Internet Routing and MPLS

N. C. State University

CSC557 ♦ Multimedia Computing and Networking

Fall 2001

Lecture # 27

"Roadmap" for Multimedia Networking

1. Introduction
   - why QoS?
   - what are the problems?
2. Basic operations
   - jitter buffers (at hosts)
   - task scheduling (at hosts)
   - packet shaping (at hosts)
   - packet dropping (at routers)
   - packet scheduling (at routers)
3. Types of service
   - Integrated Services (IntServ) and Resource Reservation Protocol (RSVP)
   - Differentiated Services (DiffServ)
4. Application-level feedback and control
   - Real-time Protocol (RTP), Real-time Control Protocol (RTCP)
5. Application signaling and device control
   - Session Description Protocol (SDP)
   - Real-time Streaming Protocol (RTSP)
   - Session Initiation Protocol (SIP)
   - Media Gateway Control Protocol (MGCP)
6. Routing
   - **Multi-protocol Label Switching (MPLS)**
   - multicasting

Today's Lecture

Roadmap

I. Internet Routing

II. Connection-oriented networks

III. MPLS overview

IV. MPLS benefits and applications

V. Encapsulation and hierarchical forwarding

VI. Label distribution

Routing Protocols

- Goal: determine paths to each destination, based on *some* optimality criterion
  - and create routing tables for packet forwarding

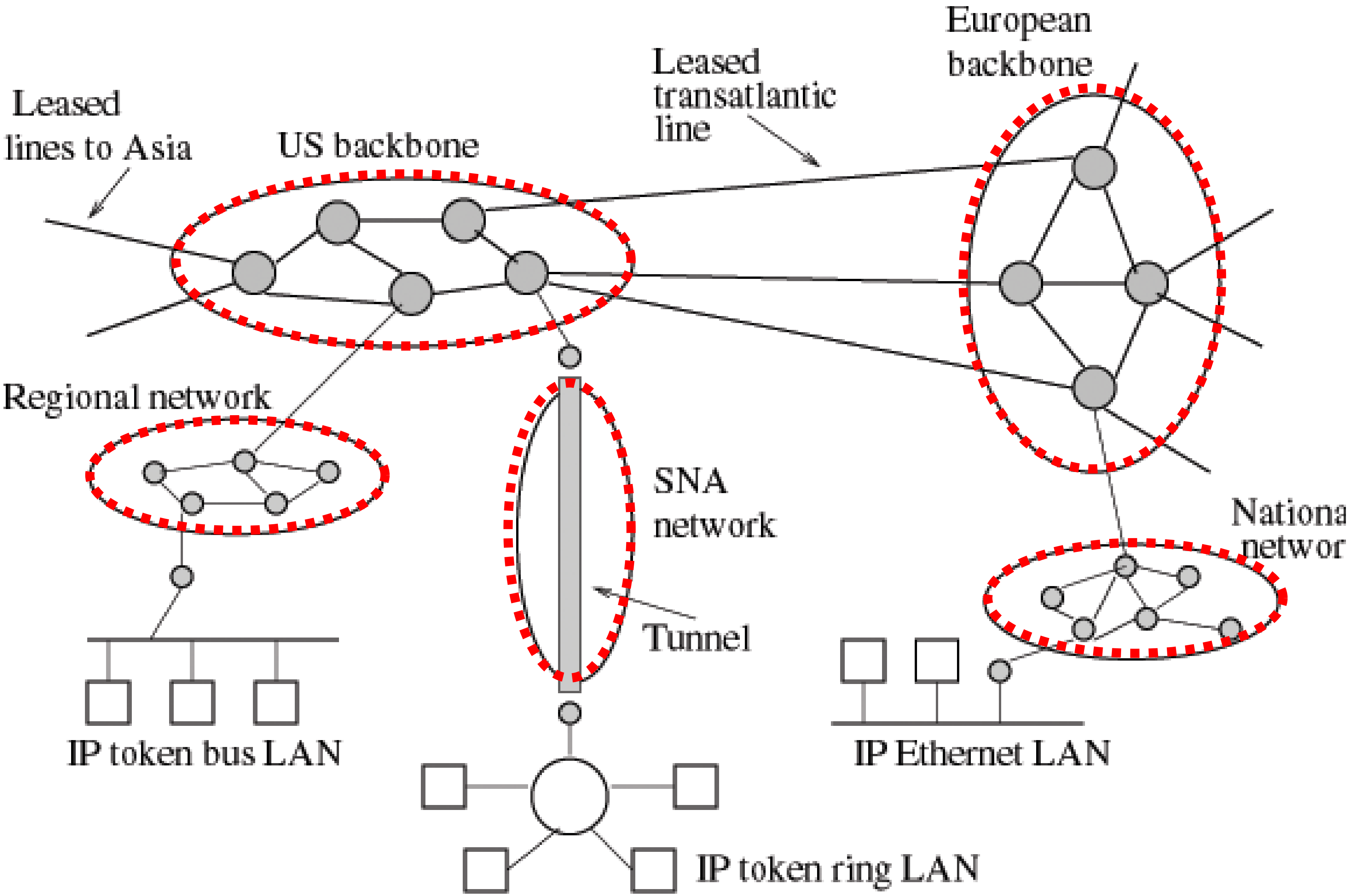
- Routing protocol
  1. routers exchange info regarding state of network
  2. run routing algorithm to compute routes (paths)

- Optimization criteria
  - what metric to use (hops, delay, cost,...)?
  - how often does it change?
  - based purely on *destination*, not on source

Requirements For Routing Algorithms

1. Simplicity
2. Scale to large Internet size (hierarchical)
3. "Safe" interconnection of different organizations
4. Adapt to changes in topology
5. Avoid routing loops
6. Efficient (low message overhead, converges quickly)
7. Paths should be (near) “optimal”, according to some criterion

Autonomous Systems



Networks and Autonomous Systems

- In the Internet today…
  - more than 10,000 ASes
  - more than 250,000 networks

Classification of Routing Algorithms

- Intra-domain vs. inter-domain
  - intra-domain: within a single administrative entity (autonomous system)
  - inter-domain: between autonomous systems
- Intra-domain routing protocols
  - single administration → simplicity
  - routing efficiency/optimality is primary objective
- Inter-domain routing protocols
  - larger geographical reach → scalability, security
  - routing policies, distribution of control
  - fault isolation

Routing Optimization

- Criteria
  - minimum number of “hops”
  - minimum propagation delay
  - minimum congestion / packet loss
  - least cost
- Constraints
  - security considerations
  - business relationships

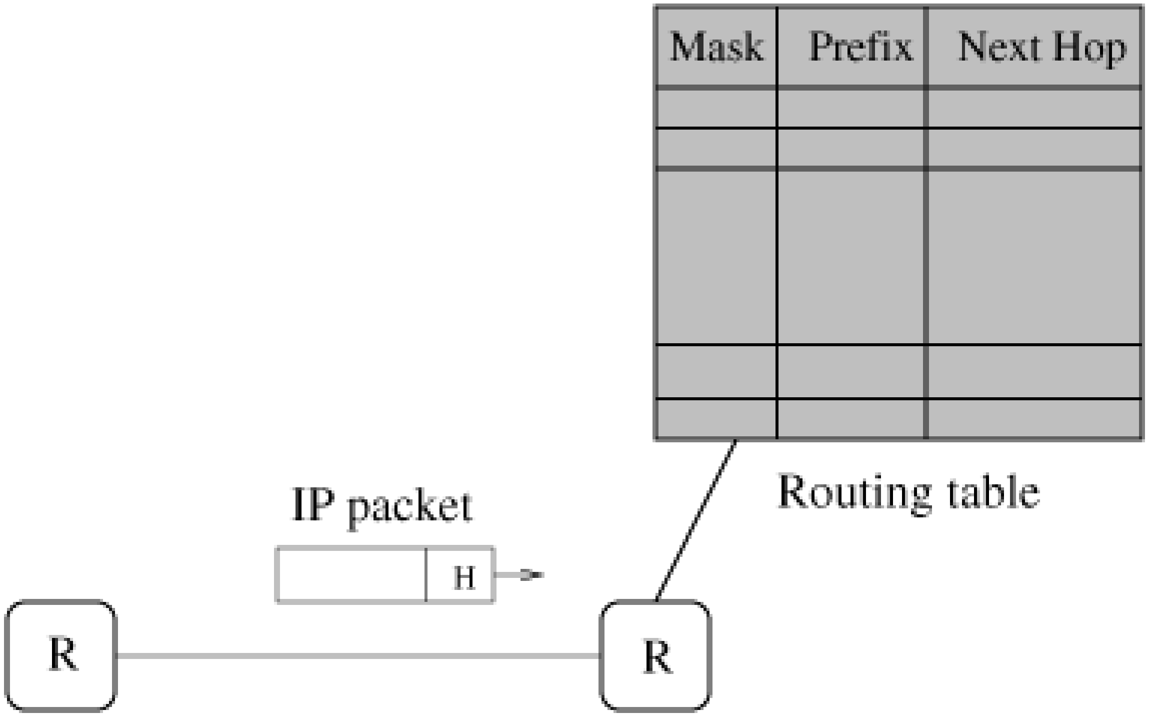
Fault Recovery and Convergence Time

- Routing algorithms will eventually converge on optimal routes
  - as long as state of the links / routers remain stable
  - no requirement for updates to be synchronized
- During convergence, non-shortest paths and loops may develop
  - "good news travels fast, bad news travels slowly"

Problems with Routing for Multimedia

1. Destination-based routing only
2. Routing recalculation / convergence / fault recovery time
3. Lack of optimization for QoS

Conventional Packet Forwarding



- As a packet travels in an IP network, each router...
  - analyzes the packet's header
  - consults the routing, or forwarding, table
  - chooses a next hop router for the packet
    - independently of any choices made for other packets

Conventional Packet Forwarding (cont'd)

- Packet headers contain many fields for varying purposes
  - only some of them are used for routing purposes
- Choosing the next hop involves two steps
  - partition the entire set of possible packets into *forwarding equivalence classes* (*FEC*s)
  - map each FEC to a next hop
- Current forwarding scheme has limitations
  - uses only destination IP address from packet
  - doesn't support QoS, traffic engineering, fast recovery from failures, …

Connection-Oriented Architectures

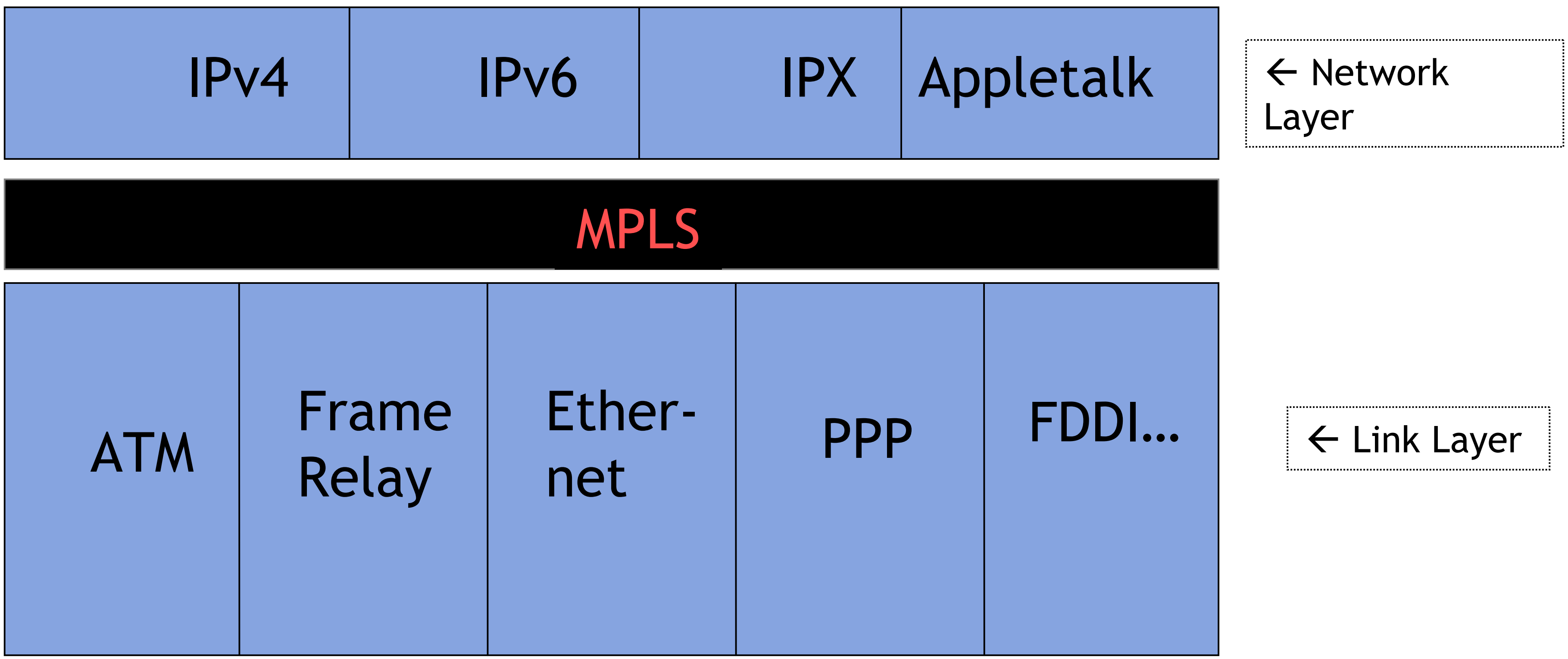
- Ex.: ATM, Frame Relay, X.25
- A *logical connection* must be set up before data is exchanged
  - the *state* of the connection is maintained at each network switch
- A *flow* is the sequence of datagrams exchanged over a TCP or UDP connection
  - multiple flows may be multiplexed into a single logical connection
- Connection-oriented architectures enable the type of services that not well-supported by conventional IP datagram routing

MPLS Networks

- A logical connection is established between two points in a pure datagram network
  - connection carries normal datagram traffic
- MPLS adds an additional header, containing a *label*
  - identifies the connection
- A hybrid architecture (advantages of both?)
  - logical connections can be used for connection-oriented services
  - normal datagram processing (forwarding) still available for datagram services

Where it Fits

- Below the network layer
  - not an end-to-end protocol



MPLS Labels and Encapsulation

- Insert in each packet a new header ("shim header")

Link Layer Header	MPLS "Shim" Header	IP Header	Payload

- A *label* = short, fixed length value
  - used to identify the FEC
- Labels have local significance only
  - adjacent routers must agree on the binding of label ↔ FEC
  - does not have to be globally unique
  - no meaning to the label; just an identifier

The MPLS Forwarding Table

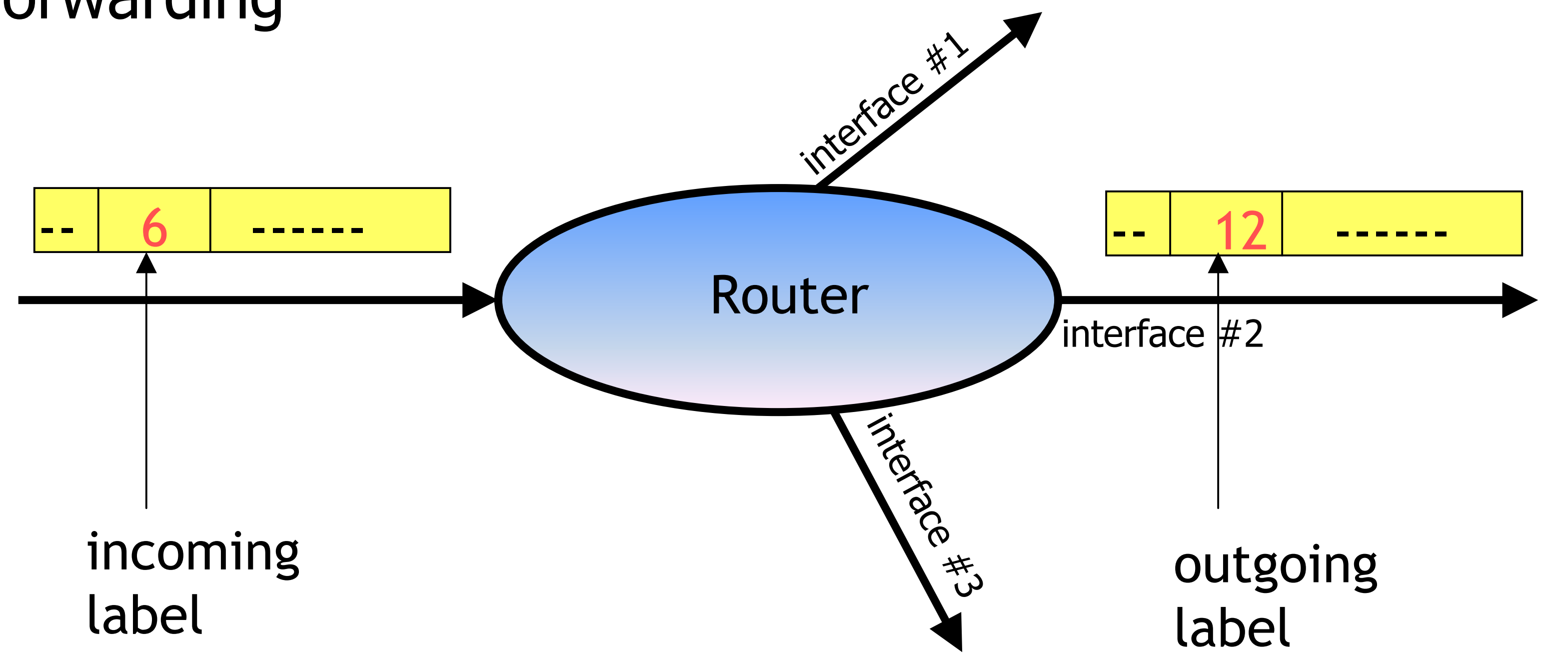
- Add a new table to router: the Label Switching Forwarding Table
  - may be other info in this table, as well (e.g., quality of service)

Forwarding Table

Incoming Packet Label	Outgoing Interface	Next Hop Address	Outgoing Packet Label	Other Require-ments
6	#2	192.0.168.100	**12**	...
...	...	...	...	...

Basic MPLS Idea

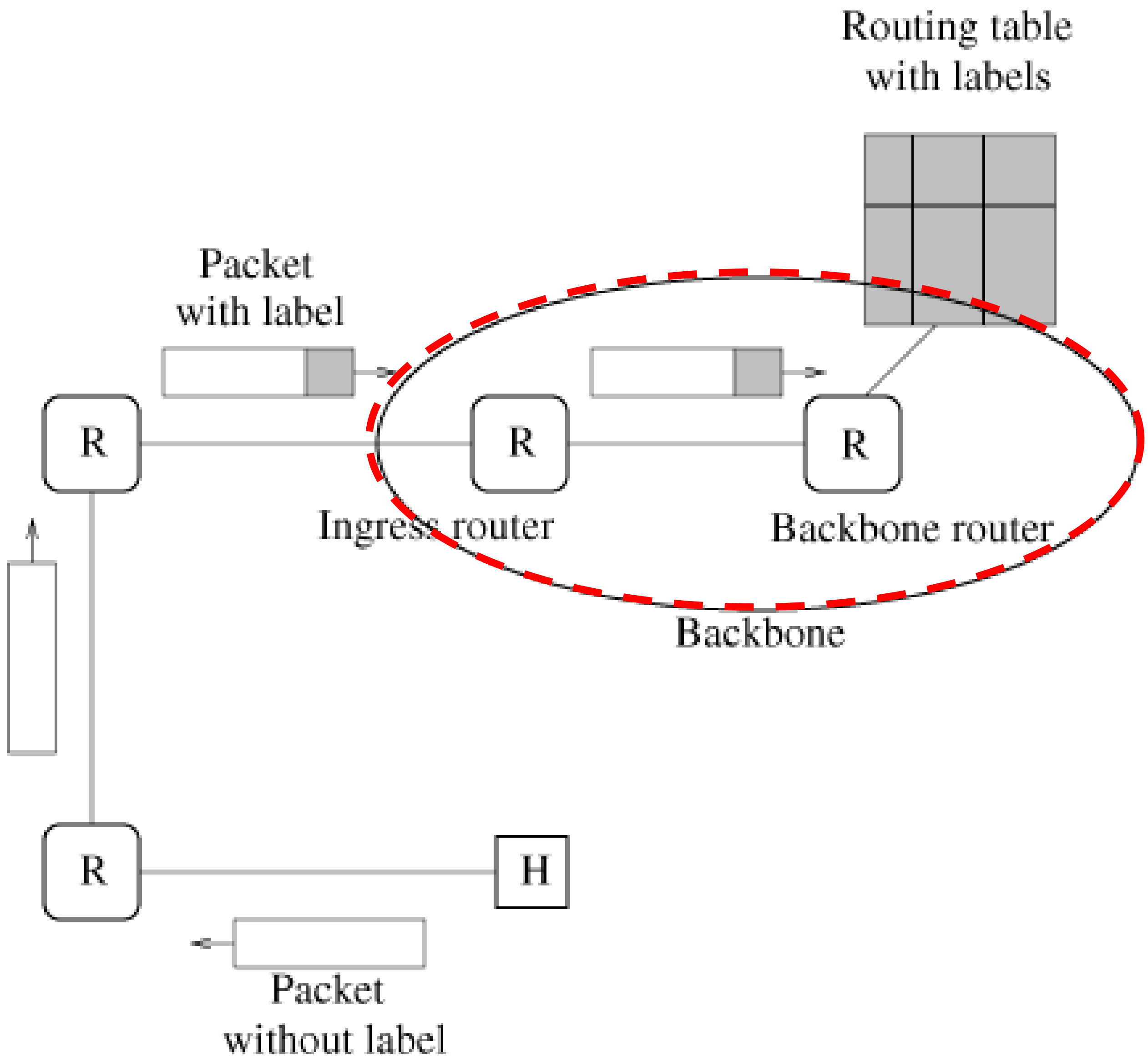
- Look at the label to pick an outgoing interface
- Then replace the incoming label with the appropriate outgoing label
- Routers that don't support MPLS do normal packet forwarding



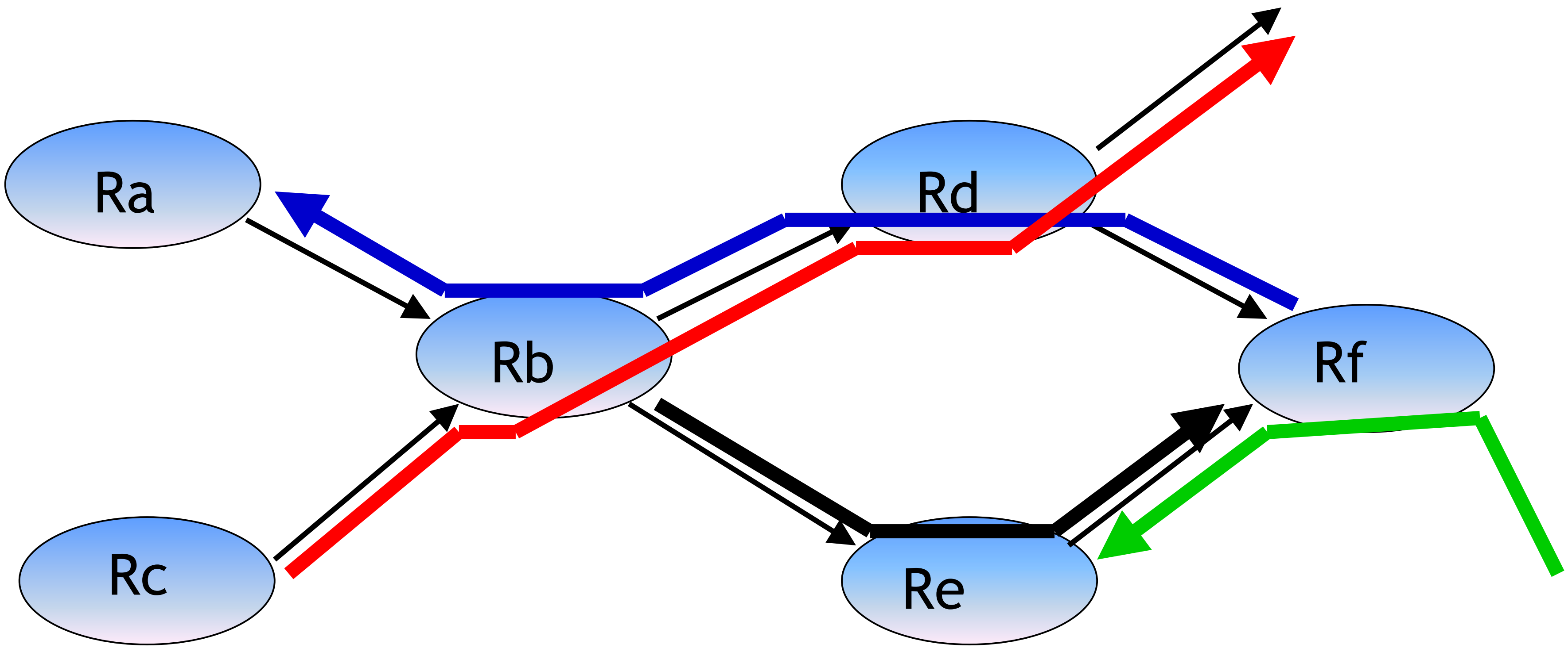
MPLS Terminology

- A *label-switched router (LSR)* can perform MPLS label-switching
- A *label-switched path (LSP)* is a consecutive sequence of LSRs that forward a packet using MPLS
- An *ingress LSR* is the first LSR on a LSP
  - determines FEC for packet from routing table
  - inserts a label (shim header) in front of the packet
  - at this point, the label is *bound* to the FEC at this router
- An *egress LSR* is the last LSR on a LSP
  - responsible for removing the label from the packet

Inserting Labels into Packets



Label-Switched Paths



- Can start and terminate in the middle of the network

Notes

- Labels are an optimization
  - packets can be routed even if labels aren't set up at all, or are set up on just parts of the path
- Assignment of a packet to an FEC is done only once, as the packet enters the MPLS network
  - subsequent hops do not need to examine the network layer header
- Important questions
  - on what basis are LSPs set up?
  - how are they set up, and how long do they last?

Some Benefits / Applications of MPLS

1. Traffic engineering, source-based QoS routing
2. Route *pinning*
3. Protection and fast rerouting
4. Hierarchical forwarding

- Also: faster packet processing at routers (= greater throughput)

1. Traffic Engineering

- Datagram networks (IP) have poor controllability
  - route on destination address, least-hop path
- MPLS provides better control
  - classification into FECs at entry to LSP may be arbitrarily complex
- Operator can use global optimization algorithms to map traffic demands to physical links
  - not possible using local optimization

2. Route Pinning

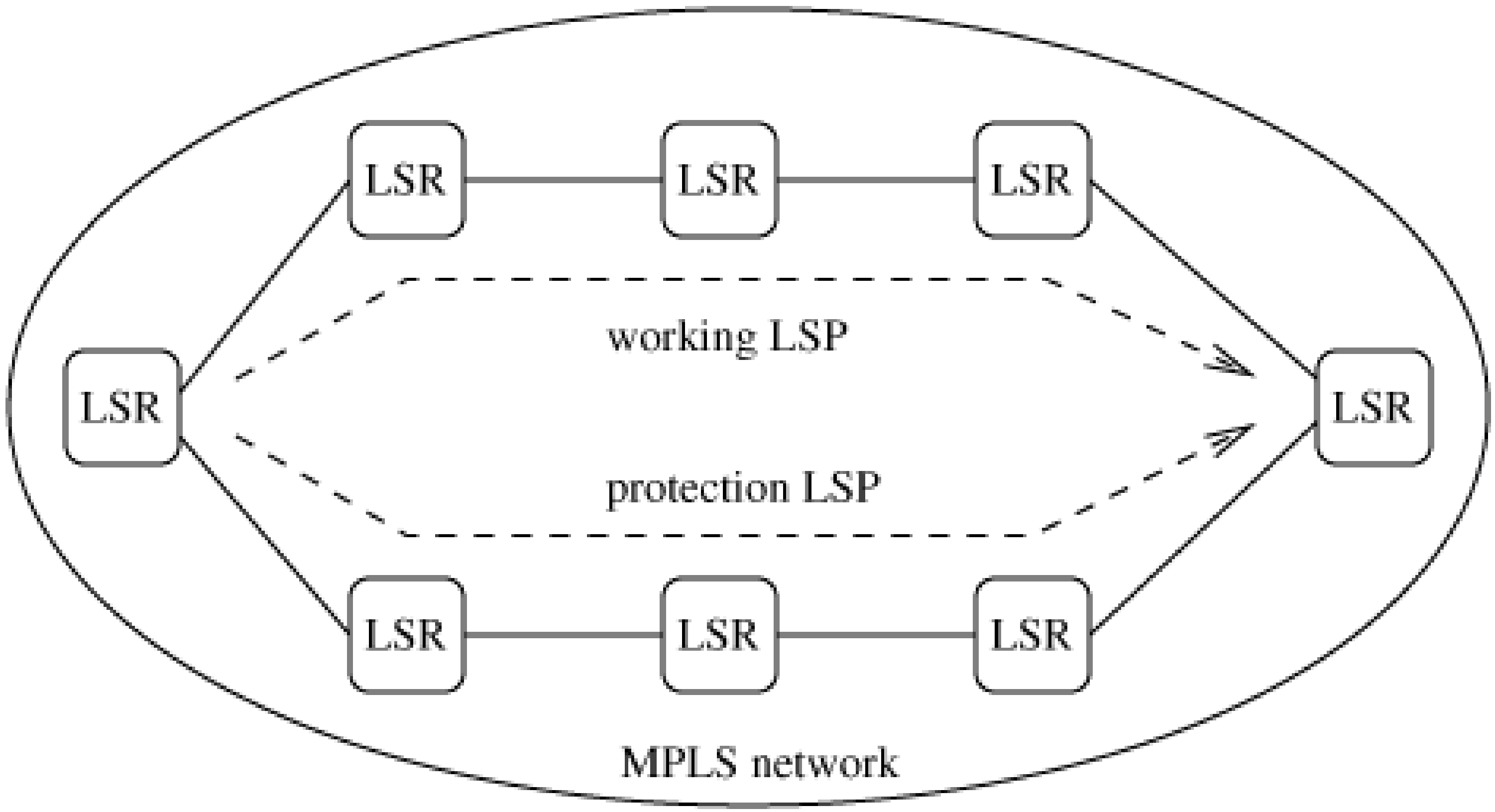
- Some applications are sensitive to path changes or disruptions
- IP routing does not guarantee routing changes won’t occur
- MPLS routing
  - LSP will not change unless there is explicit intervention
  - provides a specific and stable path (*pinned* route)

3. Protection and Fast Rerouting

- For datagram networks, routing recovers from failures by computing alternative paths
  - drawback: long latency in reconfiguring network and re-establishing connections is possible

- MPLS provides two choices
  1. dynamically establish new LSPs around points of failure
     - relatively long latency
     - for most datagram traffic
  2. switch to a "backup" LSP previously established for the failed LSP
     - relatively short latency
     - no signaling or setup required
     - much less expensive than using SONET for recovery

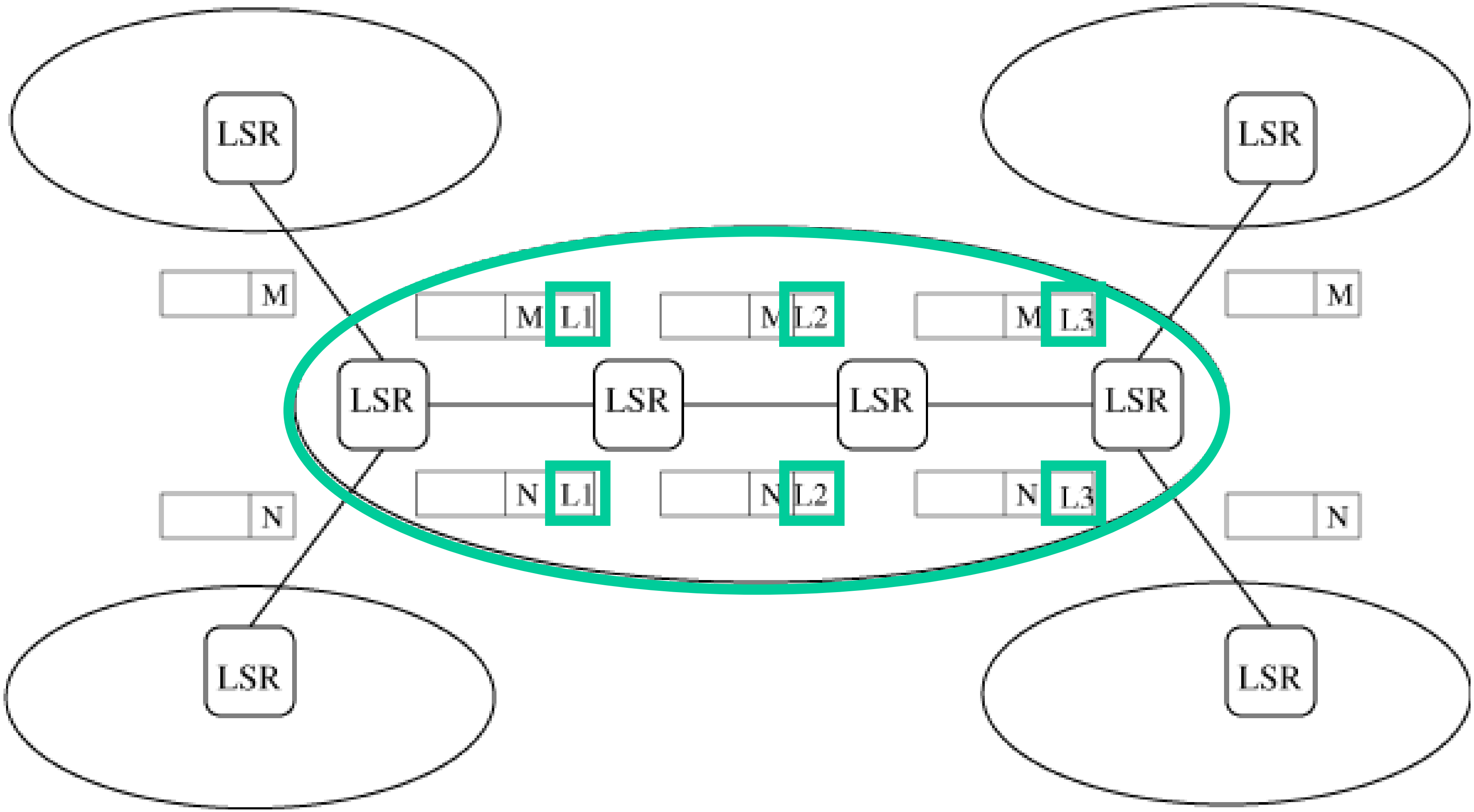
Protection Switching



4. Hierarchical Routing and the Label Stack

- A packet may carry several labels
  - organized in a LIFO stack
- Purpose: a hierarchy of label-switched paths
  - e.g., paths within paths
- The bottom of the stack is indicated by a bit in the shim header

Hierarchical Forwarding Example



Hierarchical Forwarding (cont'd)

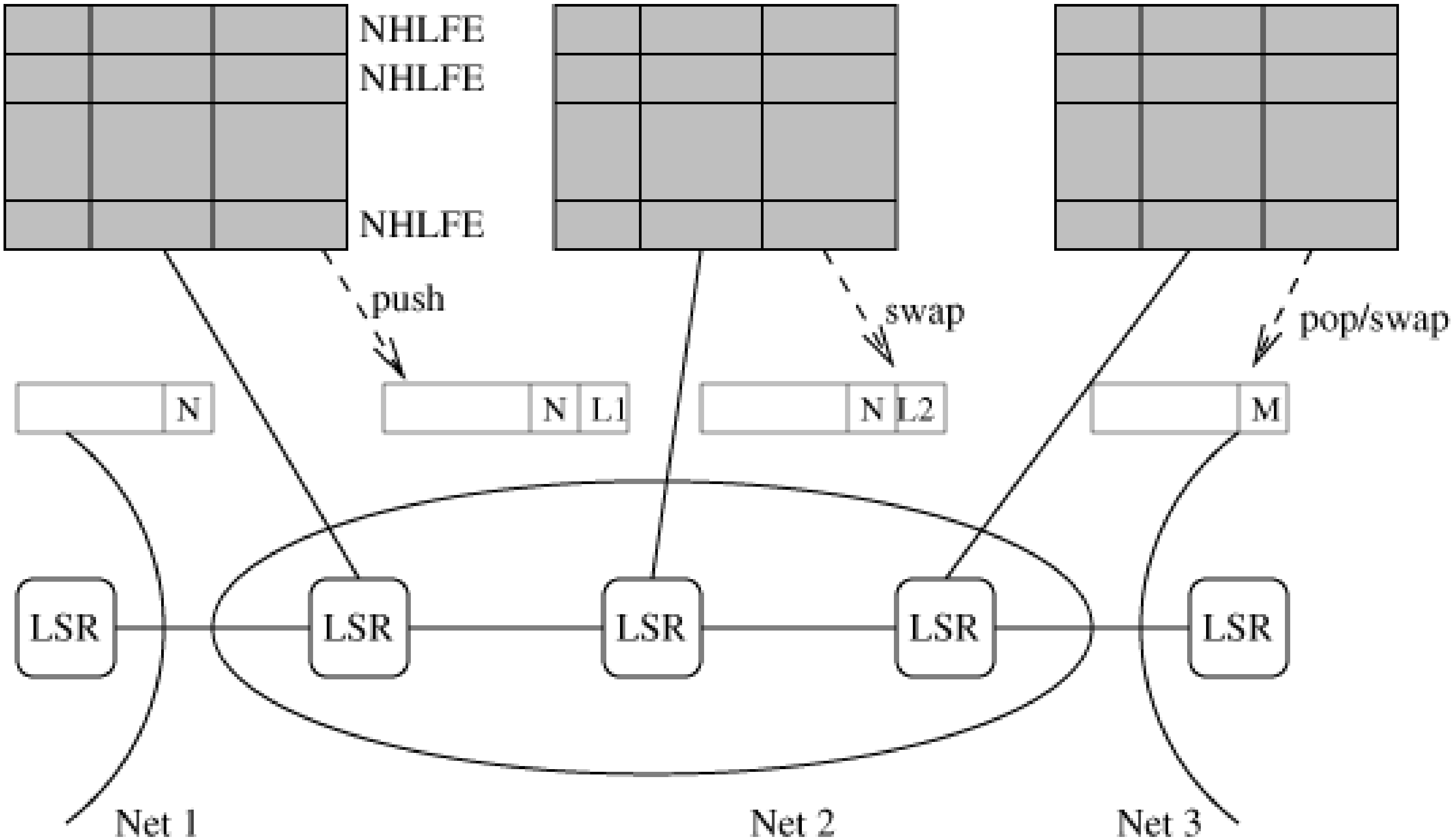
- One MPLS network can run over another
  - forwarding hierarchy can scale to very large MPLS networks, each constructed from smaller MPLS networks

- Forwarding Table entries include operation to perform on the packet's label stack
  - swap: replace label at top of stack with a specified new label
  - pop label stack (end of nested LSP reached)
  - push a new label onto the label stack (start of nested LSP)

Forwarding Table, again

Incoming Packet Label	Outgoing Interface	Next Hop Address	Operation	Outgoing Packet Label	Other Require -ments
6	#2	192.0.168.100	swap	12	…
…	…	…		…	…

Stack Operations

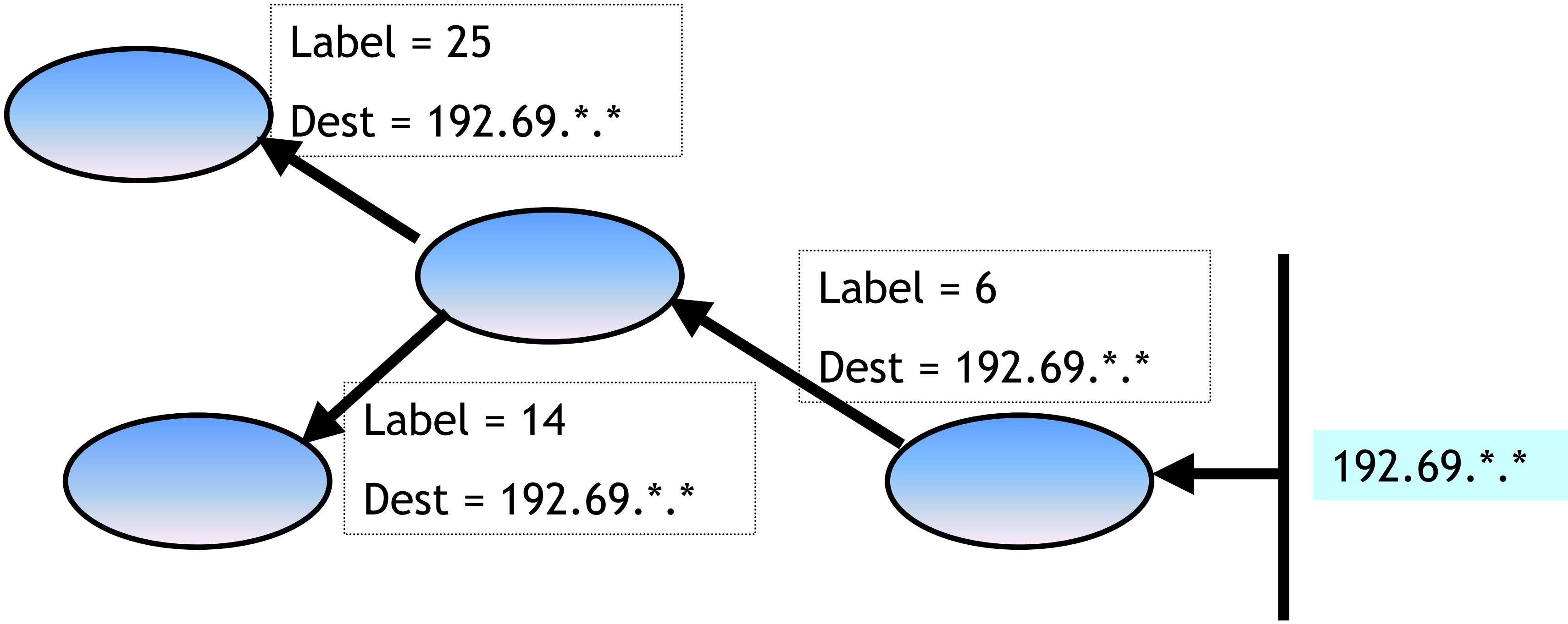


Establishing LSPs

- Before traffic forwarding can take place...
  - an LSP must be set up
  - labels / operations must be assigned at each hop
- Two LSP types (based on method used to determine the route)
  - *control-driven* LSP
  - *explicitly-routed* LSP or *constraint-based* LSP (*CR-LSP*)
- *Label distribution protocol =* method by which one LSR informs another LSR of label↔FEC bindings it has made

Constructing Label-Switched Paths

- Creating labels and FECs
  - Construct when routing table is updated
  - Or, construct when instructed by a "downstream" router
  - Or, construct when packet volume justifies



Control-Driven LSPs

- The LSP will be the same path computed by IP routing
- Main benefits
  - simplicity, compatibility with existing routing
  - speed-up of forwarding function

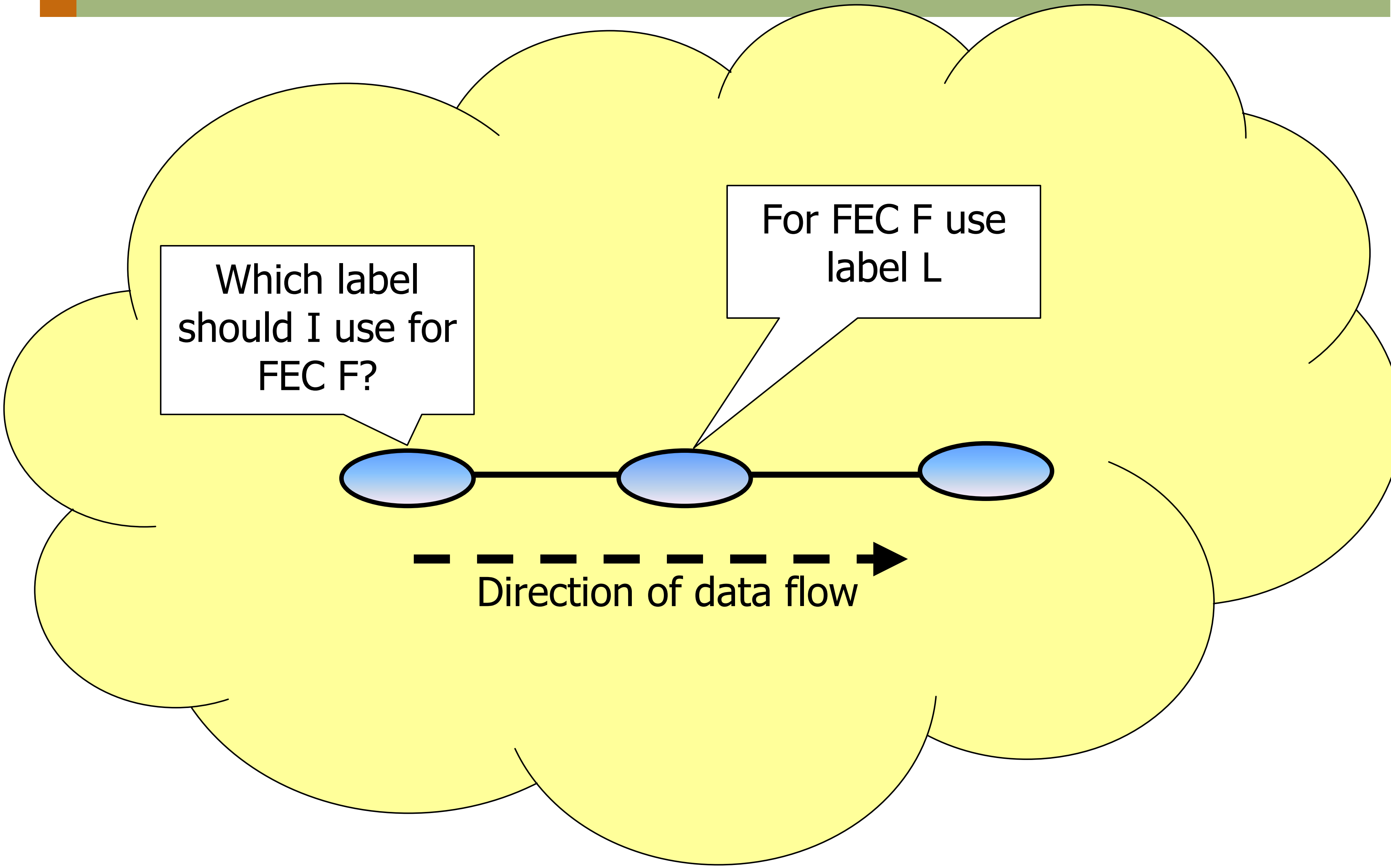
Explicitly-Routed LSPs (CR-LSPs)

- CR-LSP specified by network operator or network management application
  - route for CR-LSP is specified in the setup message
  - setup message traverses all LSRs along the specified route
  - each LSR sends label request to the next-hop LSR
- Network traffic will be directed to a path independent of what is computed by IP routing
- Benefits
  - traffic engineering
  - route pinning
  - etc.

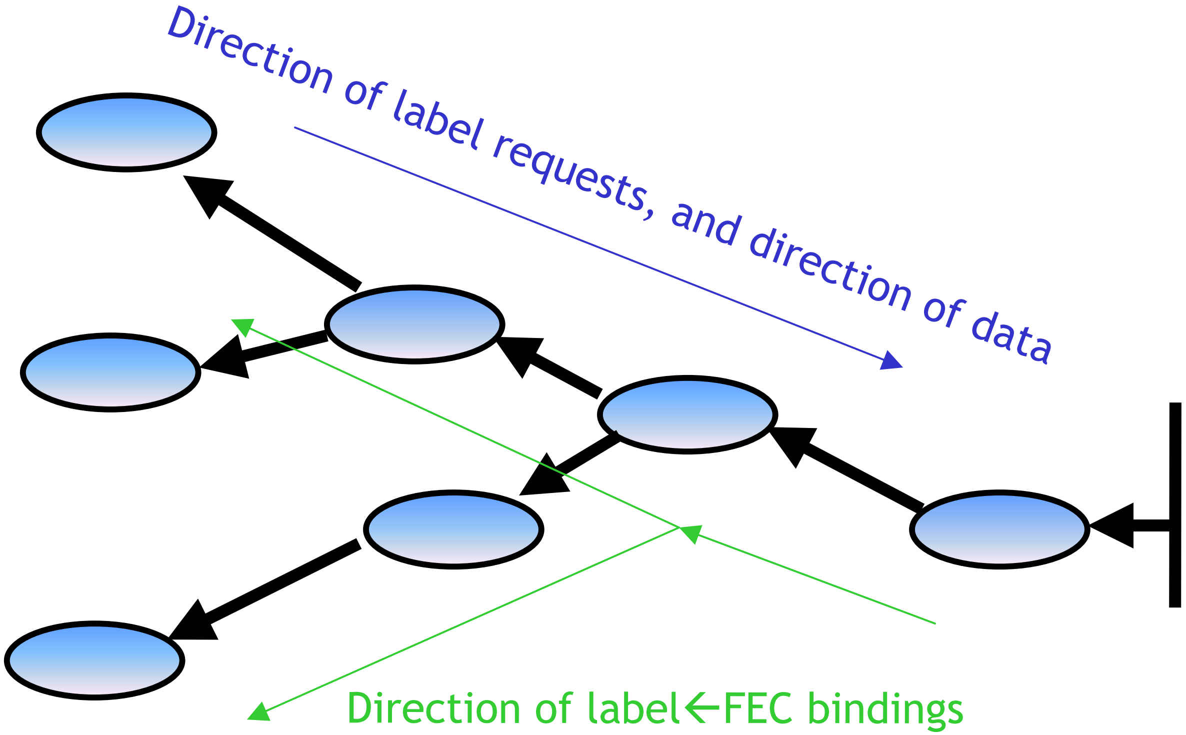
Informing Upstream Routers

- Directions
  - *downstream* = direction in which data is propagated along an LSP to the destination
  - *upstream* = opposite direction (away from destination)
- Technique #1: *on-demand (solicited)* label distribution
  - upstream LSR explicitly requests, from its next-hop LSR for a particular FEC, a label↔FEC binding for that FEC
- Technique #2: *unsolicited* label distribution
  - downstream LSR distributes its label↔FEC bindings to upstream LSRs that have not explicitly requested them
- Both techniques may be used in same network at same time

On-Demand label Distribution



Direction of Messages



Label Distribution Protocols

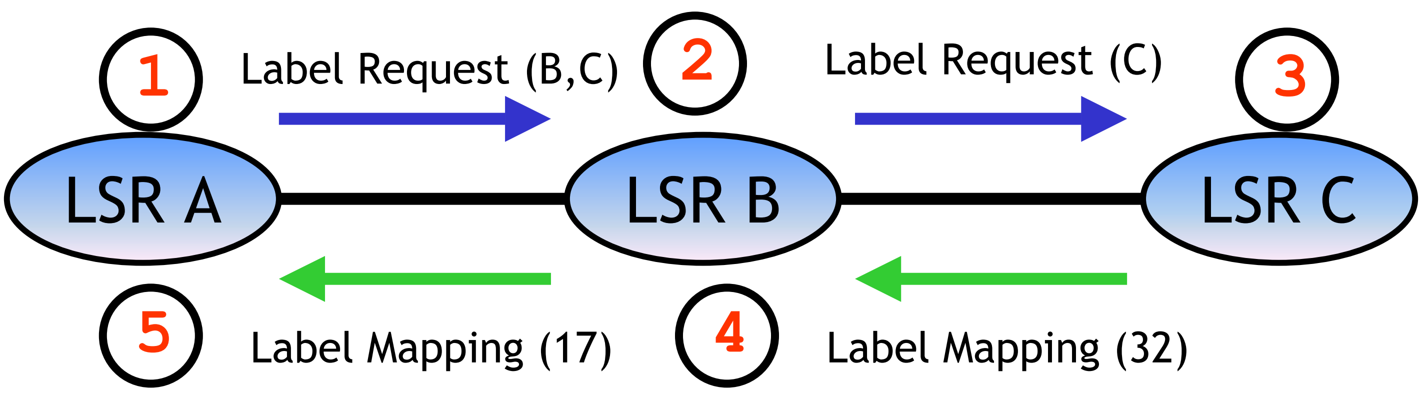
- Label Distribution Protocol (LDP)
- Explicit routing
  - Constrained-Based Routing LDP (CR-LDP)
  - RSVP extensions
- Extensions to existing routing protocols (e.g., BGP) to carry label information

Explicit (Source) Routing

- Initiating (source) node specifies the route for the LSP
  - based on load balancing, QoS, security, …

CR-LDP (Internet Draft)

- Extensions to LDP (RFC 3036)
  - carries explicit route information
  - carries traffic parameters for resource reservation
  - (carries options for CR-LSP resilience (e.g., global or local repair))



CR-LDP Steps

1. LSR A determines it needs LSP to LSR C
   - policies / other factors determine route should go through B
   - creates Label Request with route (B,C), + traffic parameters
   - reserves resources for new LSP
   - forwards Label Request to B
2. LSR B receives request and processes it
   - reserves resources (may also reduce resource amount)
   - modifies route to (C)
   - forwards to C

CR-LDP Steps (cont'd)

3. LSR C is egress of LSP
   - makes reservation
   - allocates label
   - sends Label Mapping message back to B (including reservation information)
4. LSR B receives Label Mapping
   - matches to original request
   - finalizes reservation
   - allocates incoming label
   - passes new label to A
5. LSR A receives label mapping
   - matches to request, finalizes reservation

RSVP Extensions for MPLS

- RSVP reservations
  - RSVP "PATH" message is sent from source to destination LSR
  - RSVP "RESV" message sent in reverse direction (from destination to source, in response to PATH message)
- Two MPLS-specific objects added to PATH message
  - "Label Request" object to request label bindings
  - "Explicit Route" object listing LSRs in the path
- One MPLS object added to RESV message
  - " Label" object with the label binding
- Resources are reserved for a LSP at the same time as label assignments are made
  - while RESV message propagated to source

Sources of Info

- Books
  - B. Davie and Y. Rekhter, *MPLS: Technology and Applications*, 2000
- Web
  - L-o-n-g list of web resources: http://www.mplsrc.com/articles.shtml
  - IETF
    - http://www.ietf.org/html.charters/mpls-charter.html