Dynamic Bandwidth Provisioning on the Internet

from
"New Internet and Networking Technologies and Their Application on Computational Sciences",
invited talk given at
Ho Chi Minh City, Vietnam
March 3-5, 2004
and
"New Internet and Networking Technologies for Grids and High-Performance Computing",
tutorial given at HiPC 2004, Bangalore, India
December 22nd, 2004

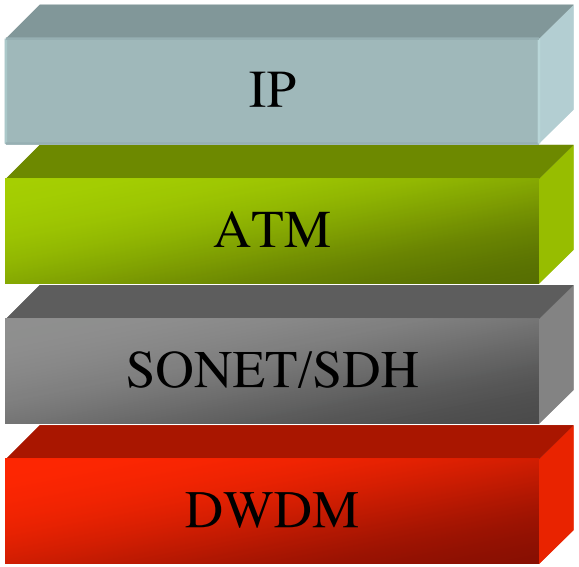
C. Pham
University of Pau, France

Limitations of the current Internet

- Bandwidth
  - Raw bandwidth is not a problem: DWDM
  - Provisioning bandwidth on demand is more problematic
- Latency
  - Mean latencies on Internet is about 80-160ms
  - Bounding latencies or ensuring lower latencies is a problem
- End-to-end performances
  - Links are getting faster and faster!
  - Why my FTP is still going so slow?
- Communication models
  - Only unicast communications are well-defined: UDP, TCP
  - Multi-parties communication models are slow to be deployed

Bandwidth provisioning

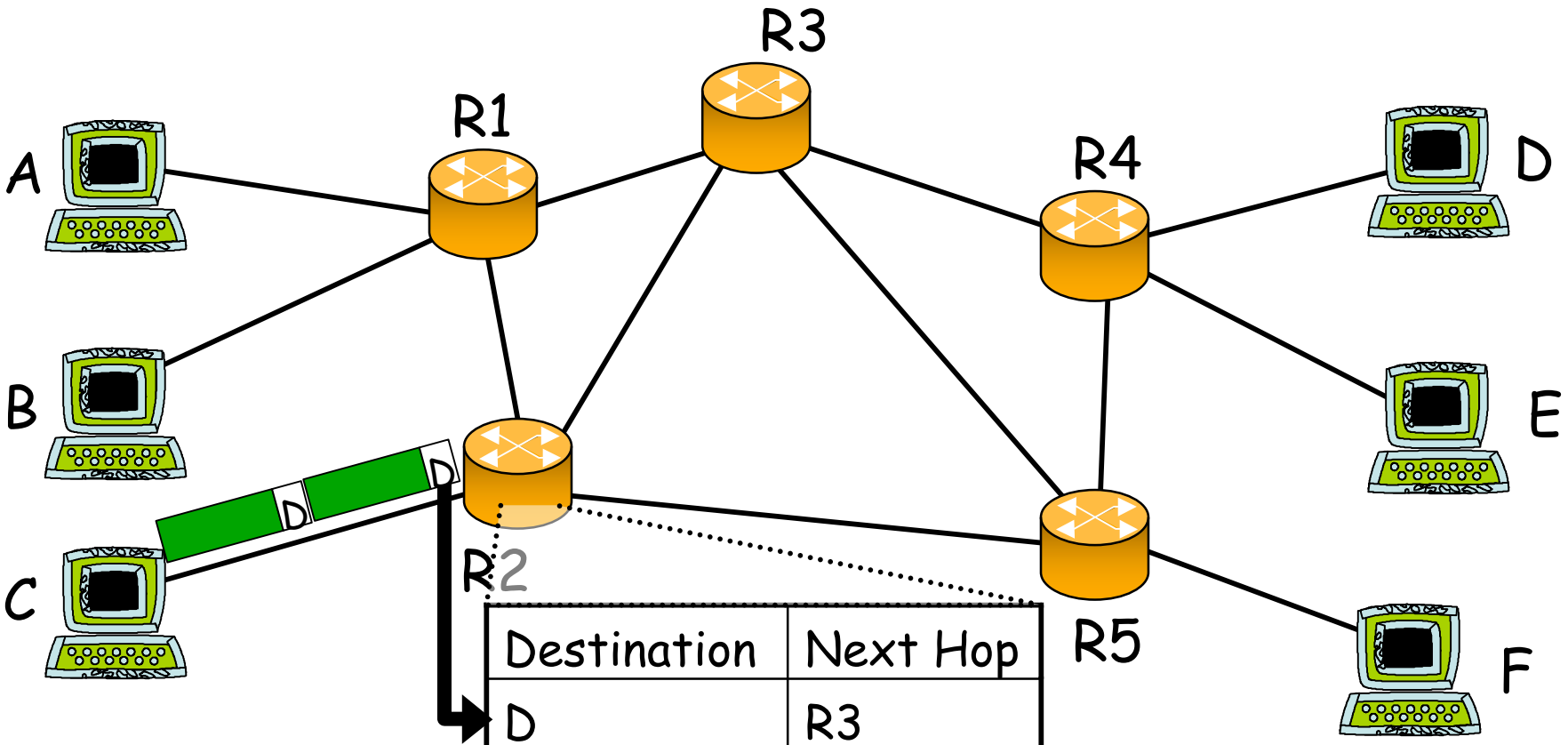
- DWDM-based optical fibers have made bandwidth very cheap in the backbone
- On the other hand, dynamic provisioning is difficult because of the complexity of the network control plane:
  - Distinct technologies
  - Many protocols layers
  - Many control software



Provider's view

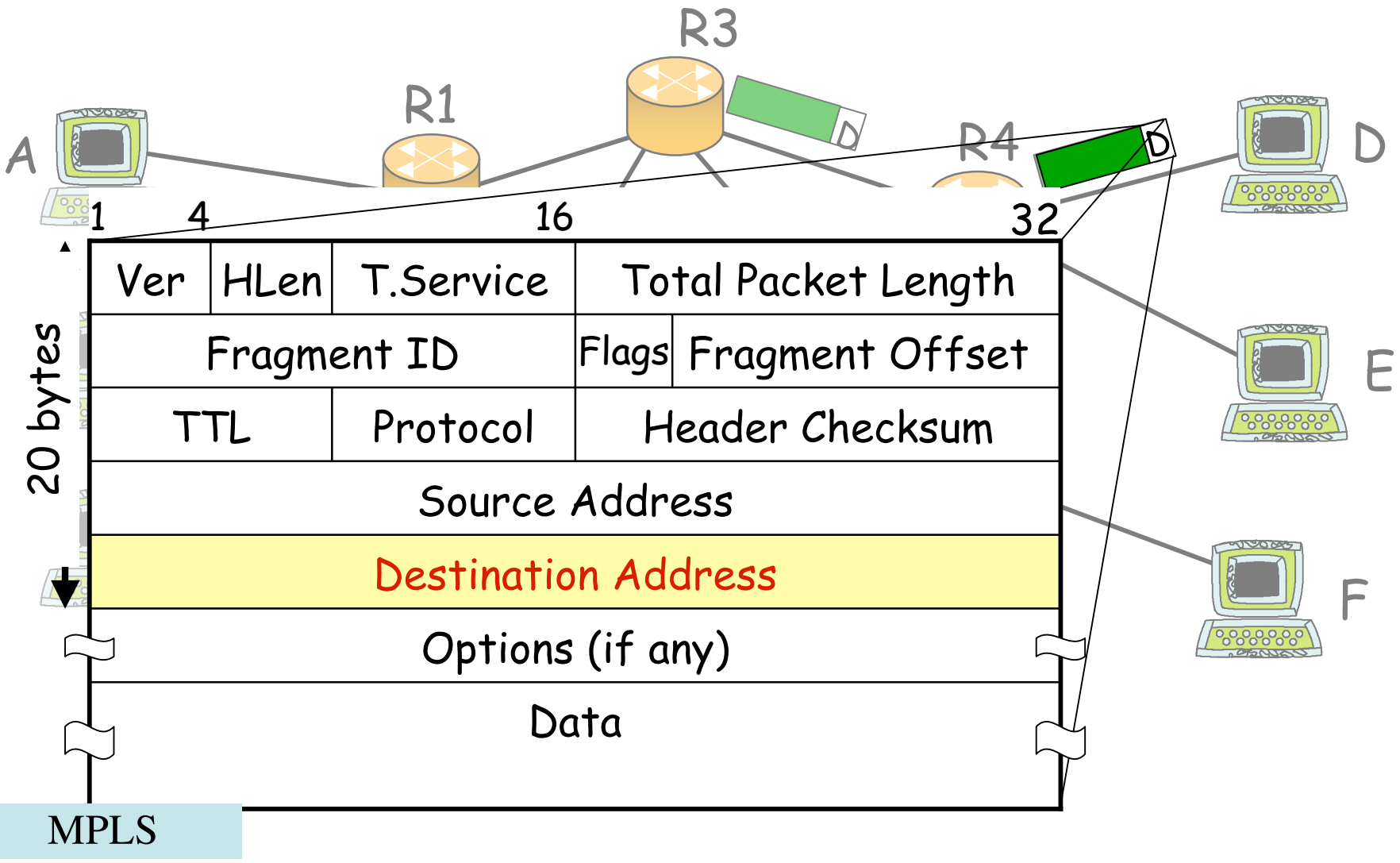
Today's setting time is several weeks/months! We want to set dynamic links within hours

Review of IP routing



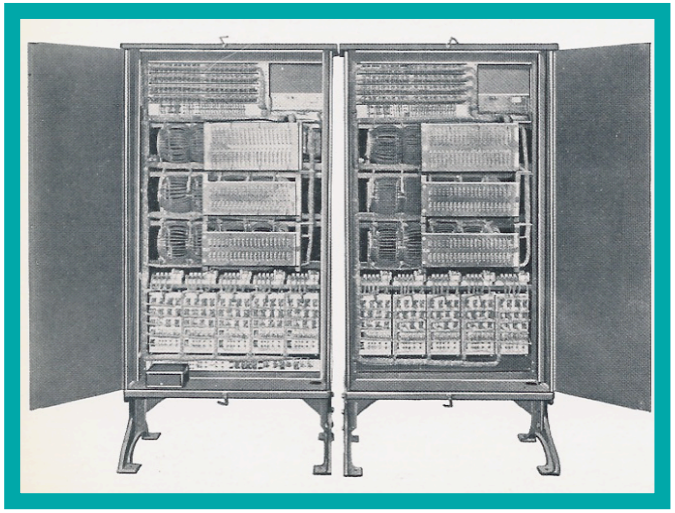
Destination	Next Hop
D	R3
E	R3
F	R5

Review of IP routing



Review of telephone network

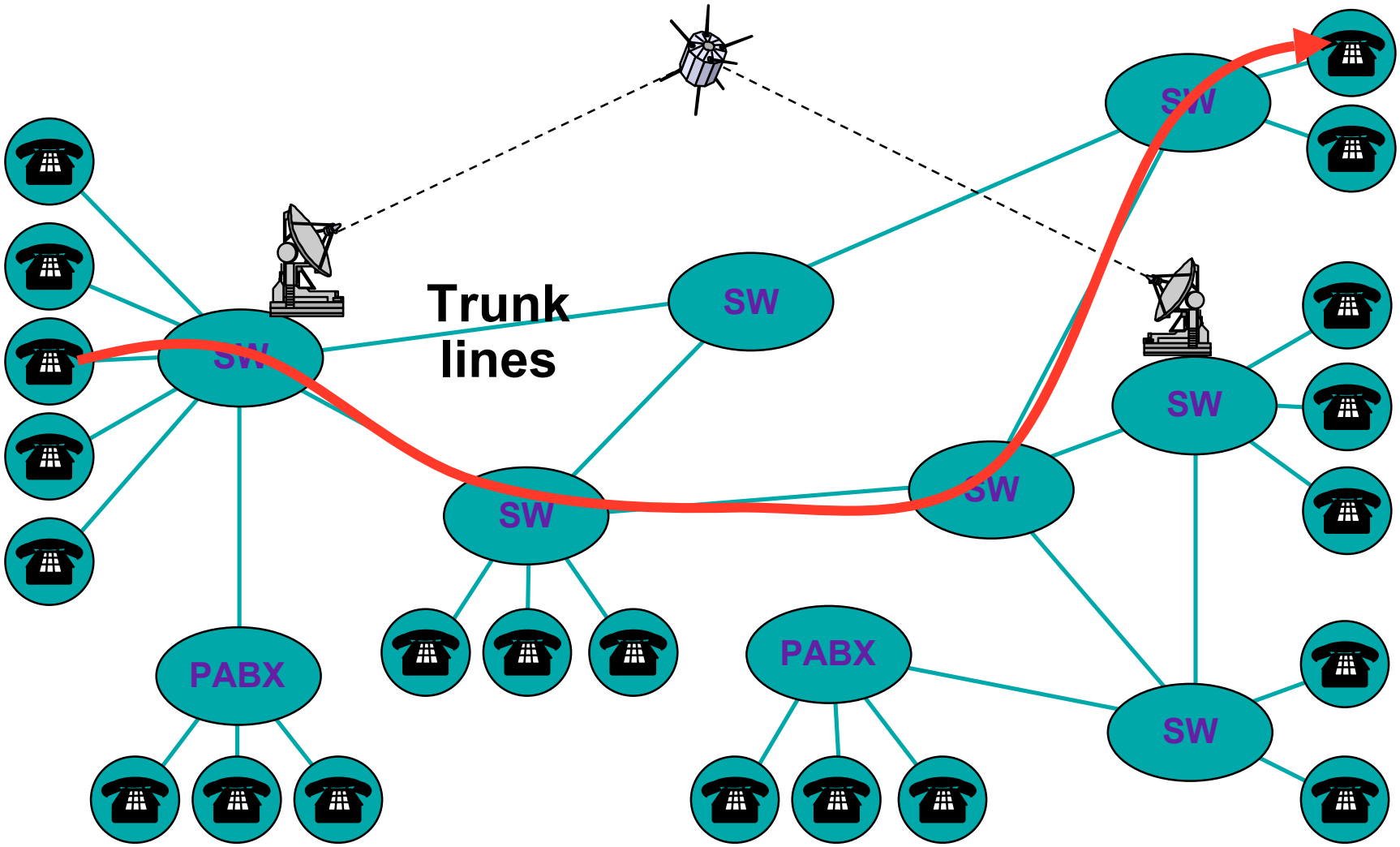
First automatic Branch Exchange
Almond B. Strowger, 1891…



Signaling replaces the operator

Source J. Tiberghien, VUB

The telephone circuit view



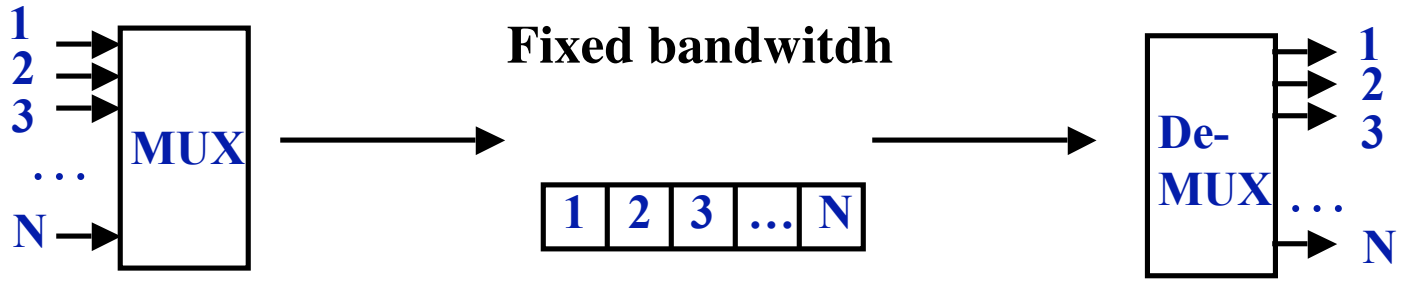
Advantages of circuits

- Provides the same path for information of the same connection: less out-of-order delivery
- Easier provisioning/reservation of network's resources: planning and management features

Time Division Circuits

- Most trunks time division multiplex voice samples
- At a central office, trunk is demultiplexed and distributed to active circuits
- Synchronous multiplexor
  - N input lines
  - Output runs N times as fast as input

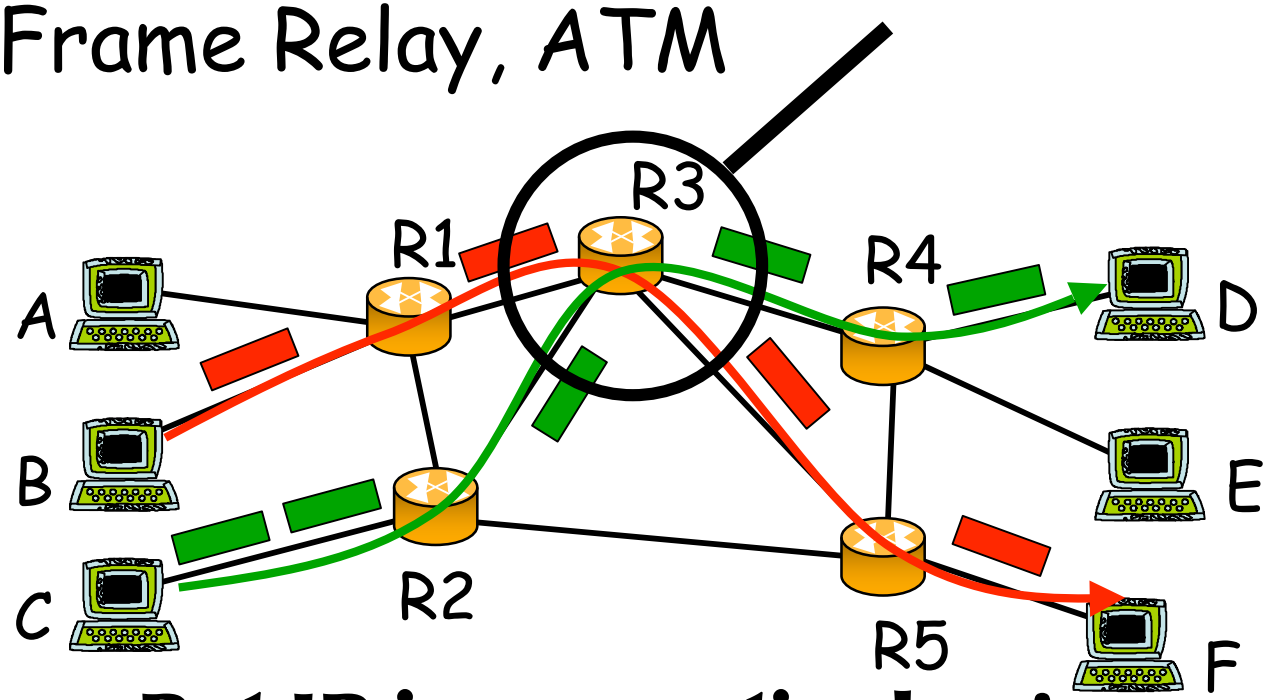
Simple, efficient, but low flexibility and wastes resources



1 sample every 125us gives a 64Kbits/s channel

Back to virtual circuits

- Virtual circuit refers to a connection oriented network/link layer: e.g. X.25, Frame Relay, ATM



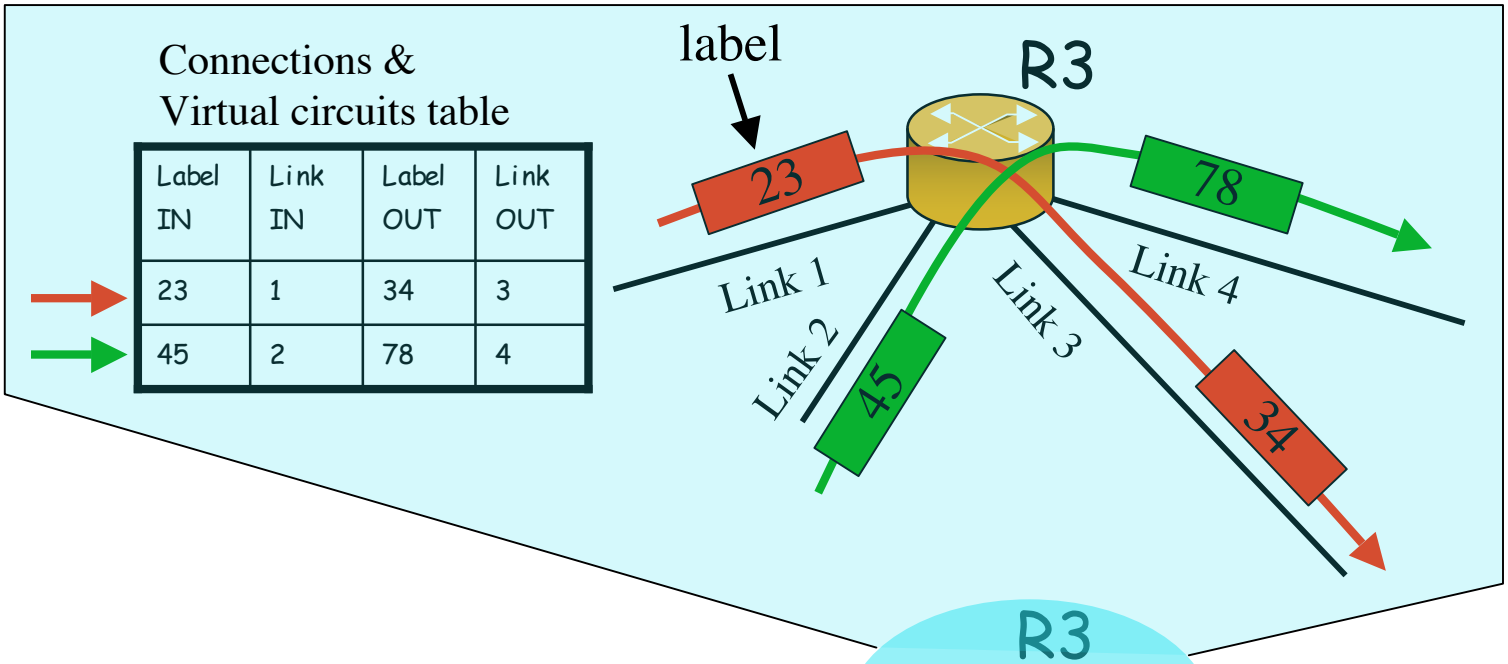
Virtual Circuit Switching: a path is defined for each connection

But IP is connectionless!

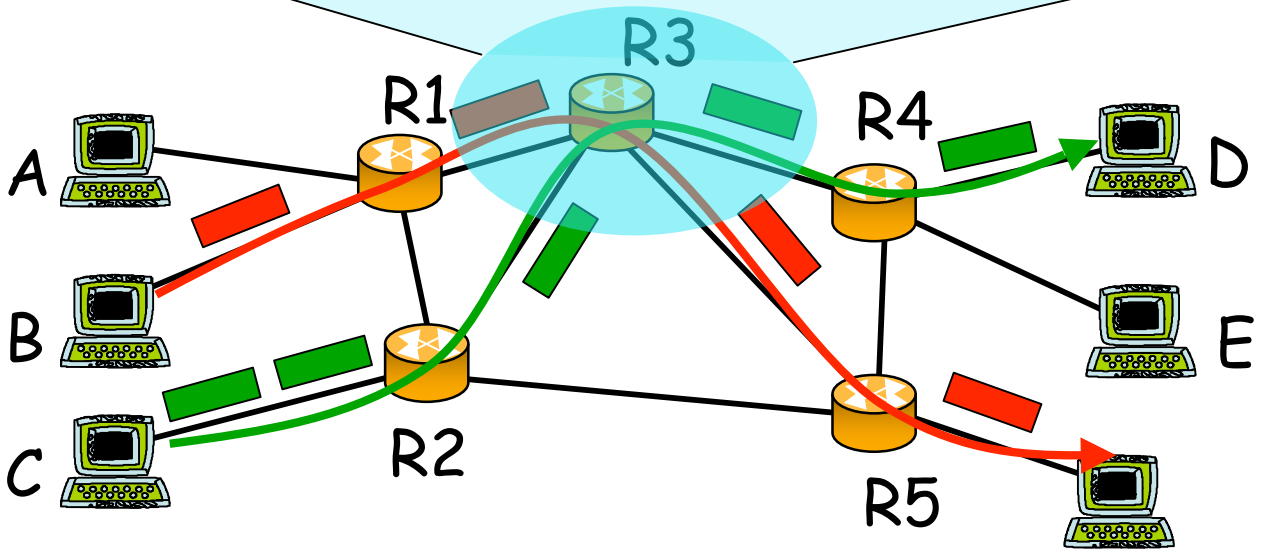
Virtual circuit principles

Connections &
Virtual circuits table

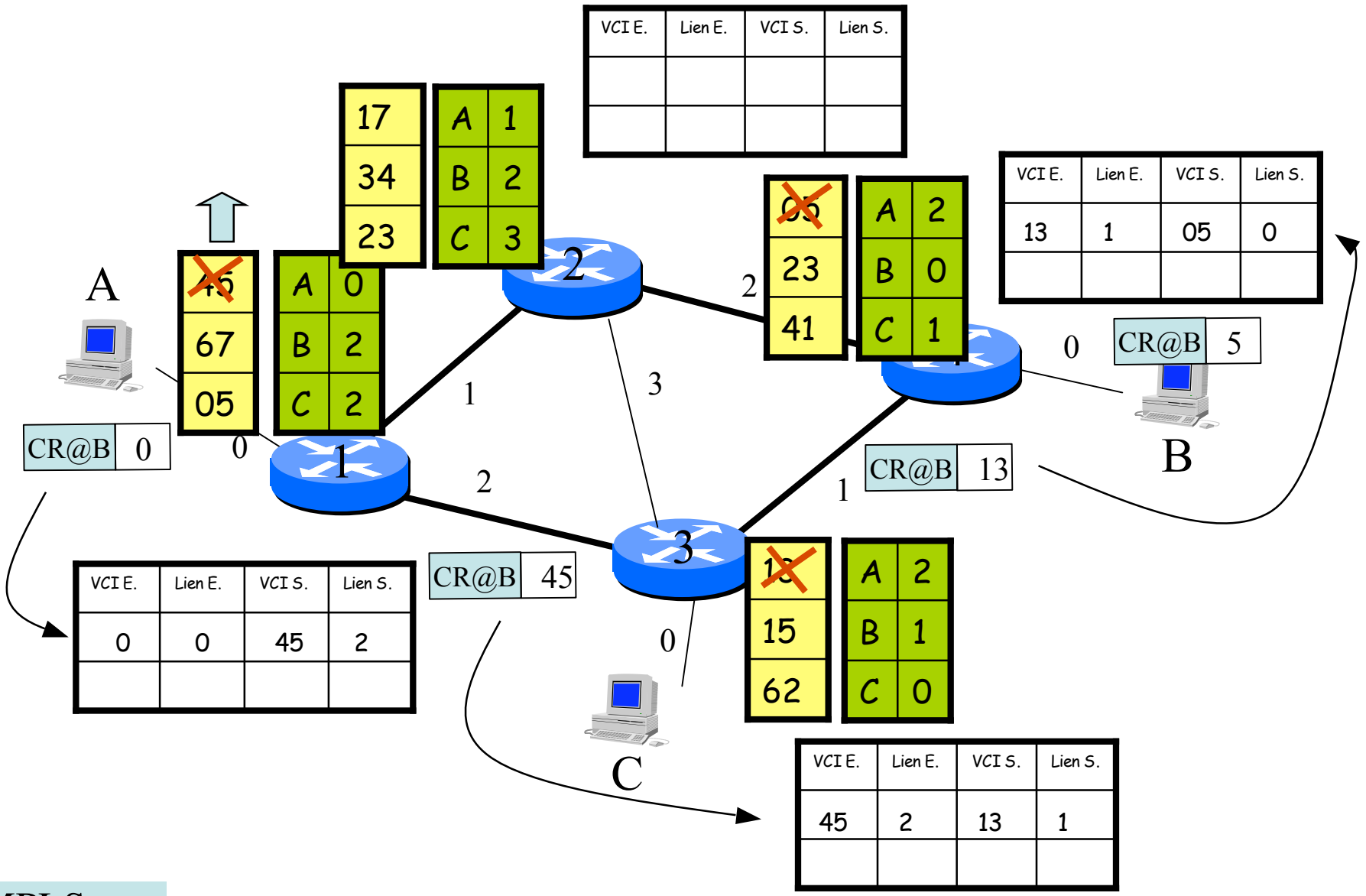
Label IN	Link IN	Label OUT	Link OUT
23	1	34	3
45	2	78	4



Virtual
Circuit
Switching

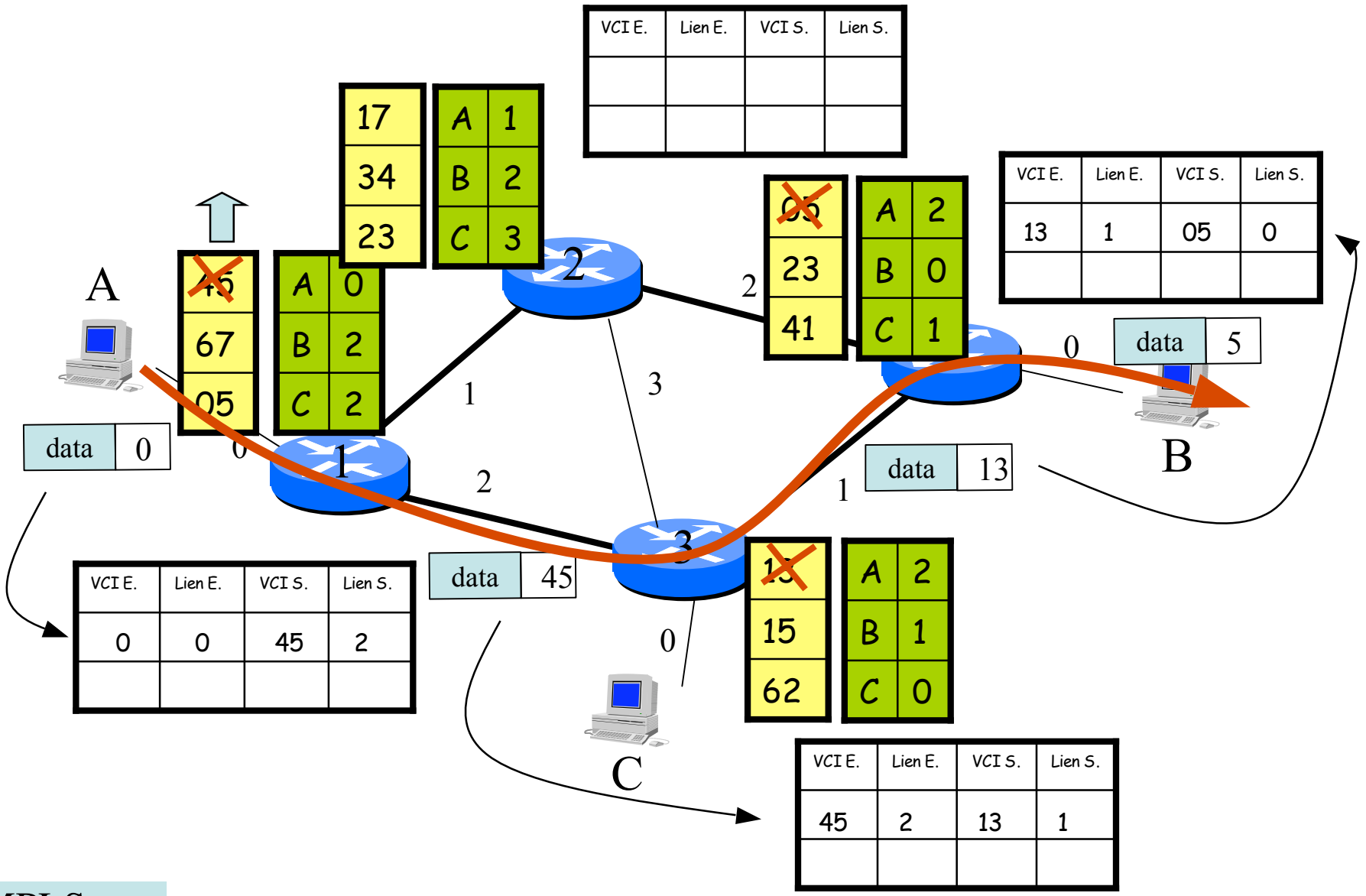


End-to-end operation (1)



MPLS

End-to-end operation (2)



Why virtual circuit?

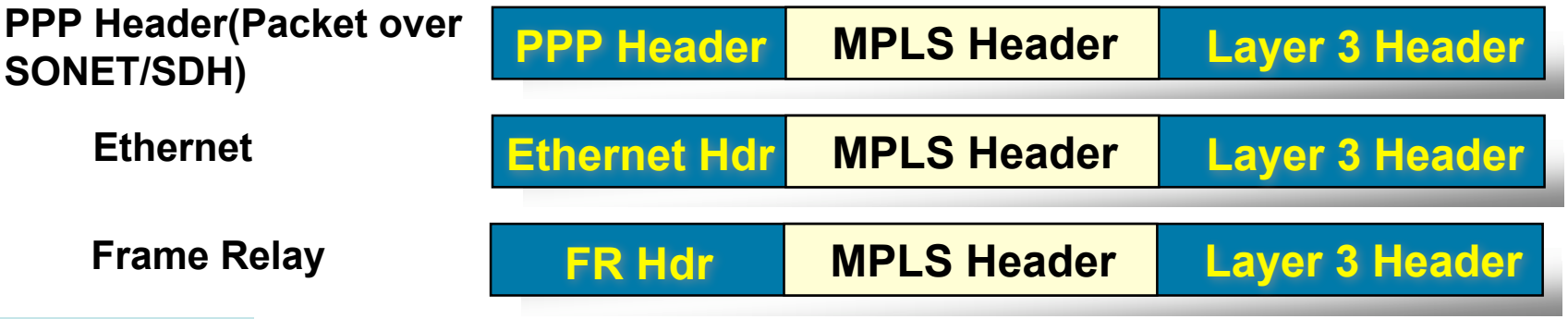
- Initially to speed up router's forwarding tasks: X.25, Frame Relay, ATM.

Now: Virtual circuits for traffic engineering!

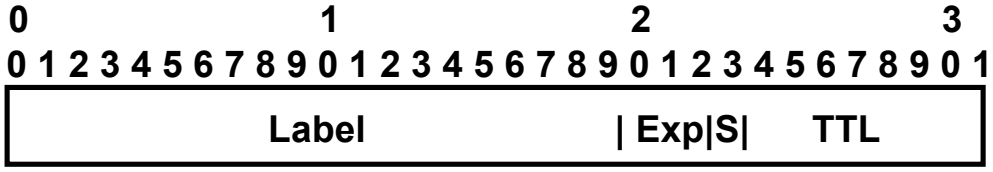
Virtual circuits in IP networks

- Multi-Protocol Label Switching
  - Fast: use label switching➔ LSR

  - Multi-Protocol: above link layer, below network layer
  - Facilitate traffic engineering



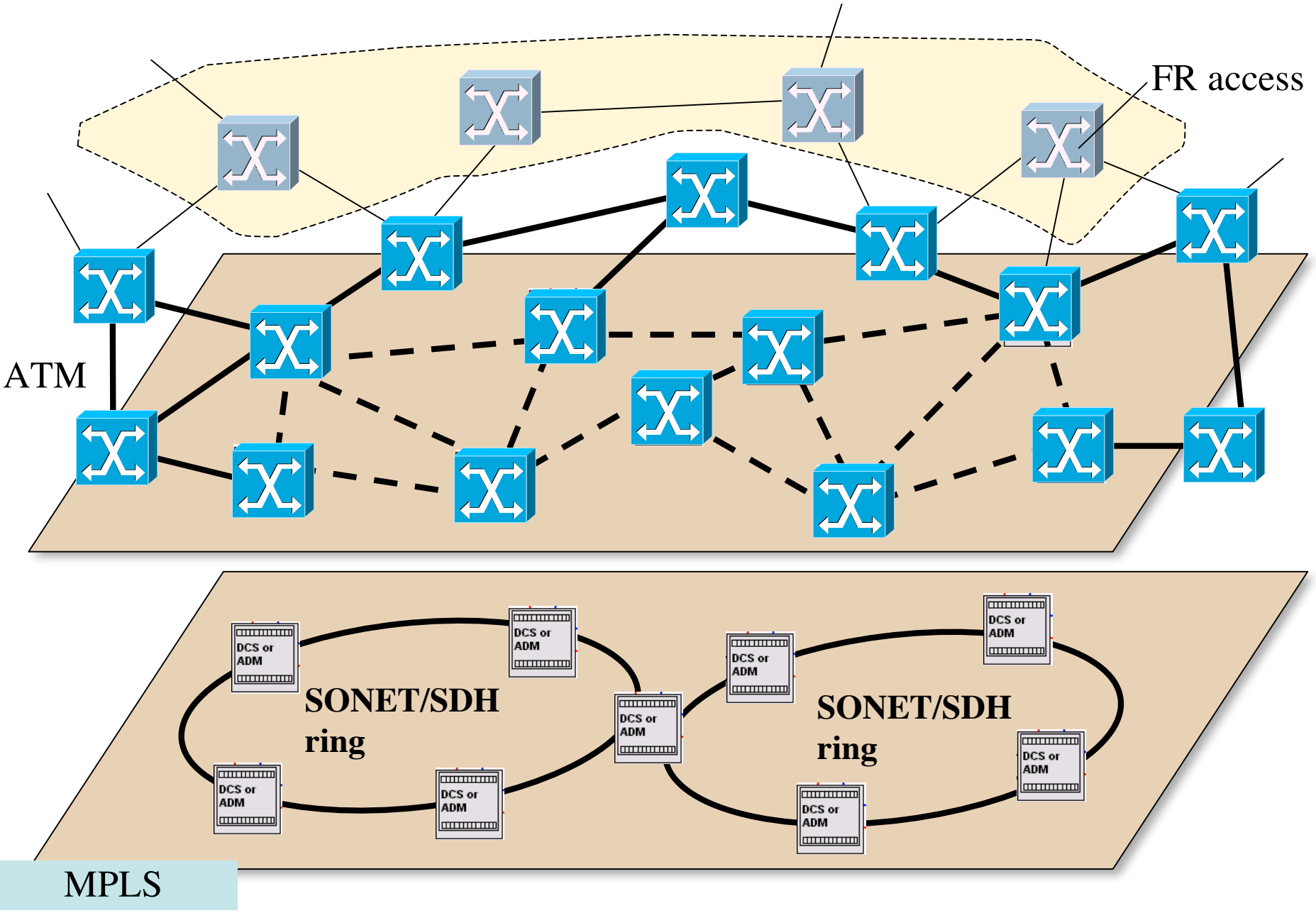
Label structure



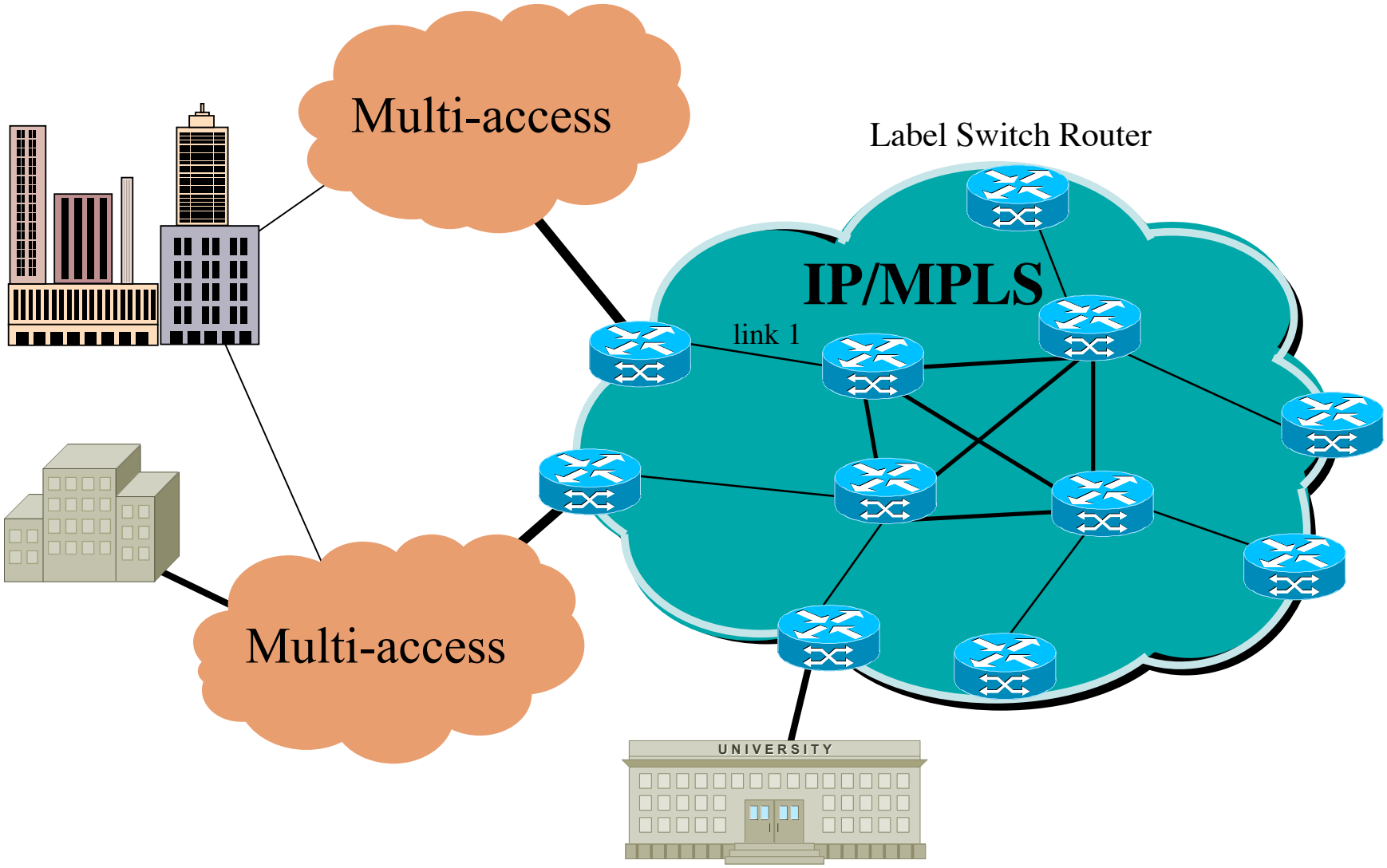
Label = 20 bits
Exp = Experimental, 3 bits
S = Bottom of stack, 1bit
TTL = Time to live, 8 bits

- More than one label is allowed -> Label Stack
- MPLS LSRs always forward packets based on the value of the label at the top of the stack

From multilayer networks…



…to IP/MPLS networks

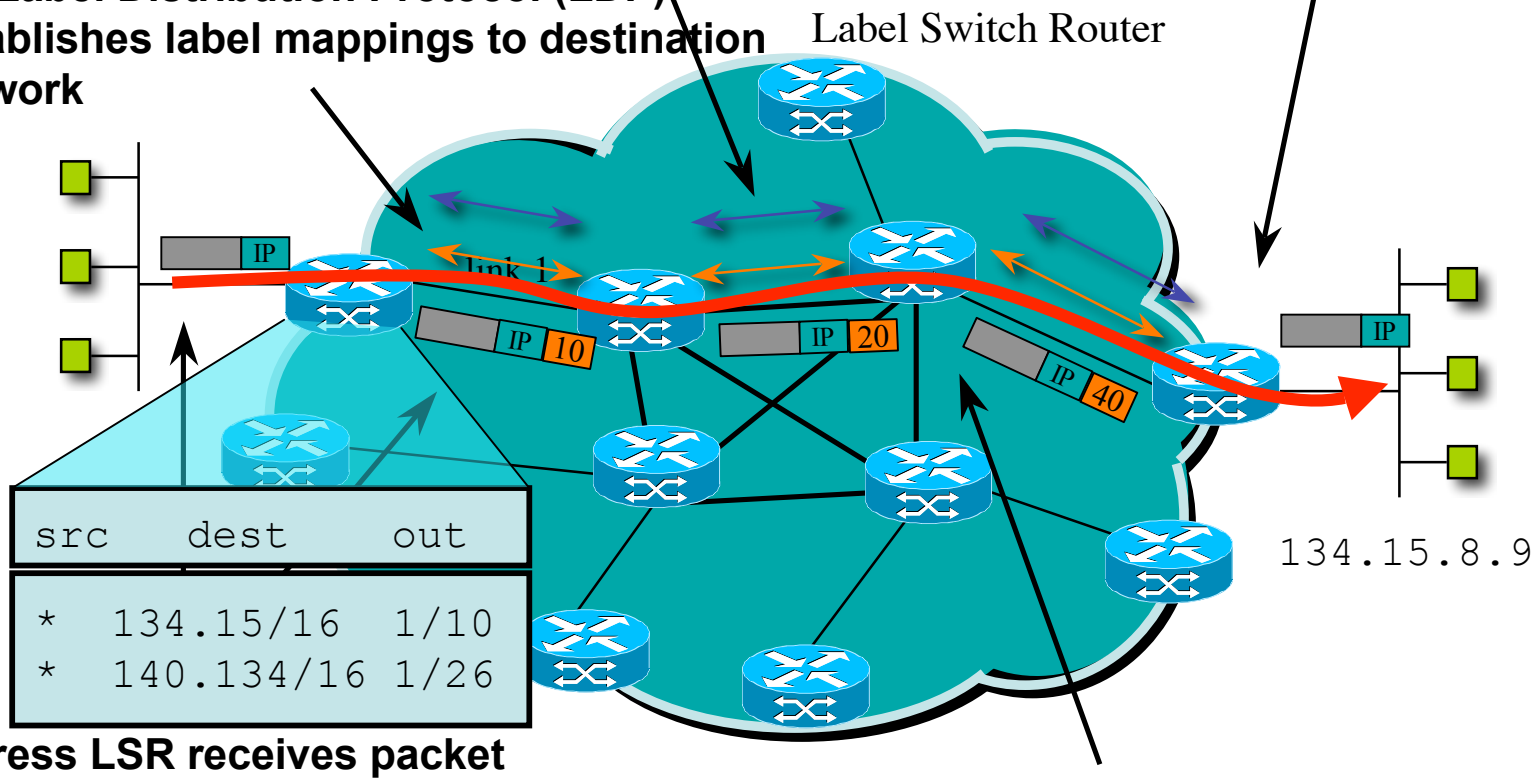


MPLS operation

1a. Routing protocols (e.g. OSPF-TE, IS-IS-TE) exchange reachability to destination networks

1b. Label Distribution Protocol (LDP) establishes label mappings to destination network

4. LSR at egress removes label and delivers packet

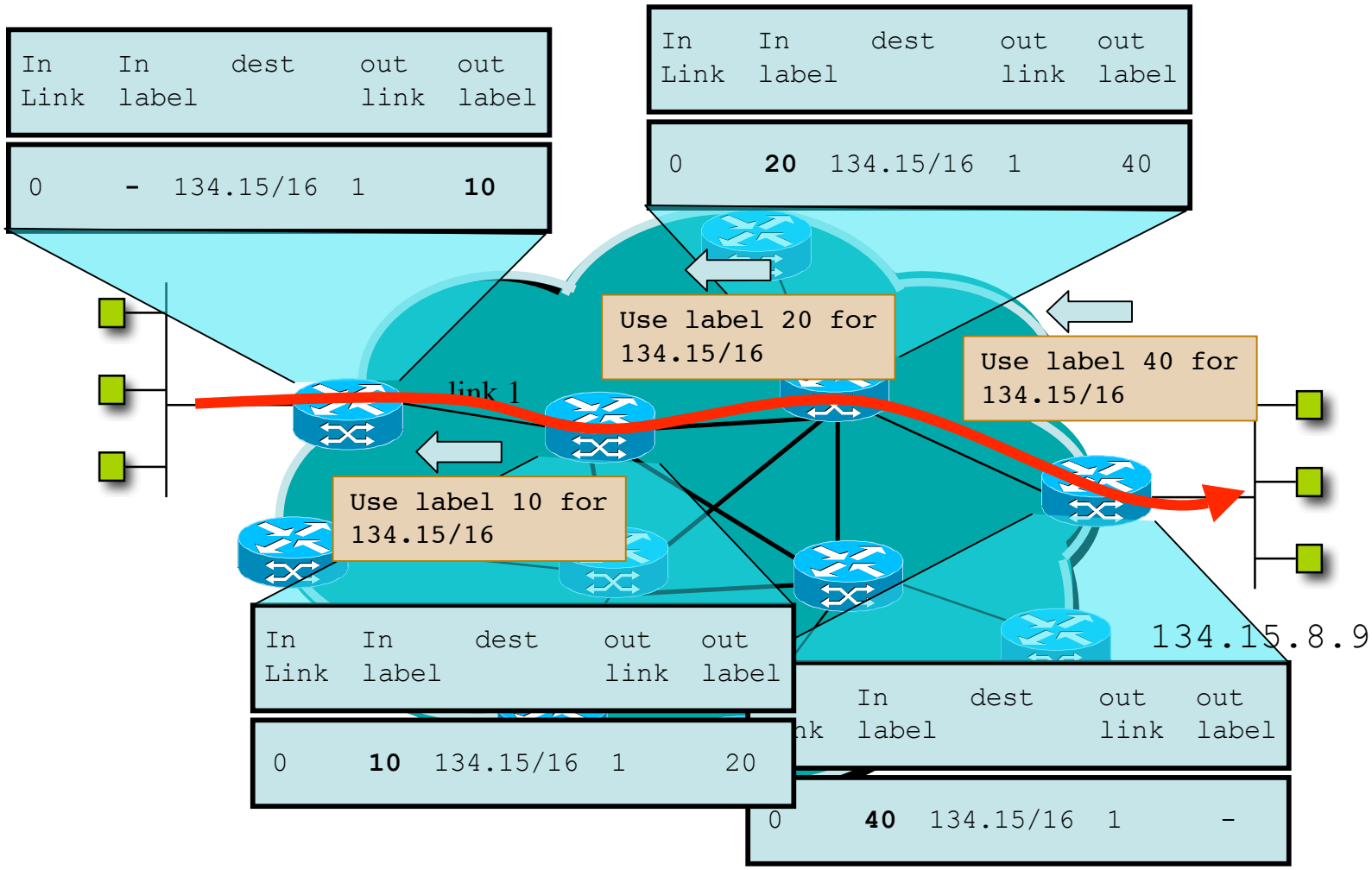


2. Ingress LSR receives packet and "label"s packets

3. LSR forwards packets using label switching

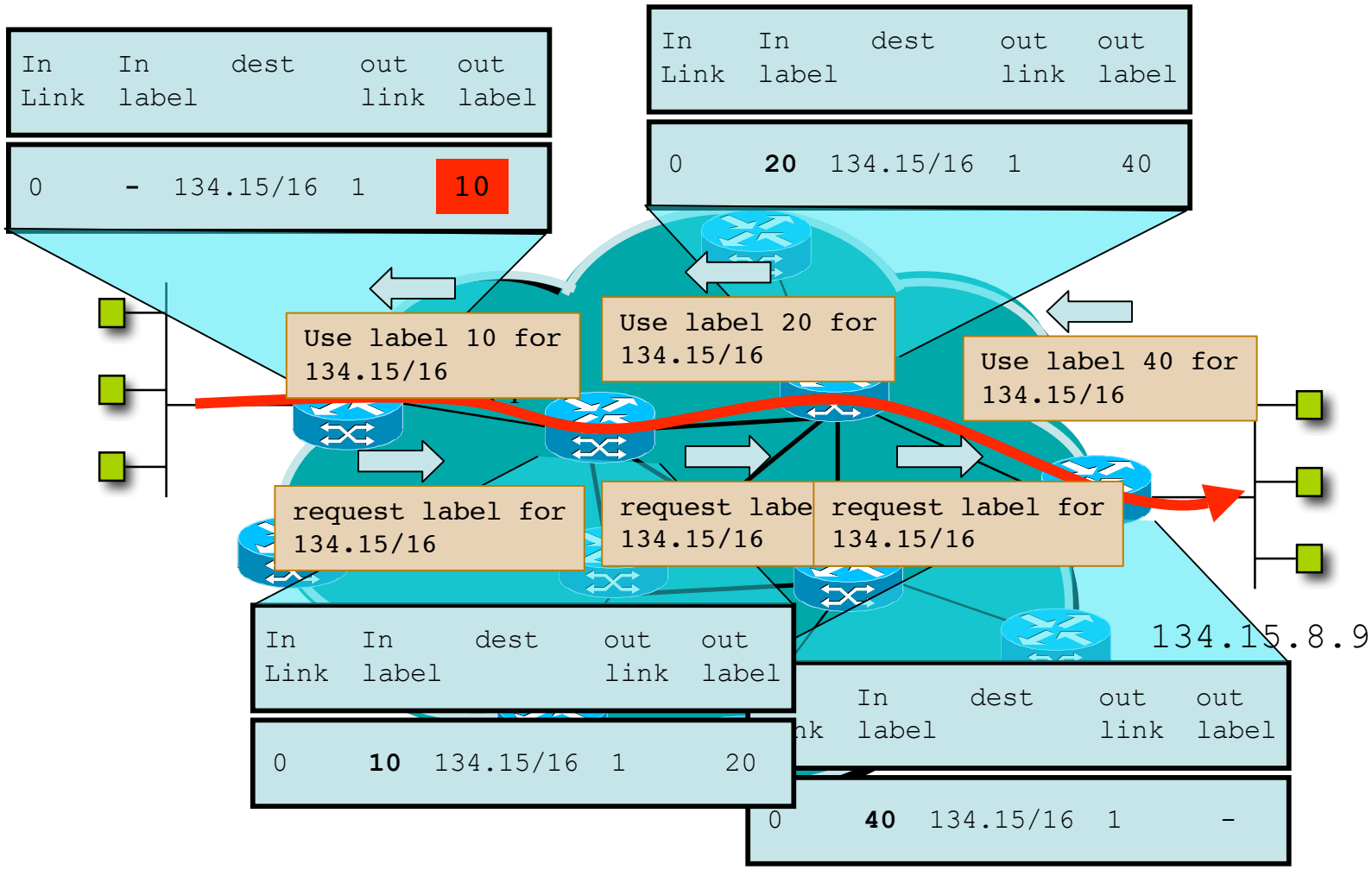
Source Yi Lin, modified C. Pham

Label Distribution



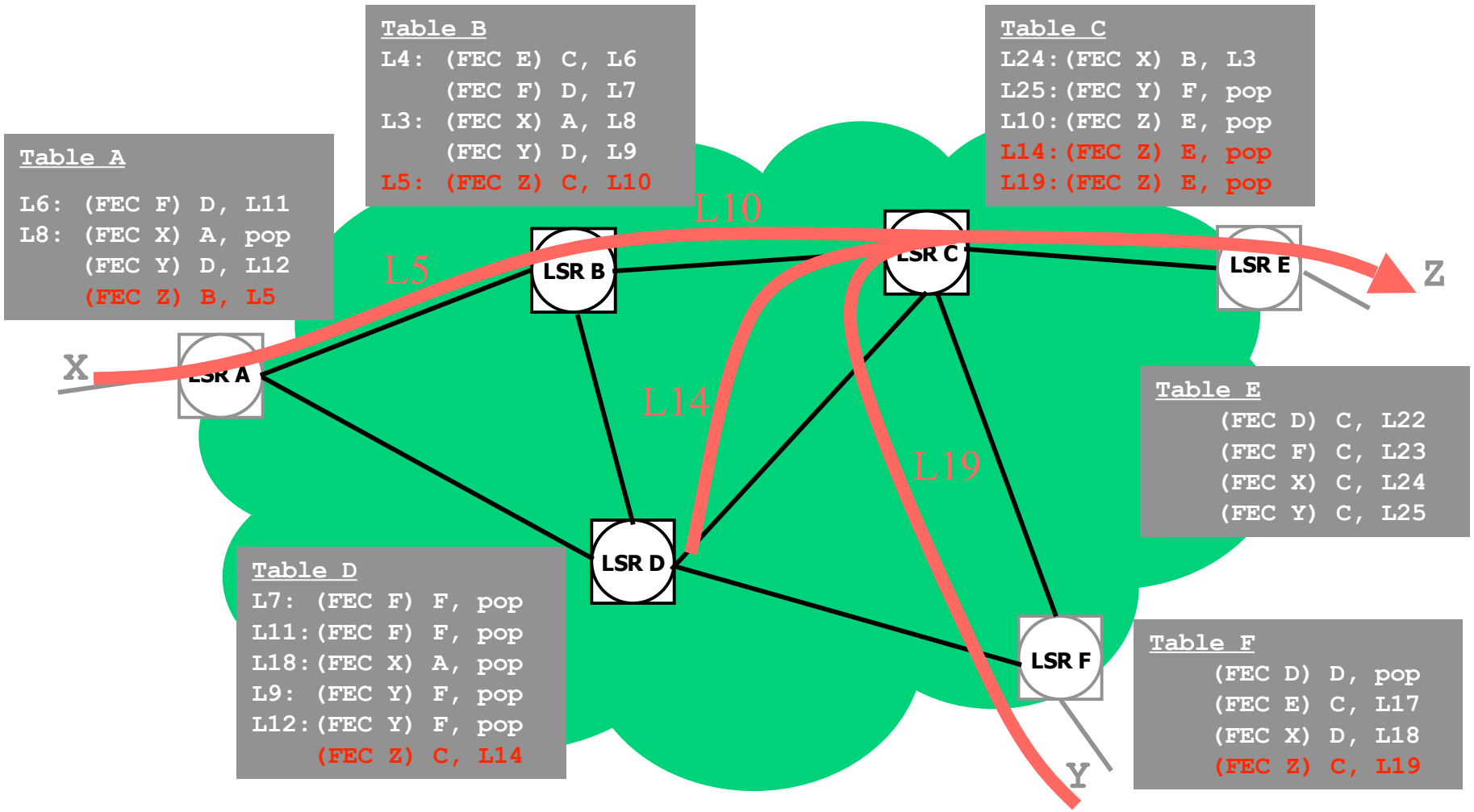
Unsolicited downstream label distribution

Label Distribution (con't)



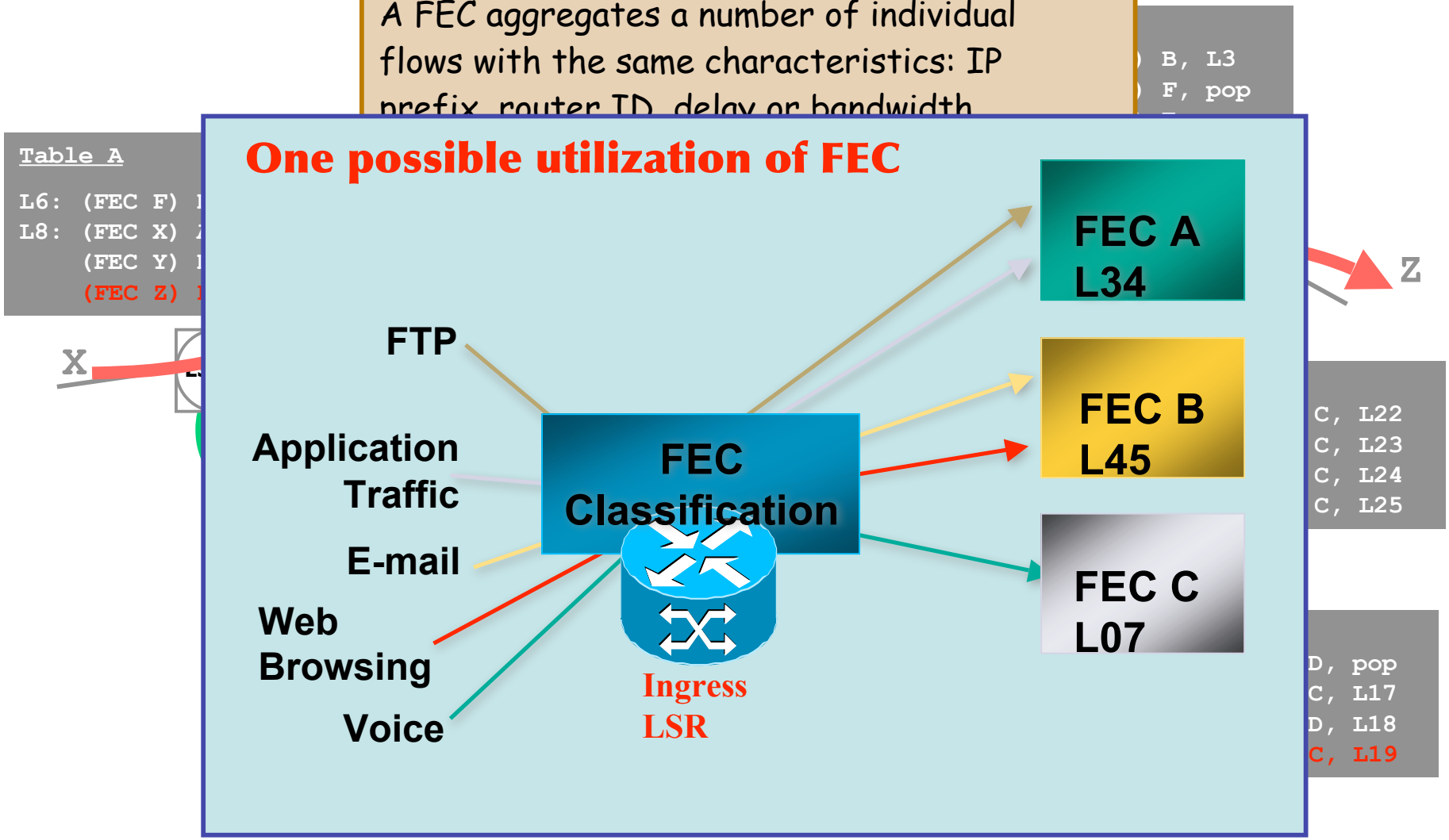
On-demand downstream label distribution

Forwarding Equivalent Class: high-level forwarding criteria



Forwarding Equivalent Class

A FEC aggregates a number of individual flows with the same characteristics: IP prefix, router ID, delay or bandwidth



Label & FEC

- Independent LSP control
  - An LSR binds a label to a FEC, whether or not the LSR has received a label from the next-hop for the FEC
  - The LSR then advertises the label to its neighbor

- Ordered LSP control
  - An LSR only binds and advertises a label for a particular FEC if:
    - it is the egress LSR for that FEC or
    - it has already received a label binding from its next-hop

Label Distribution Protocols

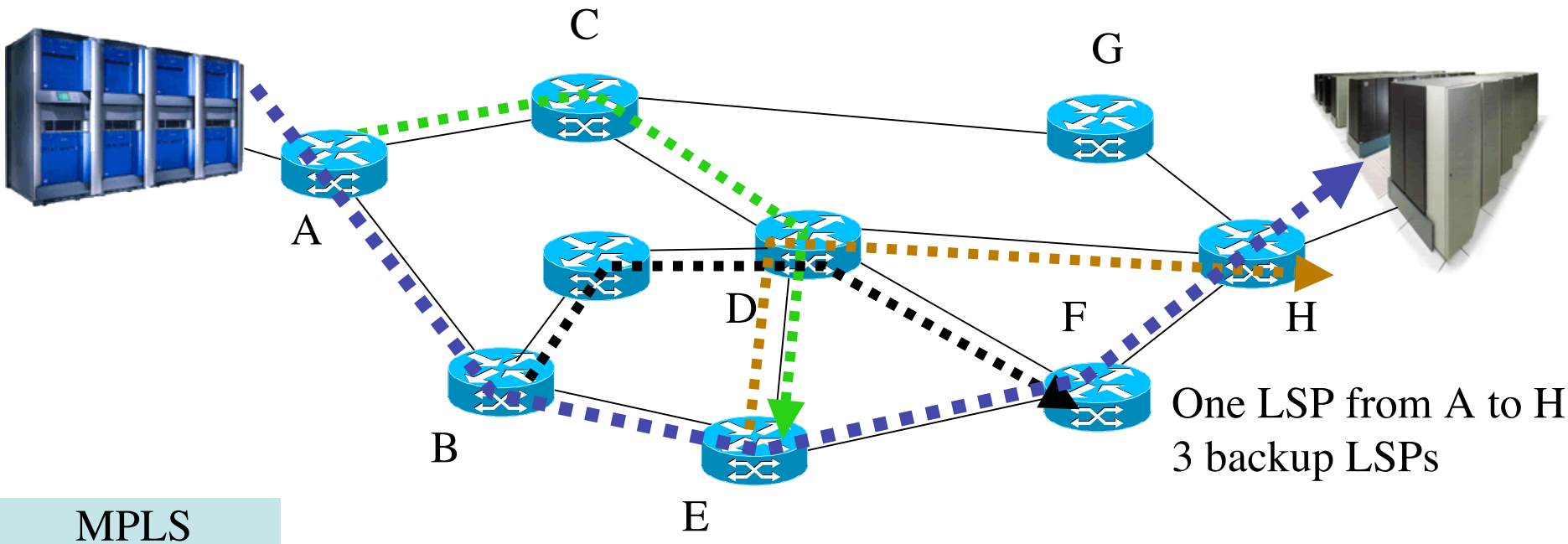
- LDP
  - Maps unicast IP destinations into labels
- RSVP, CR-LDP
  - Used in traffic engineering
- BGP
  - External labels (VPN)
- PIM
  - For multicast states label mapping

MPLS FastReroute

- Intended to provide SONET/SDH-like healing capabilities
- Selects an alternate route in tenth of ms, provides path protection
- Traditional routing protocols need minutes to converge!
- FastReroute is performed by maintaining backup LSPs

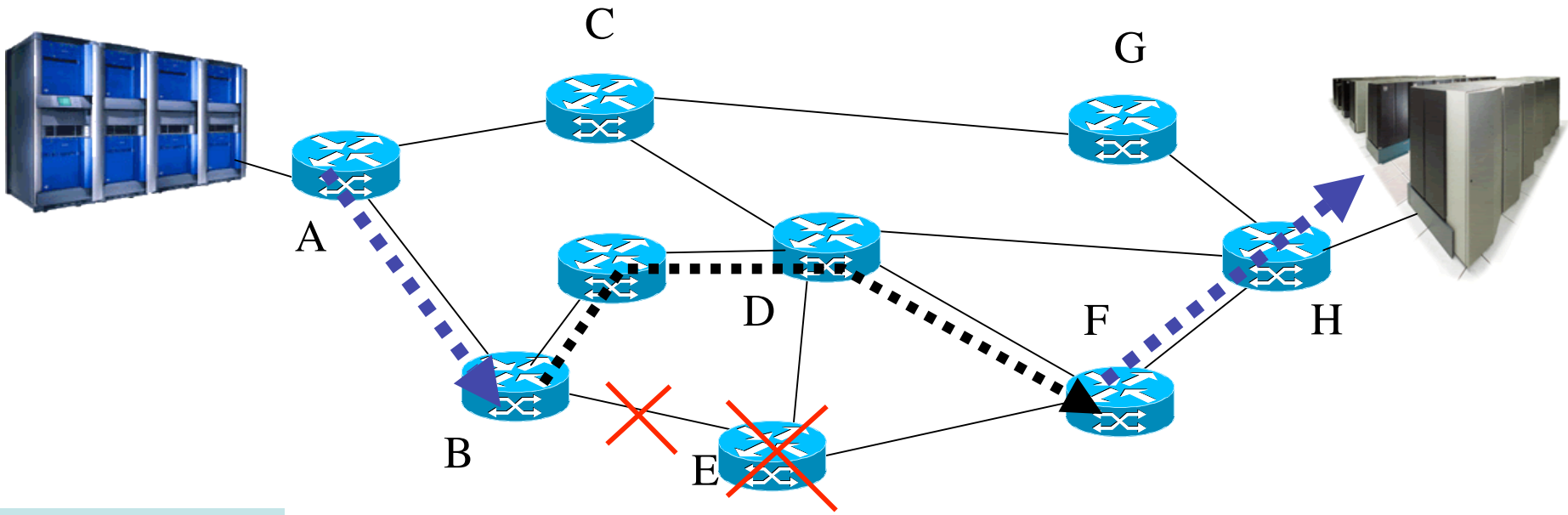
Backup LSPs

- One-to-one
- Many-to-one: more efficient but needs more configurations



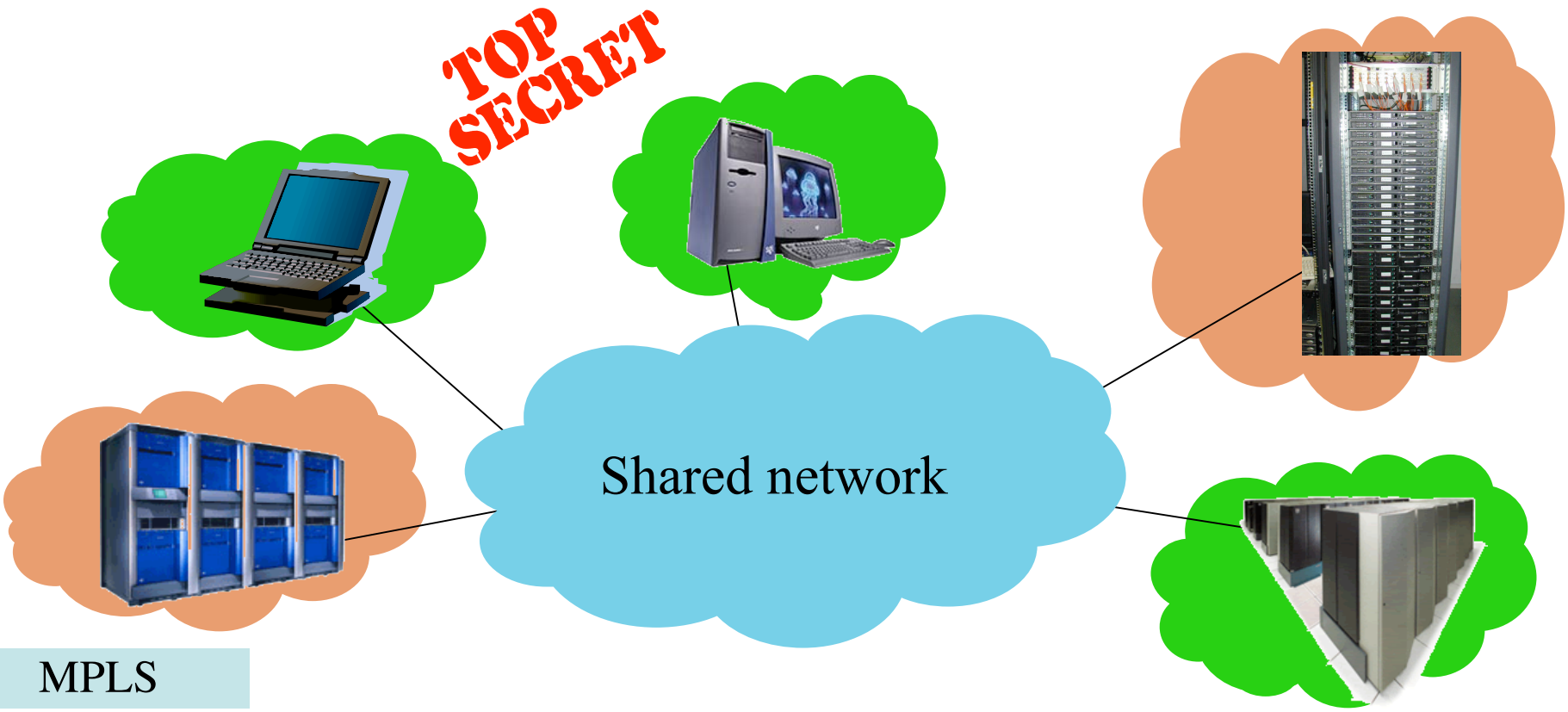
Recovery on failures

- Suppose E or link B-E is down…
- B uses detour around E with backup LSP



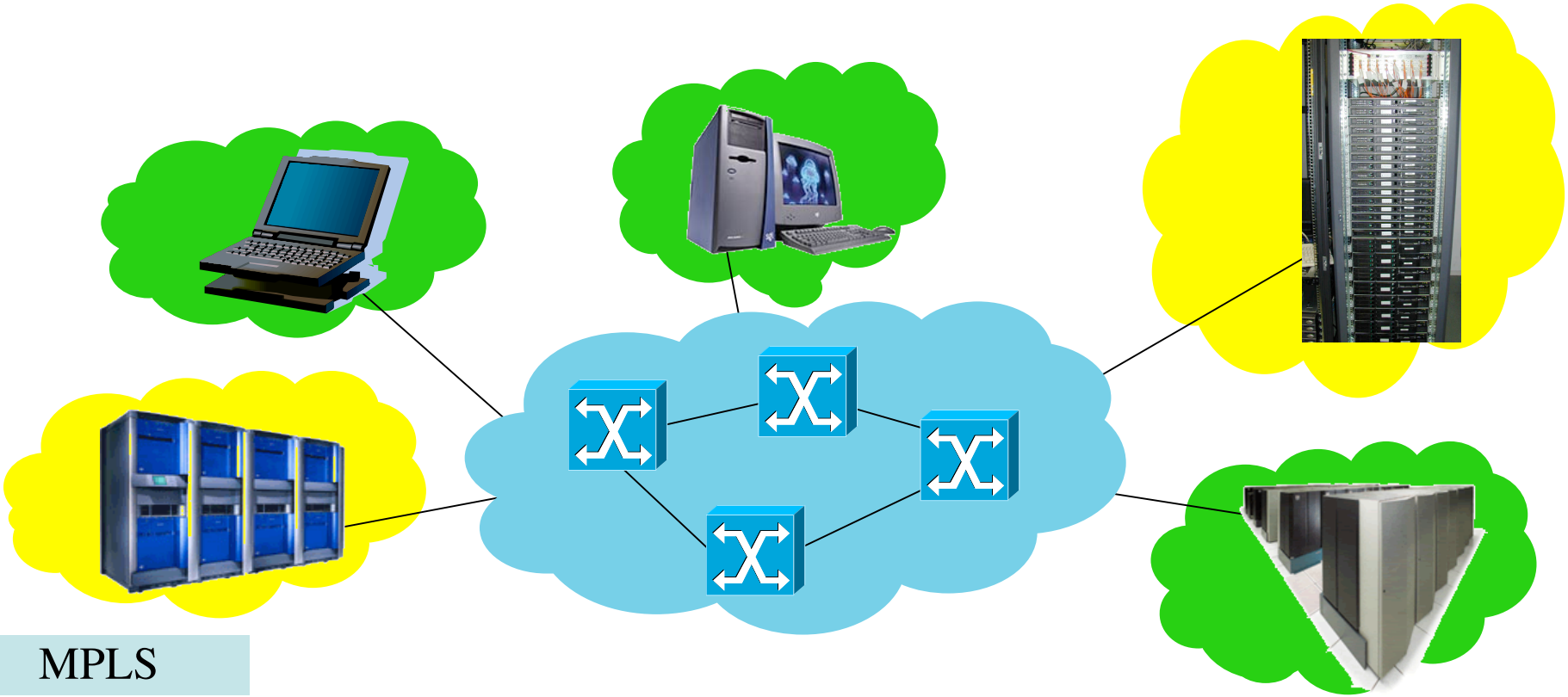
Virtual Private Networks

- Virtual Private Networks: build a secure, confidential communication on a public network infrastructure using routing, encryption technologies and controlled accesses



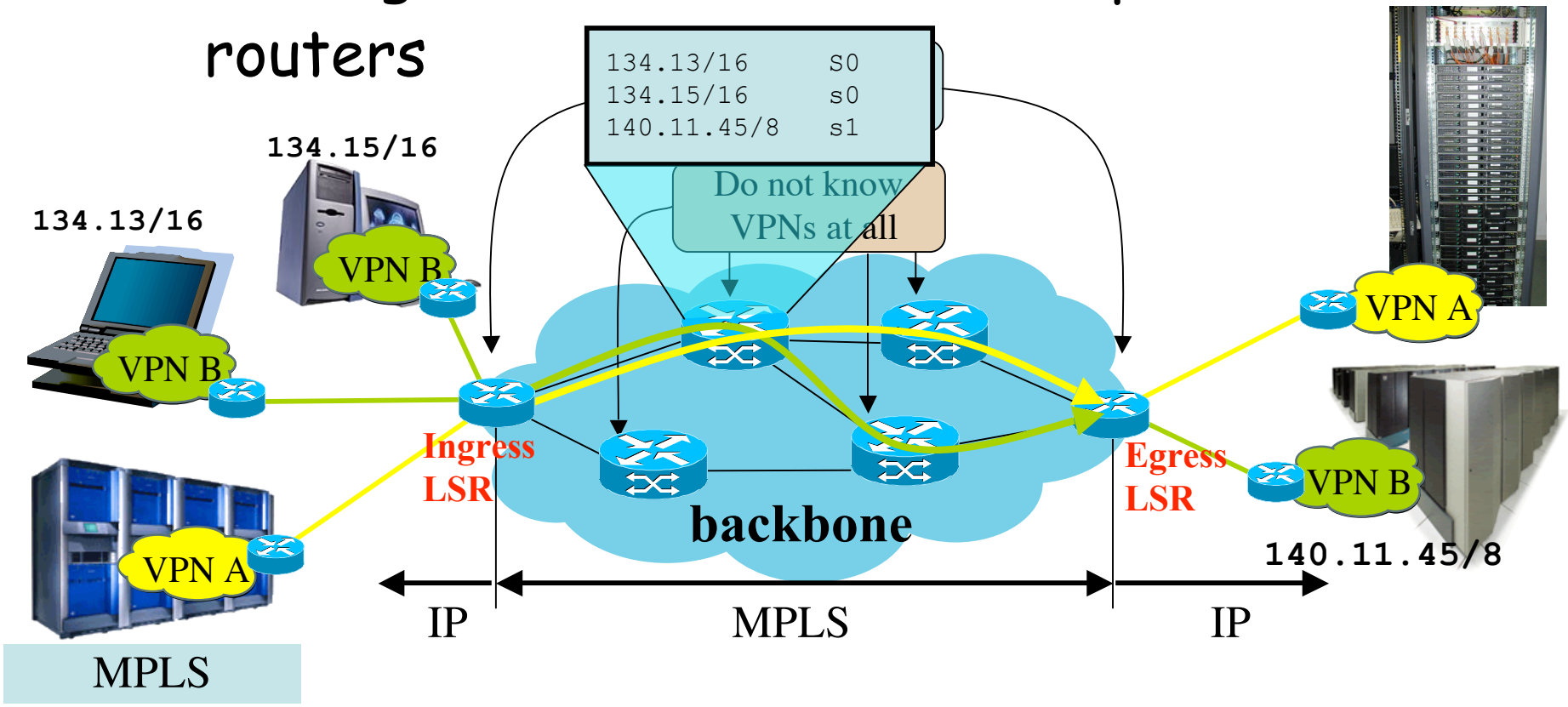
The traditional way of VPN

- Uses leases lines, Frame Relay/ATM infrastructures…

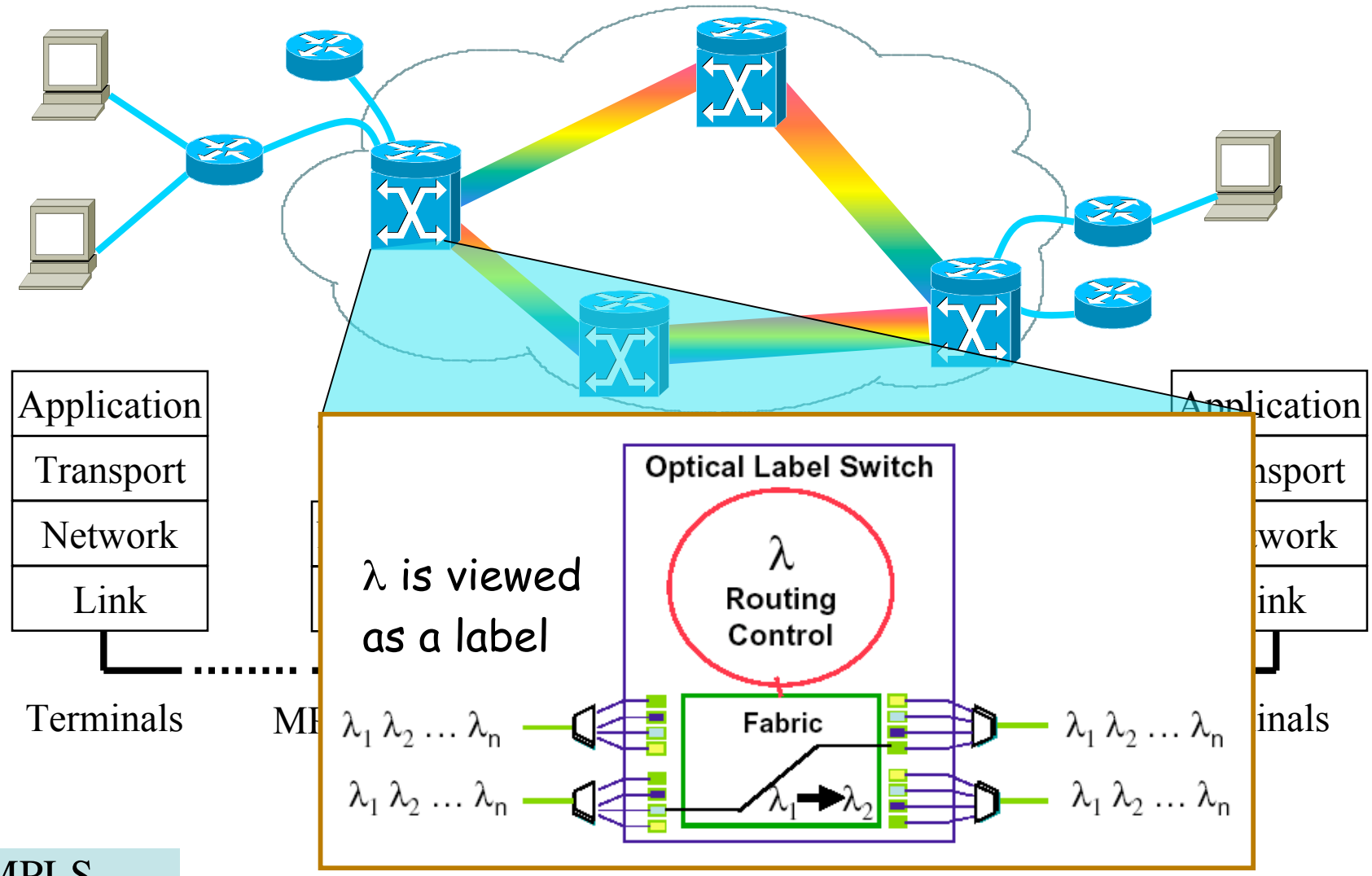


IP/MPLS & VPN

- IP/MPLS replace dedicated networks
- MPLS reduces VPN complexity by reducing routing information needed at provider's routers



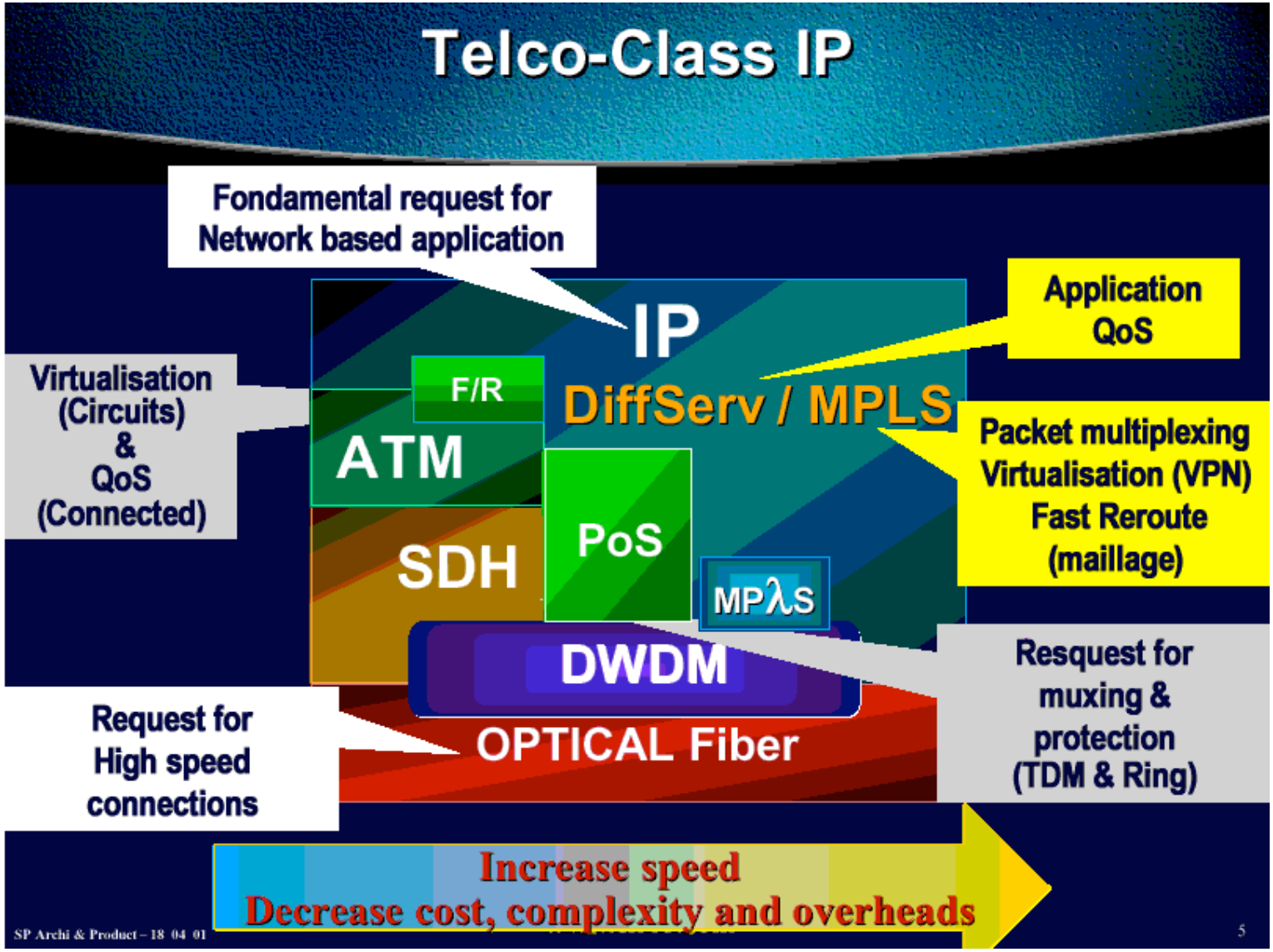
MPλS: MPLS+optical



MPLS

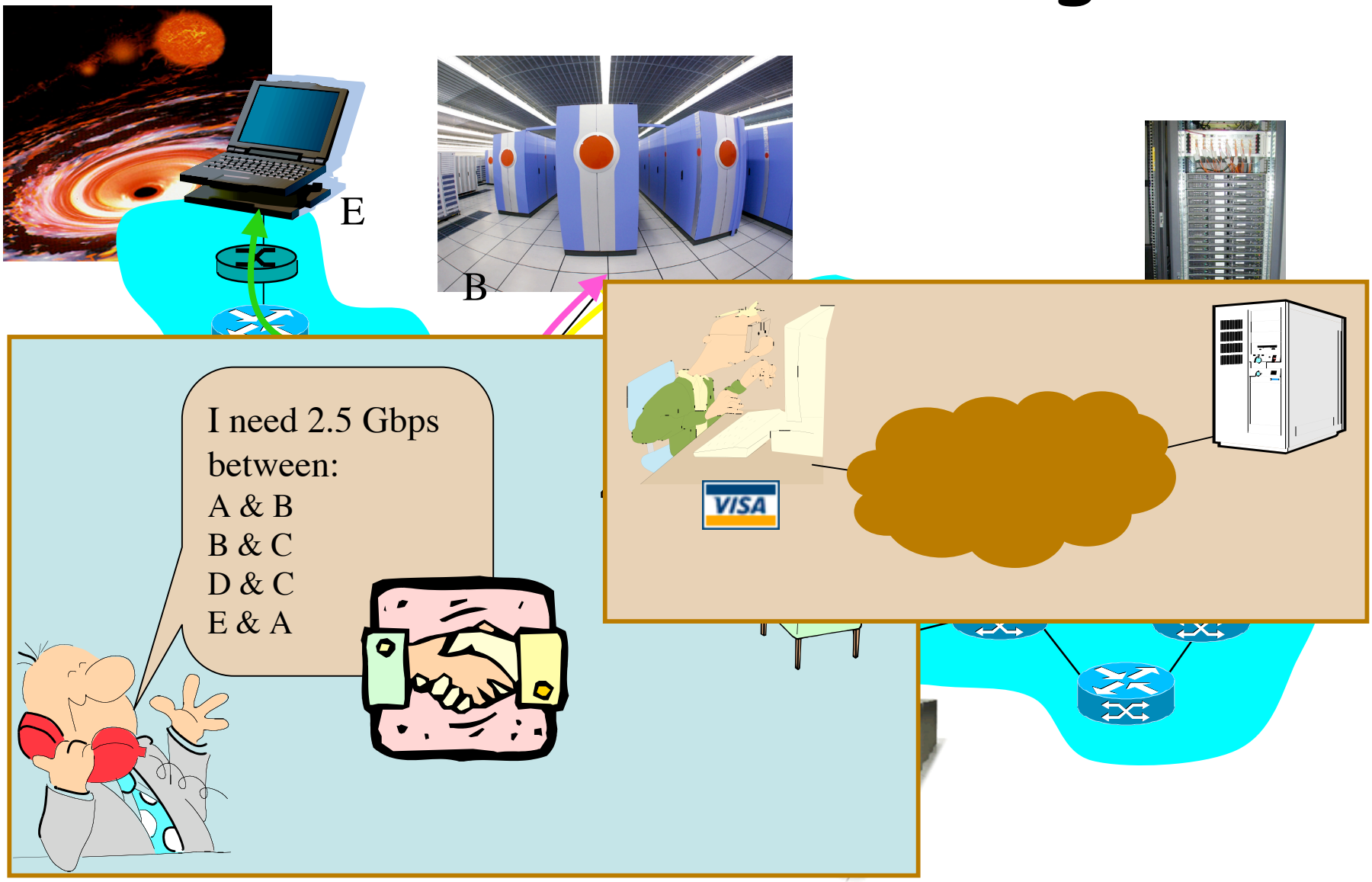
Towards IP/MPLS/DWDM

From cisco



MPLS

Ex: MPLS circuits on grids



Ex: MPLS FEC for the grid