



Capítol 3. Xarxes troncals

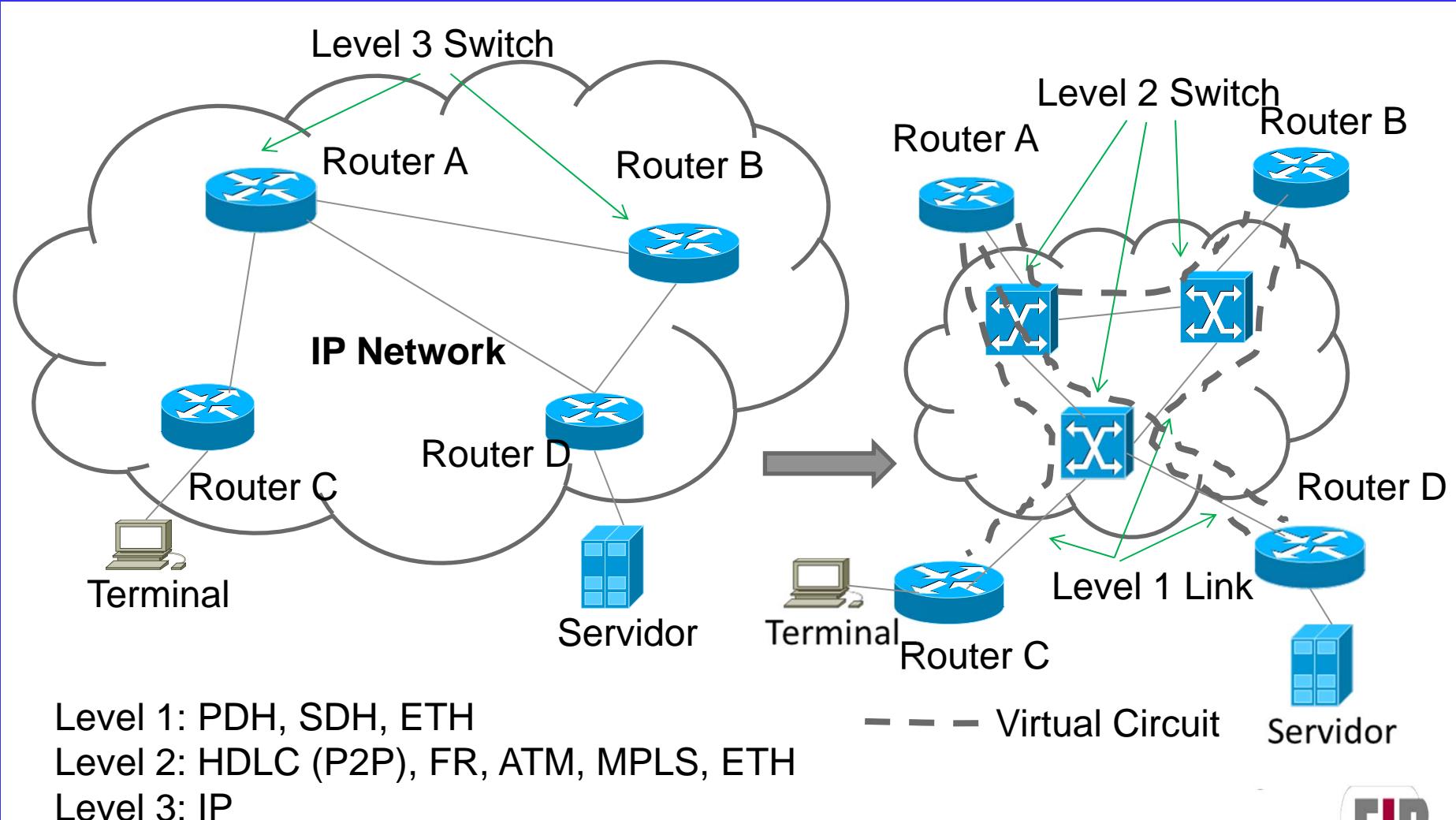
- 3.1 Commutació de trames
- 3.2 Commutació de cel·les
- 3.3 Commutació d'etiquetes
- 3.4 Carrier Ethernet
- 3.5 Control de la congestió

Core (IP) network

Departament d'Arquitectura de Computadors

Level 3: Datagram

Level 2: Virtual Circuit





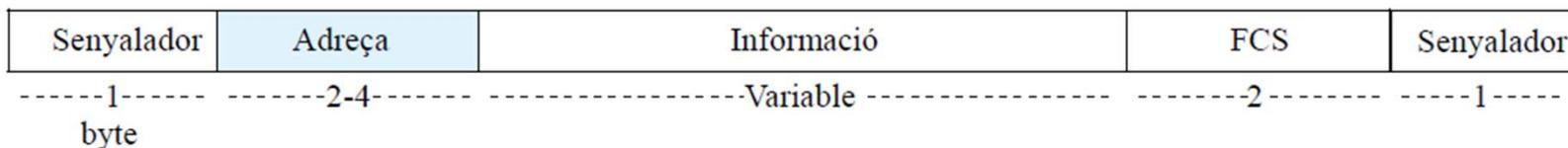
3.1 Commutació de trames Frame Relay

Source Book: Data and Computer Communication Ed 8. W. Stallings Cap. 10.7

Frame Relay

- Developed to take advantage of high data rates and low error rates
- Operates at data rates of up to 2 Mbps
- Key to achieving high data rates is to strip out most of the overhead involved with error control

Frame relay: LAPF Core



(a) Format de trama

8	7	6	5	4	3	2	1
DLCI superior						C/R	EA 0
DLCI inferior						FECN BECN	DE EA 1

(b) Camp d'adreça: 2 bytes (per defecte)

8	7	6	5	4	3	2	1
DLCI superior						C/R	EA 0
DLCI						FECN BECN	DE EA 0
DLCI						EA 0	
Control DLCI inferior o DL-CORE						D/C	EA 1

(d) Camp d'adreça: 4 bytes

8	7	6	5	4	3	2	1
DLCI superior						C/R	EA 0
DLCI						FECN BECN	DE EA 0
Control DLCI interior o DL-CORE						D/C	EA 1

(c) Camp d'adreça: 3 bytes

- EA Bit d'estensió de camp d'adreça
- C/R Bit d'ordre/resposta
- FECN Notificació de congestió explícita cap endavant
- BECN Notificació de congestió explícita cap enrere
- DLCI Identificador de connexió d'enllaç de dades
- D/C Indicador de control DLCI o DL-CORE
- DE Idoneïtat del descart



3.2 Commutació de cel·les

ATM

Source: Data and Computer Communication Ed 8. W. Stallings Cap. 11
Data and Computer Communication Ed 10. W. Stallings Cap. 9.6

Asynchronous Transfer Mode (ATM)

- A switching and multiplexing technology that employs small, fixed-length packets called cells
- A fixed-size packet ensures function could be carried out efficiently, with little delay variation
- Small cell size supports delay-intolerant interactive voice service with a small packetization delay
- Designed to provide the performance of a circuit-switching network and the flexibility and efficiency of a packet-switching network
- Standardization effort was to provide a powerful set of tools for supporting a rich QoS capability and a powerful traffic management capability

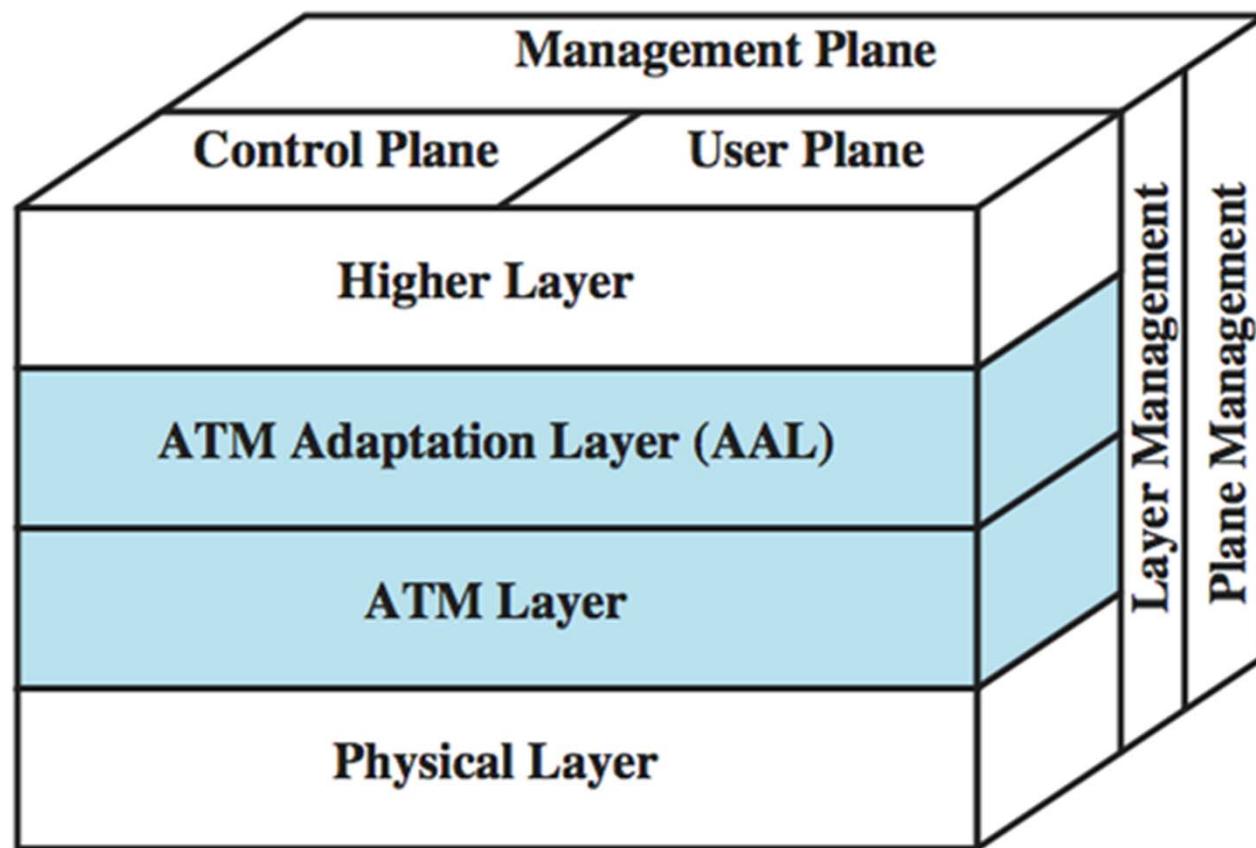
ATM

- Commonly used by telecommunications providers to implement wide area networks
- Used by many DSL implementations
- Used as a backbone network technology in numerous IP networks
- Multiprotocol Label Switching (MPLS) has reduced the role for ATM

Virtual Channel Connection (VCC)

- Logical connection in ATM
- Analogous to a virtual circuit
- Basic unit of switching in an ATM network
- Set up between two end users through the network, and a variable-rate, full duplex flow of fixed-size cells is exchanged over the connection
- Also used for user-network exchange and network-network exchange

Protocol Architecture

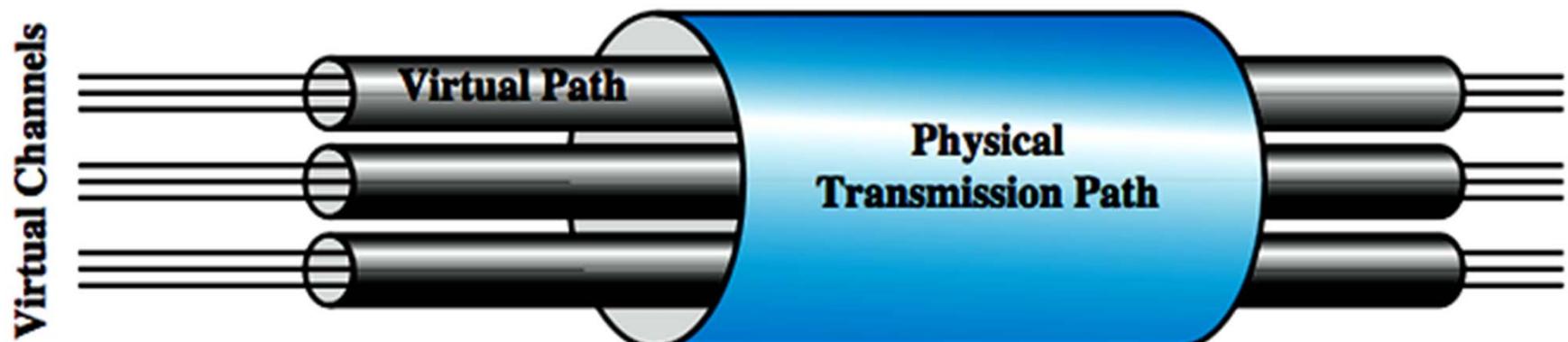


Delay in ATM networks

- End to end delay $R = Rp + Rt$
- Rp (packet delay) = $48 \times 8 / V_{ts}$
- Rt (transfer delay) = $Tt + Tp + W$
 - Tt (transmission time) $\sum t_t$ ($t_t = 53 \times 8 / V_{tn}$)
 - Tp (propagation time) $\sum t_p$ ($t_p = d / V_p$)
 - W (queue waiting time) $\sum w$ ($w = n \times t_t$)

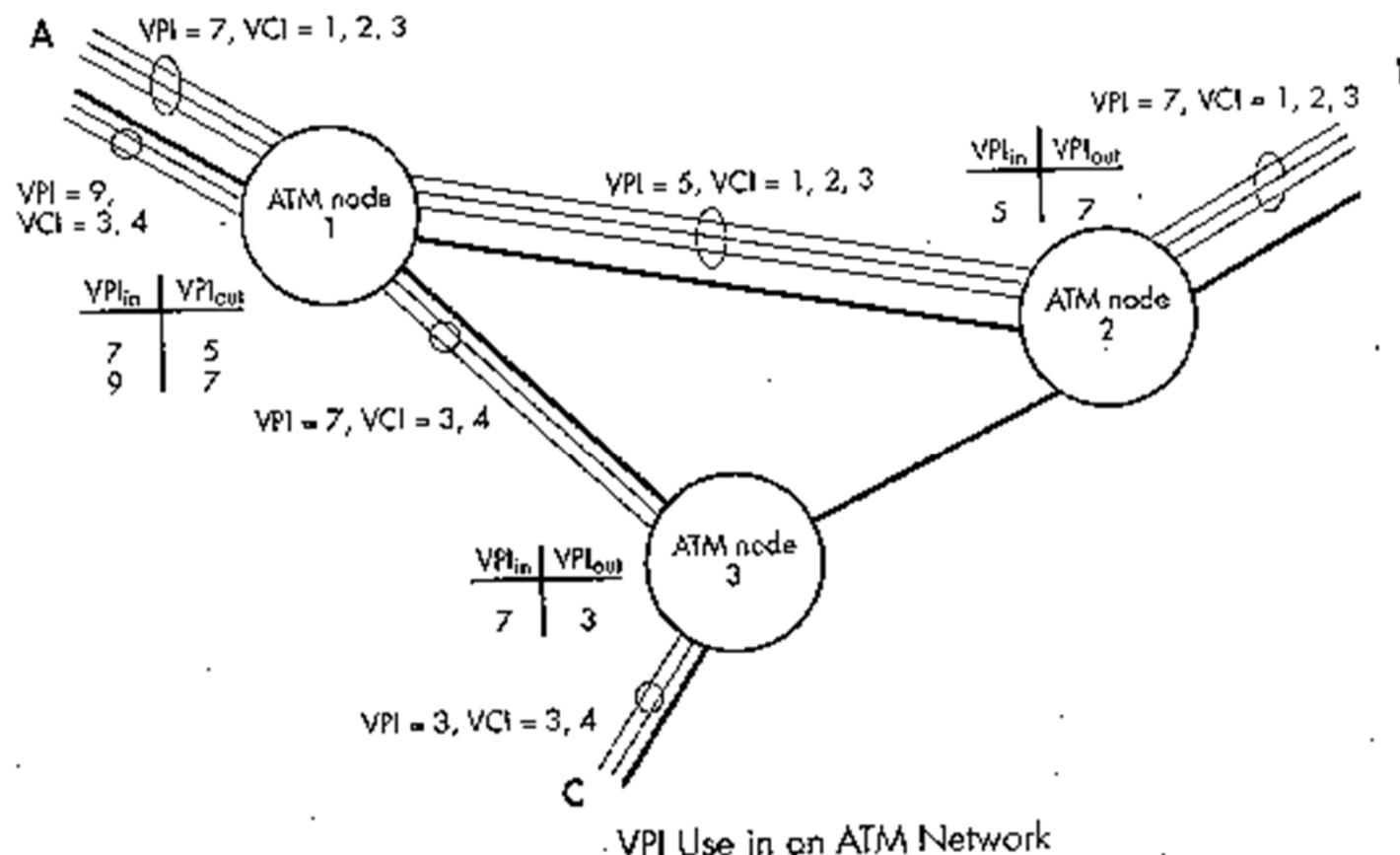
ATM Virtual Path Connection

- virtual path connection (VPC)
 - bundle of VCC with same end points

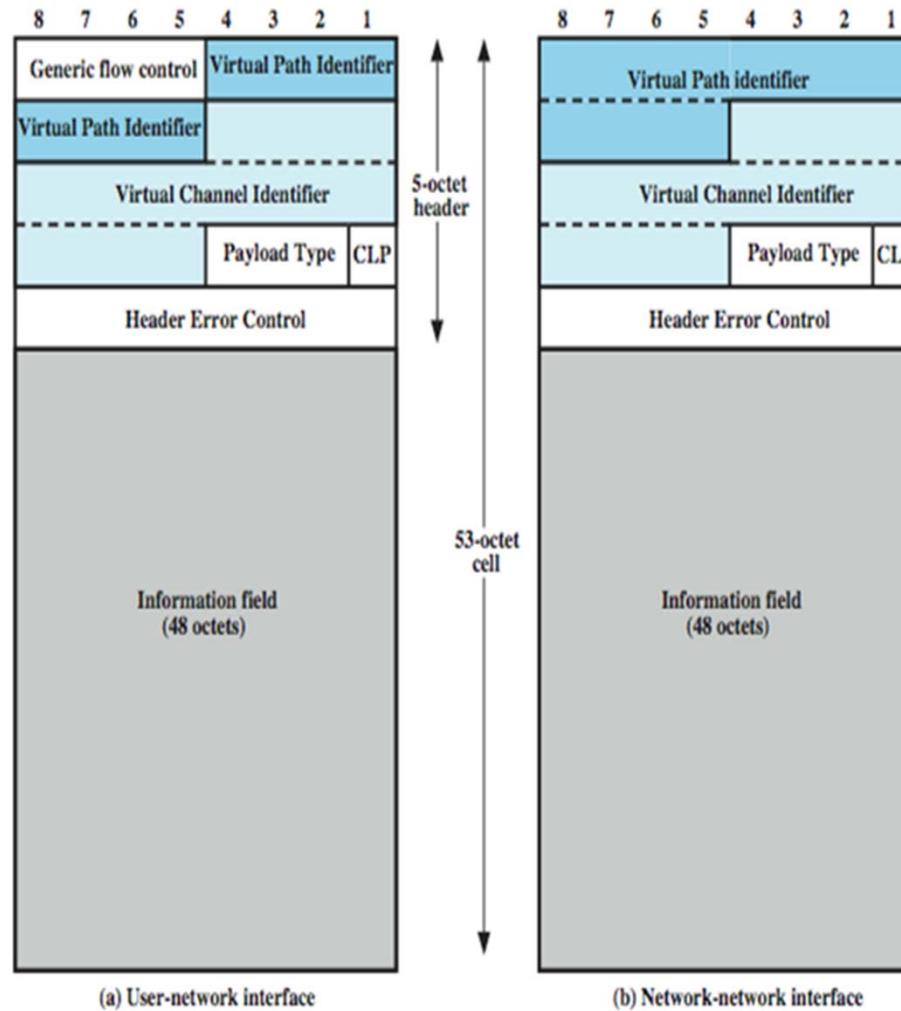


Virtual Channel and Virtual Path

- VPI and VCI relation



ATM Cells

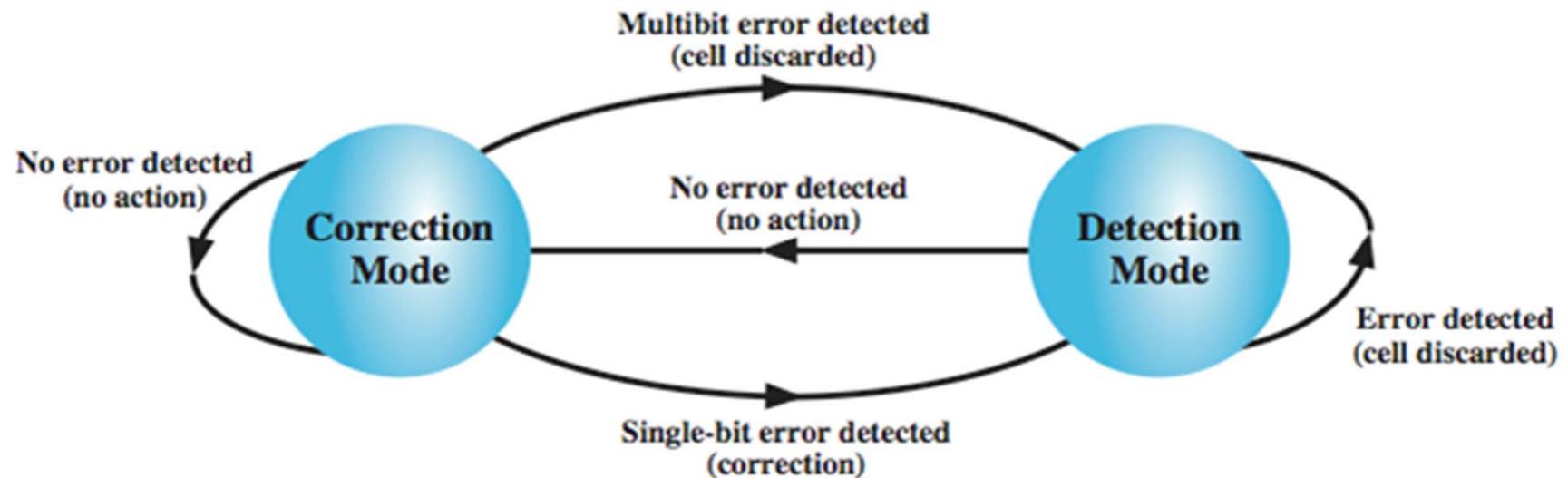


ATM Cells Format

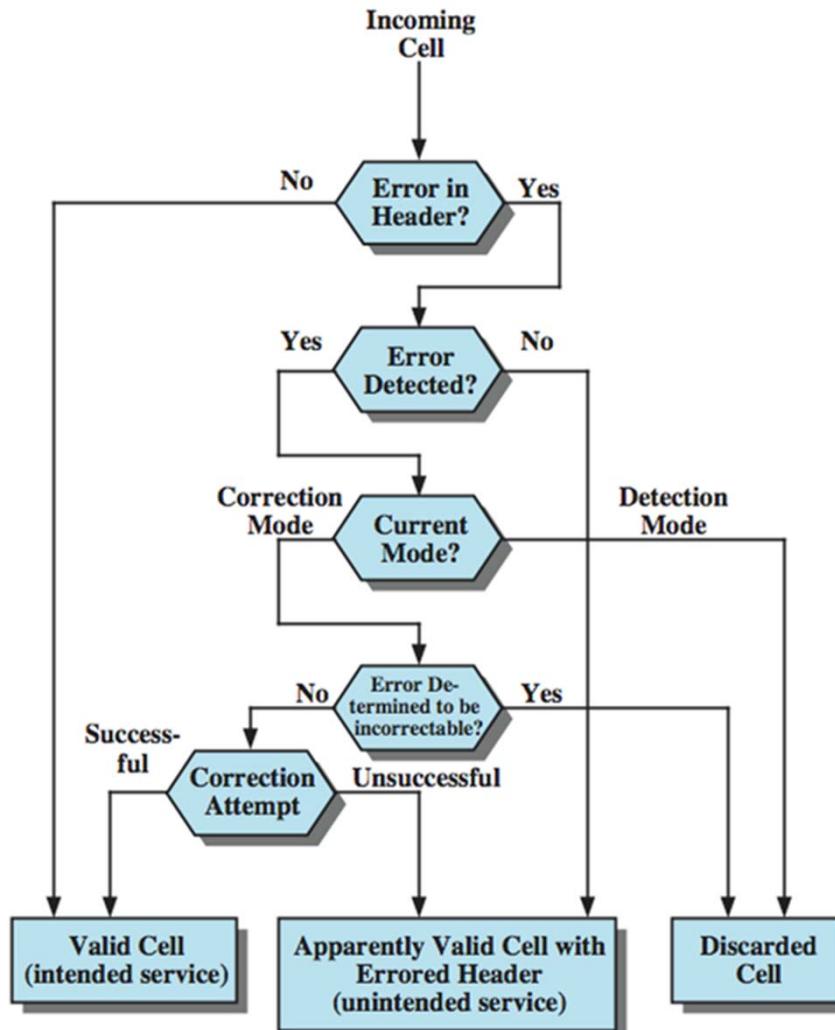
- Payload Type indicator

Payload type	Meaning
000	User data cell, no congestion, cell type 0
001	User data cell, no congestion, cell type 1
010	User data cell, congestion experienced, cell type 0
011	User data cell, congestion experienced, cell type 1
100	Maintenance information between adjacent switches
101	Maintenance information between source and destination switches
110	Resource Management cell (used for ABR congestion control)
111	Reserved for future function

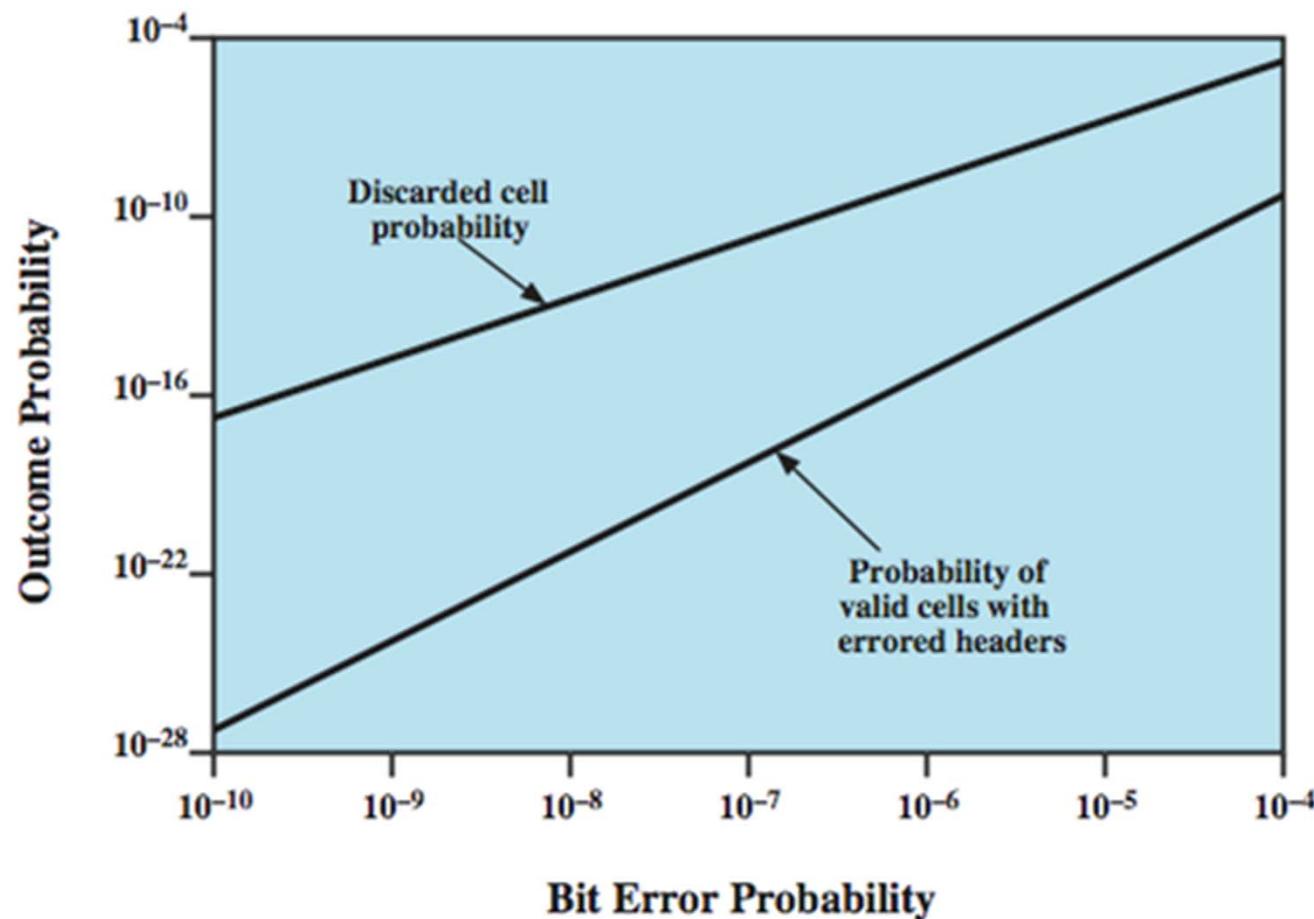
Header Error Control



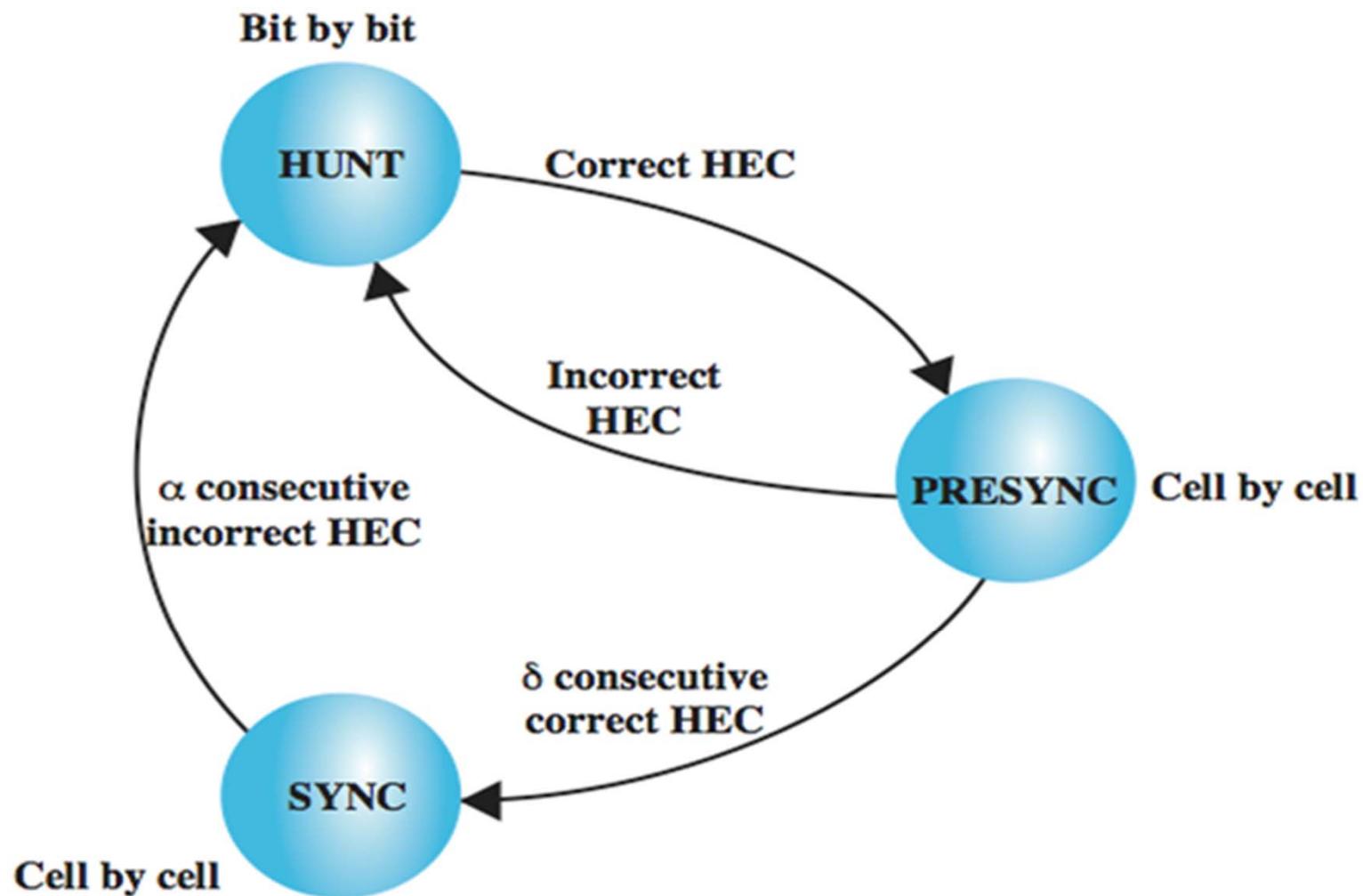
Effect of Error Cell Header



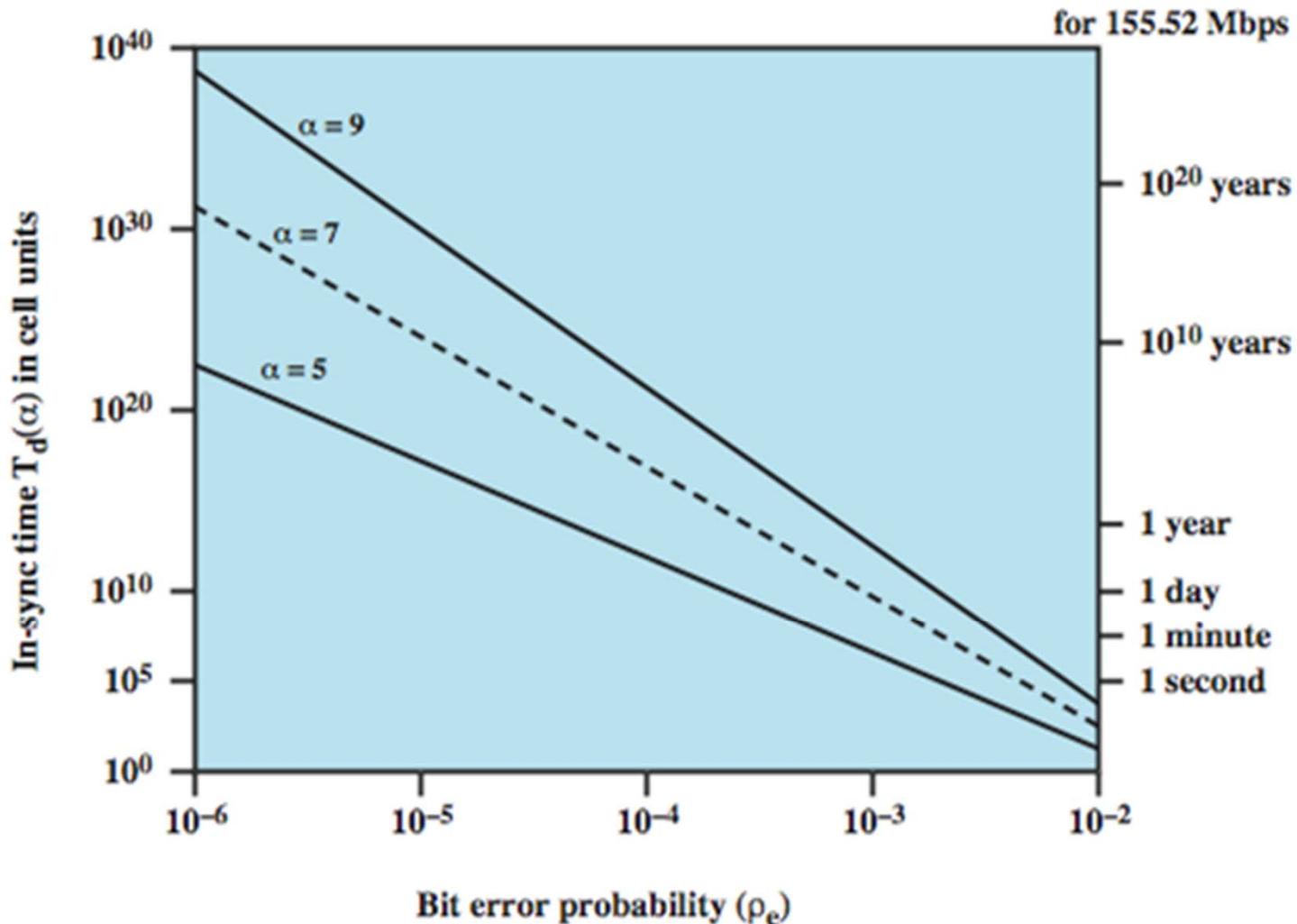
Impact of Random Bit Errors on HEC Performance



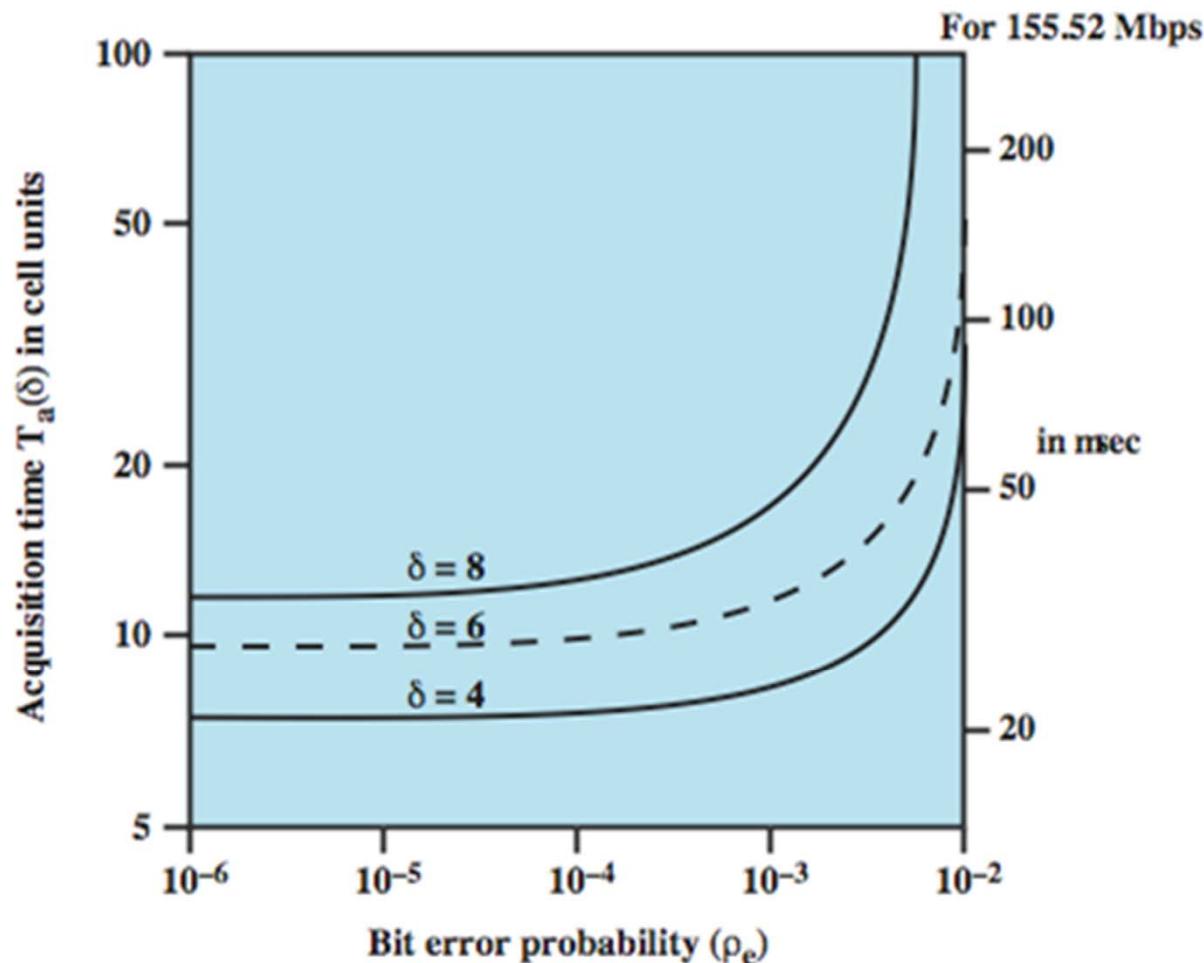
Cell Delineation State Diagram



Impact of Random Bit Errors on Cell Delineation Performance



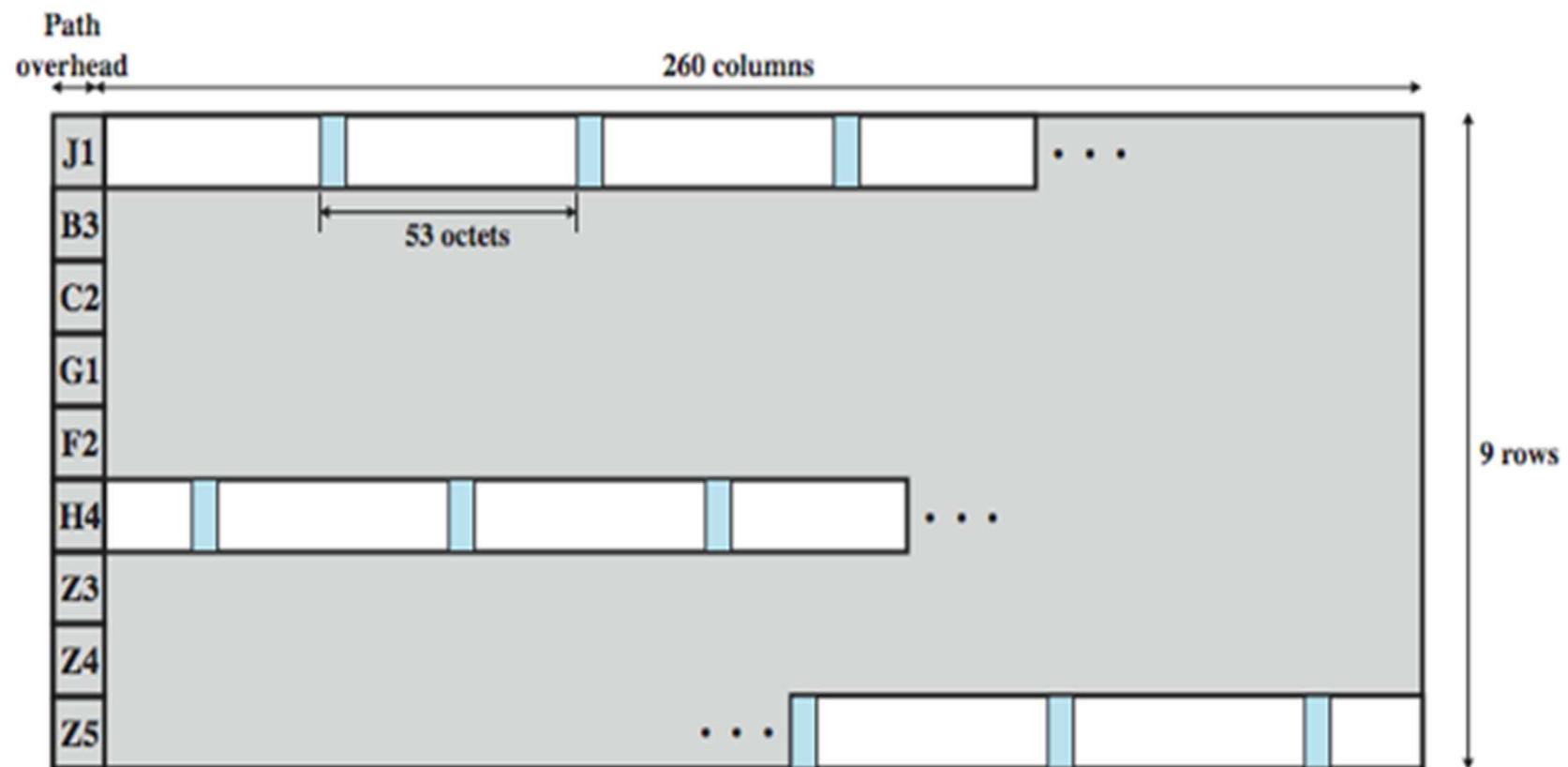
Acquisition Time vs. Bit Error Rate



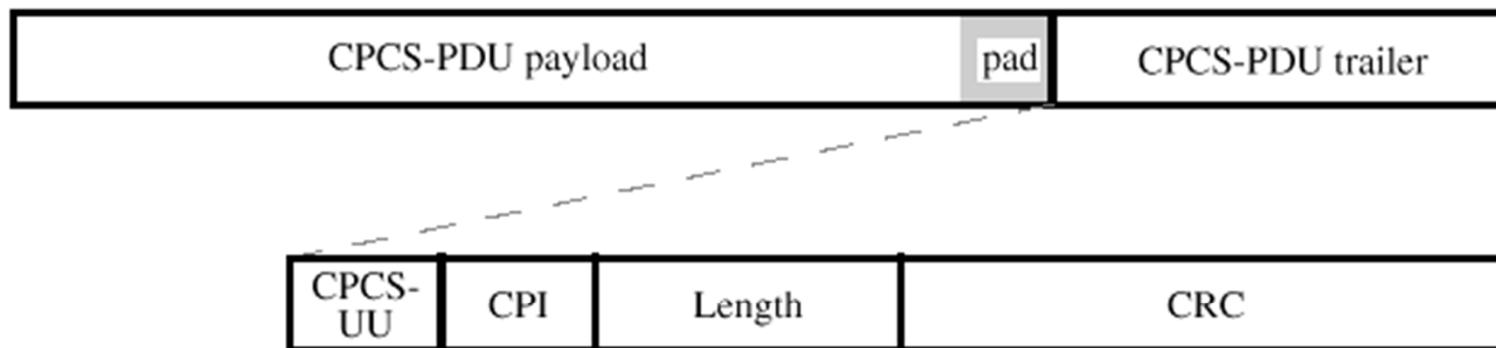
SDH Based Physical Layer

- imposes structure on ATM stream
 - eg. for 155.52Mbps
 - use STM-1 (STS-3) frame
- can carry ATM and STM payloads
- specific connections can be circuit switched using SDH channel
- SDH multiplexing techniques can combine several ATM streams

STM-1 Payload for SDH-Based ATM Cell Transmission



AAL5 CPCS-PDU



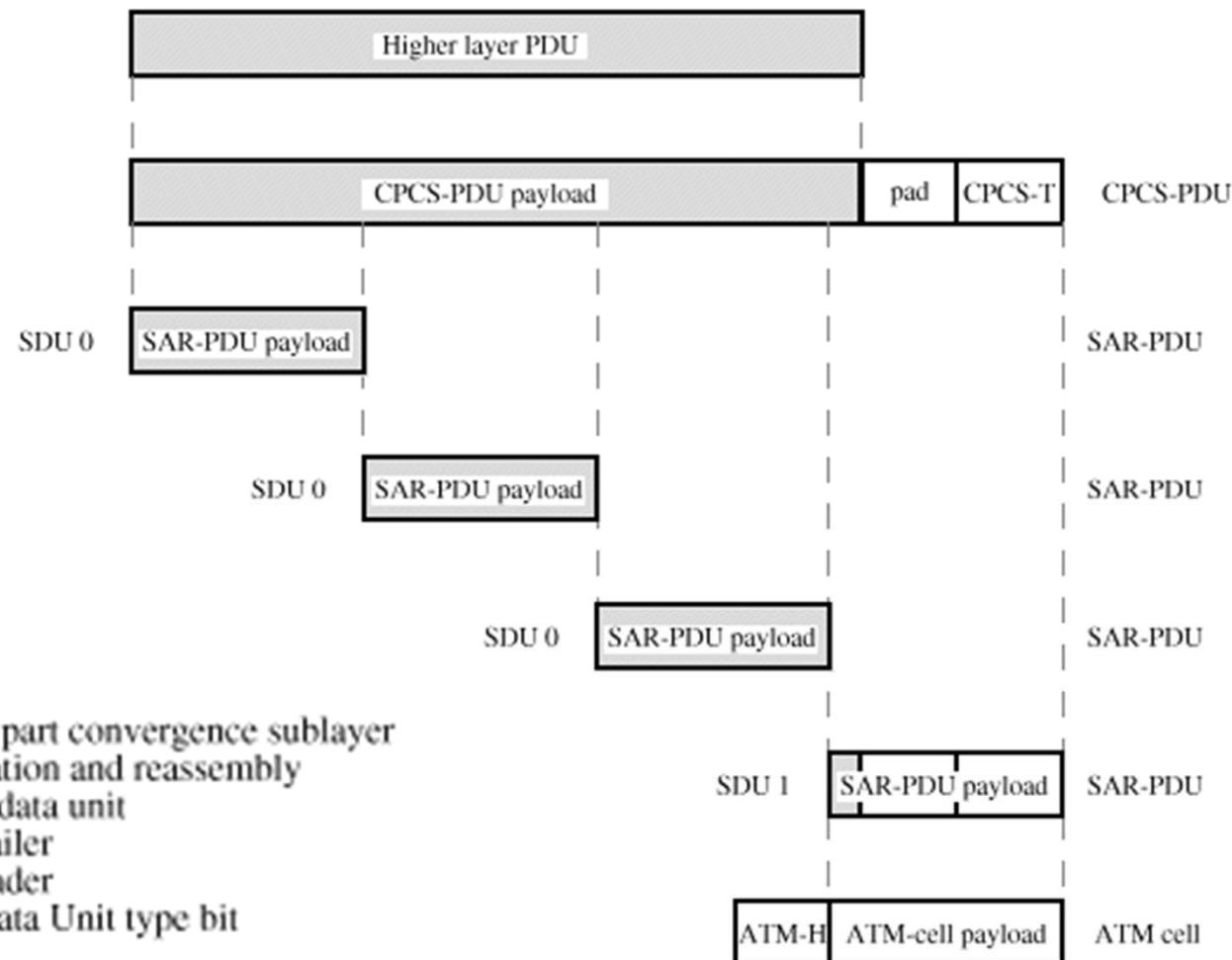
CPCS-UU = CPCS user-to-user indication (1 octet)

CPI = common part indicator (1 octet)

Length = length of CPCS-PDU payload (2 octets)

CRC = cyclic redundancy check (4 octets)

AAL5 transmission



ATM Service Categories

Real time - limit amount/variation of delay

- Constant bit rate (CBR)
- Real time variable bit rate (rt-VBR)

Non-real time - for bursty traffic

- Non-real time variable bit rate (nrt-VBR)
- Available bit rate (ABR)
- Unspecified bit rate (UBR)
- Guaranteed frame rate (GFR)



3.3 Commutació d'etiquetes MPLS

Source book: Data and Computer Communication Ed 10. W. Stallings Cap. 23

Role of MPLS

- Efficient technique for forwarding and routing packets
- Designed with IP networks in mind
 - *Can be used with any link-level protocol*
- Fixed-length label encapsulates an IP packet or a data link frame
- MPLS label contains all information needed to perform routing, delivery, QoS, and traffic management functions
- Is connection oriented

Traffic Engineering

- Ability to define routes dynamically, plan resource commitments on the basis of known demand, and optimize network utilization
- Effective use can substantially increase usable network capacity
- ATM provided strong traffic engineering capabilities prior to MPLS
- With basic IP there is a primitive form

MPLS

- Is aware of flows with QoS requirements
- Possible to set up routes on the basis of flows
- Paths can be rerouted intelligently

Key MPLS Terms

Forwarding equivalence class (FEC) A group of IP packets that are forwarded in the same manner (e.g., over the same path, with the same forwarding treatment).

Frame merge Label merging, when it is applied to operation over frame-based media, so that the potential problem of cell interleave is not an issue.

Label merging The replacement of multiple incoming labels for a particular FEC with a single outgoing label.

Label swap The basic forwarding operation consisting of looking up an incoming label to determine the outgoing label, encapsulation, port, and other data handling information.

Label swapping A forwarding paradigm allowing streamlined forwarding of data by using labels to identify classes of data packets that are treated indistinguishably when forwarding.

Label switched hop The hop between two MPLS nodes, on which forwarding is done using labels.

Label switched path The path through one or more LSRs at one level of the hierarchy followed by a packets in a particular FEC.

Label switching router (LSR) An MPLS node that is capable of forwarding native L3 packets.

Label stack An ordered set of labels.

Merge point A node at which label merging is done.

MPLS domain A contiguous set of nodes that operate MPLS routing and forwarding and that are also in one Routing or Administrative Domain

MPLS edge node An MPLS node that connects an MPLS domain with a node that is outside of the domain, either because it does not run MPLS, and/or because it is in a different domain. Note that if an LSR has a neighboring host that is not running MPLS, then that LSR is an MPLS edge node.

MPLS egress node An MPLS edge node in its role in handling traffic as it leaves an MPLS domain.

MPLS ingress node An MPLS edge node in its role in handling traffic as it enters an MPLS domain.

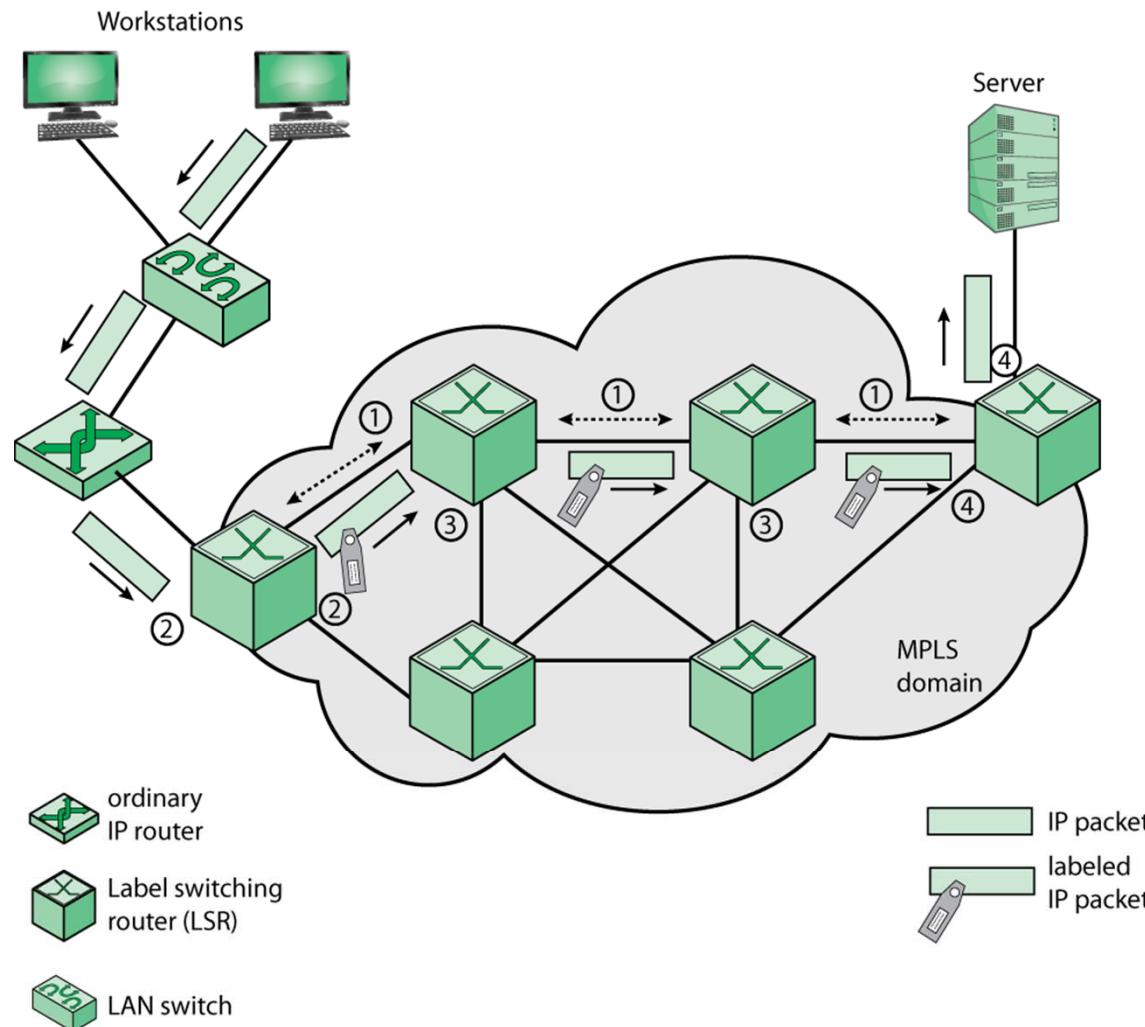
MPLS label A short, fixed-length physically contiguous identifier that is used to identify a FEC, usually of local significance. A label is carried in a packet header.

MPLS node A node that is running MPLS. An MPLS node will be aware of MPLS control protocols, will operate one or more L3 routing protocols, and will be capable of forwarding packets based on labels. An MPLS node may optionally be also capable of forwarding native L3 packets.

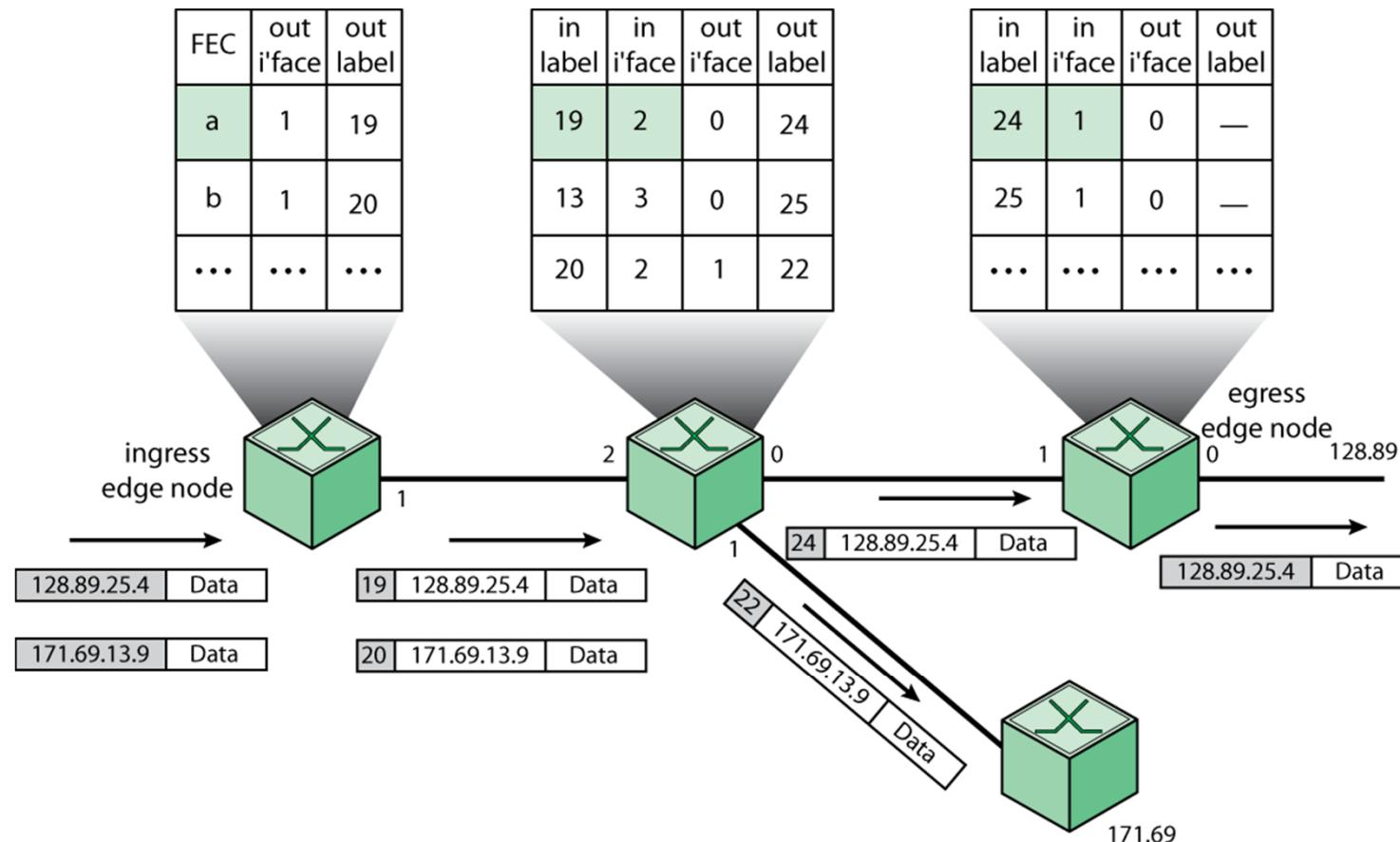
MPLS Operation

- Label switching routers (LSRs)
 - *Nodes capable of switching and routing packets on the basis of label*
- Labels define a flow of packets between two endpoints
- Assignment of a particular packet is done when the packet enters the network of MPLS routers
- Connection-oriented technology

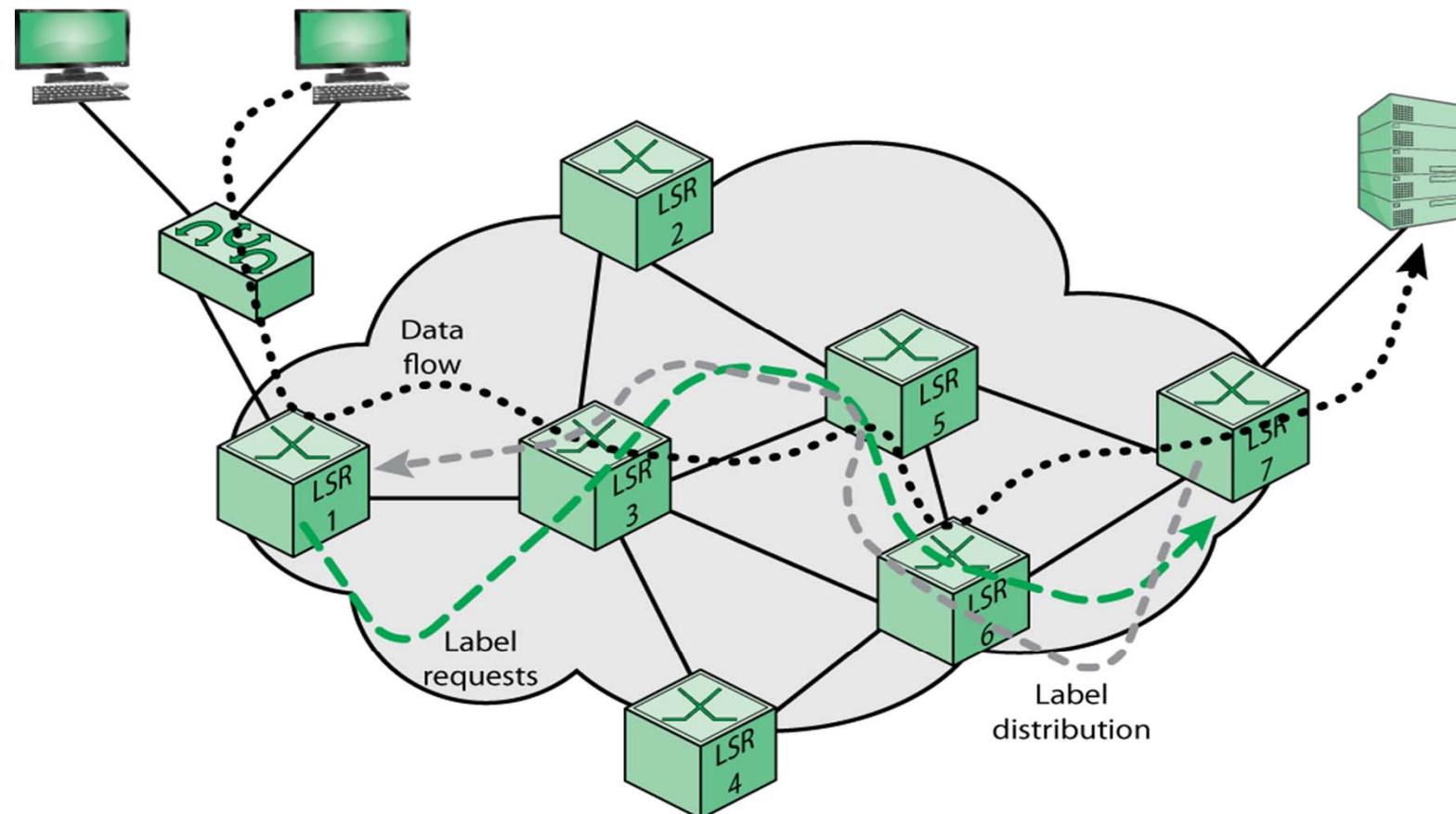
MPLS Operation



MPLS Packet Forwarding



LSP Creation and Packet Forwarding

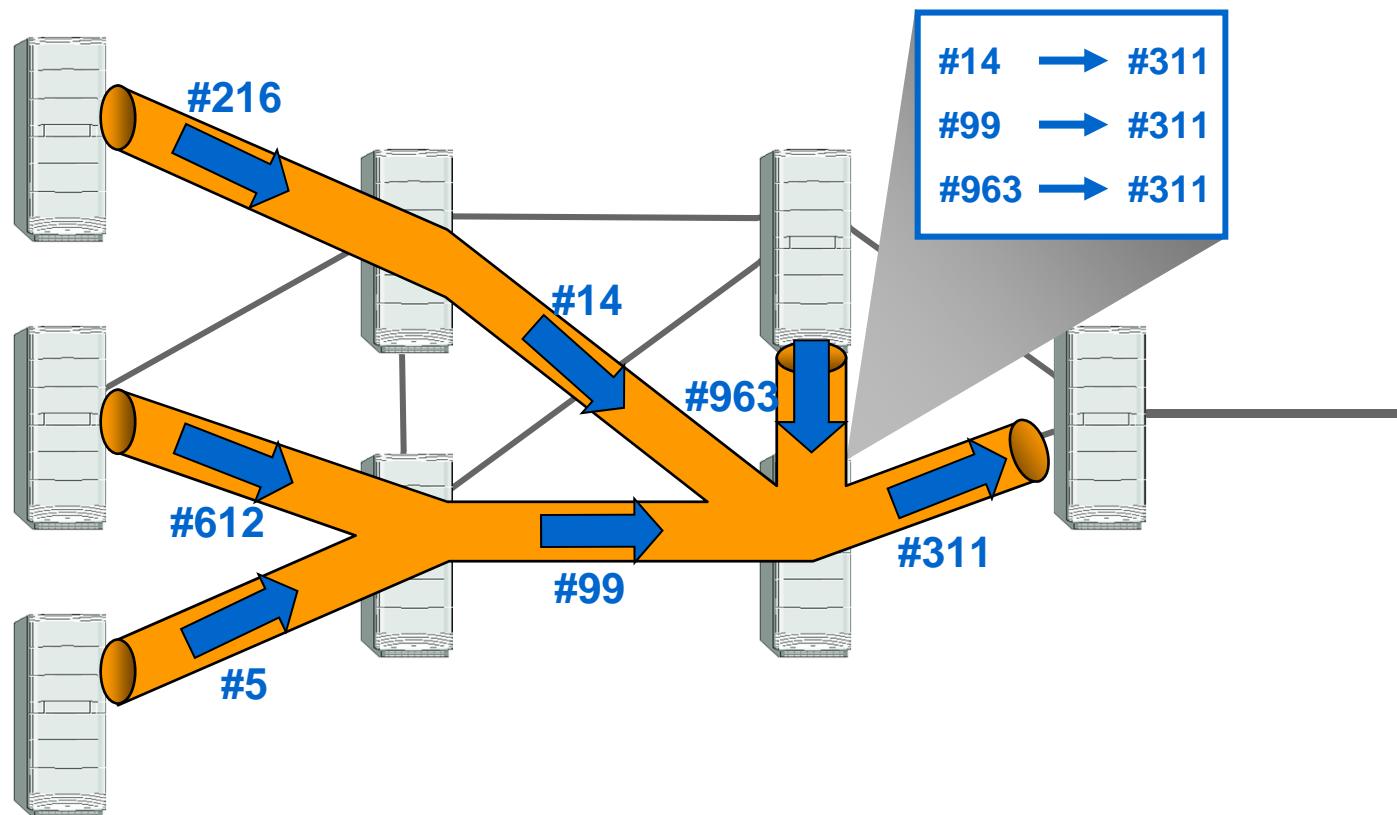


Label Stacking

- One of the most powerful features of MPLS
 - Processing is always based on the top label
 - At any LSR a label may be removed or added
- Allows creation of tunnels
 - Tunnel refers to traffic routing being determined by labels
- Provides considerable flexibility
- Unlimited stacking

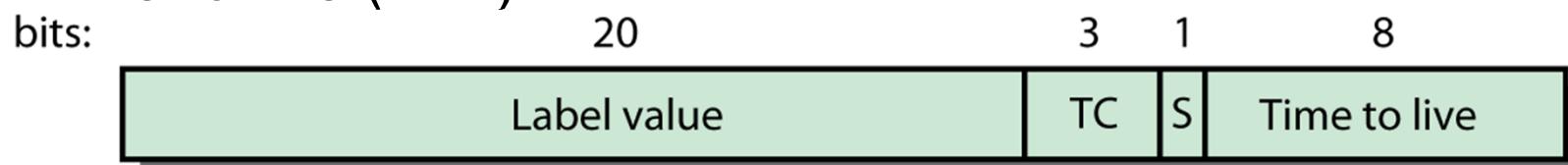


Merging LSP



Label Format

- defined in RFC 3032
 - 32-bit field consisting of:
 - Label value
 - Traffic class (TC)
 - S
 - Time to live (TTL)



TC = traffic class

S = bottom of stack bit

Figure 21.4 MPLS Label Format

Label Placement

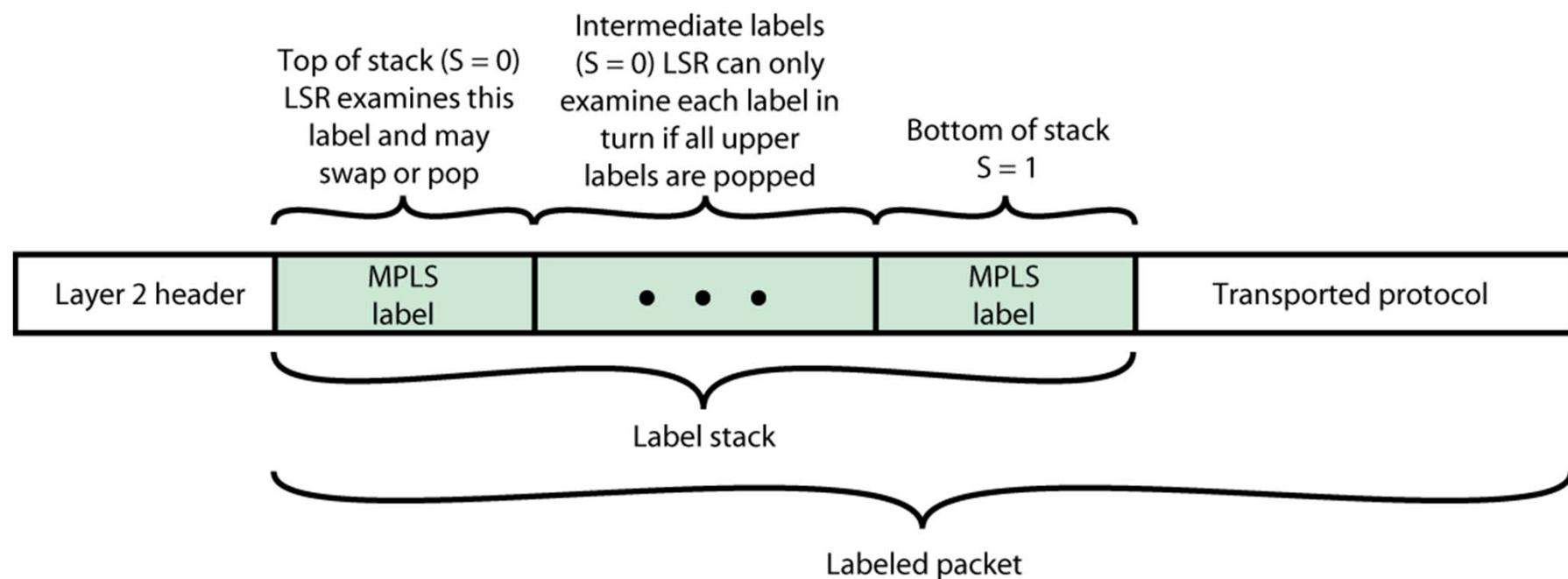
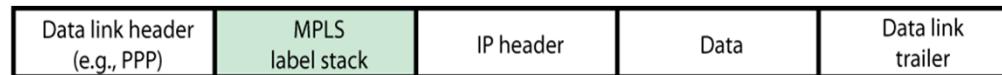


Figure 21.5 Encapsulation for Labeled Packet

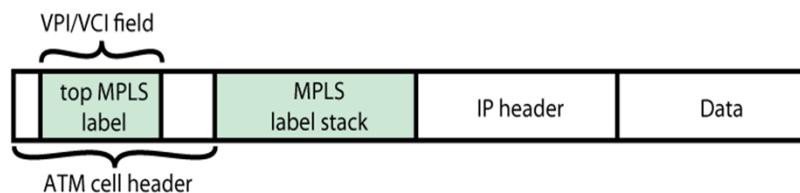
Label Stack



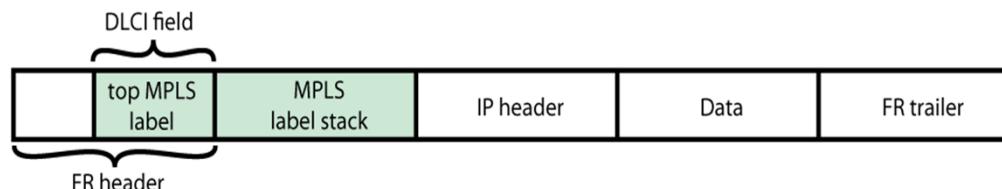
(a) Data link frame



(b) IEEE 802 MAC frame

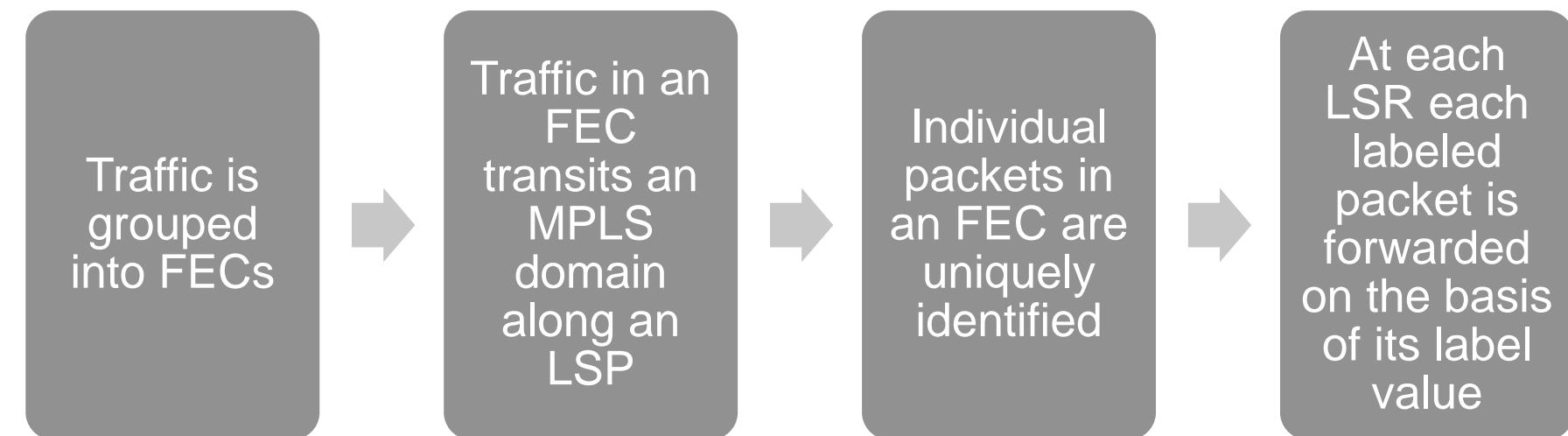


(c) ATM cell



(d) Frame relay frame

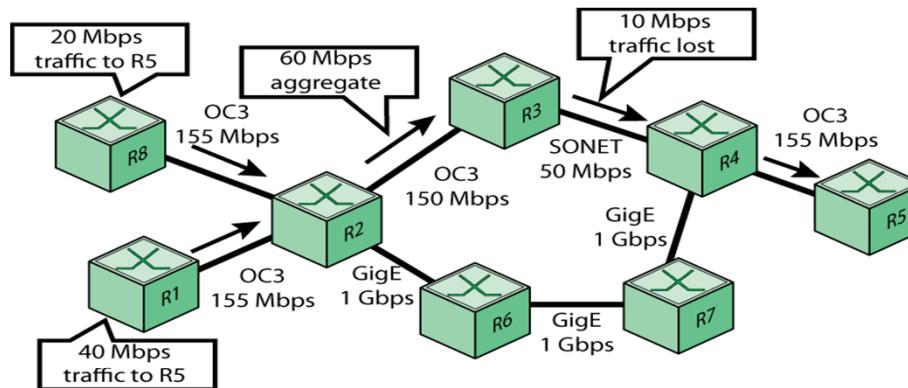
FECs, LSPs, and Labels



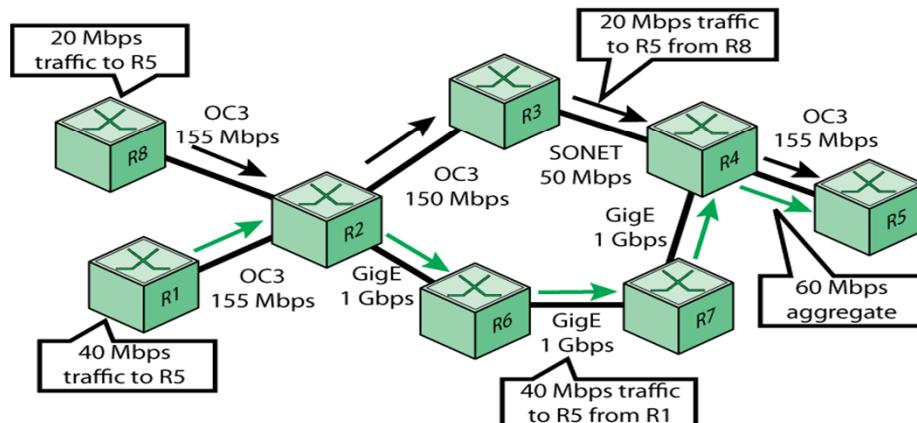
Traffic Engineering

- RFC 2702
- Allocate traffic to the network to maximize utilization of the network capacity
- Ensure the most desirable route through the network while meeting QoS requirements

Example of Traffic Engineering



(a) A shortest-path solution

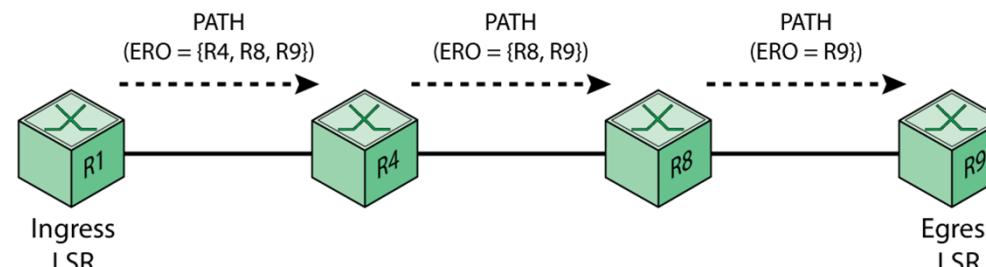


(b) A traffic-engineered solution

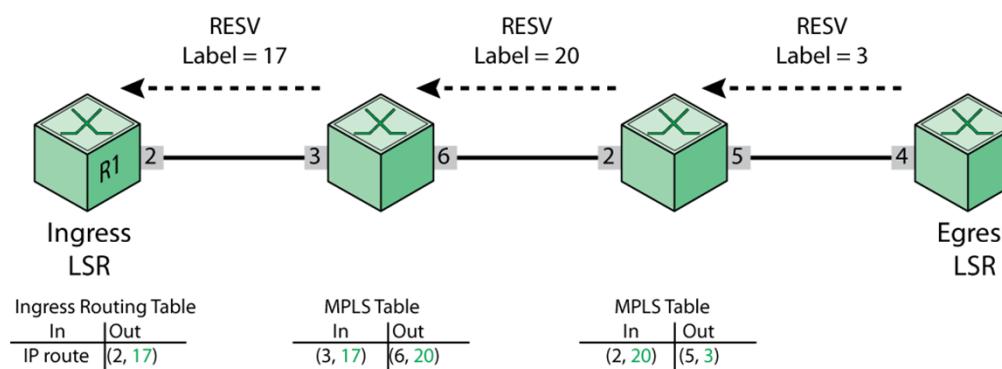
Elements of MPLS Traffic Engineering (MPLS TE)

- Information distribution
 - A link state protocol is necessary to discover the topology of the network
- Path calculation
 - Shortest path through a network that meets the resource requirements of the traffic flow
- Path setup
 - Signaling protocol to reserve the resources for a traffic flow and to establish the LSP
- Traffic forwarding
 - Accomplished with MPLS using the LSP

RSVP – TE Operation



(a) Use of PATH message



(b) Use of RESV message

Virtual Private Network (VPN)

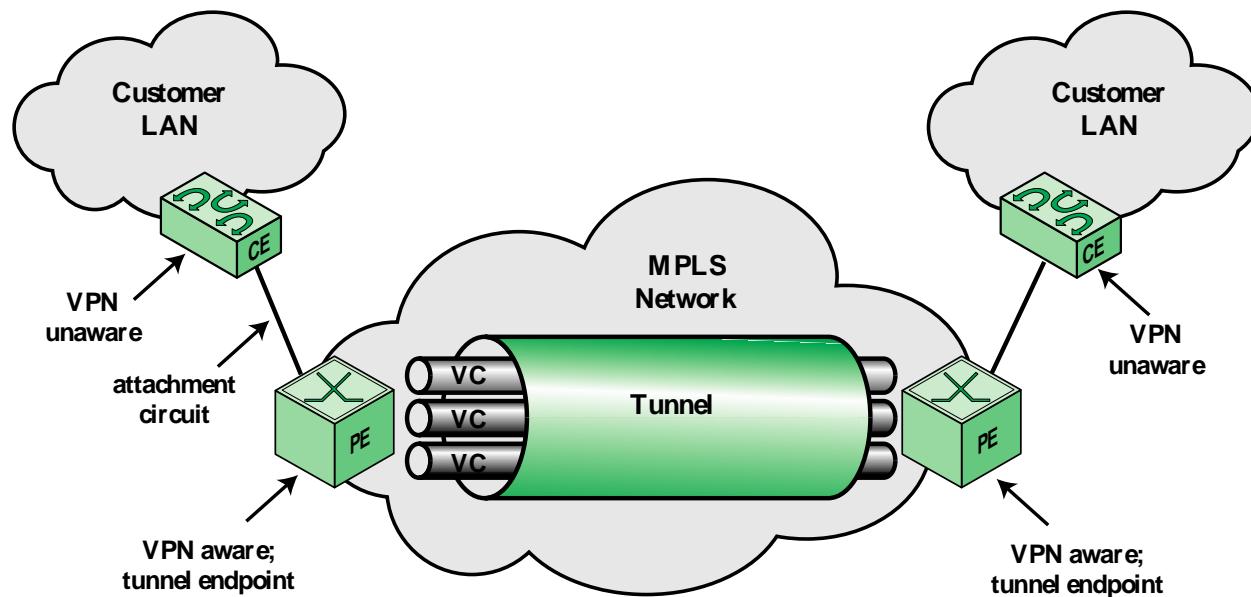
- Private network configured within a public network in order to take advantage of management facilities of larger networks
- Traffic designated as VPN traffic can only go from a VPN source to a destination in the same VPN

Widely used by enterprises to:

- Create wide area networks (WANs)
- Provide site-to-site communications to branch offices
- Allow mobile user to dial up their company LANs



Layer 2 VPN

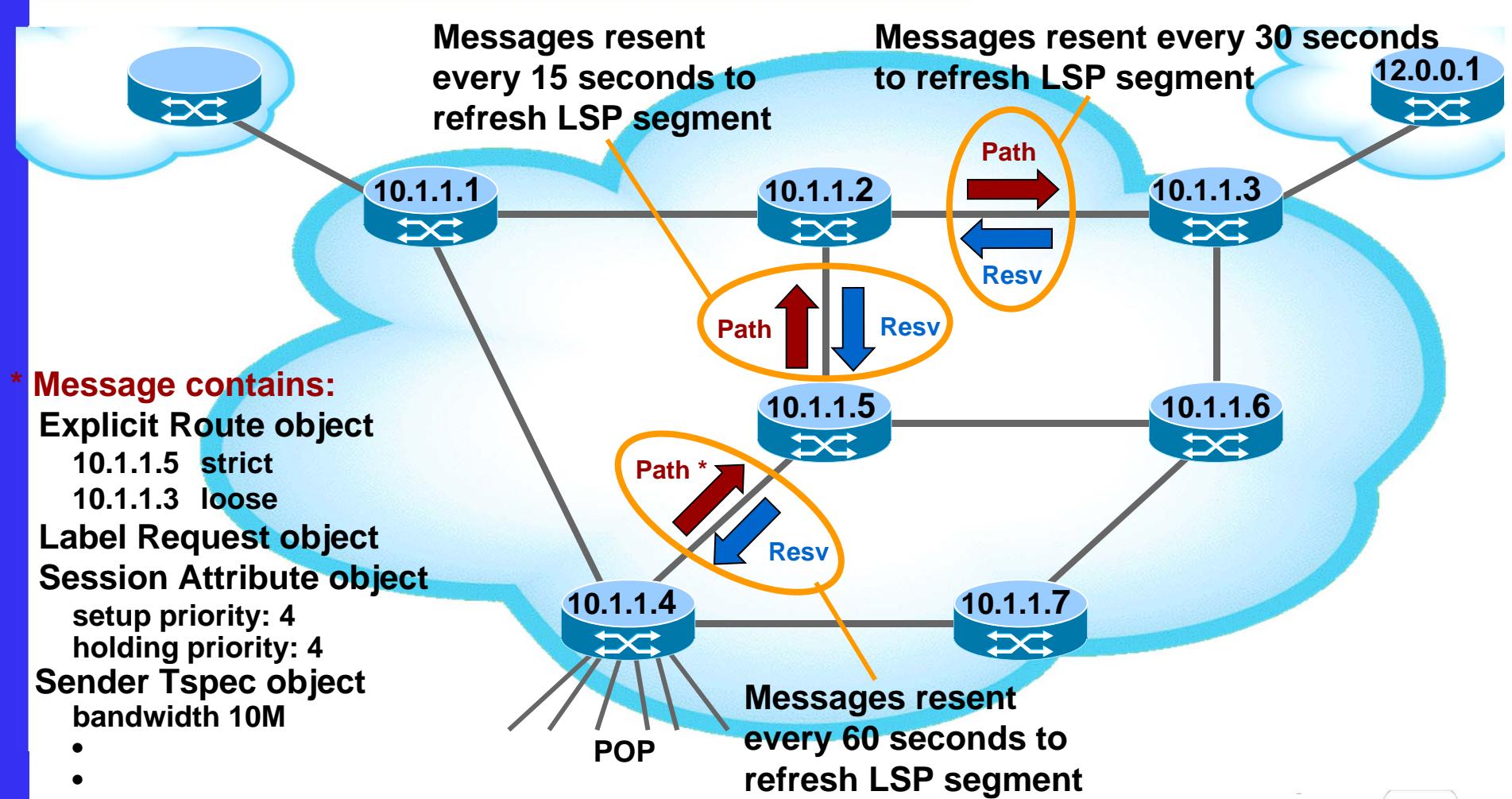


 Label switching router (LSR)

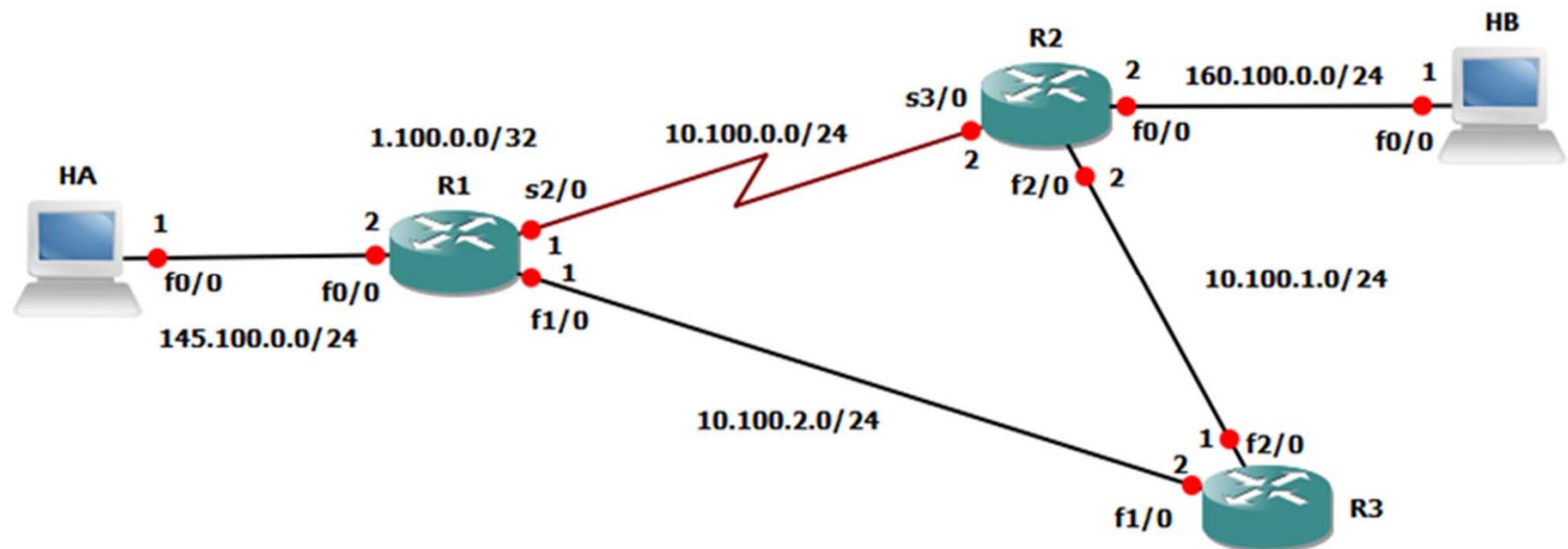
 LAN switch

CE = customer edge
PE = provider edge
VC = virtual channel
VPN = virtual private network

RSVP-TE details: Example



Real example MPLS



Programació Router

R1

```

ip cef
mpls label protocol ldp
interface Loopback0
  ip address 1.100.1.1 255.255.255.255
interface FastEthernet0/0
  ip address 145.100.0.2 255.255.255.0
interface FastEthernet1/0
  ip address 10.100.2.1 255.255.255.0
mpls ip
interface Serial2/0
  ip address 10.100.0.1 255.255.255.0
  mpls ip
router ospf 1
  passive-interface FastEthernet0/0
  network 1.100.1.1 0.0.0.0 area 0
  network 10.100.0.0 0.0.0.255 area 0
  network 10.100.2.0 0.0.0.255 area 0
  network 145.100.0.0 0.0.0.255 area 0

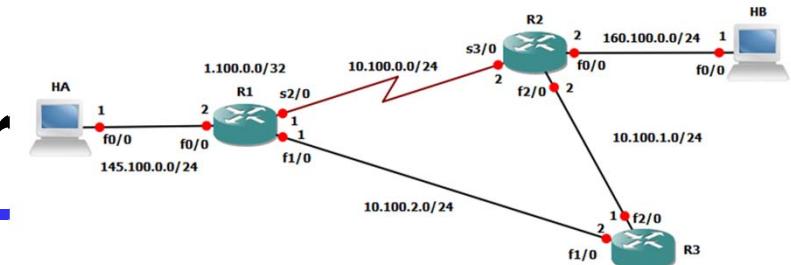
```

R2

```

ip cef
mpls label protocol ldp
interface Loopback0
  ip address 1.100.1.2 255.255.255.255
interface FastEthernet0/0
  ip address 160.100.0.2 255.255.255.0
interface FastEthernet2/0
  ip address 10.100.1.2 255.255.255.0
mpls ip
interface Serial3/0
  ip address 10.100.0.2 255.255.255.0
  mpls ip
router ospf 1
  passive-interface FastEthernet0/0
  network 1.100.1.2 0.0.0.0 area 0
  network 10.100.0.0 0.0.0.255 area 0
  network 10.100.1.0 0.0.0.255 area 0
  network 160.100.0.0 0.0.0.255 area 0

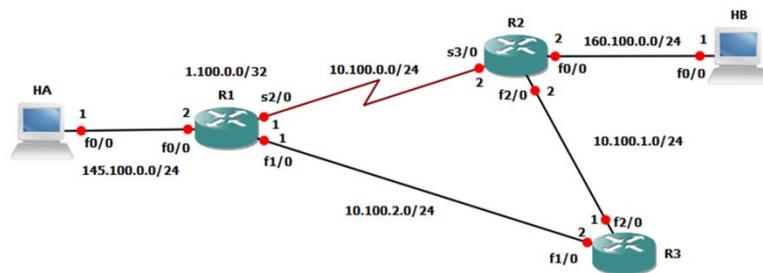
```



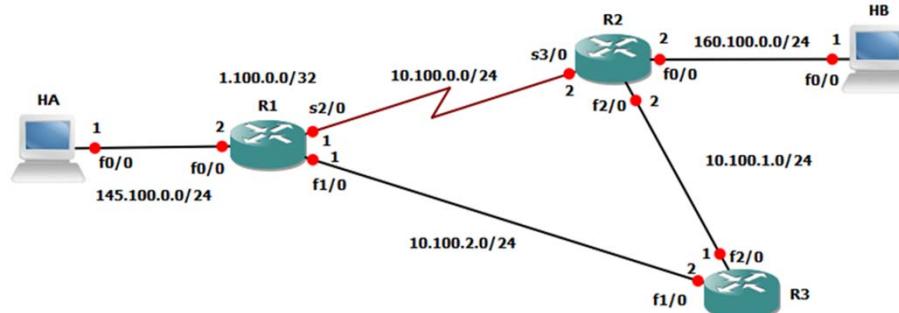
Routing table MPLS

R1#show ip route

- 1.0.0.0/32 is subnetted, 3 subnets
 - C 1.100.1.1 is directly connected, Loopback0
 - O 1.100.1.2 [110/3] via 10.100.2.2, 00:00:59, FastEthernet1/0
 - O 1.100.1.3 [110/2] via 10.100.2.2, 00:00:59, FastEthernet1/0
- 145.100.0.0/24 is subnetted, 1 subnets
 - C 145.100.0.0 is directly connected, FastEthernet0/0
- 160.100.0.0/24 is subnetted, 1 subnets
 - O 160.100.0.0 [110/3] via 10.100.2.2, 00:00:59, FastEthernet1/0
- 10.0.0.0/24 is subnetted, 3 subnets
 - C 10.100.2.0 is directly connected, FastEthernet1/0
 - C 10.100.0.0 is directly connected, Serial2/0
 - O 10.100.1.0 [110/2] via 10.100.2.2, 00:00:59, FastEthernet1/0



Label table MPLS

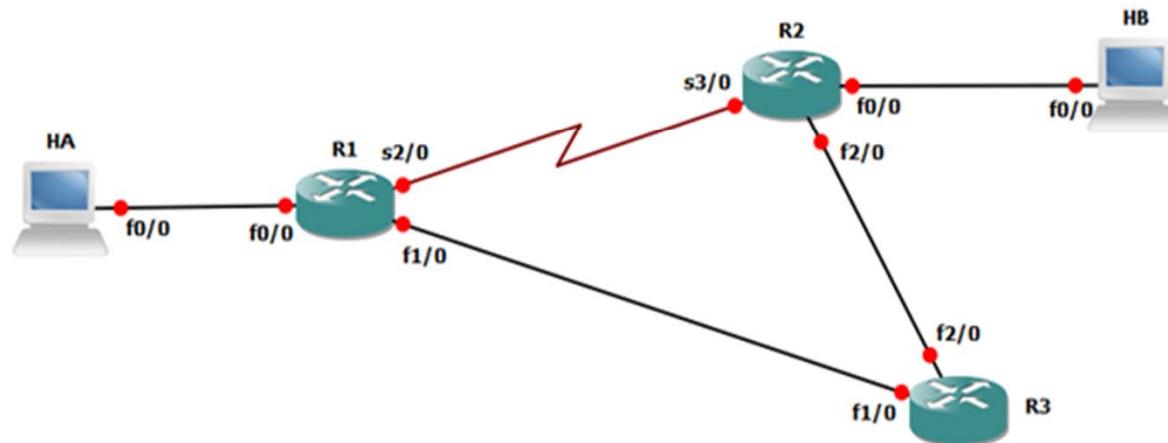


R1#show mpls forwarding-table

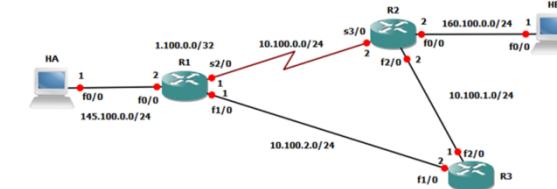
Local tag	Outgoing tag or VC	Prefix or Tunnel Id	Bytes tag switched	Outgoing interface	Next Hop
16	17	1.100.1.2/32	0	Fa1/0	10.100.2.2
17	Pop tag	10.100.1.0/24	0	Fa1/0	10.100.2.2
18	20	160.100.0.0/24	0	Fa1/0	10.100.2.2
19	Pop tag	1.100.1.3/32	0	Fa1/0	10.100.2.2

MPLS TE example

Tunel1 entre R1 y R2 a 50 Kbps dynamic. esto quiere decir que pasará por donde le diga OSPF. Seguramente R1-R3-R2. Si quiero que pase por R1-R2 deberé ponerlo explicito.
Tunel 2 entre R2-R3-R1 a 100 Kbps explicito por lo que pasará siempre R2-R3-R1
Se puede ver con "show ip route" en cada router



Programació Routers



R1

```

ip cef
mpls label protocol ldp
mpls traffic-eng tunnels
interface Loopback0
  ip address 1.100.1.1 255.255.255.255
interface Tunnel1
  ip unnumbered Loopback0
  tunnel destination 1.100.1.2
  tunnel mode mpls traffic-eng
  tunnel mpls traffic-eng autoroute
  announce
    tunnel mpls traffic-eng bandwidth 50
    tunnel mpls traffic-eng path-option 1
    dynamic
  interface FastEthernet0/0
    ip address 145.100.0.2 255.255.255.0
  
```

```

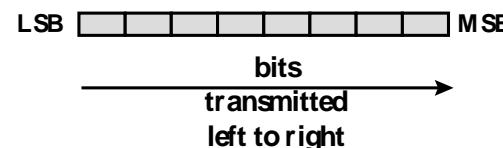
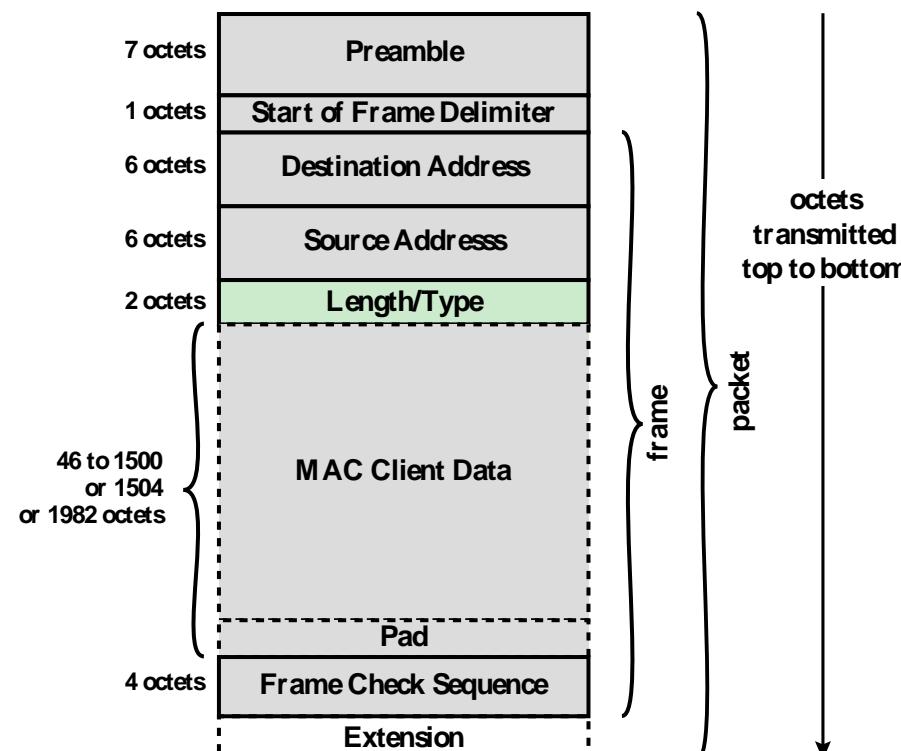
  interface FastEthernet1/0
    ip address 10.100.2.1 255.255.255.0
  mpls ip
  mpls traffic-eng tunnels
  ip rsvp bandwidth 1000
  interface Serial2/0
    ip address 10.100.0.1 255.255.255.0
  mpls ip
  mpls traffic-eng tunnels
  clock rate 56000
  ip rsvp bandwidth 100
  router ospf 1
    mpls traffic-eng router-id Loopback0
    mpls traffic-eng area 0
    passive-interface FastEthernet0/0
    network 1.100.1.1 0.0.0.0 area 0
    network 10.100.0.0 0.0.0.255 area 0
    network 10.100.2.0 0.0.0.255 area 0
    network 145.100.0.0 0.0.0.255 area 0
  
```



2.4 Carrier Ethernet

Source book: Data and Computer Communication Ed 10. W. Stallings Cap. 12

Ethernet MAC Frame format



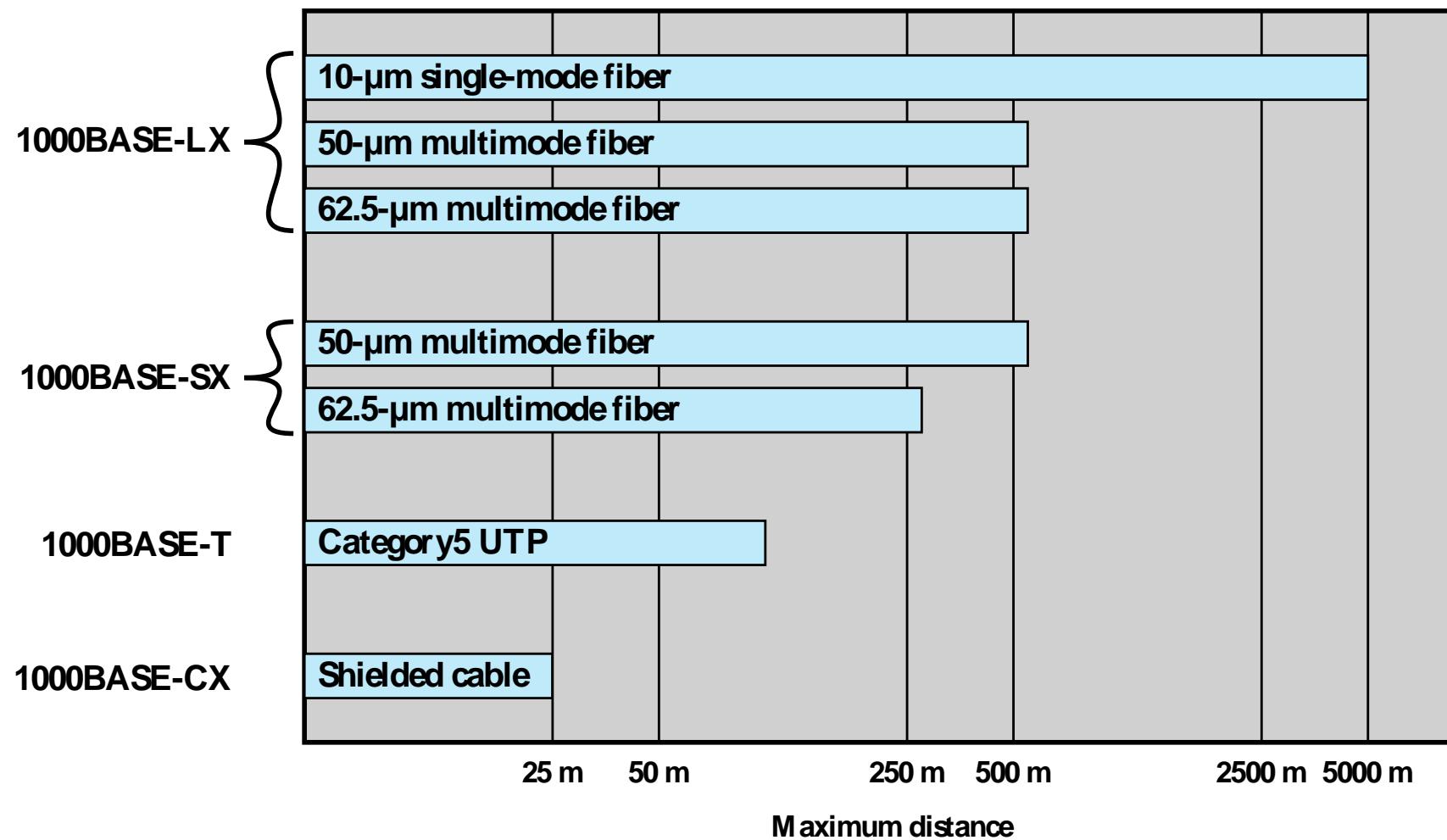
Full Duplex Operation

- Traditional Ethernet half duplex
- Using full-duplex, station can transmit and receive simultaneously
- 100-Mbps Ethernet in full-duplex mode, giving a theoretical transfer rate of 200 Mbps
- Stations must have full-duplex adapter cards
- And must use switching hub
 - Each station constitutes separate collision domain
 - CSMA/CD algorithm no longer needed
 - 802.3 MAC frame format used

Gigabit Ethernet for carrier

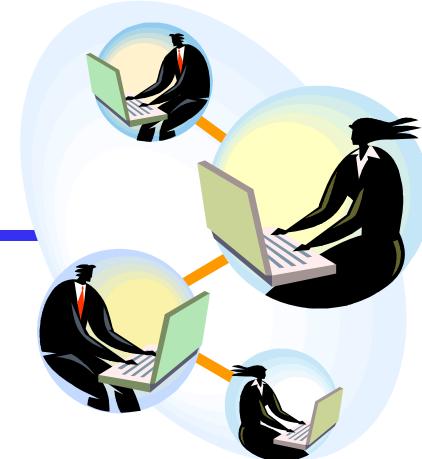
- Carrier extension
 - At least 4096 bit-times long (512 for 10/100)
- Frame bursting
- Not needed if using a switched hub to provide dedicated media access

Gigabit Ethernet Medium Options

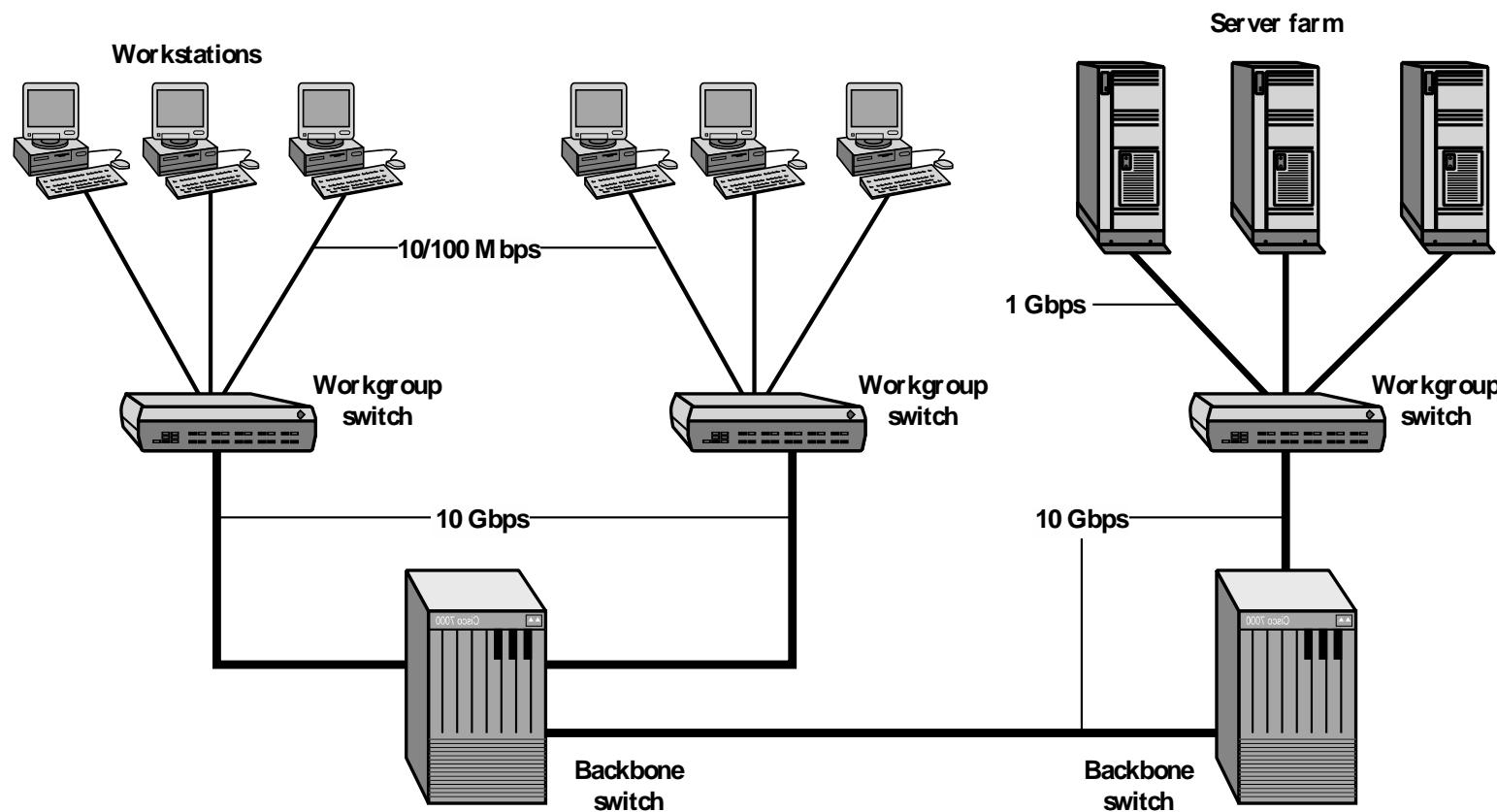


10Gbps Ethernet

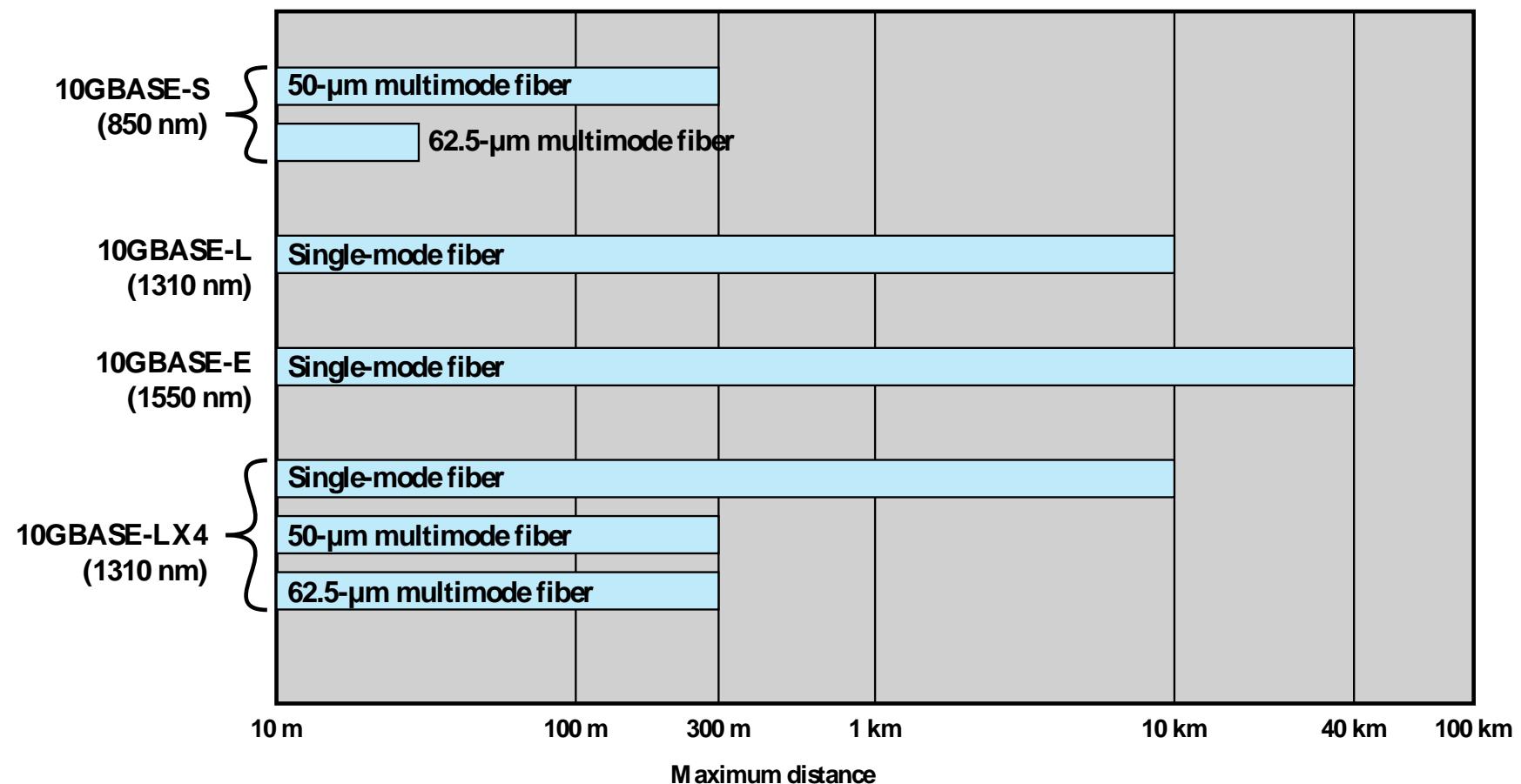
- Growing interest in 10Gbps Ethernet
 - High-speed backbone use
 - Future wider deployment
- Alternative to ATM and other WAN technologies
- Uniform technology for LAN, MAN, or WAN
- Advantages of 10Gbps Ethernet
 - No expensive, bandwidth-consuming conversion between Ethernet packets and ATM cells
 - IP and Ethernet together offers QoS and traffic policing approach ATM
 - Have a variety of standard optical interfaces



10Gbps Ethernet exemple



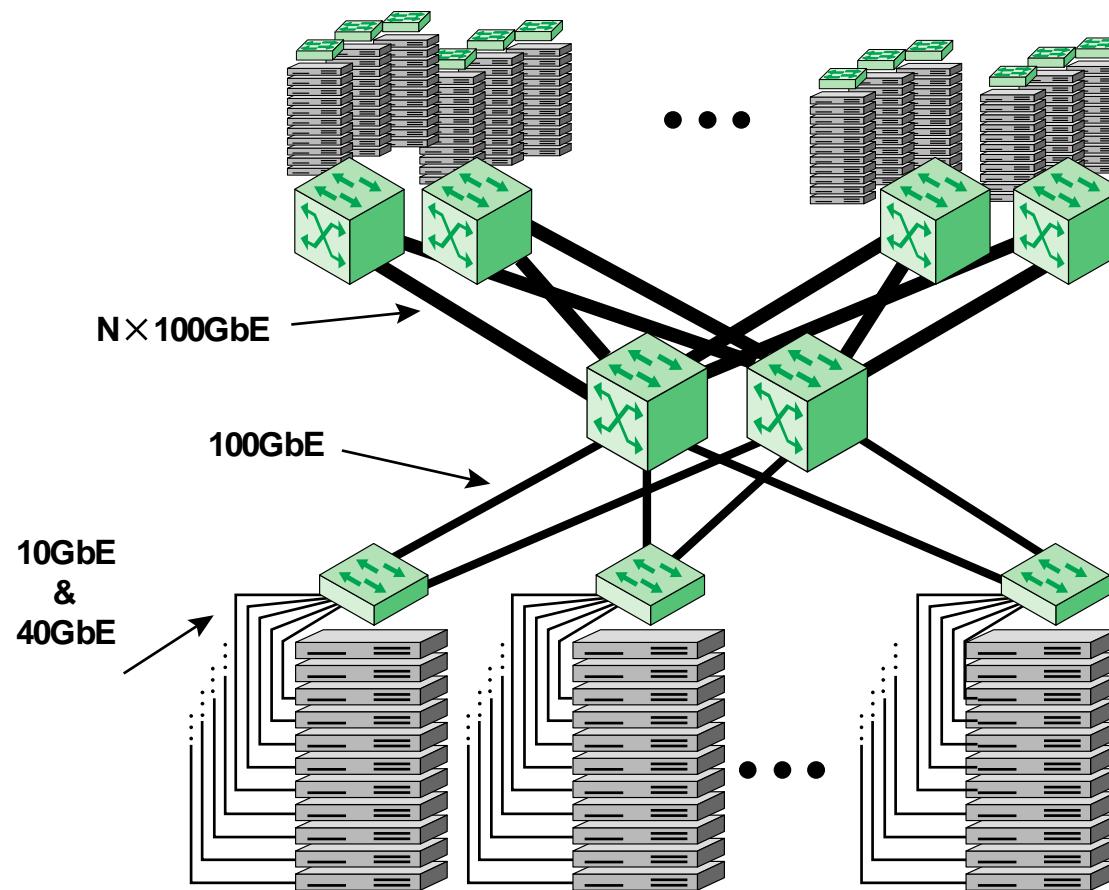
10 Gbps Ethernet Distance Options



100-Gbps Ethernet

- Preferred technology for wired LAN
- Preferred carrier for bridging wireless technologies into local Ethernet networks
- Cost-effective, reliable and interoperable
- Popularity of Ethernet technology:
 - Availability of cost-effective products
 - Reliable and interoperable network products
 - Variety of vendors

Example 100 Gbps Ethernet



Media Options for 40-Gbps and 100-Gbps Ethernet

	40 Gbps	100 Gbps
1m backplane	40GBASE-KR4	
10 m copper	40GBASE-CR4	1000GBASE-CR10
100 m multimode fiber	40GBASE-SR4	1000GBASE-SR10
10 km single mode fiber	40GBASE-LR4	1000GBASE-LR4
40 km single mode fiber		1000GBASE-ER4

Naming nomenclature:

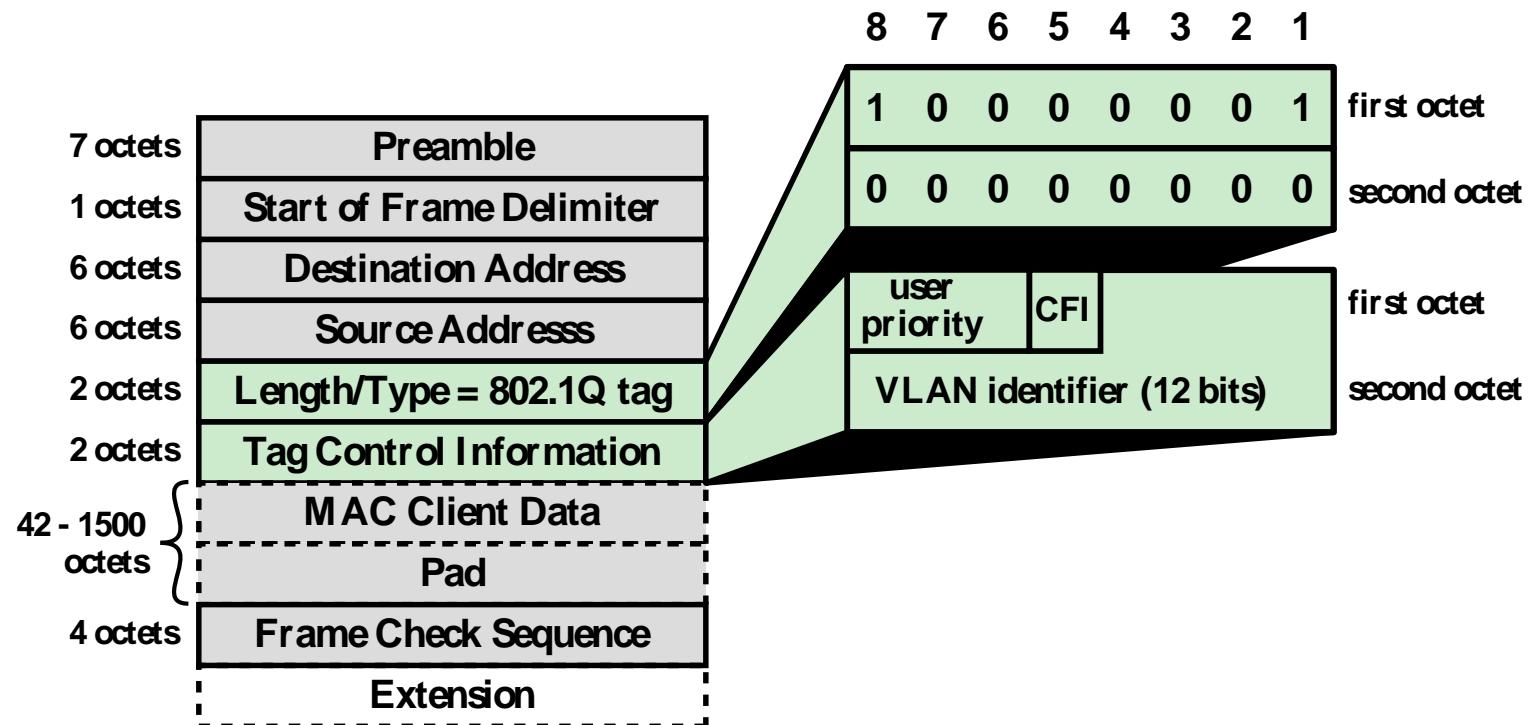
Copper: K = backplane; C = cable assembly

Optical: S = short reach (100m); L - long reach (10 km); E = extended long reach (40 km)

Coding scheme: R = 64B/66B block coding

Final number: number of lanes (copper wires or fiber wavelengths)

Tagged IEEE 802.3 MAC Frame Format

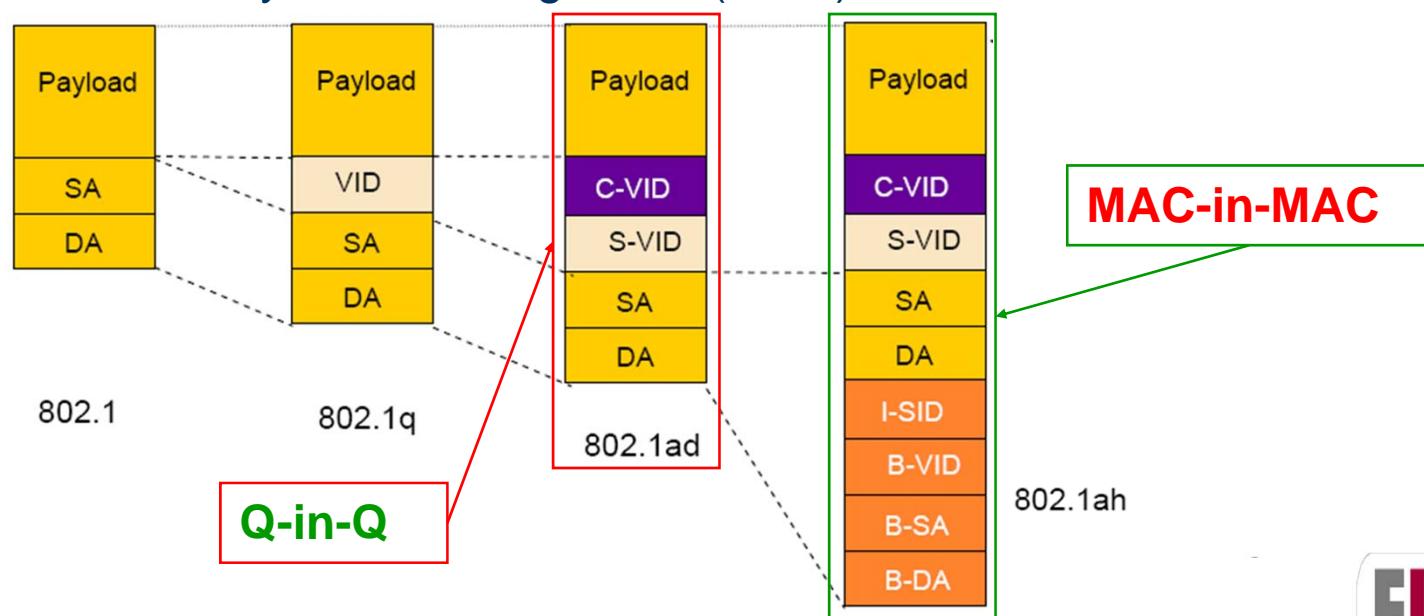


CFI = Canonical Format Indicator

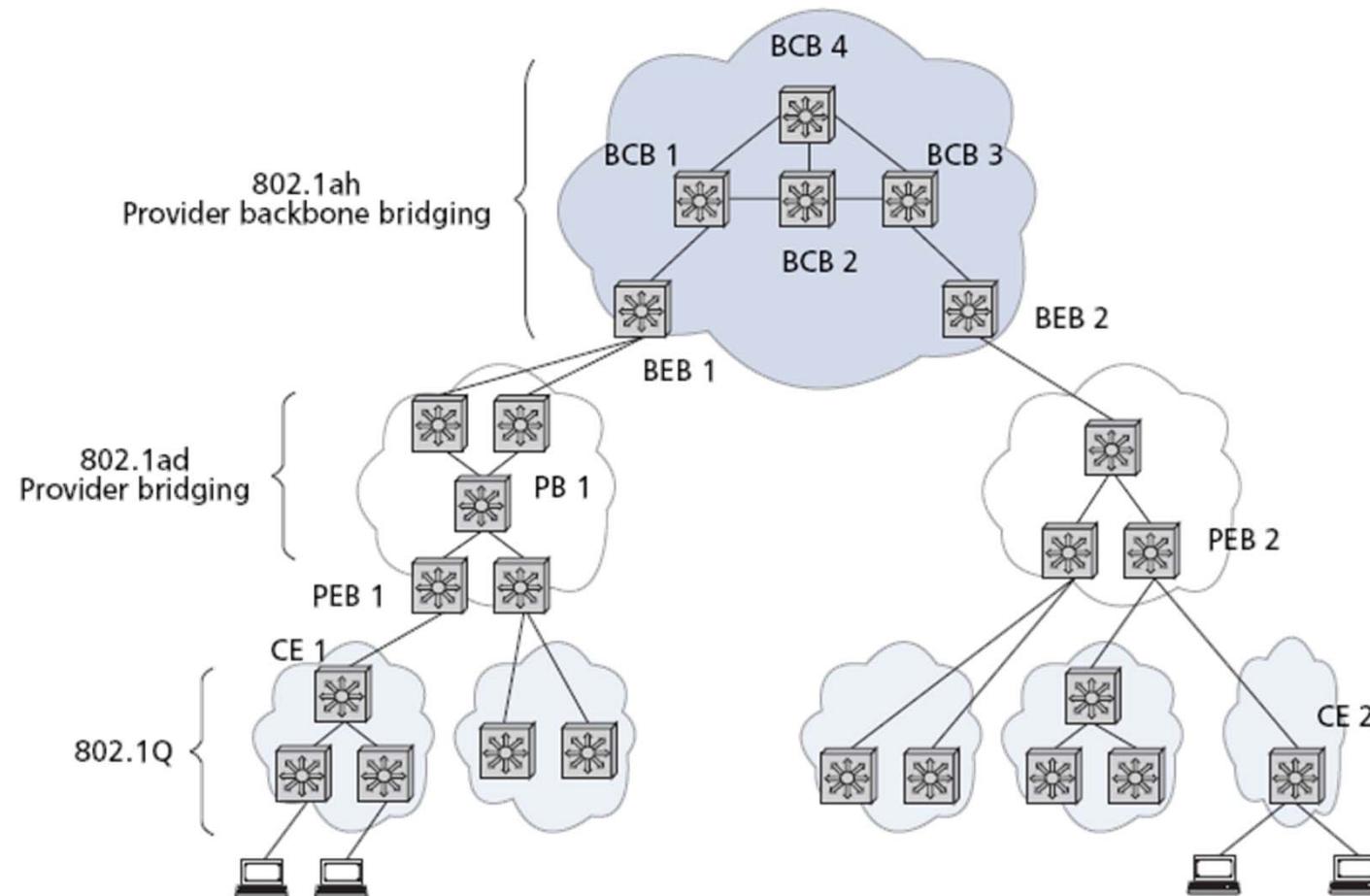
VLAN = virtual local area network

Provider Backbone Bridge Traffic Engineering (PBB-TE)

- IEEE has developed a number of standards providing enhancements to the original Ethernet standards. PBB-TE adapts Ethernet technology to carrier class transport networks
 - 802.1Q: Virtual LAN
 - 802.1ad: Provider Bridging
 - 802.1ah: Provider Backbone Bridging
 - 802.1ag: Connectivity Fault Management (OAM)



The way to PBB-TE

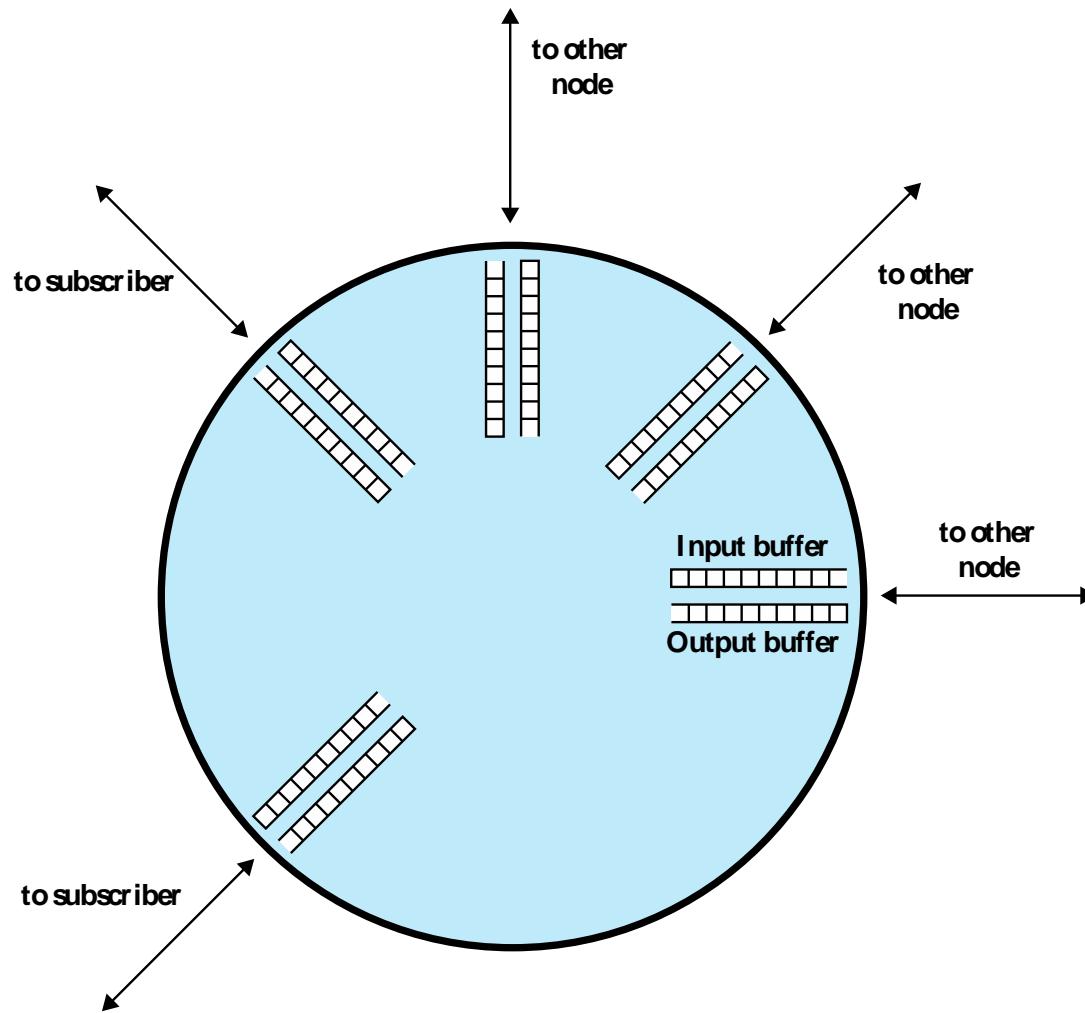




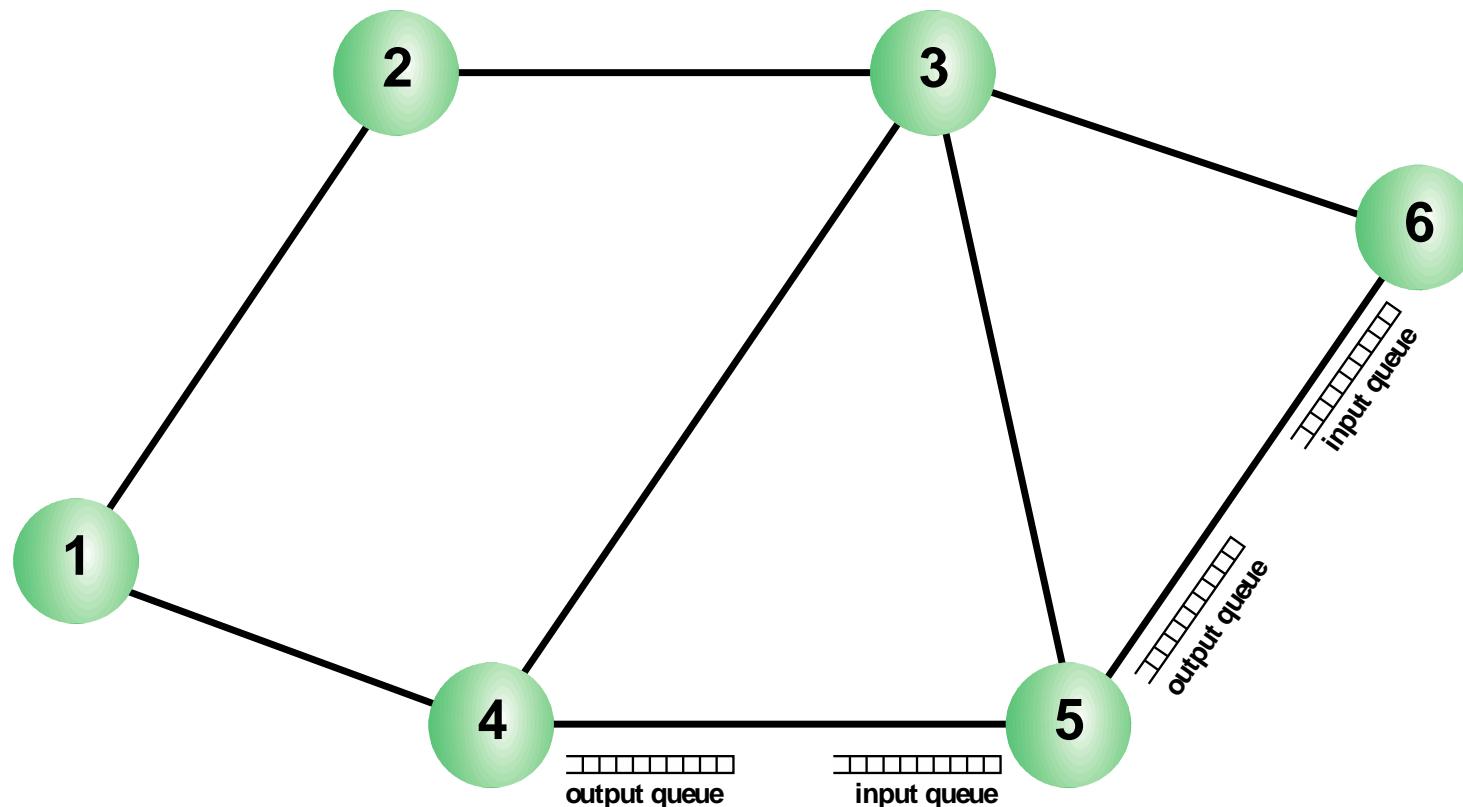
2.5 Control de la congestió en xarxes de commutació nivell 2

Source book: Data and Computer Communication Ed 10. W. Stallings Cap. 20
: Data and Computer Communication Ed 8. W. Stallings Cap. 13

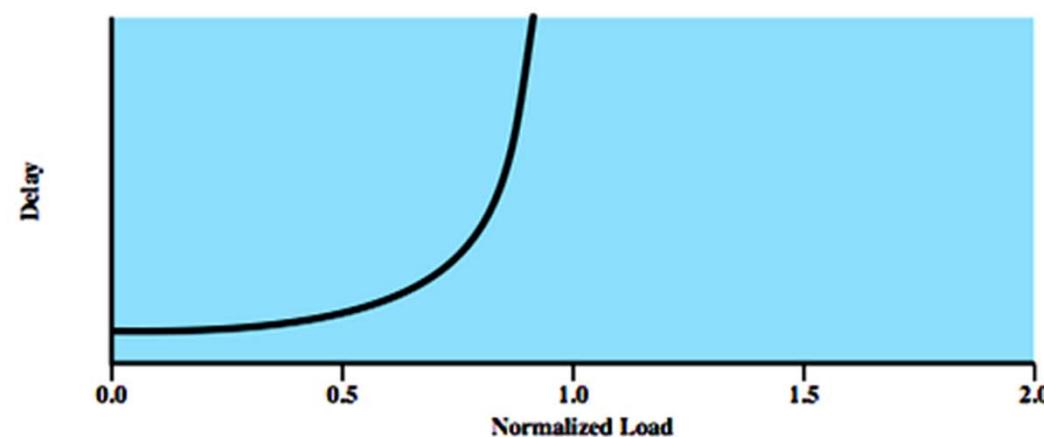
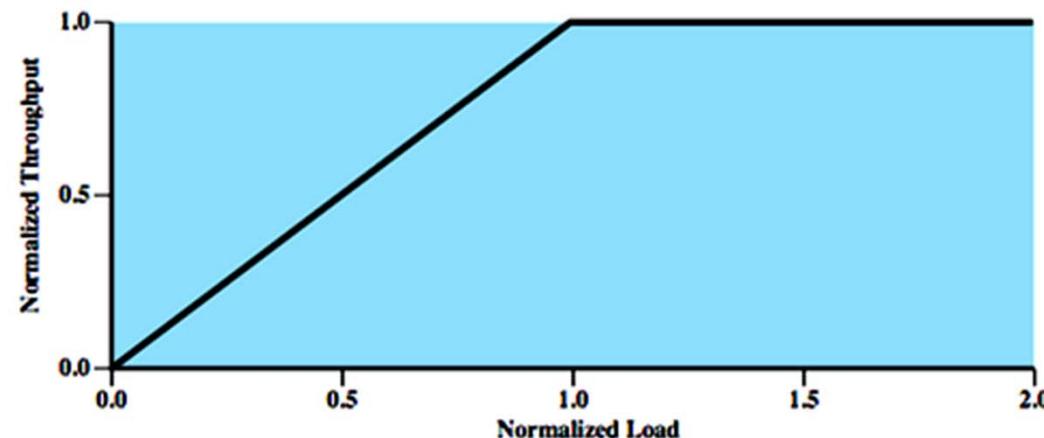
Queues at Node



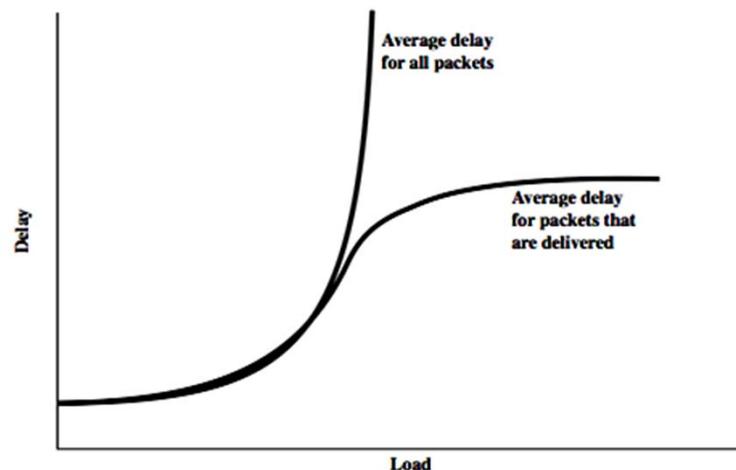
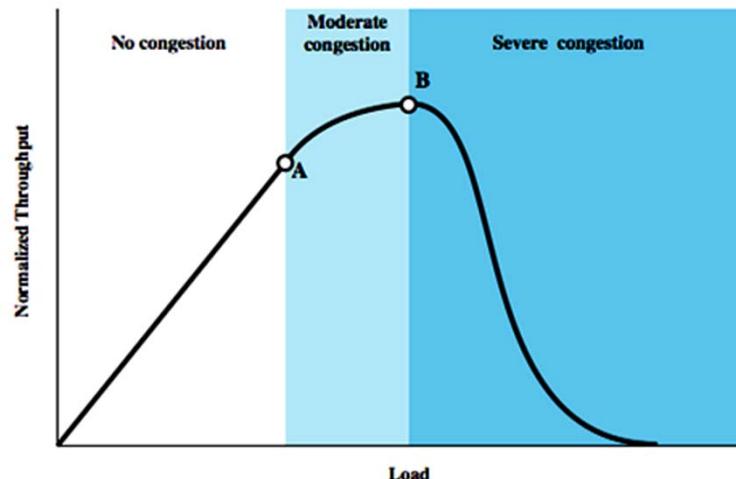
Interaction of Queues in a data network



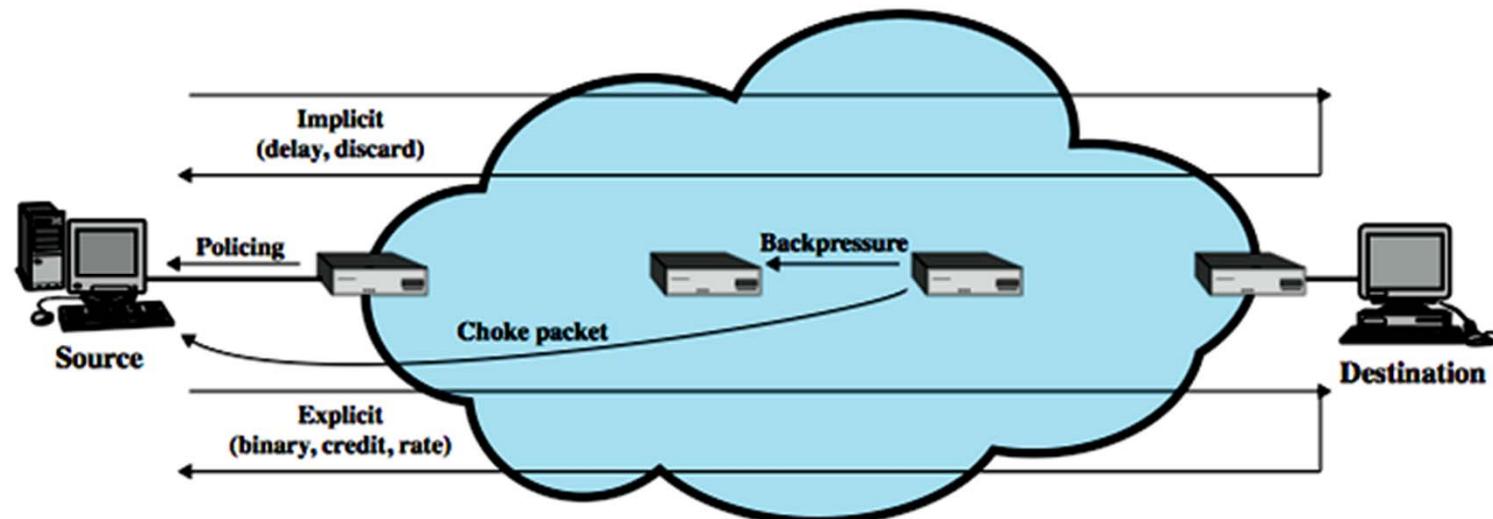
Ideal Network Utilization



Effects of Congestion - No Control



Mechanisms for Congestion Control



Backpressure

- If node becomes congested it can slow down or stop flow of packets from other nodes
- Can be exerted on the basis of links or logical connections
- Flow restriction propagates backward to sources, which are restricted in the flow of new packets into the network
- Can be selectively applied to logical connections so that the flow from one node to the next is only restricted or halted on some connections

Choke Packet

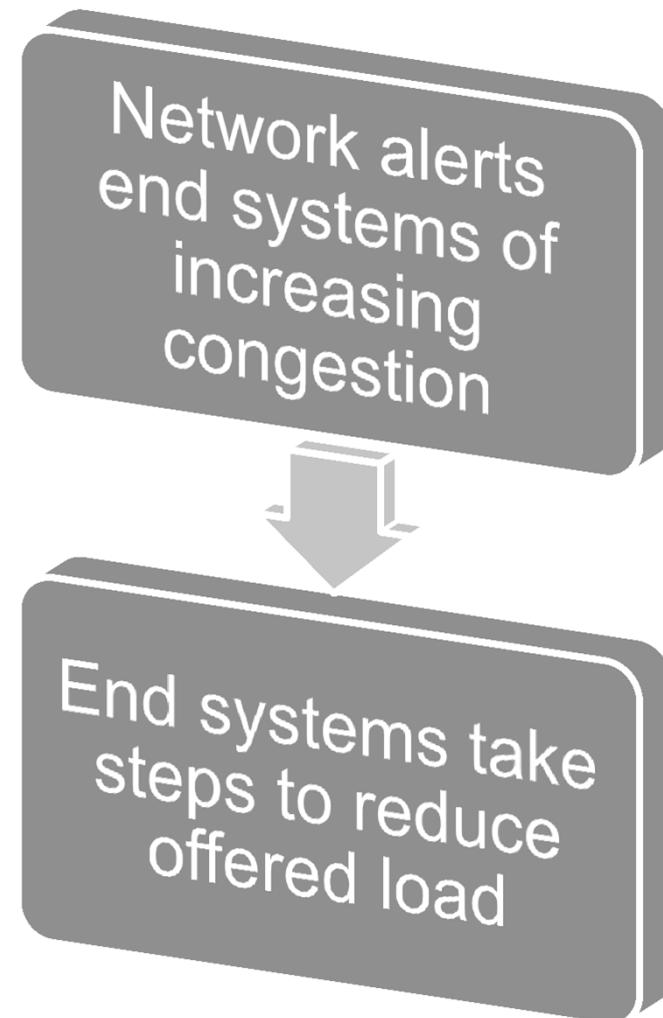
- A control packet
 - Generated at congested node
 - Sent back to source node
- An example is the Internet Control Message Protocol (ICMP) Source Quench packet
 - *From router or destination end system*
 - *Source cuts back until it no longer receives quench messages*
 - *Message is issued for every discarded packet*
 - *Message may also be issued for anticipated congestion*
- Is a crude technique for controlling congestion

Implicit Congestion Signaling

- With network congestion:
 - Transmission delay increases
 - Packets may be discarded
- Source can detect congestion and reduce flow
- Responsibility of end systems
- Effective on connectionless (datagram) networks
- Also used in connection-oriented networks
 - LAPF control is capable of detecting lost frames

Explicit Congestion Signaling

- Backward
 - Congestion avoidance notification in opposite direction to packet required
- Forward
 - Congestion avoidance notification in same direction as packet required



Explicit Signaling Categories

Binary

- A bit set in a packet indicates congestion

Credit based

- Indicates how many packets source may send
- Common for end-to-end flow control

Rate based

- Supply explicit data rate limit
- Nodes along path may request rate reduction

Traffic Management

Fairness

- Provide equal treatment of various flows

Quality of service

- Different treatment for different connections

Reservations

- Traffic contract between user and network
- Excess traffic discarded or handled on a best-effort basis

Traffic Shaping/Traffic Policing

- Two important tools in network management:
 - Traffic shaping
 - Concerned with traffic leaving the switch
 - Reduces packet clumping
 - Produces an output packet stream that is less bursty and with a more regular flow of packets
 - Traffic policing
 - Concerned with traffic entering the switch
 - Packets that don't conform may be treated in one of the following ways:
 - Give the packet lower priority compared to packets in other output queues
 - Label the packet as nonconforming by setting the appropriate bits in a header
 - Discard the packet



Token Bucket / Leaky Bucket

- Widely used traffic management tool
- Advantages:
 - Many traffic sources can be defined easily and accurately
 - Provides a concise description of the load to be imposed by a flow, enabling the service to determine easily the resource requirement
 - Provides the input parameters to a policing function

Token bucket

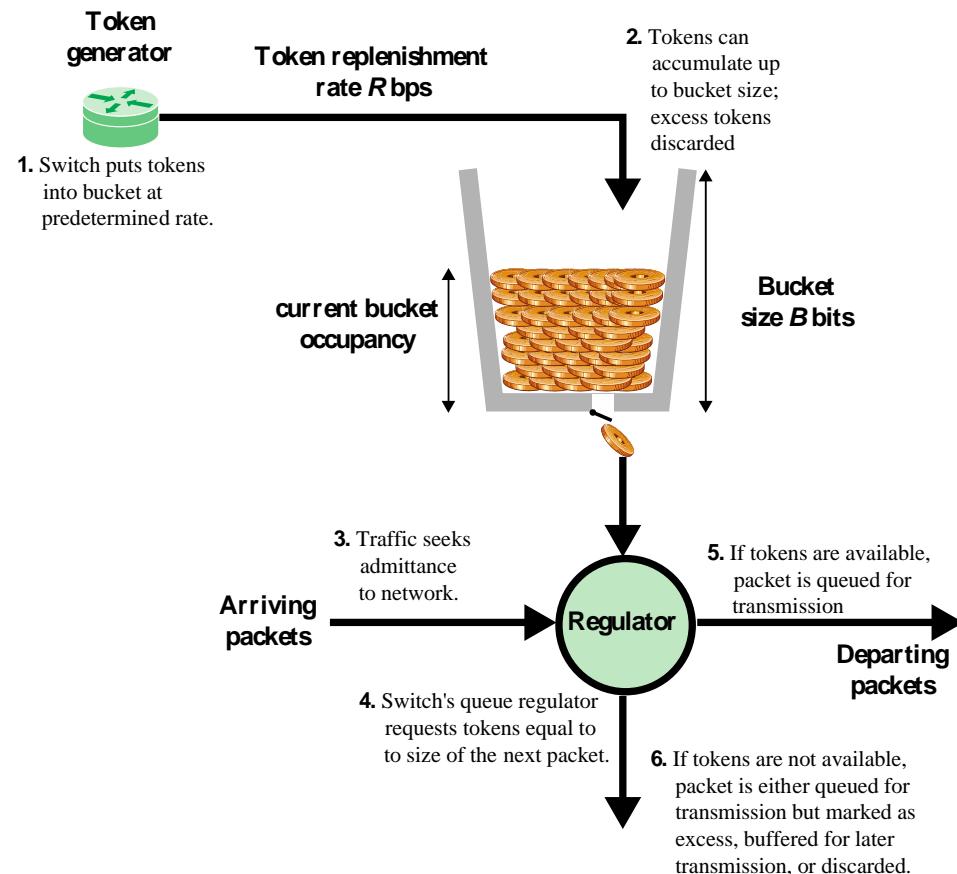


Figure 20.6 Token Bucket Scheme

Leaky bucket

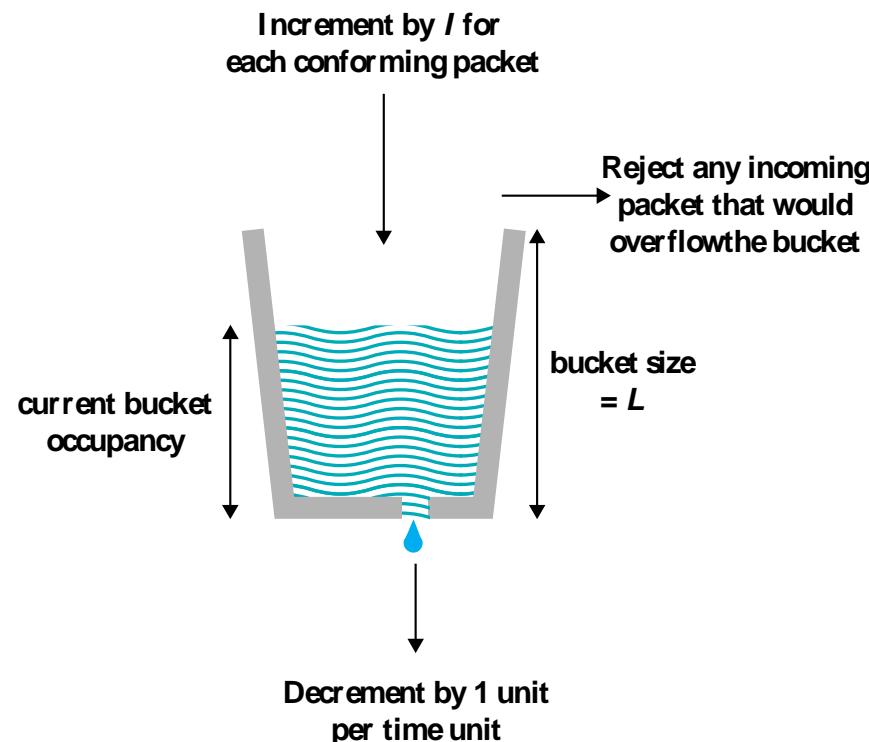
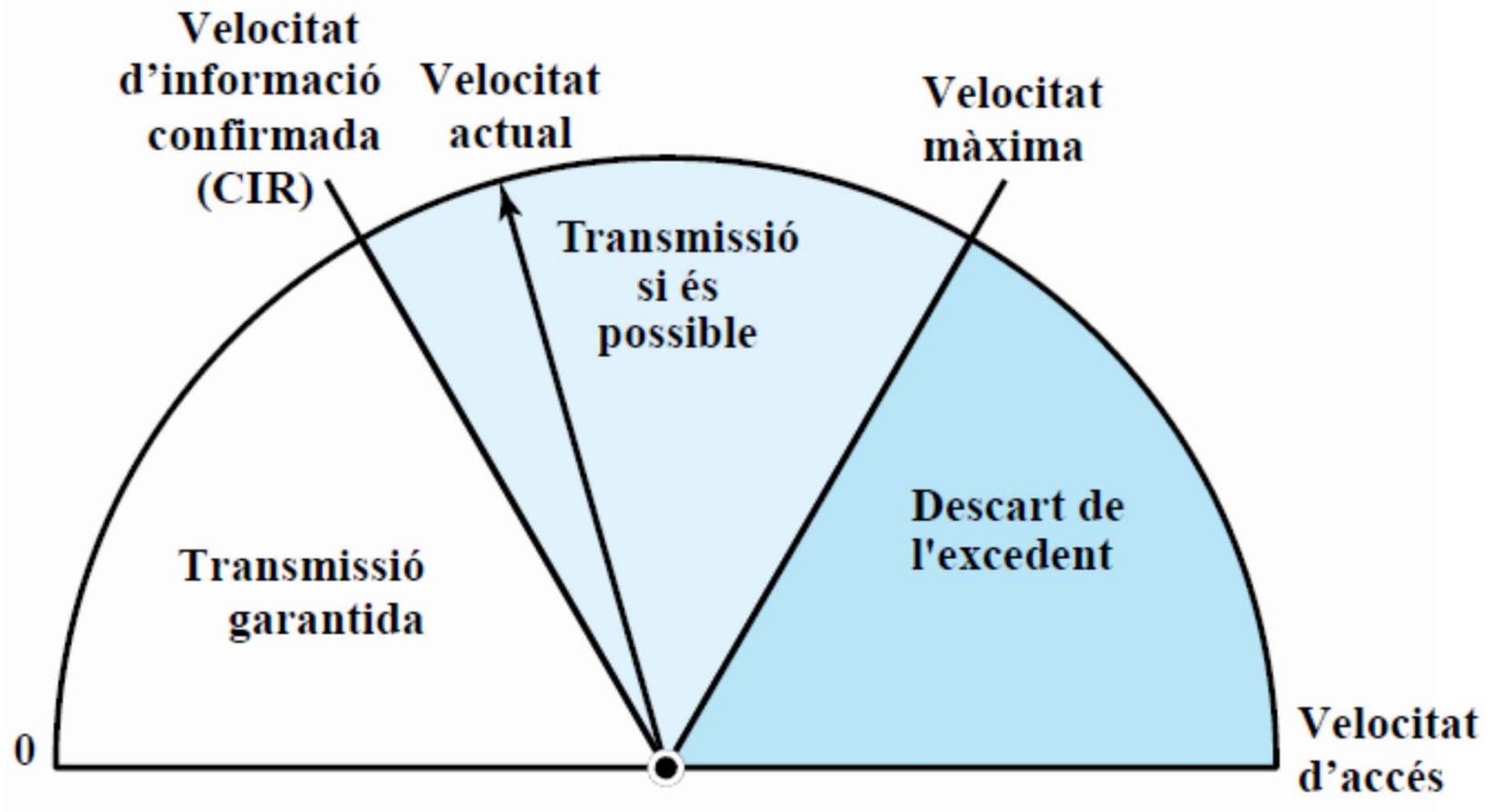


Figure 20.7 Leaky Bucket Algorithm

Example: FR congestion control

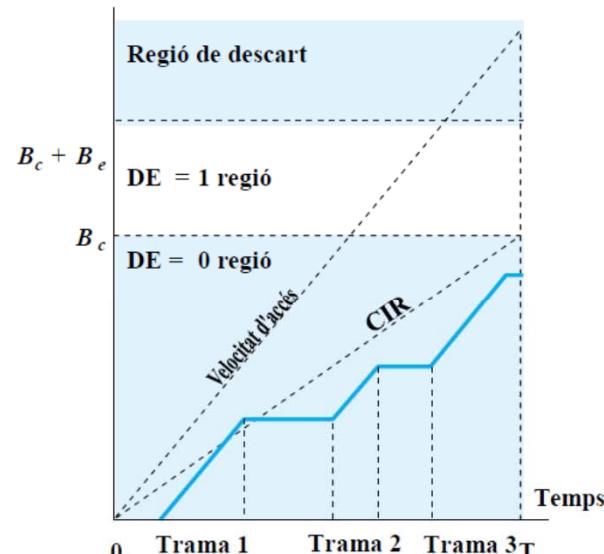
Tècnica	Tipus	Funció	Elements clau
Control de descart	Estratègia de descart	Proporciona una guia per a la xarxa sobre les trames que cal descartar	Bit DE
Notificació de la congestió explícita cap enrere	Elisió de la congestió	Proporciona una guia per als sistemes finals sobre la congestió a la xarxa	Bit BECN o missatge CLLM
Notificació de la congestió explícita cap endavant	Elisió de la congestió	Proporciona una guia per als sistemes finals sobre la congestió a la xarxa	Bit FECN
Notificació de la congestió implícita	Recuperació de la congestió	El sistema final infereix la congestió de la pèrdua de trames	Números de seqüència a PDU de capa superior

CIR strategy

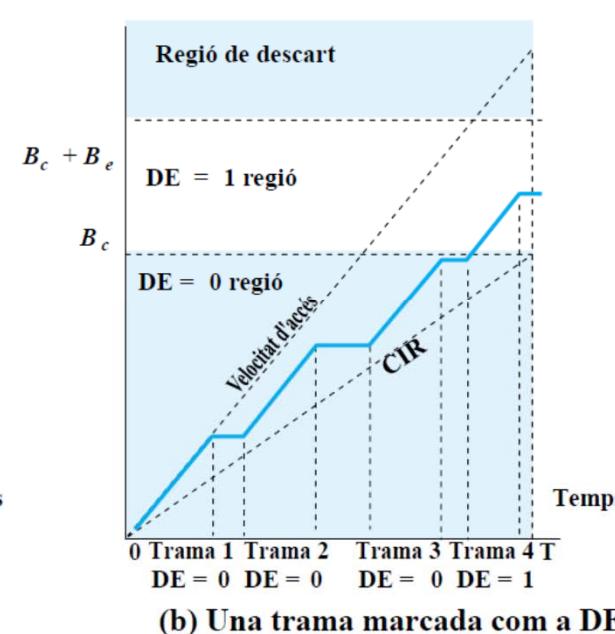


Congestion control

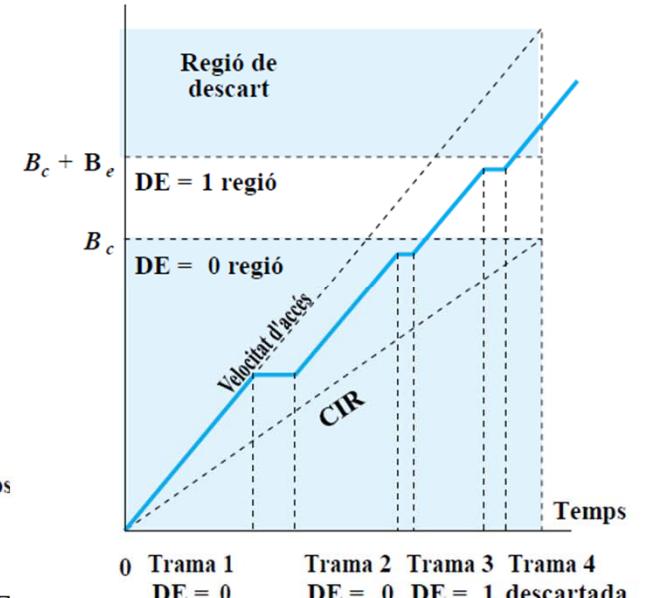
Nombre de bits transmesos



Nombre de bits transmesos



Nombre de bits transmesos



$$T = \frac{B_c}{CIR}$$