ccamp)**
- **Internet Traffic Engineering (tewg)**
- **Provider Provisioned Virtual Private Networks (ppvpn)**
- **IP over Optics (ipo)**
- **IP over Resilient Packet Rings (iporpr)**
- **General Switch Management Protocol (gsmp)**

See <http://www.ietf.org/html.charters/foo-charter.html>,
where `foo` is the working group acronym.

16

MPLS = MultiProtocol Label Switching

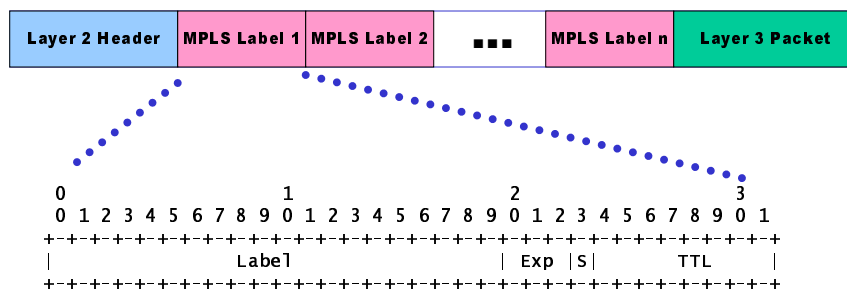


- A “Layer 2.5” tunneling protocol
- Based on ATM-like notion of “label swapping”
- A simple way of labeling each network layer packet
- Independent of Link Layer
- Independent of Network Layer
- Used to set up “Label-switched paths” (LSP), similar to ATM PVCs

RFC 3031 : Multiprotocol Label Switching Architecture

17

Generic MPLS Encapsulation



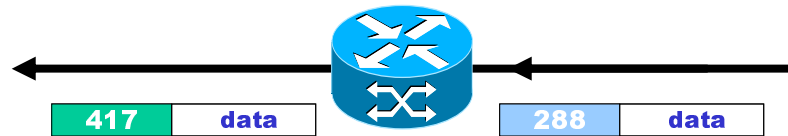
Often called a “shim”
(or “sham”) header

RFC 3032. MPLS
Label Stack Encoding

- **Label:** Label Value, 20 bits
- **Exp:** Experimental, 3 bits
- **S:** Bottom of Stack, 1 bit
- **TTL:** Time to Live, 8 bits

18

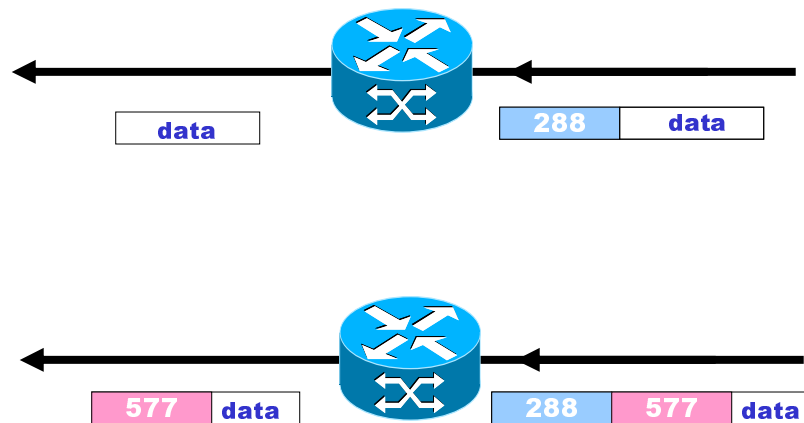
Forwarding via Label Swapping



Labels are short, fixed-length values.

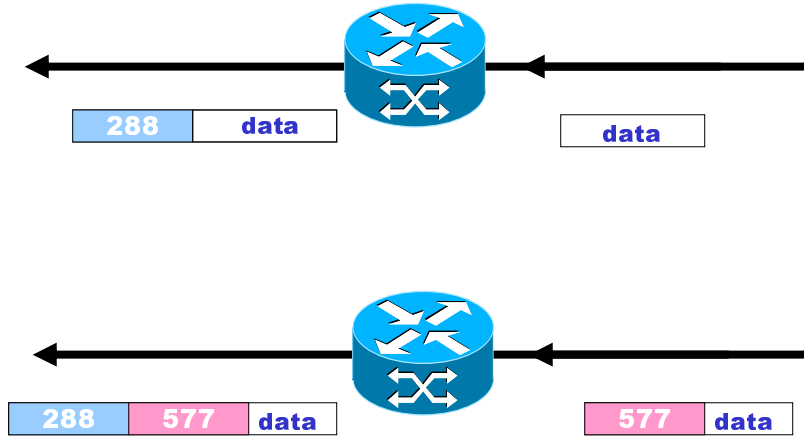
19

Popping Labels



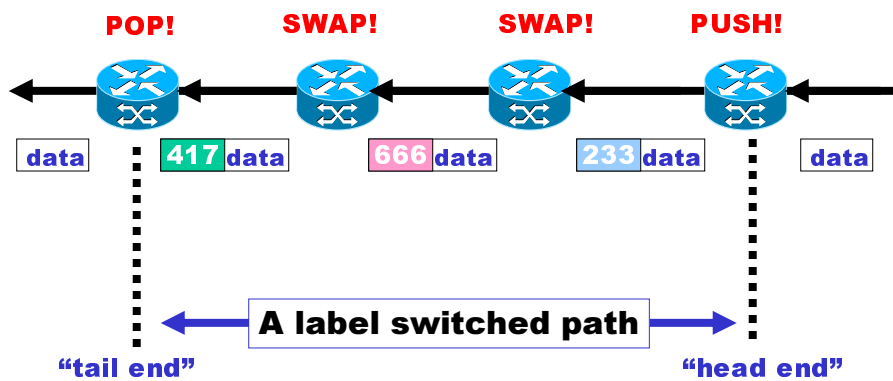
20

Pushing Labels



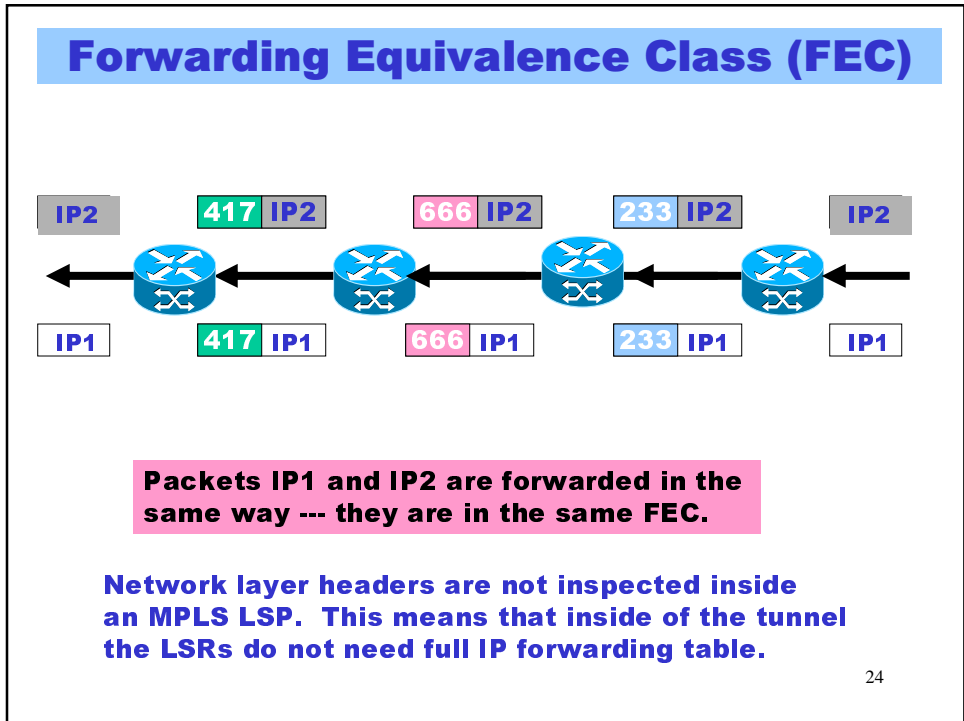
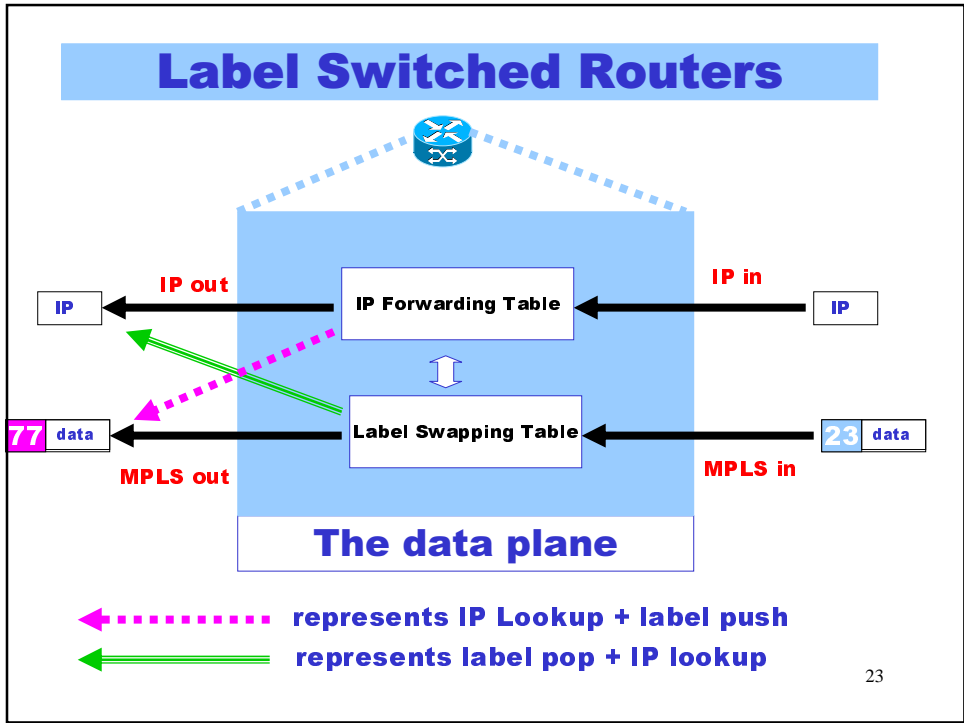
21

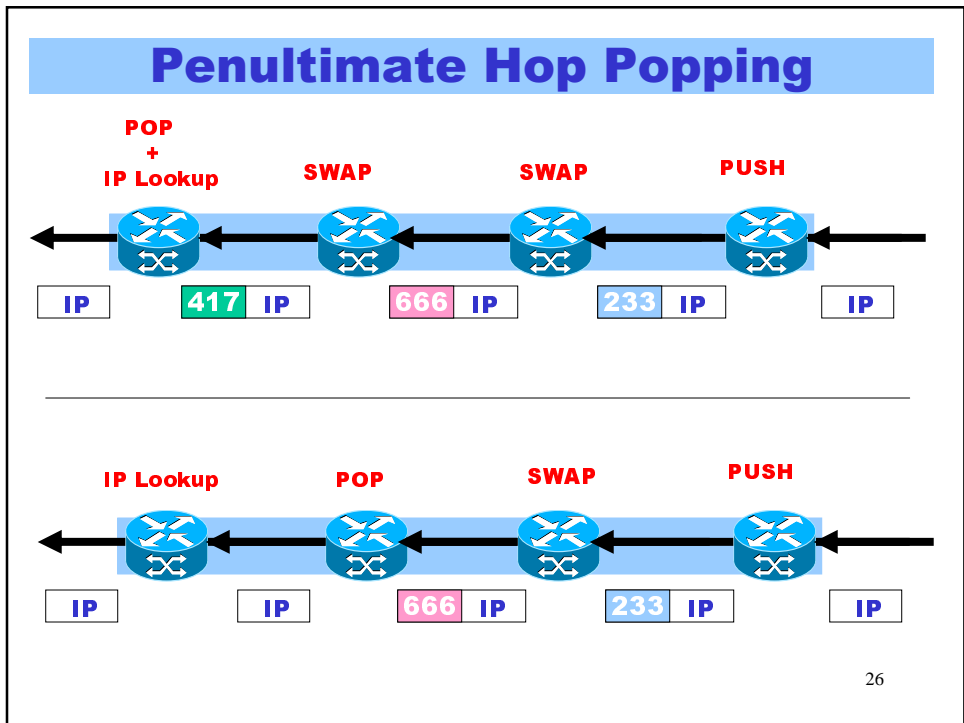
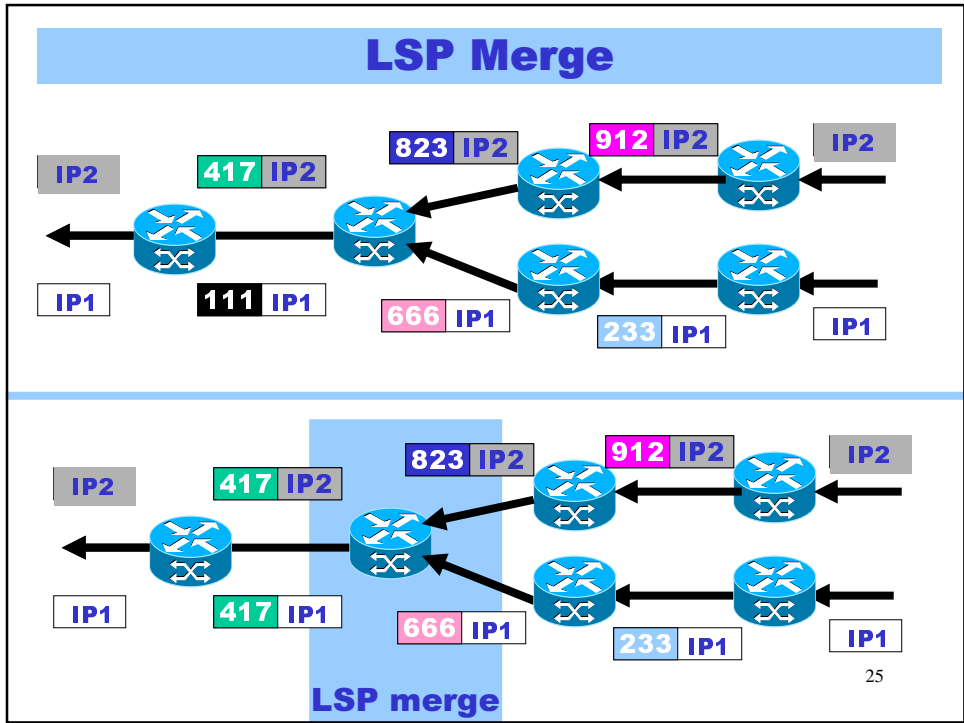
A Label Switched Path (LSP)

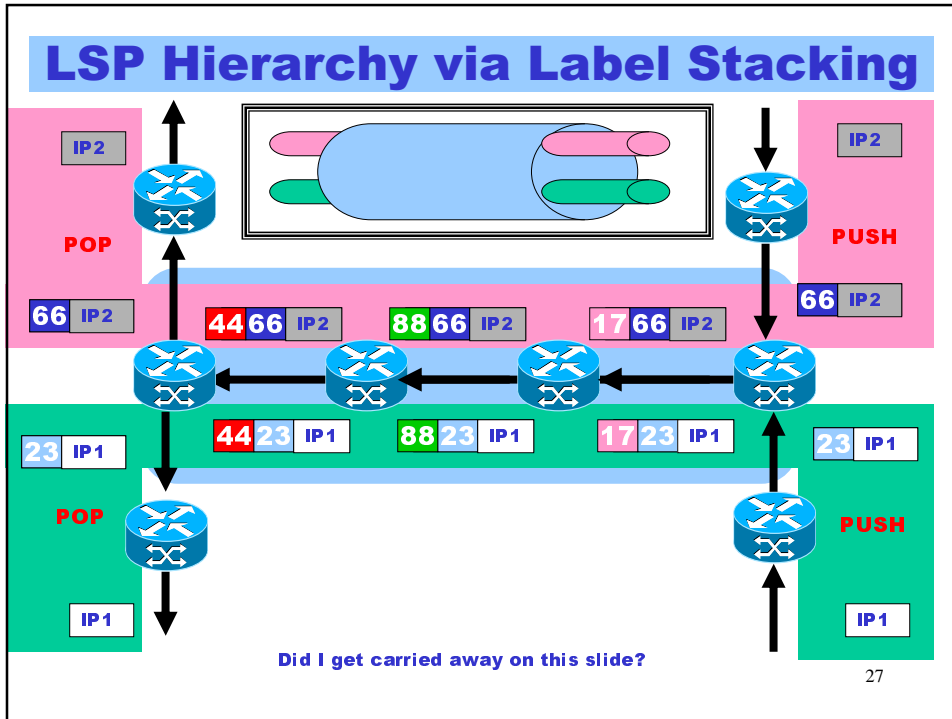


Often called an MPLS tunnel: payload headers are not inspected inside of an LSP. Payload could be MPLS ...

22







27

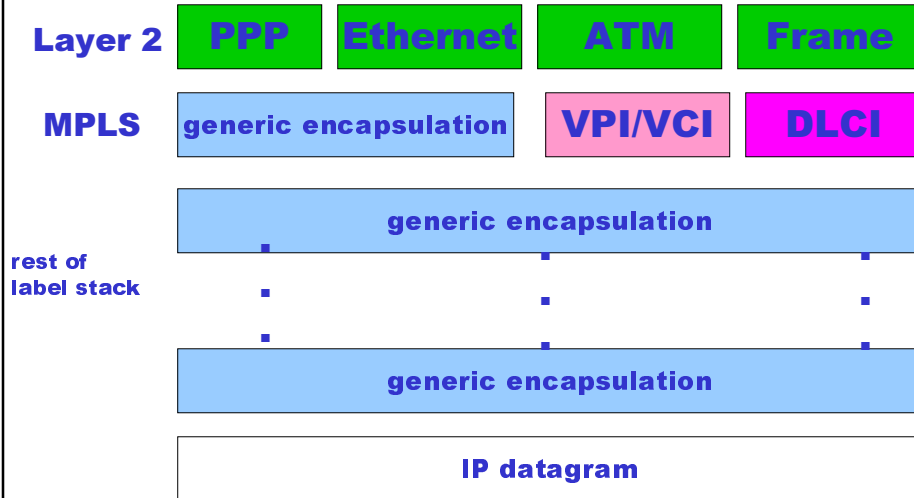
MPLS Tunnels come at a cost

- **ICMP messages may be generated in the middle of a tunnel, but the source address of the “bad packet” may not be in the IP forwarding table of the LSR!**
 - **TTL expired:** traceroute depends on this!
 - **MTU exceeded:** Path MTU Discovery (RFC1191) depends on this!

None of the proposed solutions are without their own problems...

28

MPLS also supports “native encapsulation”



29

But Native Labels May Cause Big Headaches

- **No TTL!**
 - Loop detection?
 - Loop prevention?
- **LSP merge may not be supported**
 - Label bindings cannot flow from destination to source, but must be requested at source

MPLS was initially designed to exploit the existence of ATM hardware and reduce the complexity of overlay networks. But IP/MPLS with native ATM labels results in a large number of problems and complications.

30

Basic MPLS Control Plane

**MPLS control plane
=
IP control plane
+
*label distribution***

Label distribution protocols are needed to
(1) create label↔FEC bindings
(2) distribute bindings to neighbors,
(3) maintain consistent label swapping tables

31

Label Distribution: Option I

“Piggyback” label information on top of existing IP routing protocol

Good Points

- **Guarantees consistency of IP forwarding tables and MPLS label swapping tables**
- **No “new” protocol required**

Bad Points

- **Allows only traditional destination-based, hop-by-hop forwarding paths**
- **Some IP routing protocols are not suitable**
 - **Need explicit binding of label to FEC**
 - **Link state protocols (OSPF, ISIS) are implicit, and so are not good piggyback candidates**
 - **Distance vector (RIP) and path vector (BGP) are good candidates. Example: BGP+**

32

Label Distribution: Option II

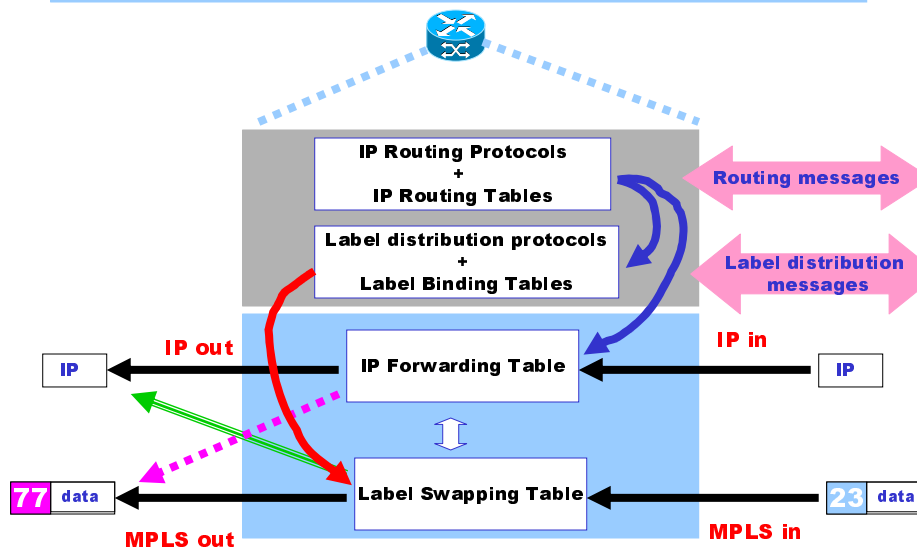
Create new label distribution protocol(s)

- Good Points**
 - Compatible with “Link State” routing protocols
 - Not limited to destination-based, hop-by-hop forwarding paths
- Bad Points**
 - Additional complexity of new protocol and interactions with existing protocols
 - Transient inconsistencies between IP forwarding tables and MPLS label swapping tables

Examples: LDP (IETF) and TDP (Cisco proprietary)

33

The Control Plane



34

Label Distribution with BGP

Carrying Label Information in BGP-4 draft-ietf-mpls-bgp4-05.txt (1/2001)

Associates a label (or label stack)
with the BGP next hop.

Uses multiprotocol features of BGP:

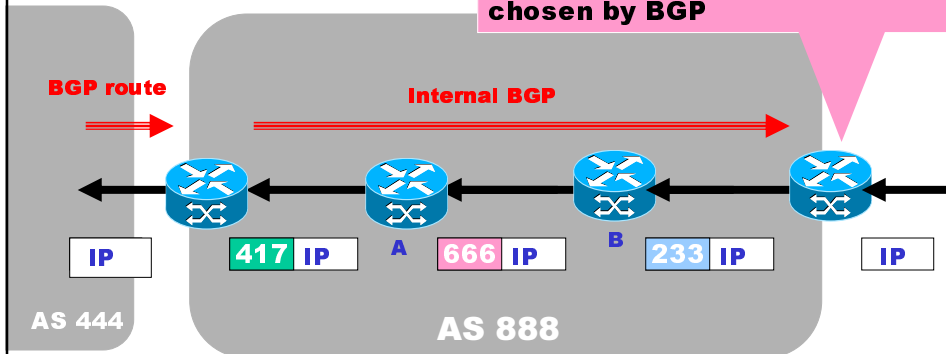
RFC 2283. Multiprotocol Extensions for BGP-4

So routes *with* labels are in a different
address space than a vanilla routes (no labels)

35

BGP piggyback not required for simple iBGP optimization

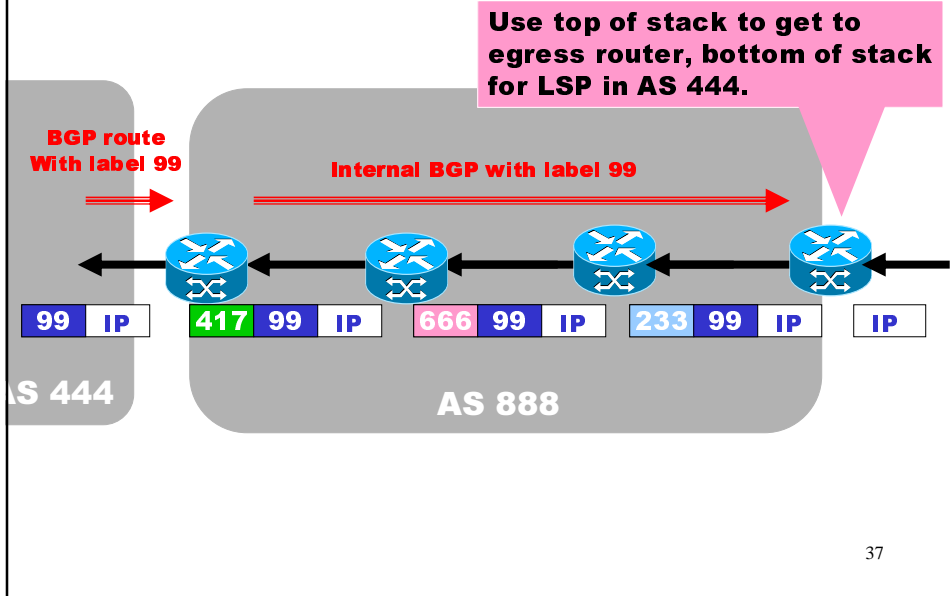
Map traffic to the LSP that
terminates at the egress router
chosen by BGP



Routers A and B do not need full routing tables.
They only need IGP routes (and label bindings).

36

BGP piggyback allows Interdomain LSPs



MPLS tunnels can decrease size of core routing state

- Core routers need only IGP routes and LSPs for IGP routes
- Implies less route oscillation
- Implies less memory
- Implies less CPU usage

Are these *really* problems?

BUT: still need route reflectors to avoid full mesh and/or to reduce BGP table size at border routers

BUT: since your core routers do not have full tables you now have all of the MPLS problems associated with ICMP source unknown (TTL, MTU, traceroute ...)

38

Label Distribution Protocol (LDP)

RFC 3036. LDP Specification. (1/2001)

- Dynamic distribution of label binding information
- **Supports only vanilla IP hop-by-hop paths**
- LSR discovery
- Reliable transport with TCP
- Incremental maintenance of label swapping tables (only deltas are exchanged)
- Designed to be extensible with Type-Length-Value (TLV) coding of messages
- Modes of behavior that are negotiated during session initialization
 - Label retention (liberal or conservative)
 - LSP control (ordered or independent)
 - Label assignment (unsolicited or on-demand)

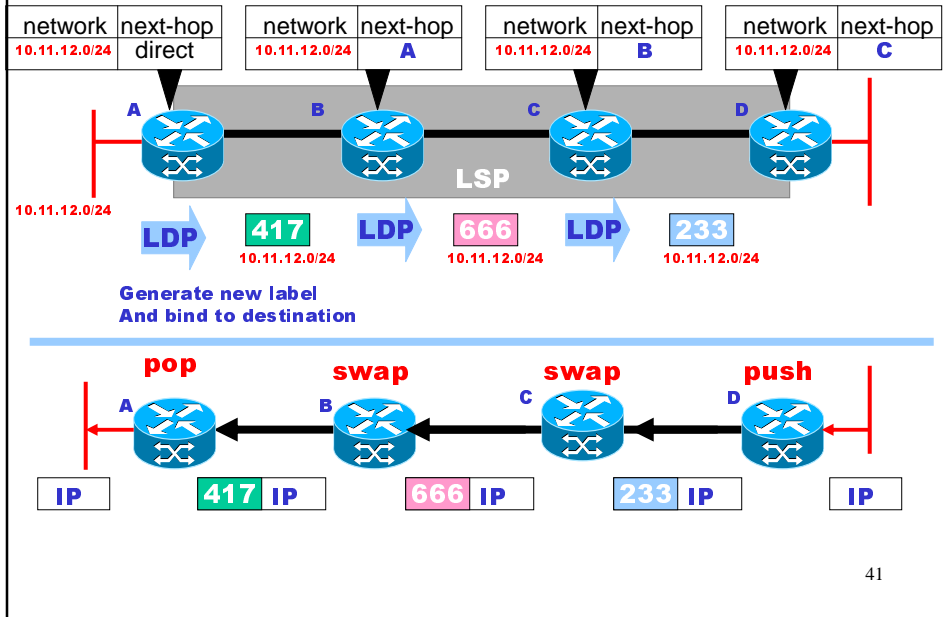
39

LDP Message Categories

- **Discovery messages:** used to announce and maintain the presence of an LSR in a network.
- **Session messages:** used to establish, maintain, and terminate sessions between LDP peers.
- **Advertisement messages:** used to create, change, and delete label mappings for FECs.
- **Notification messages:** used to provide advisory information and to signal error information.

40

LDP and Hop-by-Hop routing



41

MPLS Traffic Engineering

“The optimization goals of traffic engineering are To enhance the performance of IP traffic while utilizing Network resources economically and reliably.”

Intra-Domain

**A Framework for Internet Traffic Engineering
Draft-ietf-tewg-framework-02.txt**

“A major goal of Internet Traffic Engineering is to facilitate efficient and reliable network operations while simultaneously optimizing network resource utilization and performance.”

Intra-Domain

**RFC 2702
Requirements for Traffic Engineering over MPLS**

42

TE May Require Going Beyond Hop-by-Hop Routing

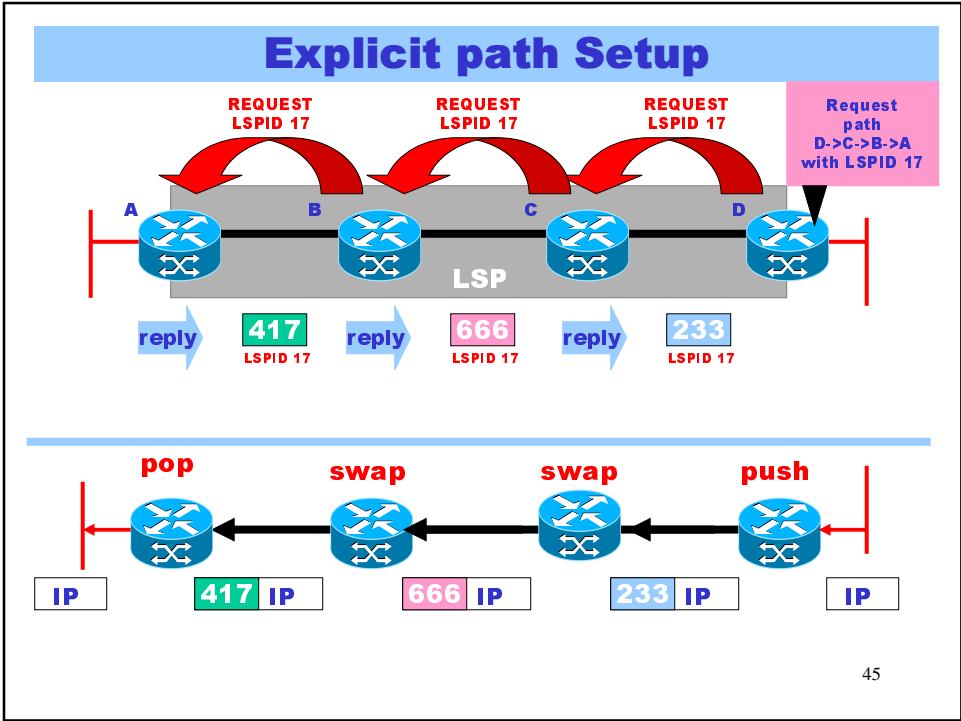
- **Explicit routes**
 - Allow traffic sources to set up paths
- **Constraint based routing**
 - Chose only best paths that do no violate some constraints
 - Needs explicit routing
 - May need resource reservation
- **Traffic classification**
 - Map traffic to appropriate LSPs

43

Hop-by-Hop vs. Explicit Routes

- | | |
|--|--|
| • Distributed control | • Originates at source |
| • LSP trees rooted at destination | • Paths from sources to destinations |
| • Destination based forwarding | • Traffic to path mapping based on what configuration commands your vendor(s) provide |

44



Constraint Based Routing

Basic components

1. Specify path constraints
2. Extend topology database to include resource and constraint information
3. Find paths that do not violate constraints and optimize some metric
4. Signal to reserve resources along path
5. Set up LSP along path (with explicit route)
6. Map ingress traffic to the appropriate LSPs

Note: (3) could be offline, or online (perhaps an extension to OSPF)

Problem here: OSPF areas hide information for scalability. So these extensions work best only within an area...

Extend Link State Protocols (IS-IS, OSPF)

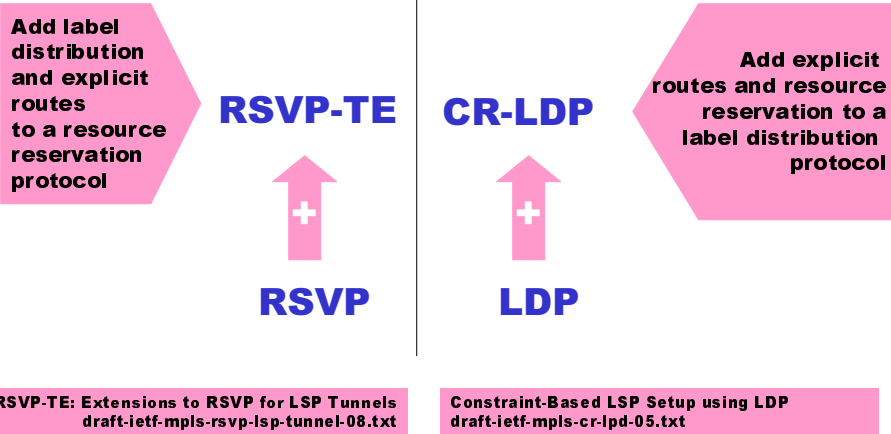
Extend RSVP or LDP or both!

Problem here: what is the "correct" resource model for IP services?

46

Resource Reservation + Label Distribution

Two emerging/competing/dueling approaches:



47

RSVP-TE vs. CR-LPD

RSVP-TE

- **Soft state periodically refreshed**
- **IntServ QoS model**

CR-LDP

- **State maintained incrementally**
- **New QoS model derived from ATM and Frame Relay**

And the QoS model determines the additional information attached to links and nodes and distributed with extended link state protocols...

And what about that other Internet QoS model, DiffServ?

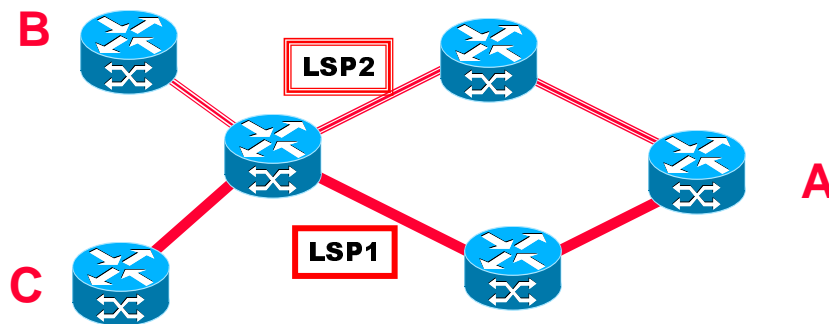
48

A closer look at CR-LDP

- Defines new Type-Length-Value (TLV) encodings and procedures for
 - Explicit routing (strict and loose)
 - Route pinning (nail down some segments of a loosely routed path)
 - Traffic parameter specification
 - Peak rate
 - Committed rate
 - Weight
 - Resource class or color
 - LSP preemption (reroute existing paths to accommodate a new path)
 - LSP Identifiers (LSPIDs)

49

The Fish Revisited

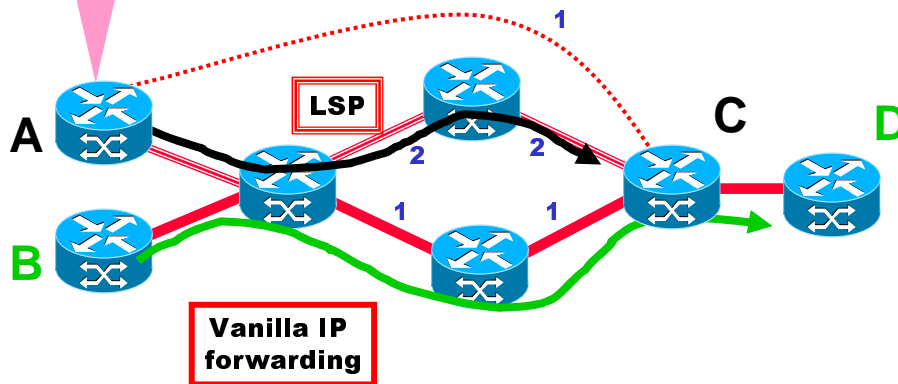


Need at least one explicit route to A

50

Use Shortest Paths to get beyond Shortest Paths!

The IP routing protocol at LSR A is configured to (privately) see A -> C LSP as one link with weight 1.



51

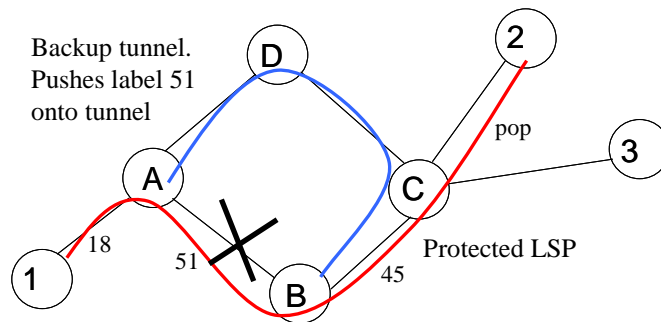
MPLS Fast Reroute

- **Using MPLS TE to improve availability**
 - RSVP-TE creates backup tunnels
 - On failure of protected LSP, packets are shoved down backup LSP tunnel
 - Switchover is faster than waiting for CSPF to calculate and signal a new LSP
- **For local repair (link or node) can recover ~100ms or better**
 - Backup LSP is already in place, so as soon as the failure is detected locally the headend just needs to reprogram the label FIB

52

Link Protection

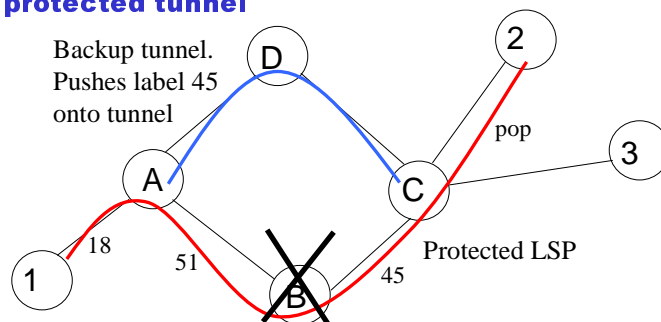
- **Create backup LSP around link to Next Hop**
- **With or without reservation**
 - **Can also backup normal LDP LSP**



53

Node Protection

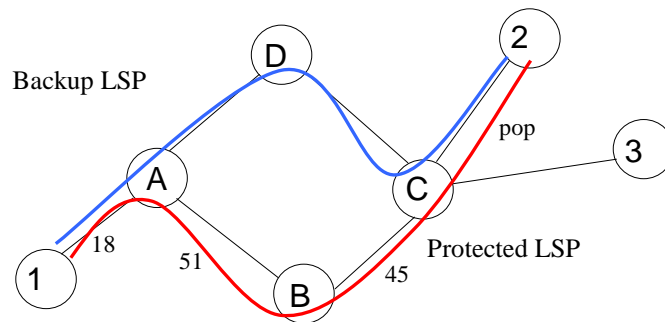
- **Create backup tunnel LSP for two hops away (next-next hop)**
- **Backs up RSVP-TE tunnel**
 - **Learns labels from RESV recorded route of protected tunnel**



54

Path Protection

- **Create an end-to-end diverse backup tunnel**
- **Slower than local protection – have to wait for headend to detect failure**



55

MPLS TE: Is it worth the cost?

- **Much of the traffic across a (transit) ISPs network is interdomain traffic**
 - Congestion is most common on peering links
 - The current work on MPLS TE does not apply to interdomain links! (Actually, it does not even work well across OSPF areas...)
- **MPLS TE is probably most valuable when IP services require more than best effort**
 - VPNs with SLAs?
 - Supporting differentiated services?

56