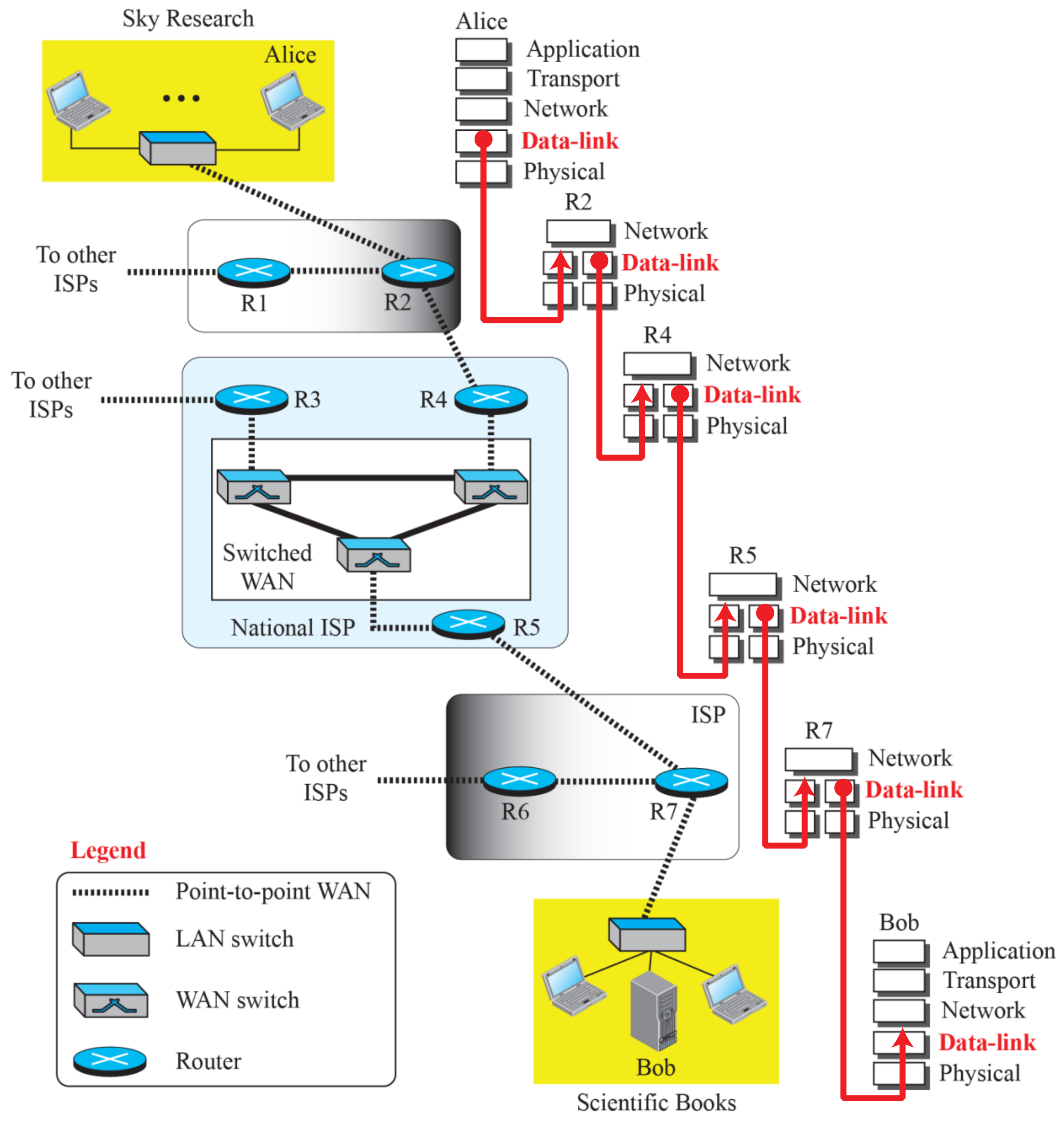
UNIT-1

Data Link Layer

INTRODUCTION

The Internet is a combination of networks glued together by connecting devices (routers or switches). If a datagram is to travel from a host to another host, it needs to pass through these networks.

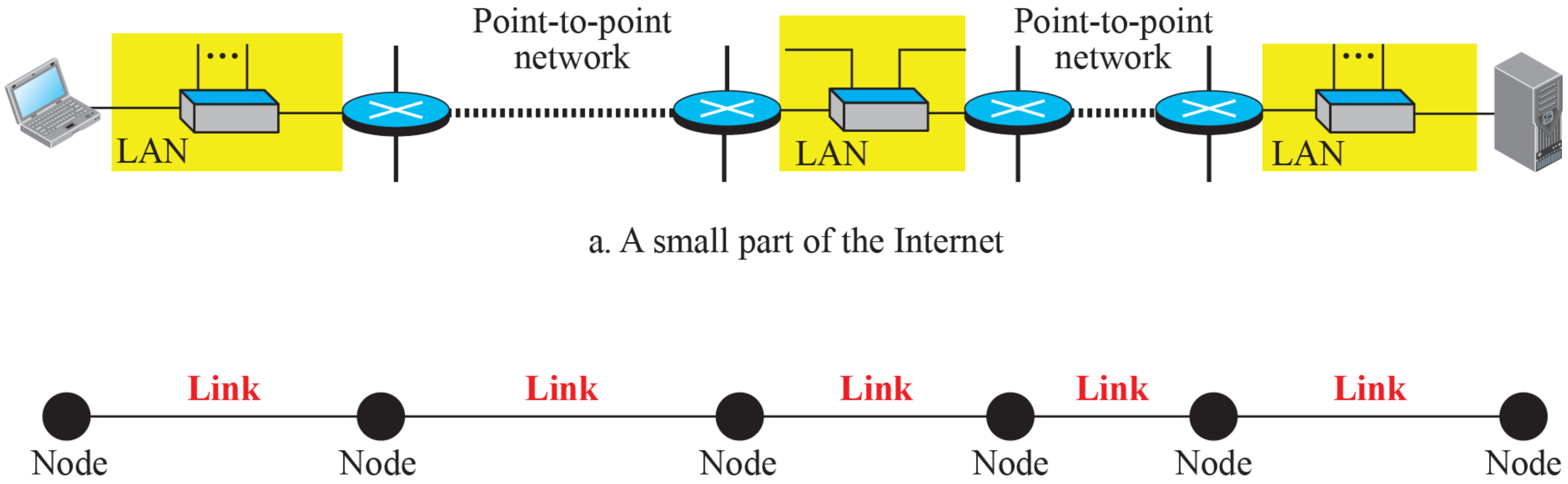
Communication at the data-link layer



Nodes and Links

Although communication at the application, transport, and network layers is end-to-end, communication at the data-link layer is node-to-node. As we have learned in the previous chapters, a data unit from one point in the Internet needs to pass through many networks (LANs and WANs) to reach another point. Theses LANs and WANs are connected by routers. It is customary to refer to the two end hosts and the routers as nodes and the networks in between as links.

Nodes and Links



a. A small part of the Internet

b. Nodes and links

Two Types of Links

Although two nodes are physically connected by a transmission medium such as cable or air, we need to remember that the data-link layer controls how the medium is used. We can have a data-link layer that uses the whole capacity of the medium; we can also have a data-link layer that uses only part of the capacity of the link. In other words, we can have a

- Point-to-Point link
- Broadcast link.

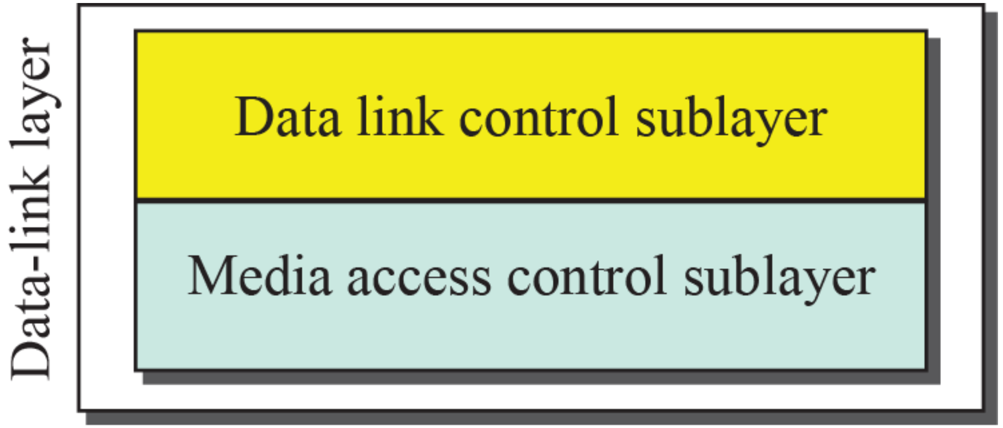
Two Sublayers

To better understand the functionality of and the services provided by the link layer, we can divide the data-link layer into two sublayers:

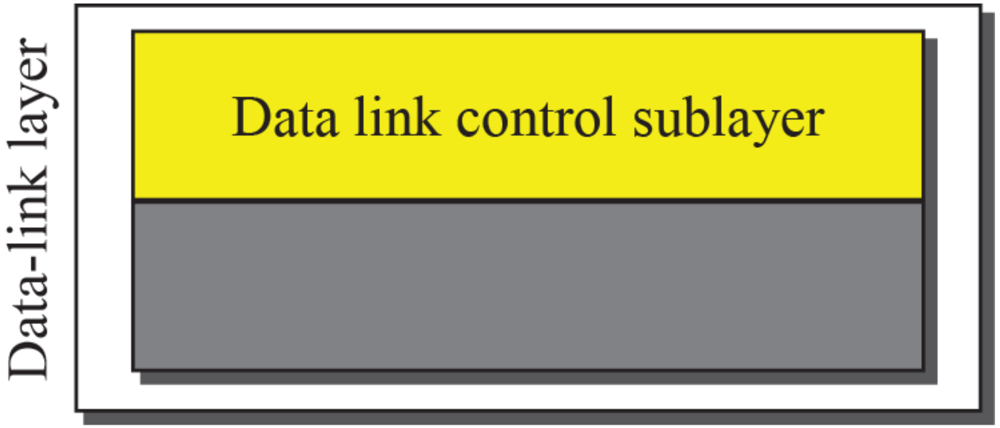
- Data Link Control (DLC) and
- Media Access Control (MAC).

The data link control sublayer deals with all issues common to both point-to-point and broadcast links;
the media access control sublayer deals only with issues specific to broadcast links.

Dividing the data-link layer into two sublayers



a. Data-link layer of a broadcast link

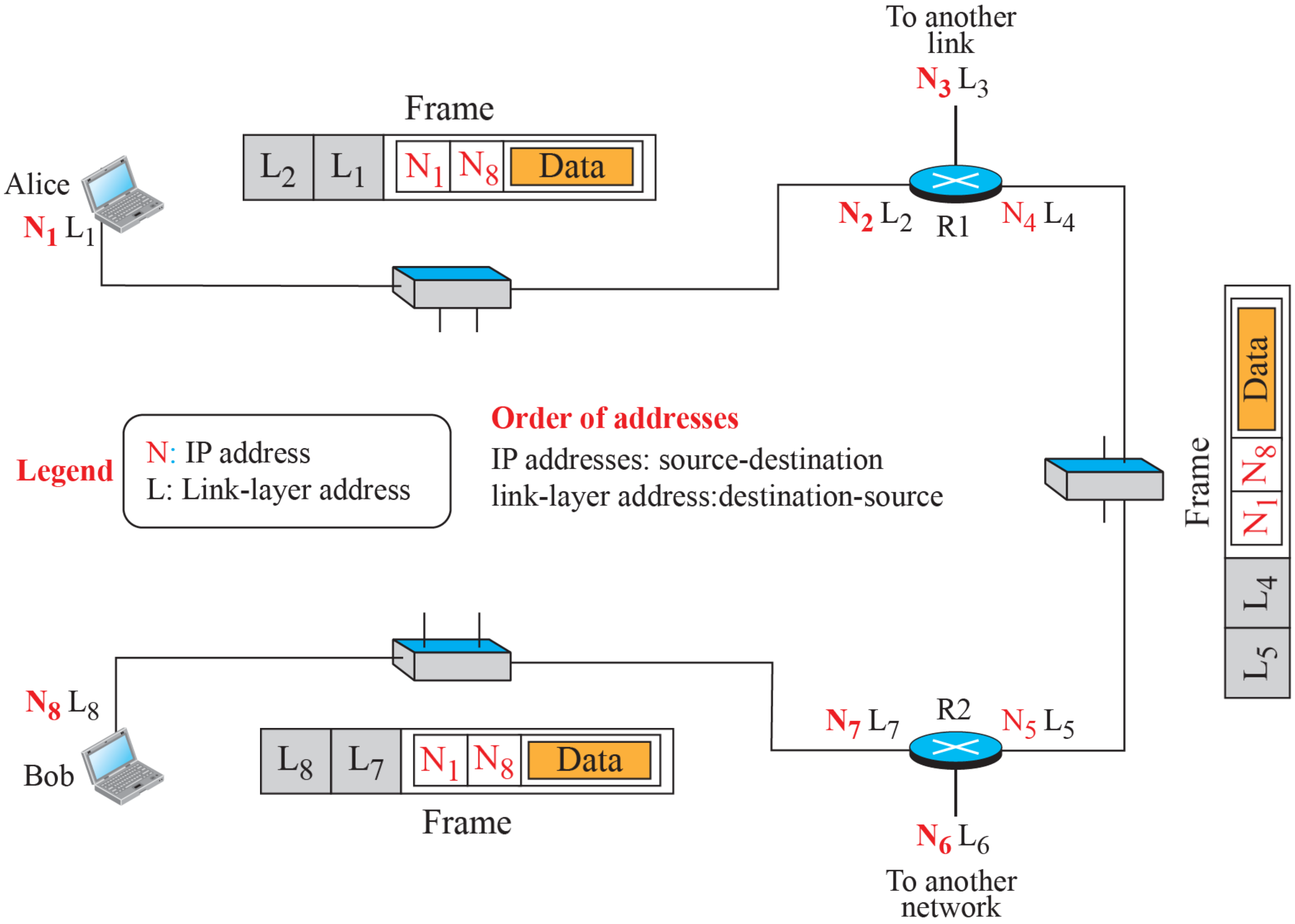


b. Data-link layer of a point-to-point link

DATA LINK CONTROL (DLC)

The data link control deals with procedures for communication between two adjacent nodes. Data link control (DLC) functions include addressing, framing, flow and error control, and error detection and correction.

IP addresses and link-layer addresses in a small internet



Legend

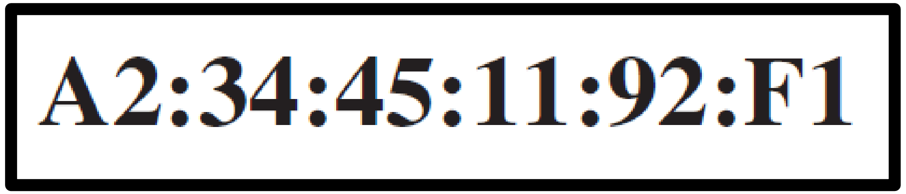
N: IP address
L: Link-layer address

Order of addresses

IP addresses: source-destination
link-layer address:destination-source

LINK-LAYER ADDRESSING

The link-layer addresses in the most common LAN, Ethernet, are 48 bits (six bytes) that are presented as 12 hexadecimal digits separated by colons; for example, the following is a link-layer address of a computer.



A2:34:45:11:92:F1

Three Types of addresses

Unicast Address

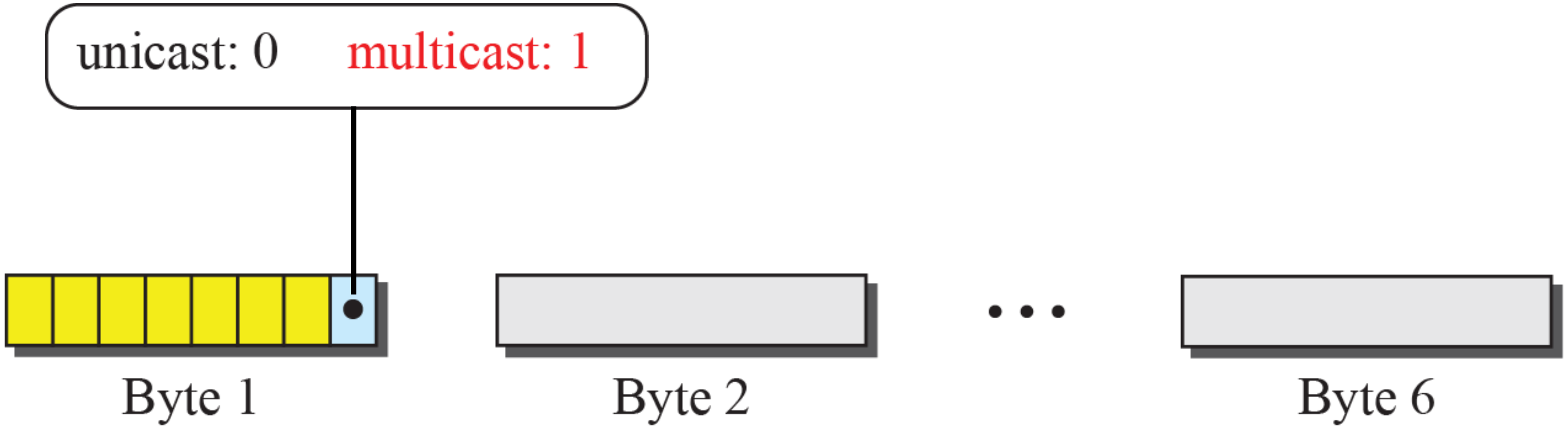
Each host or each interface of a router is assigned a unicast address. Unicasting means one-to-one communication. A frame with a unicast address destination is destined only for one entity in the link.

Multicast Address

Some link-layer protocols define multicast addresses. Multicasting means one-to-many communication. However, the jurisdiction is local (inside the link).

Broadcast Address

Some link-layer protocols define a broadcast address. Broadcasting means one-to-all communication. A frame with a destination broadcast address is sent to all entities in the link. (**FF:FF:FF:FF:FF:FF**)



Define the type of the following destination addresses:

a. 4A:30:10:21:10:1A

b. 47:20:1B:2E:08:EE

c. FF:FF:FF:FF:FF:FF

Solution

To find the type of the address, we need to look at the second hexadecimal digit from the left. If it is even, the address is unicast. If it is odd, the address is multicast. If all digits are Fs, the address is broadcast. Therefore, we have the following:

a. This is a unicast address because A in binary is 1010 (even).

b. This is a multicast address because 7 in binary is 0111 (odd).

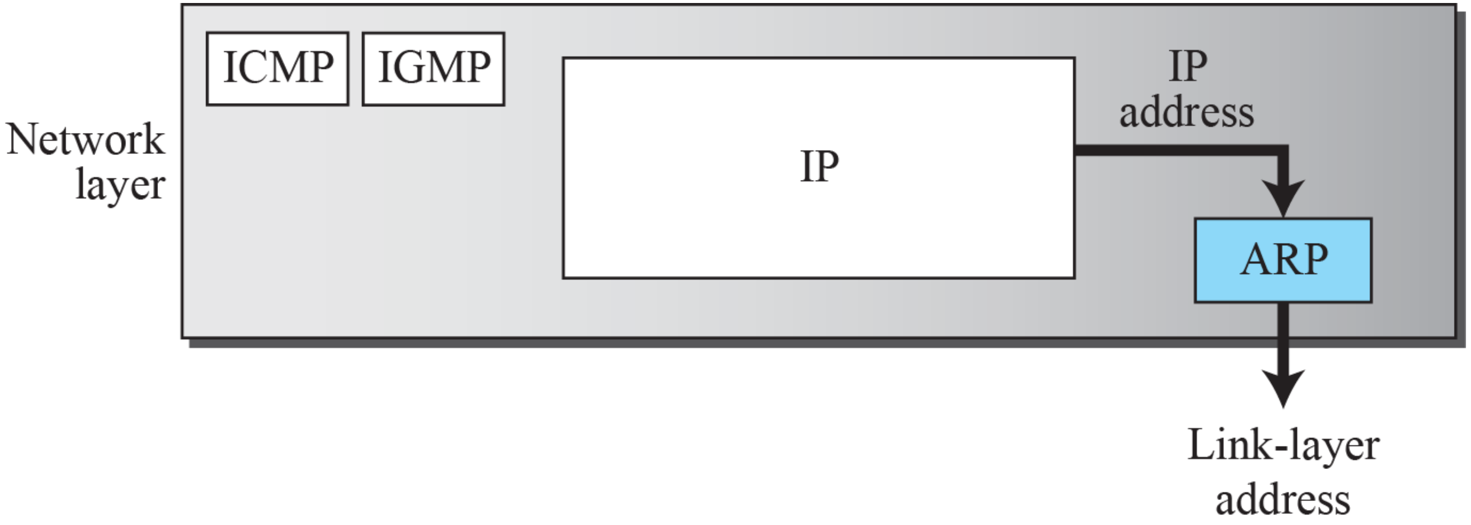
c. This is a broadcast address because all digits are Fs in hexadecimal.

Address Resolution Protocol (ARP)

- Address Resolution Protocol (ARP)
  - Packet Format
- An Example
  - Activities at the Alice Site
  - Activities at Routers
  - Activities at Bob’s Site

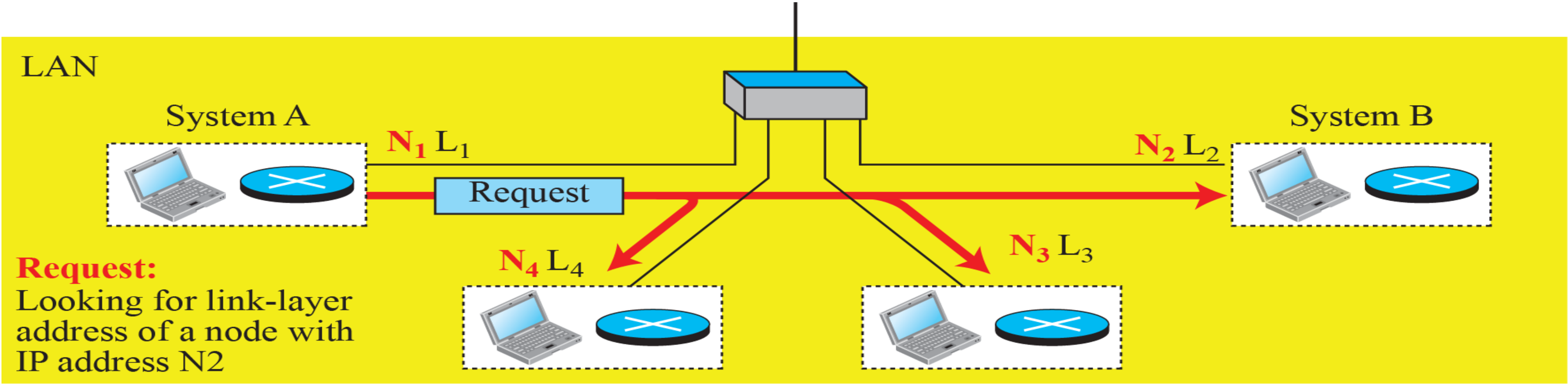
Position of ARP in TCP/IP protocol suite

- The IP address of the next node is not helpful in moving a frame through a link; but we need the link-layer address of the next node. This is the time when the **Address Resolution Protocol (ARP)** becomes helpful.
- It belongs to the network layer, ARP accepts an IP address from the IP protocol, maps the address to the corresponding link-layer address, and passes it to the data-link layer.



ARP operation

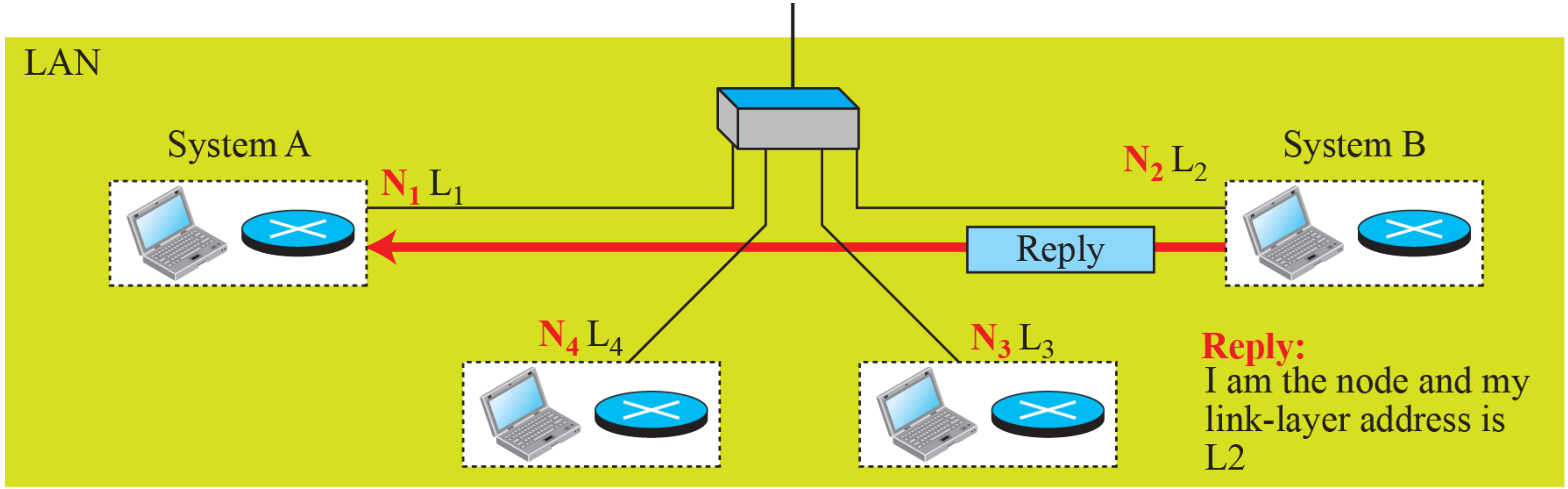
Anytime a host or a router needs to find the link-layer address of another host or router in its network, it sends an **ARP request packet**. The packet includes the link-layer and IP addresses of the sender and the IP address of the receiver. Because the sender does not know the link-layer address of the receiver, the query is broadcast over the link using the link-layer broadcast address.



a. ARP request is broadcast

ARP operation

Every host or router on the network receives and processes the ARP request packet, but only the intended recipient recognizes its IP address and sends back an ARP response packet. The response packet contains the recipient’s IP and link-layer addresses. The packet is unicast directly to the node that sent the request packet.

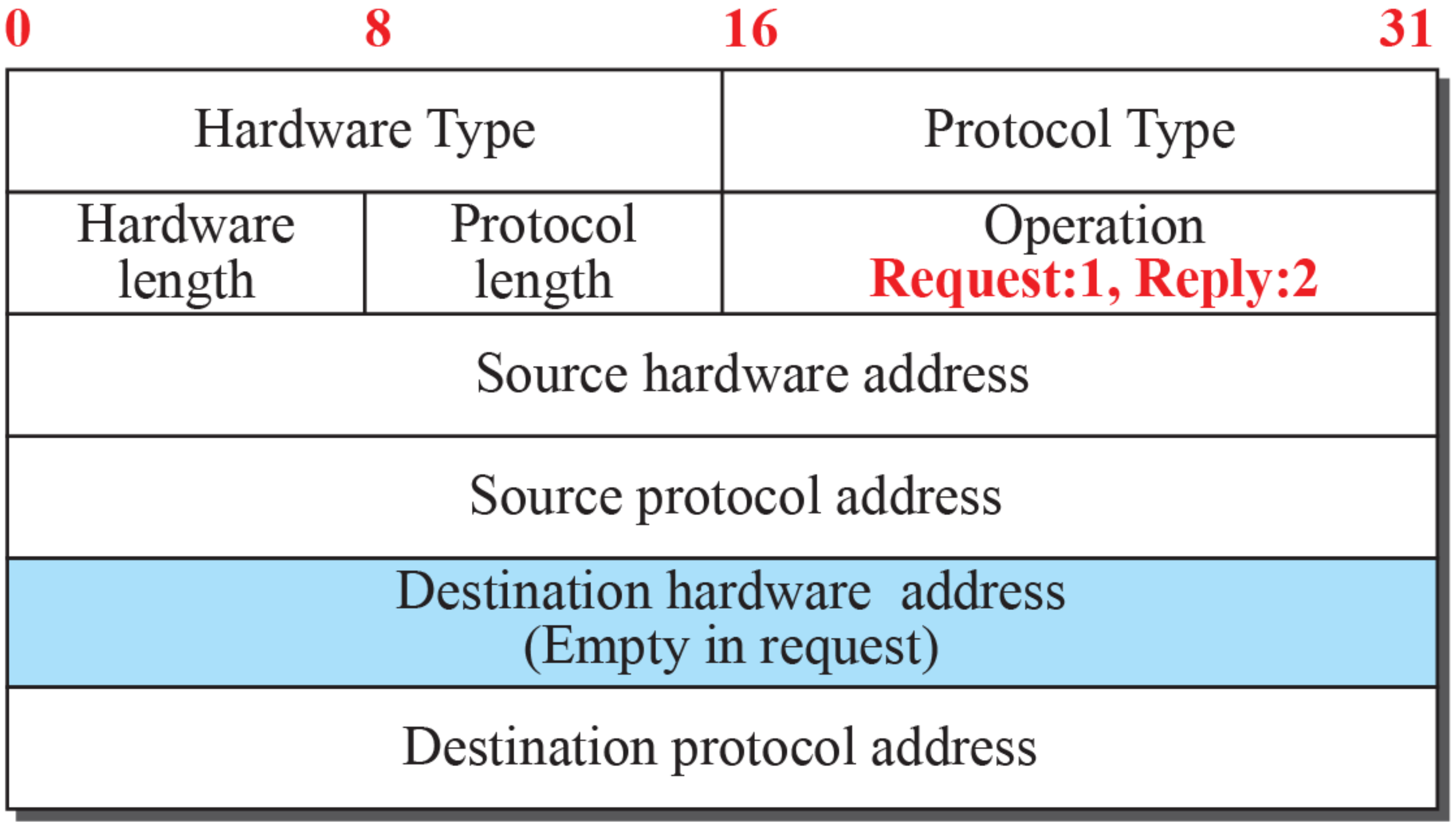


b. ARP reply is unicast

ARP packet format

Hardware: LAN or WAN protocol

Protocol: Network-layer protocol



0 8 16 31

Hardware Type		Protocol Type
Hardware length	Protocol length	Operation **Request:1, Reply:2**
Source hardware address		
Source protocol address		
Destination hardware address (Empty in request)		
Destination protocol address		

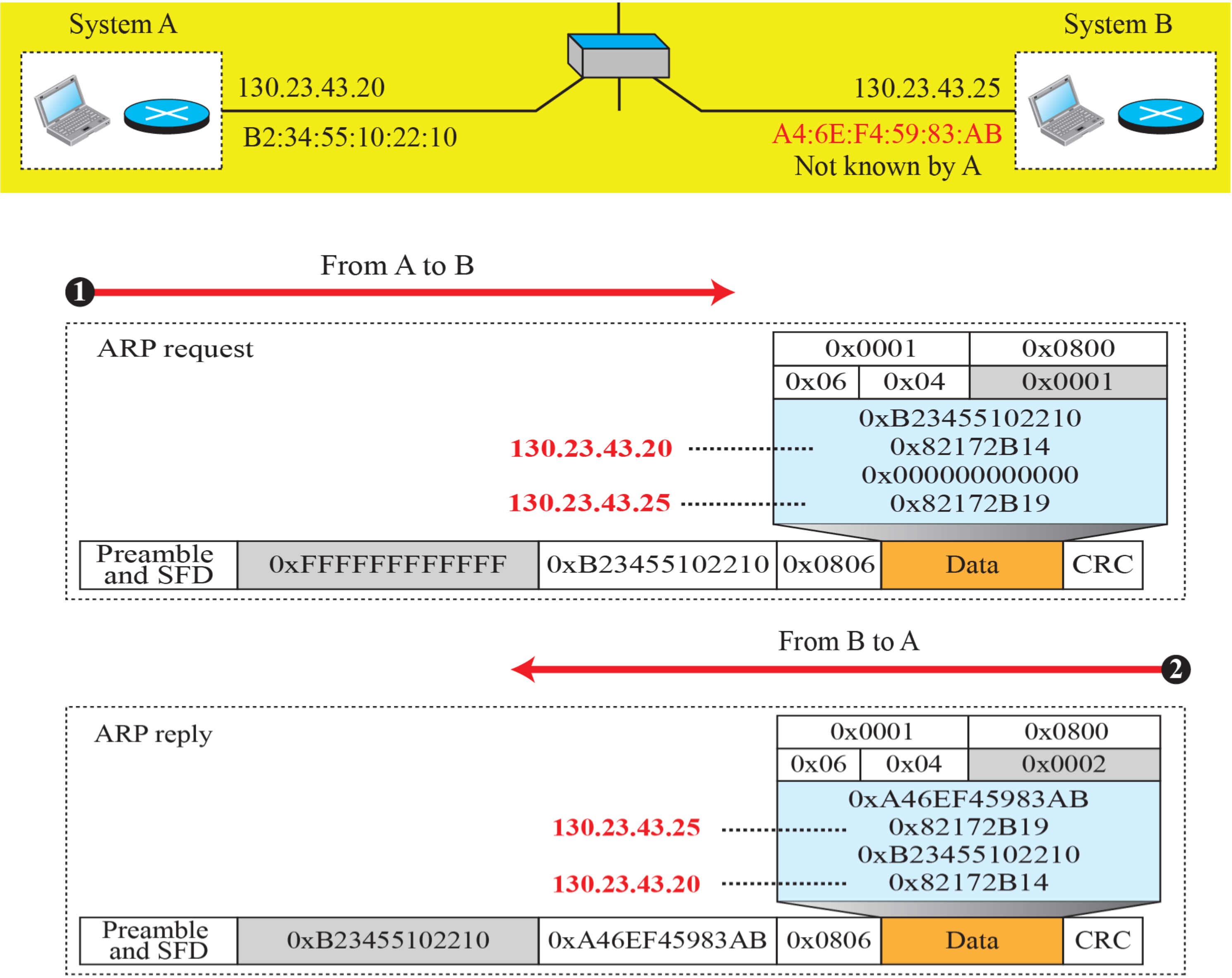
The *hardware type* field defines the type of the link-layer protocol; Ethernet is given the type 1.

The *protocol type* field defines the network-layer protocol: IPv4 protocol is $(0800)_{16}$.

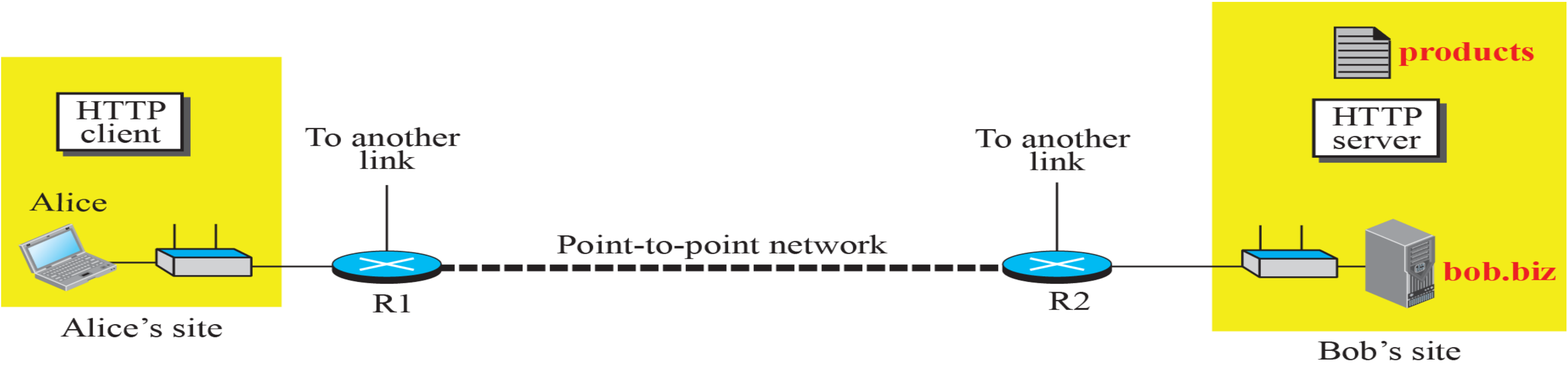
The source hardware and source protocol addresses are variable-length fields defining the link-layer and network-layer addresses of the sender.

The destination hardware address and destination protocol address fields define the receiver link-layer and network-layer addresses.

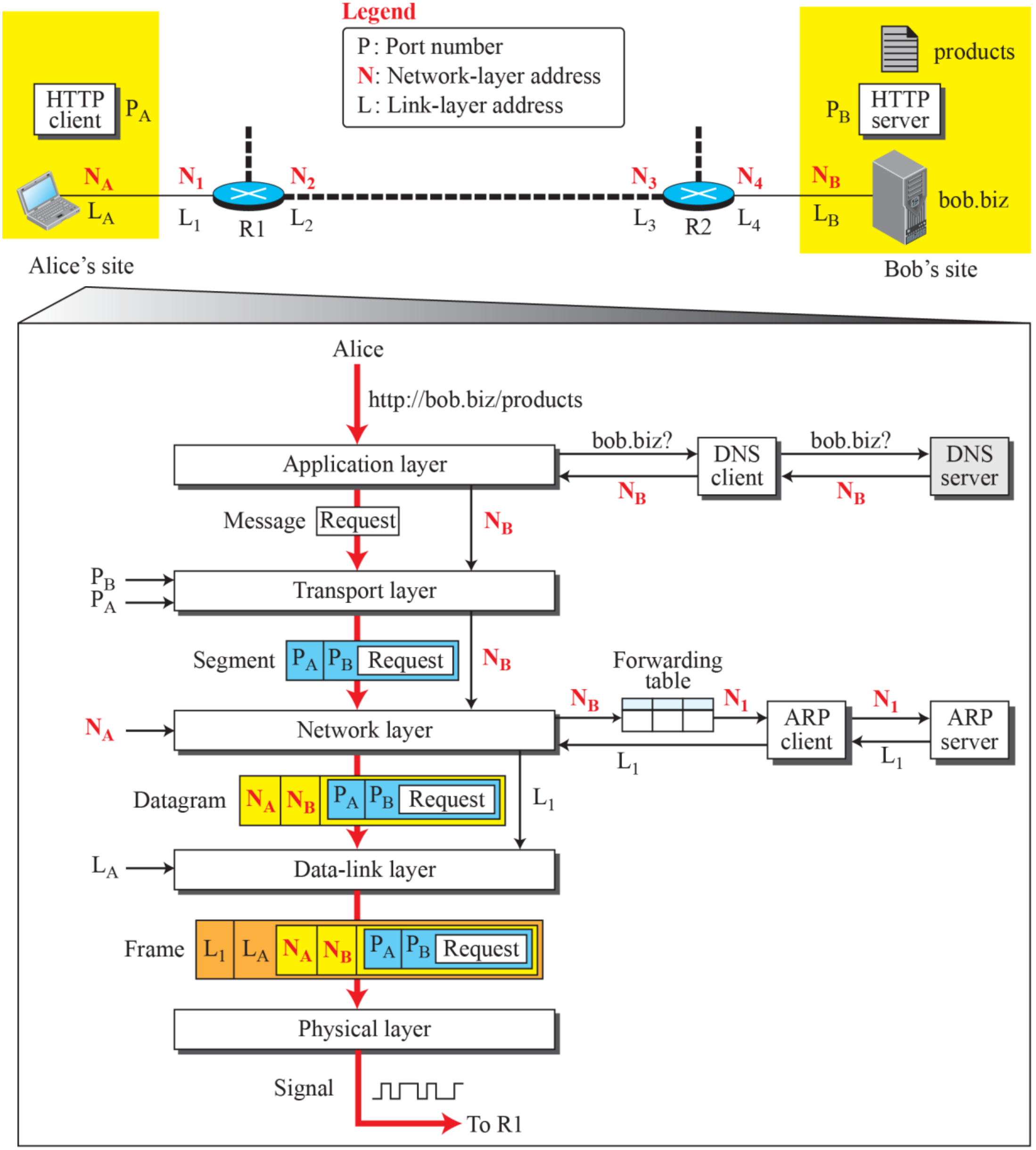
Example


System A
System B
130.23.43.20
B2:34:55:10:22:10
130.23.43.25
A4:6E:F4:59:83:AB
Not known by A
From A to B
1
ARP request
0x0001
0x0800
0x06
0x04
0x0001
0xB23455102210
130.23.43.20
0x82172B14
0x000000000000
130.23.43.25
0x82172B19
Preamble and SFD
0xFFFFFFFFFFFF
0xB23455102210
0x0806
Data
CRC
From B to A
2
ARP reply
0x0001
0x0800
0x06
0x04
0x0002
0xA46EF45983AB
130.23.43.25
0x82172B19
0xB23455102210
130.23.43.20
0x82172B14
Preamble and SFD
0xB23455102210
0xA46EF45983AB
0x0806
Data
CRC

An Example of Communication

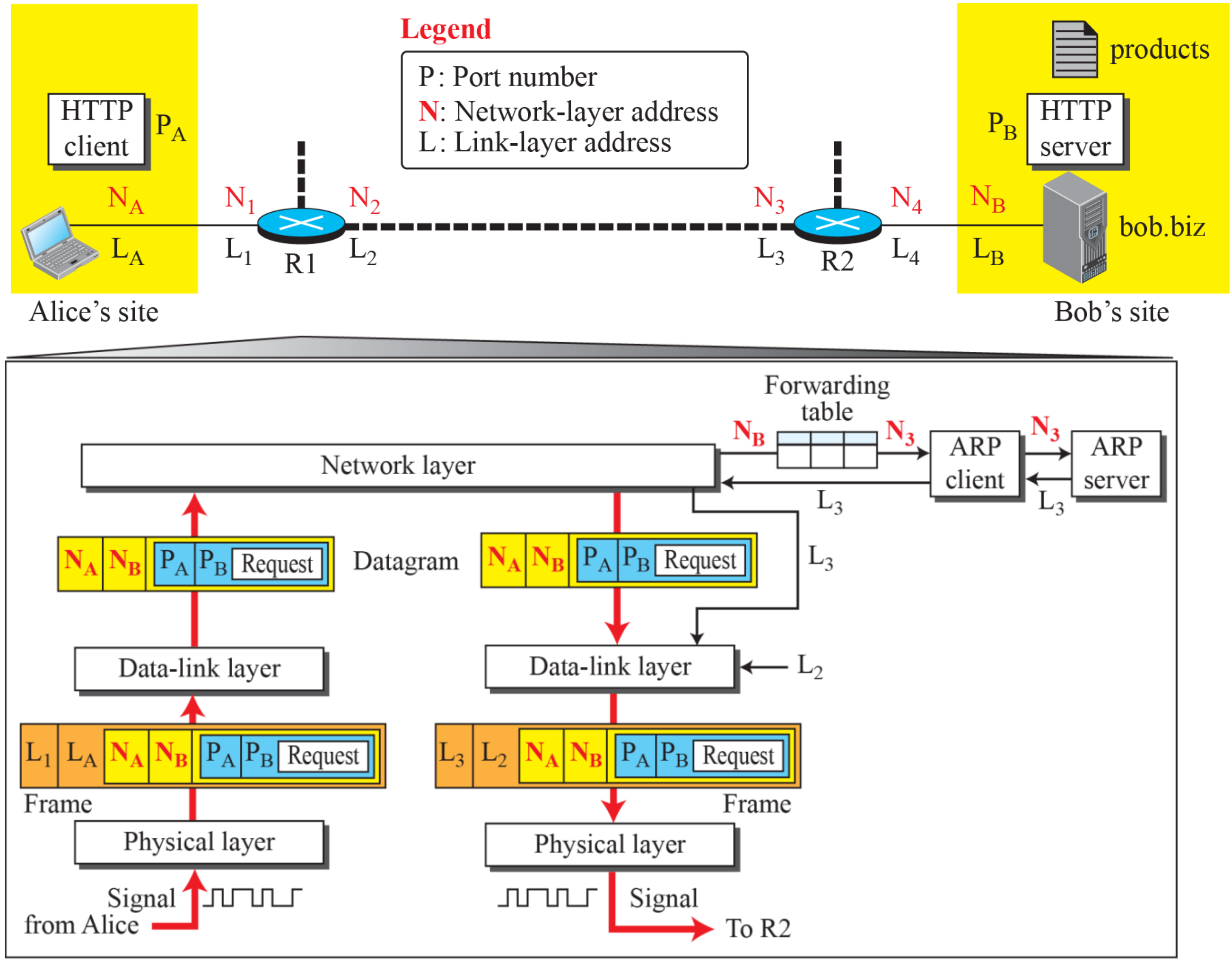


Flow of packets at Alice's computer



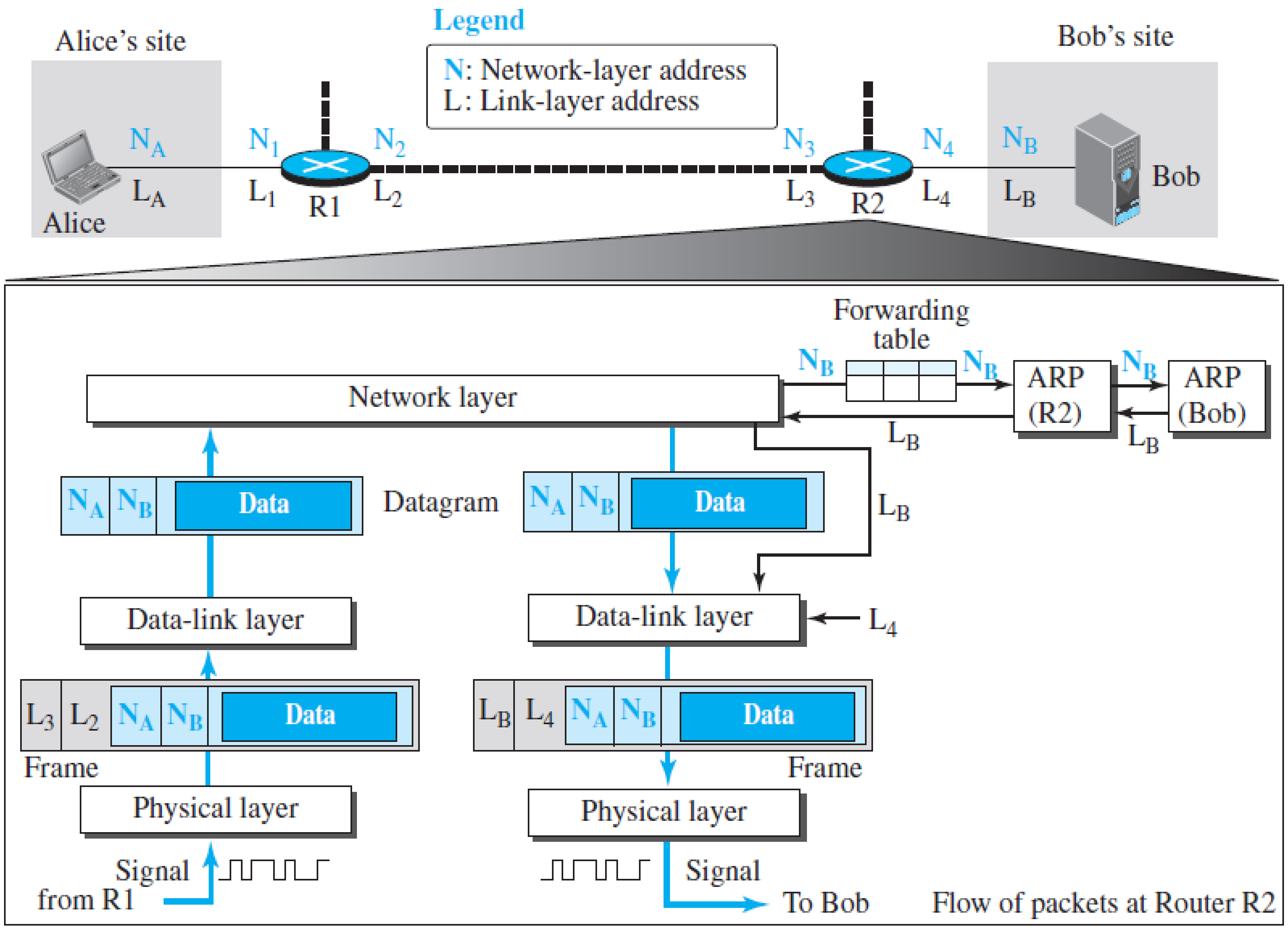
Flow of packets at Alice's computer

Flow of activities at router R1

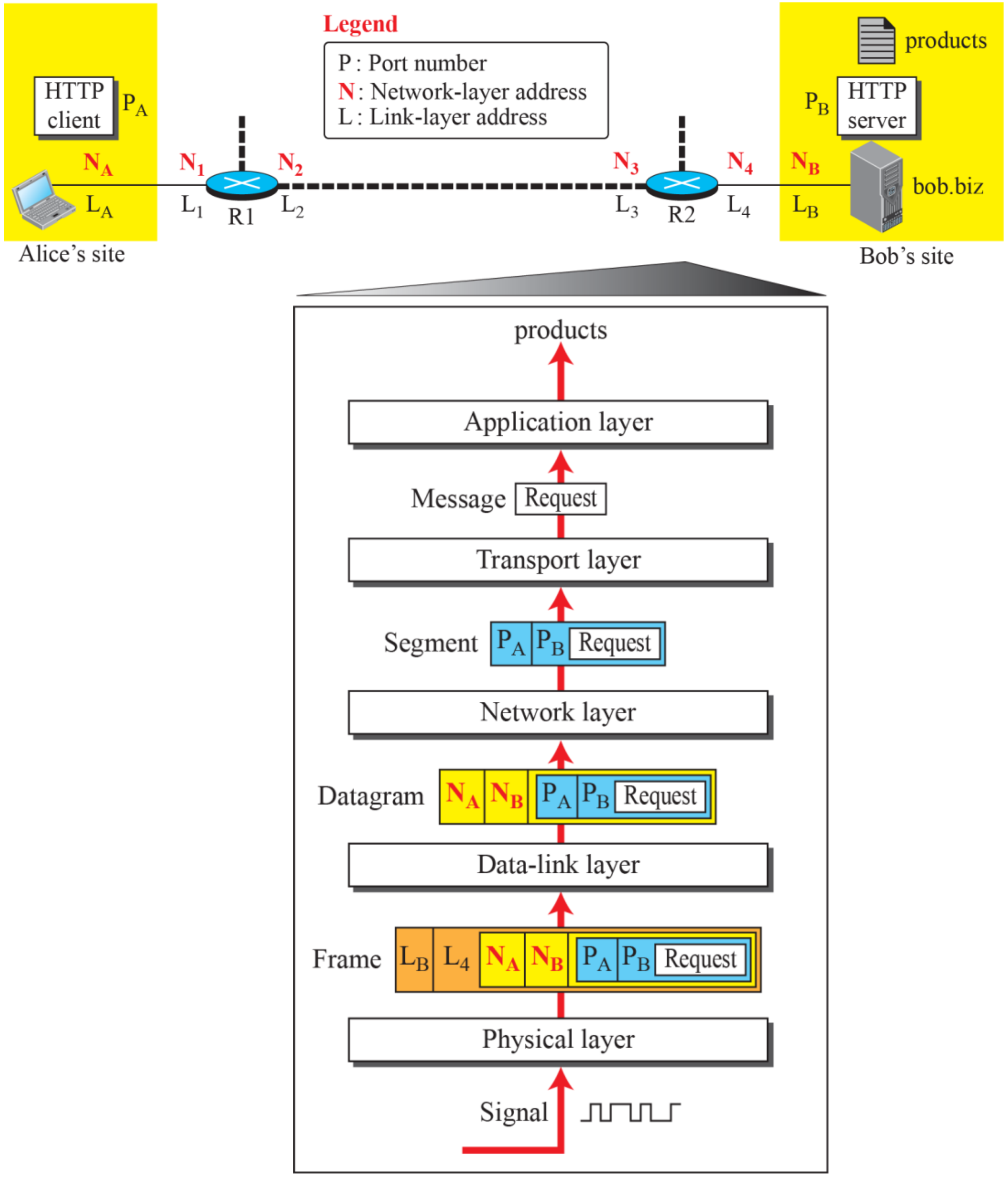


Flow of packets at Router R1

Flow of activities at router R2



Activities at Bob’s site



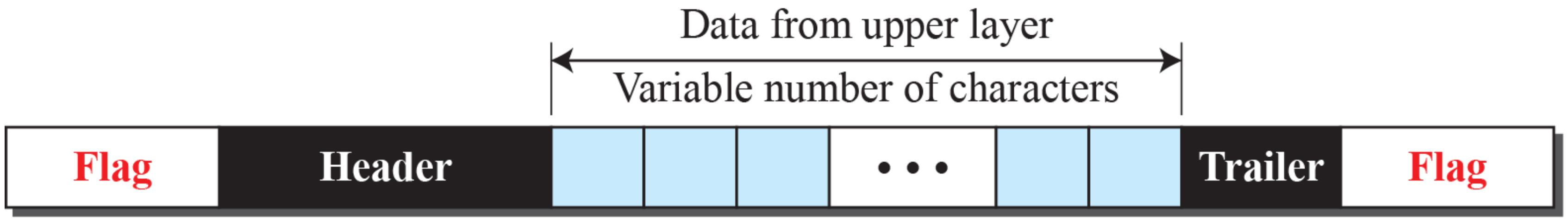
Flow of packets at Bob’s computer

Framing

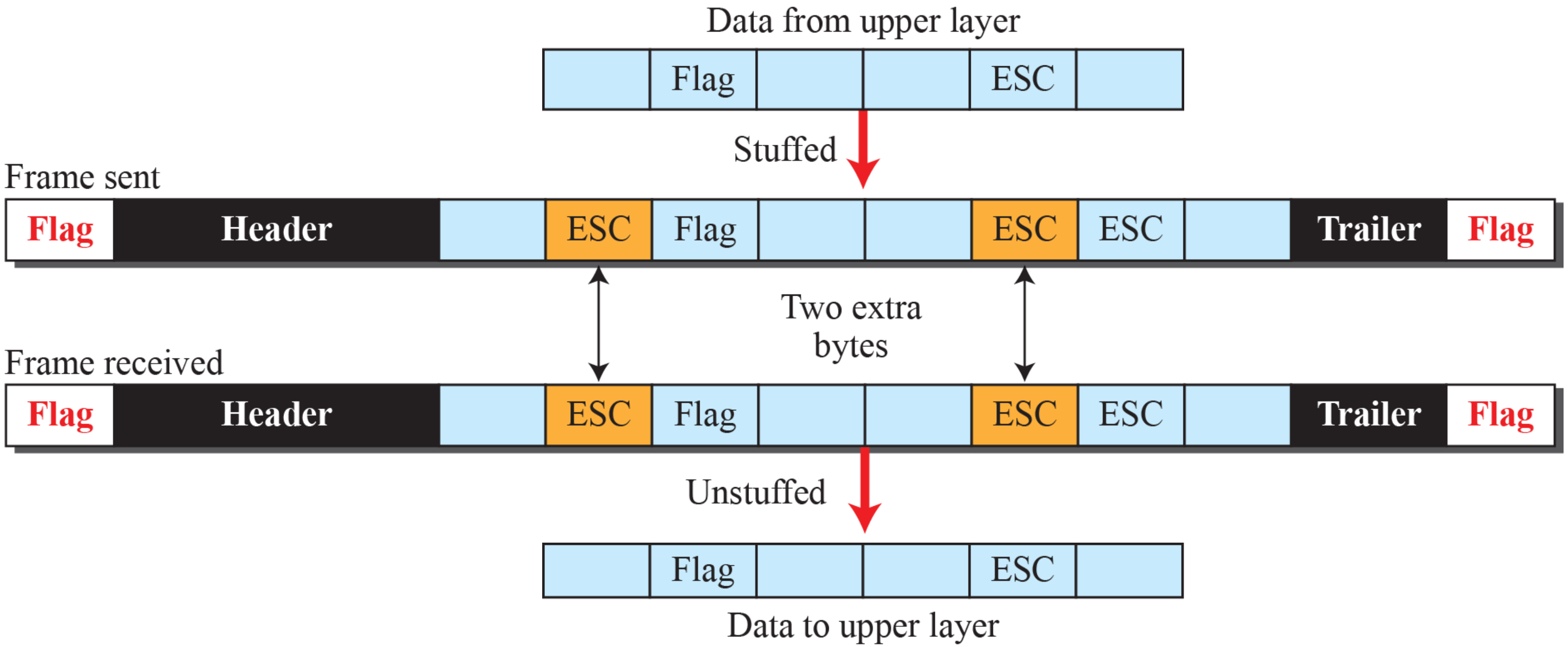
Data transmission in the physical layer means moving bits in the form of a signal from the source to the destination. The data-link layer, on the other hand, needs to pack bits into frames, so that each frame is distinguishable from another.

- Frame Size
  - Character-Oriented Framing
  - Bit-Oriented Framing

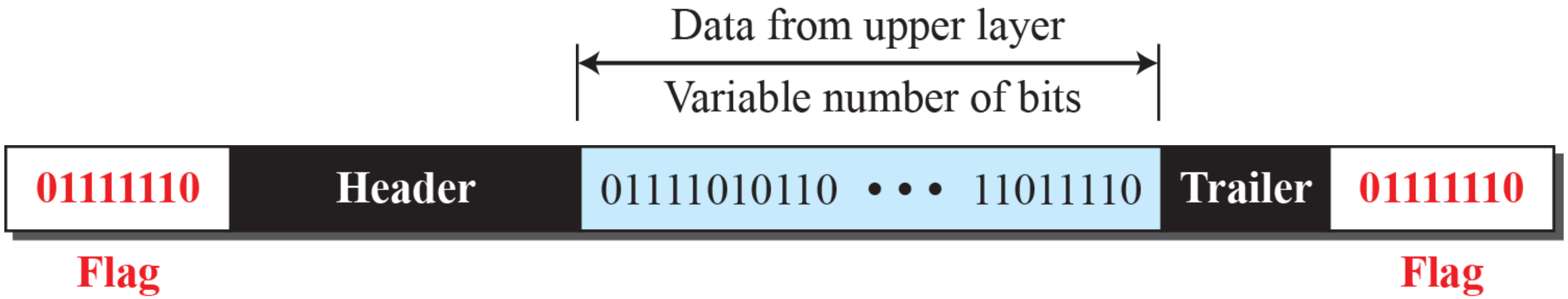
A frame in a character-oriented protocol



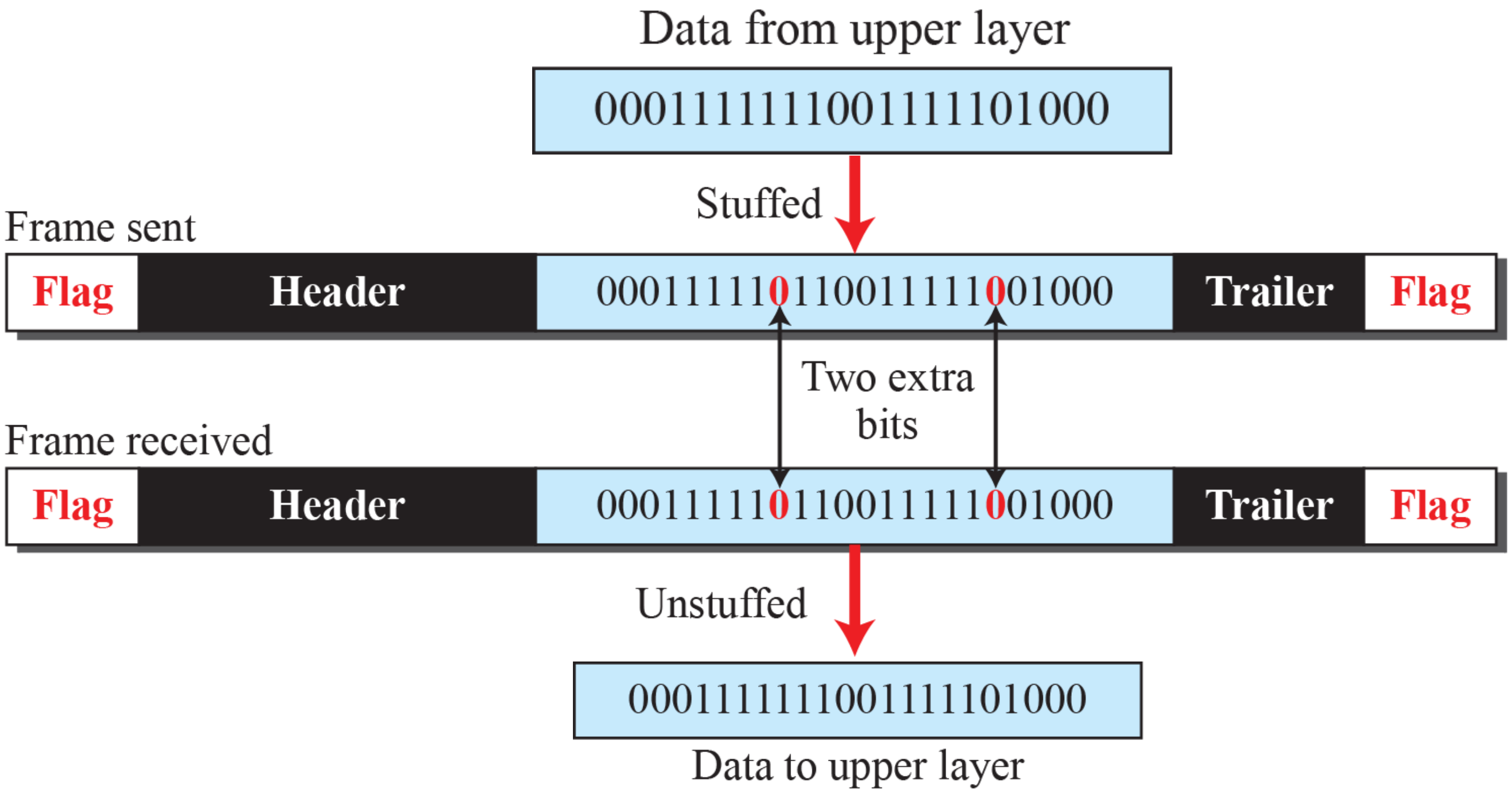
Byte stuffing and unstuffing



A frame in a bit-oriented protocol



Bit stuffing and unstuffing



Flow and Error Control

One of the responsibilities of the data-link control sublayer is flow and error control at the data-link layer.

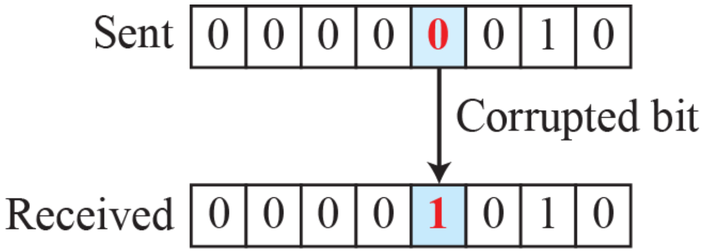
- ❑ Flow Control
- ❑ Error Control

Error Detection and Correction

At the data-link layer, if a frame is corrupted between the two nodes, it needs to be corrected before it continues its journey to other nodes. However, most link-layer protocols simply discard the frame and let the upper-layer protocols handle the retransmission of the frame. Some wireless protocols, however, try to correct the corrupted frame.

- Introduction
  - Types of Errors
  - Redundancy
  - Detection versus Correction
  - Coding
- Cyclic Codes
  - Cyclic Redundancy Check
  - Polynomials
  - Requirement
  - Performance
  - Advantages of Cyclic Codes

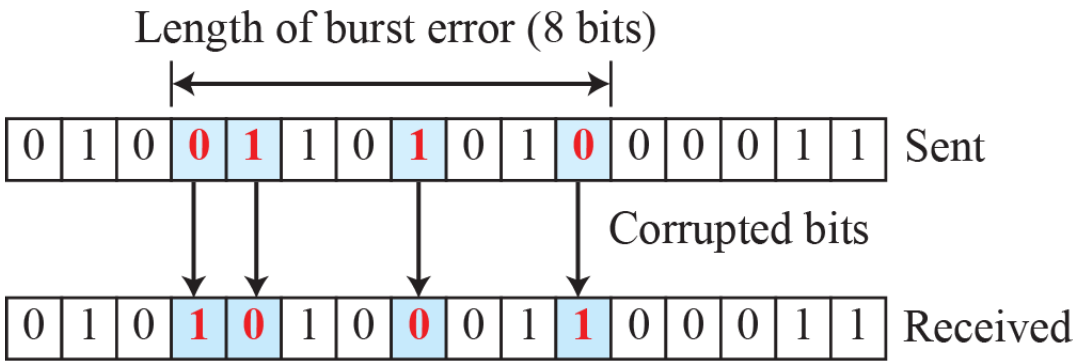
Single-bit and burst error



Sent: 0 0 0 0 **0** 0 1 0

Received: 0 0 0 0 **1** 0 1 0

a. Single-bit error

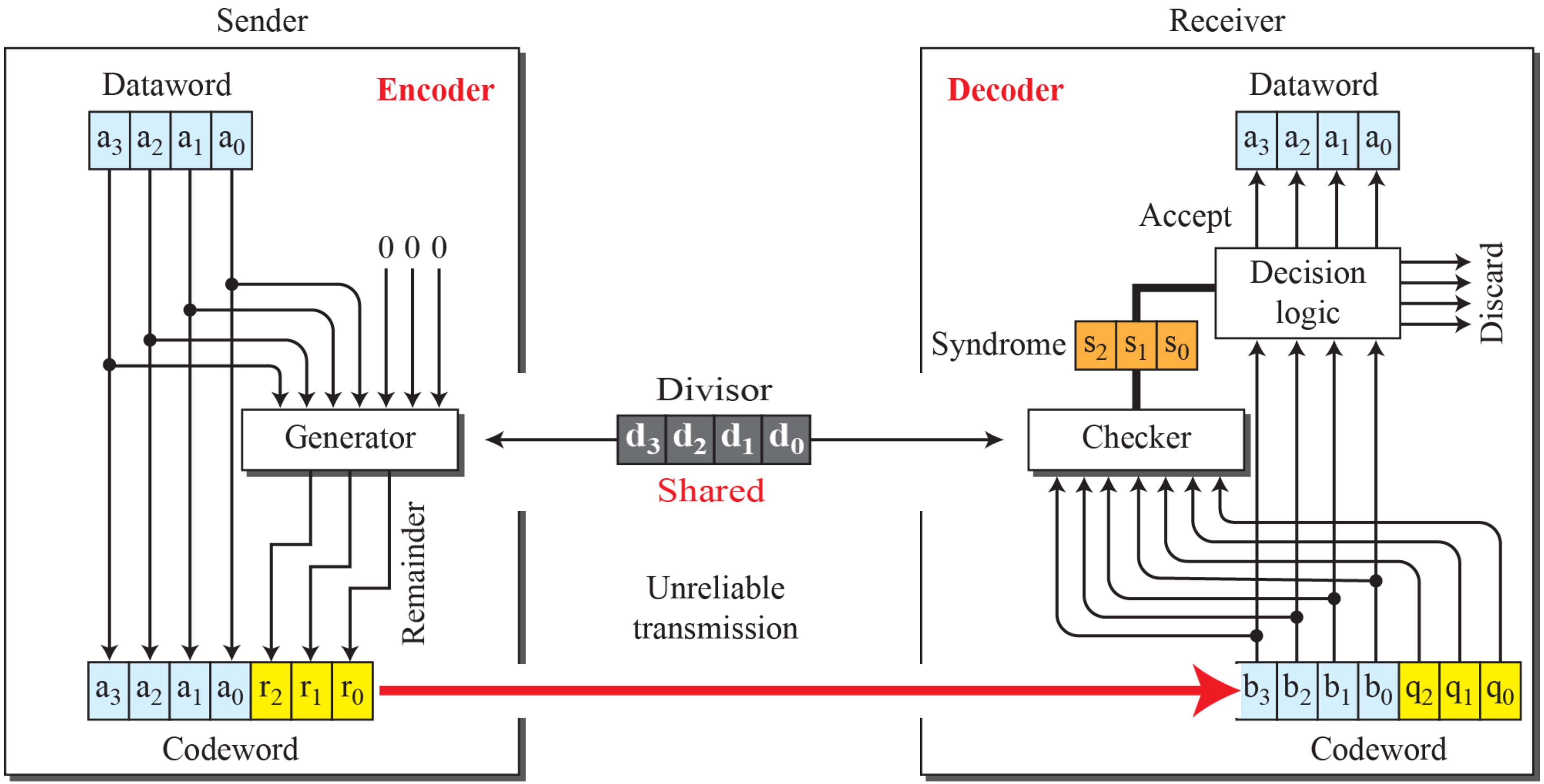


Sent: 0 1 0 **0** **1** 1 0 **1** 0 1 **0** 0 0 0 1 1

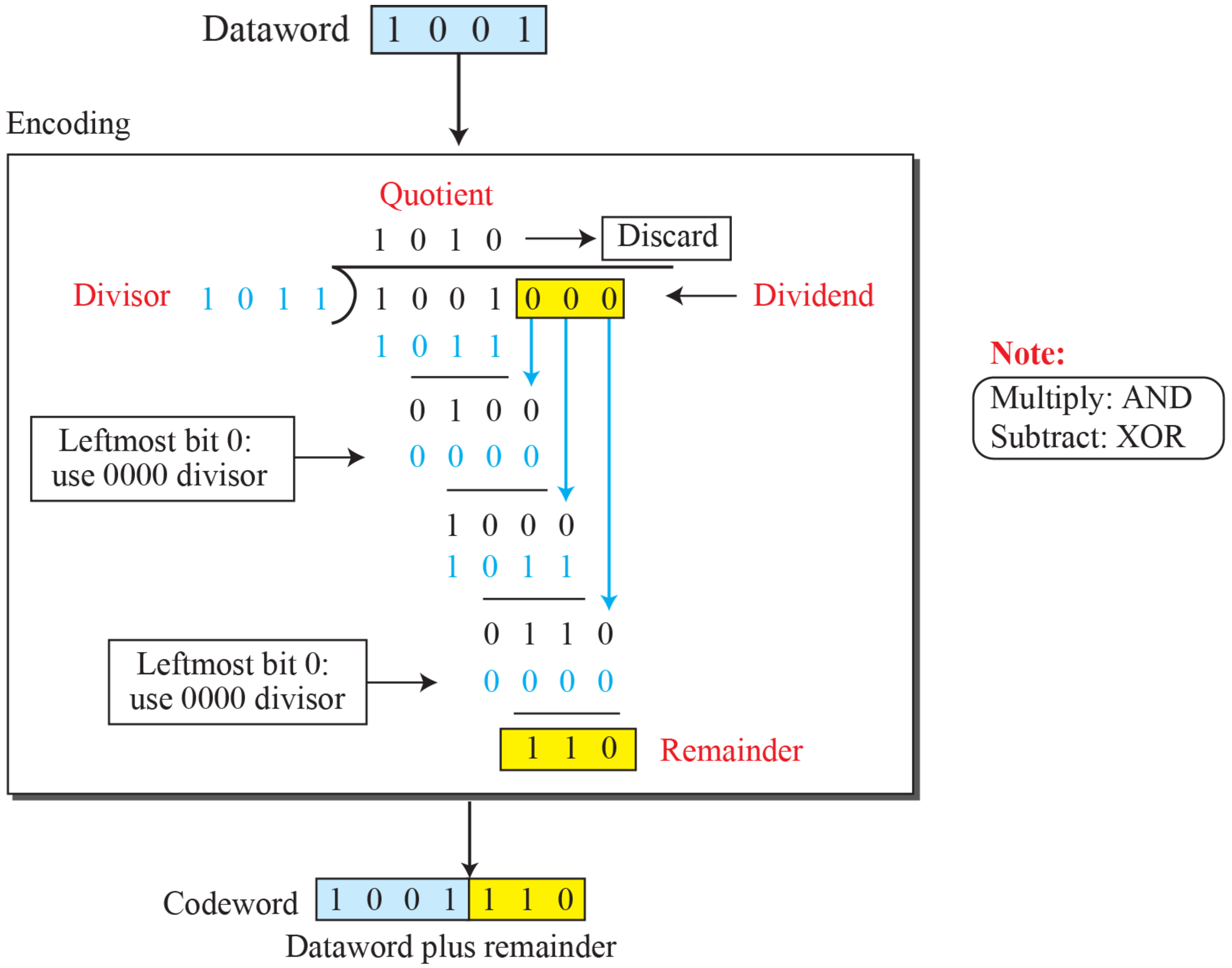
Received: 0 1 0 **1** **0** 1 0 **0** 0 1 **1** 0 0 0 1 1

b. Burst error

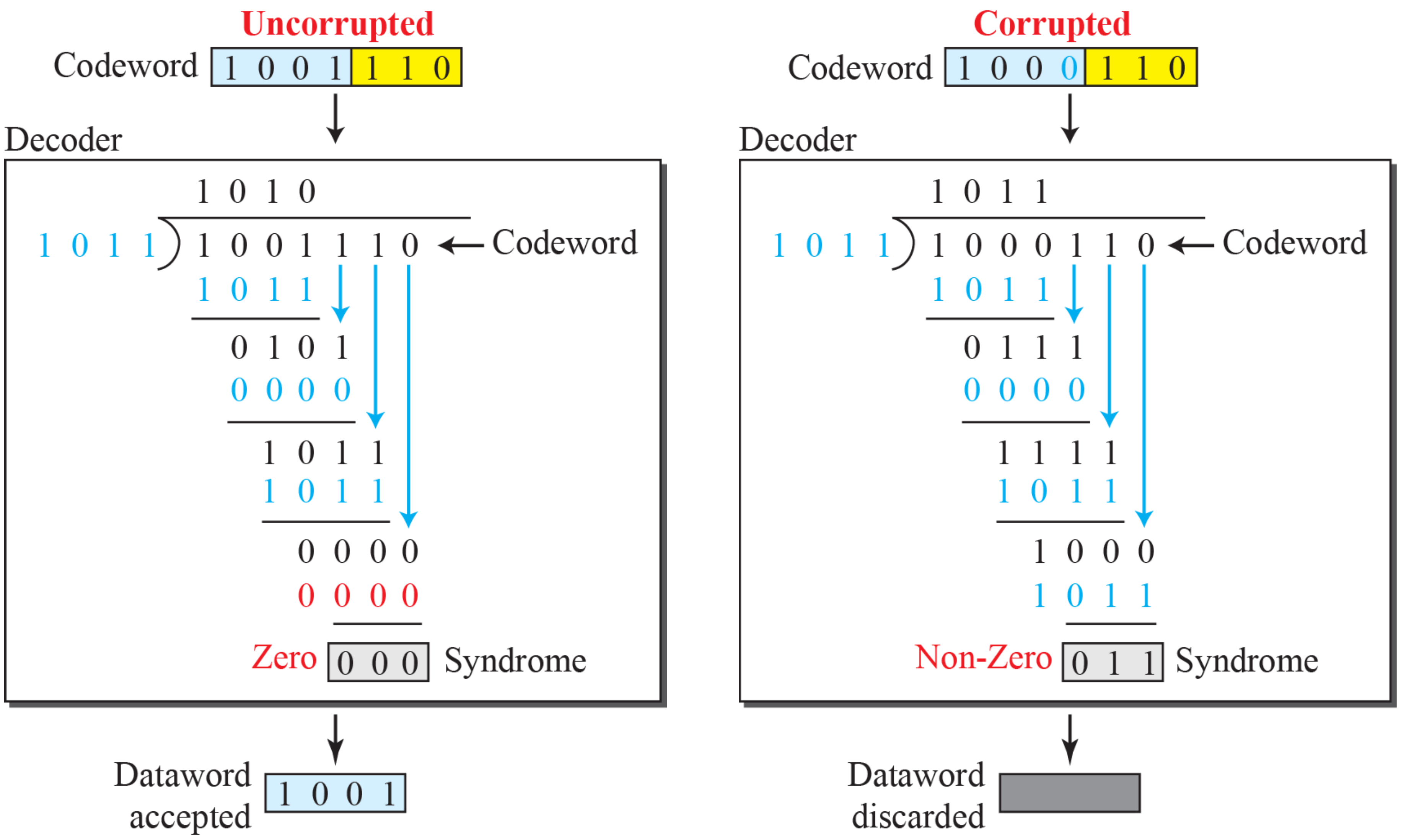
CRC encoder and decoder



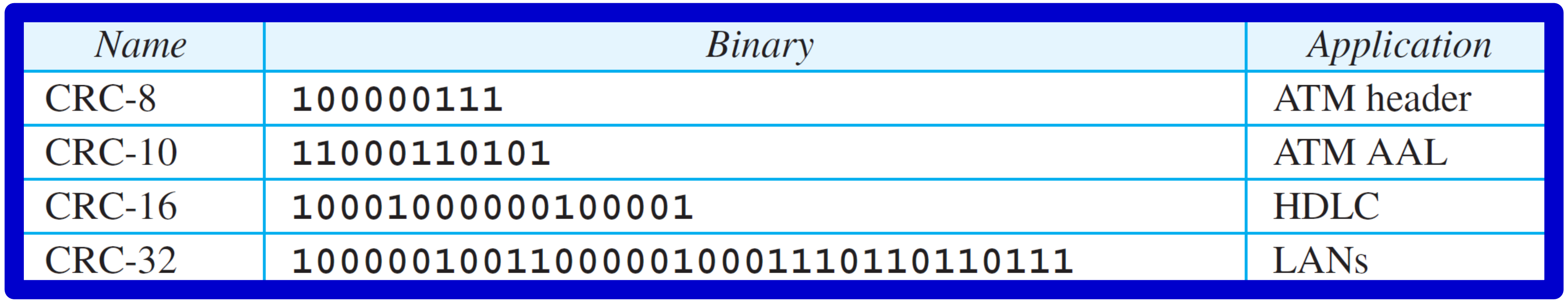
Division in CRC encoder



Division in the CRC decoder for two cases



Standard polynomials

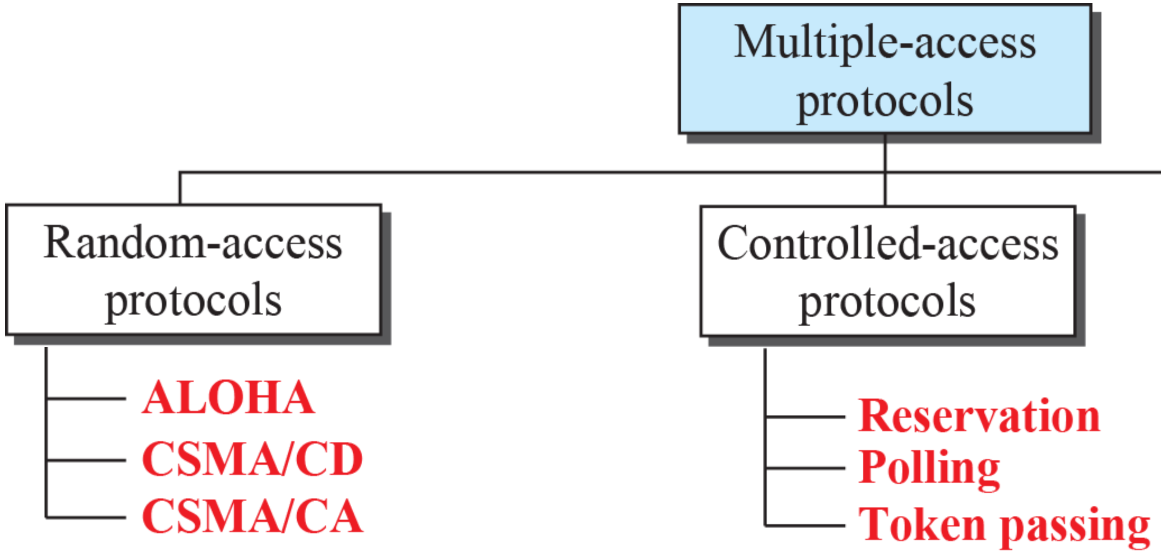


Name	*Binary*	*Application*
CRC-8	100000111	ATM header
CRC-10	11000110101	ATM AAL
CRC-16	10001000000100001	HDLC
CRC-32	100000100110000010001110110110111	LANs

5-3 MULTIPLE ACCESS PROTOCOLS

We said that the data-link layer is divided into two sublayers: data link control (DLC) and media access control (MAC). We discussed DLC in the previous section; we talk about MAC in this section.

Taxonomy of multiple-access protocols



Random Access

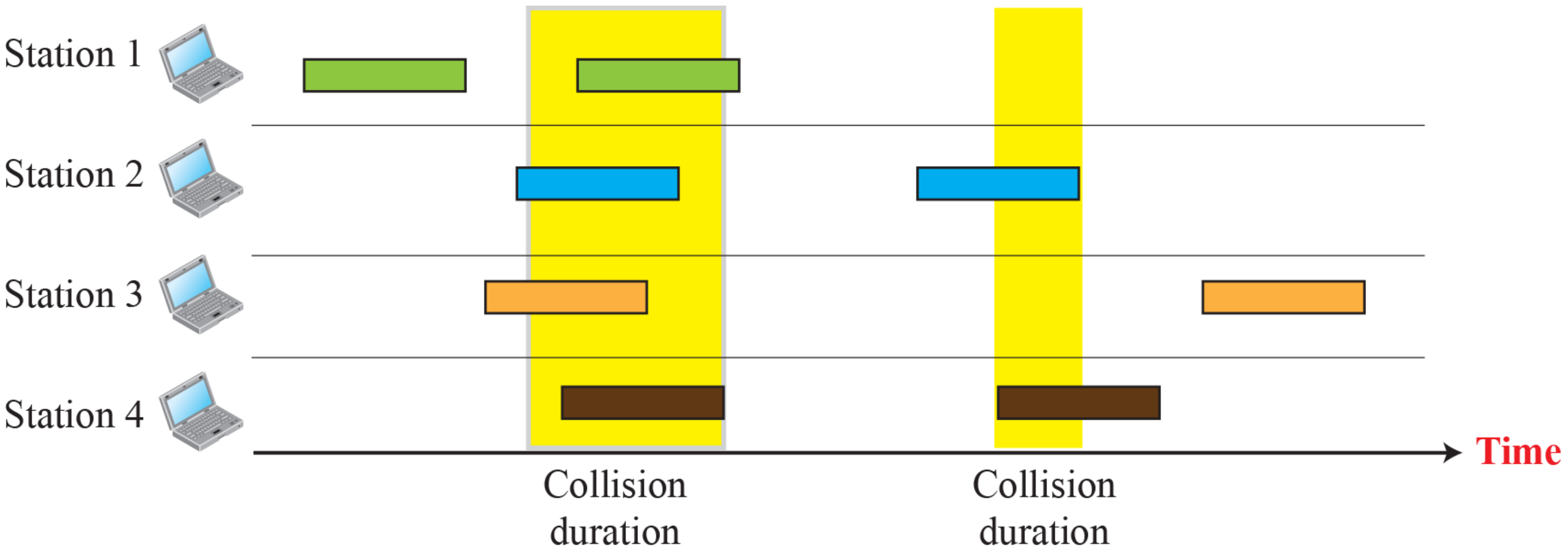
In random-access or contention methods, no station is superior to another station and none is assigned the control over another. At each instance, a station that has data to send uses a procedure defined by the protocol to make a decision on whether or not to send. This decision depends on the state of the medium (idle or busy). In other words, each station can transmit when it desires on the condition that it follows the predefined procedure, including the testing of the state of the medium.

- ALOHA
  - Pure ALOHA
  - Slotted ALOHA
- CSMA
  - Vulnerable Time
  - Persistence Methods
- CSMA/CD
  - Minimum Frame Size
  - Procedure
  - Energy Level
  - Throughput
  - Traditional Ethernet
- CSMA/CA

Pure ALOHA network

- **Pure ALOHA,** the earliest random access method, was developed at the University of Hawaii in early 1970.
- It was designed for a radio (wireless) LAN, but it can be used on any shared medium.
- The idea is that each station sends a frame whenever it has a frame to send
- there is the possibility of collision between frames from different stations
- The pure ALOHA protocol relies on acknowledgments from the receiver.
- A collision involves two or more stations. If all these stations try to resend their frames after the time-out, the frames will collide again.

Frames in a pure ALOHA network



- Pure ALOHA dictates that when the time-out period passes, each station waits a random amount of time before resending its frame. The randomness will help avoid more collisions. We call this time the *backoff time T_B.*
- The time-out period is equal to the maximum possible round-trip propagation delay, which is twice the amount of time required to send a frame between the two most widely separated stations or the average time required to send out a frame ($2 \times T_p$) or ($2 \times T_{fr}$) .
- After a maximum number of retransmission attempts K_{max}, a station must give up and try later.
- The formula for T_B depends on the implementation. One common formula is the ***binary exponential backoff.*** In this method, for each retransmission, a multiplier $R = 0$ to $2^K - 1$ is randomly chosen and multiplied by T_p *or* T_{fr}

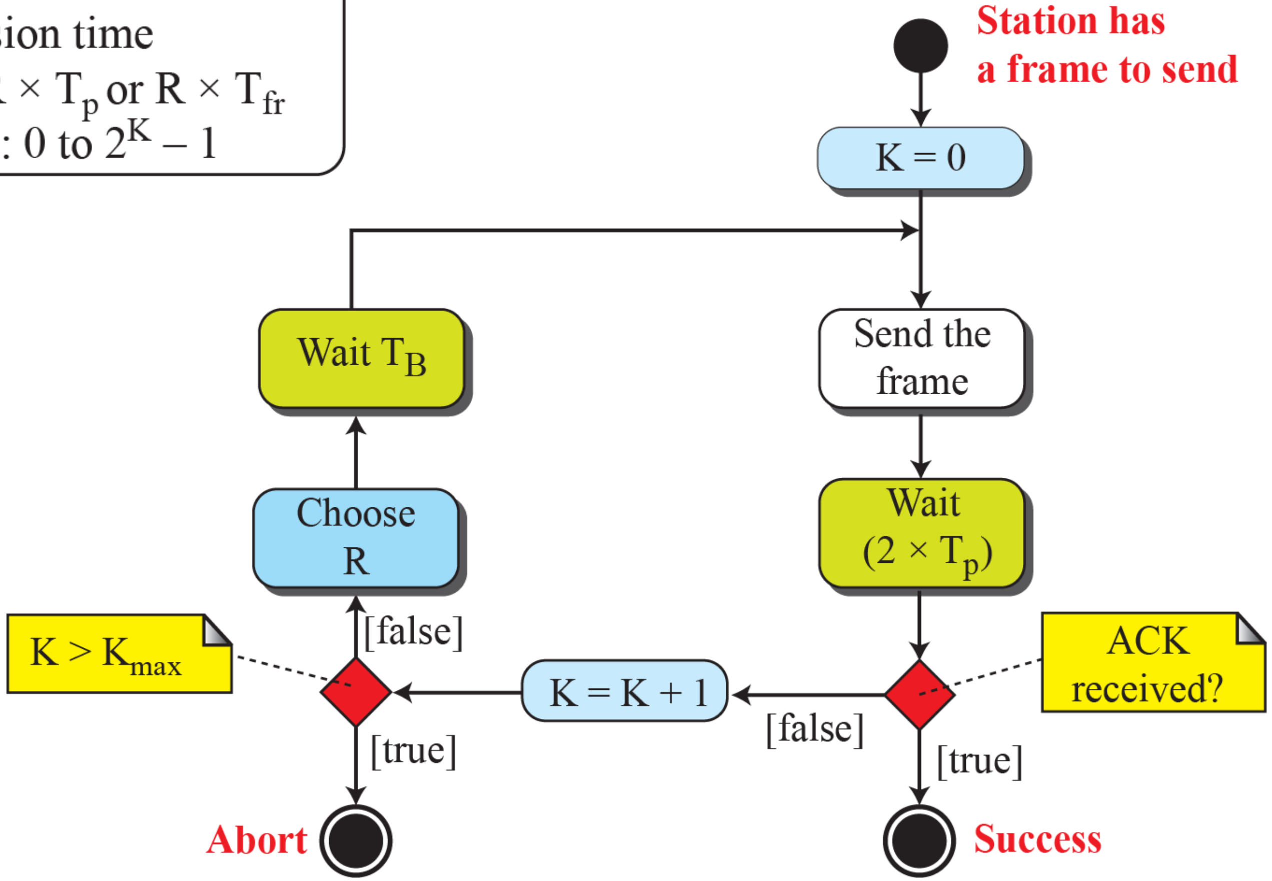
Procedure for pure ALOHA protocol

Legend

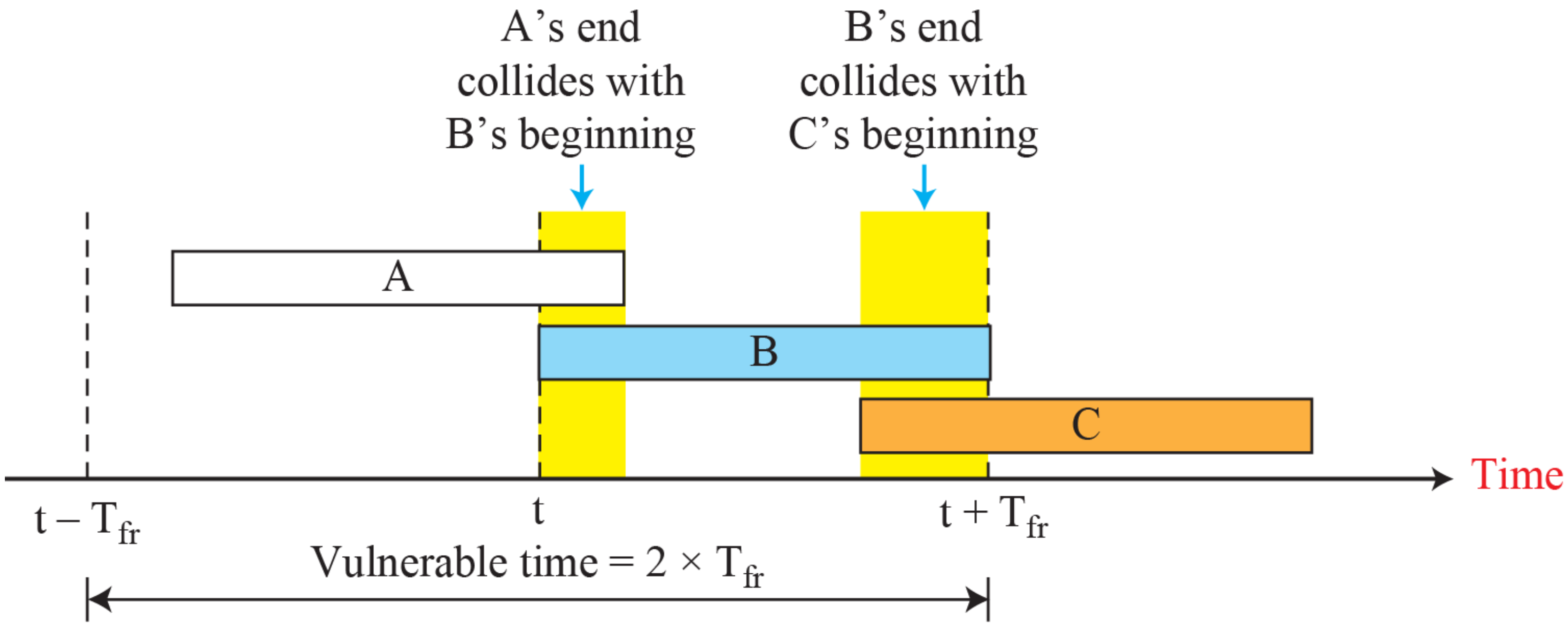
K : Number of attempts
T_p : Maximum propagation time
T_{fr}: Average transmission time
T_B: (Back-off time): $R \times T_p$ or $R \times T_{fr}$
R : (Random number): 0 to $2^K - 1$

Maximum Throughput $S = G \times e^{-2G}$
when G=1/2 → $S_{max} = 0.184$

Vulnerable time = $2 * T_{fr}$



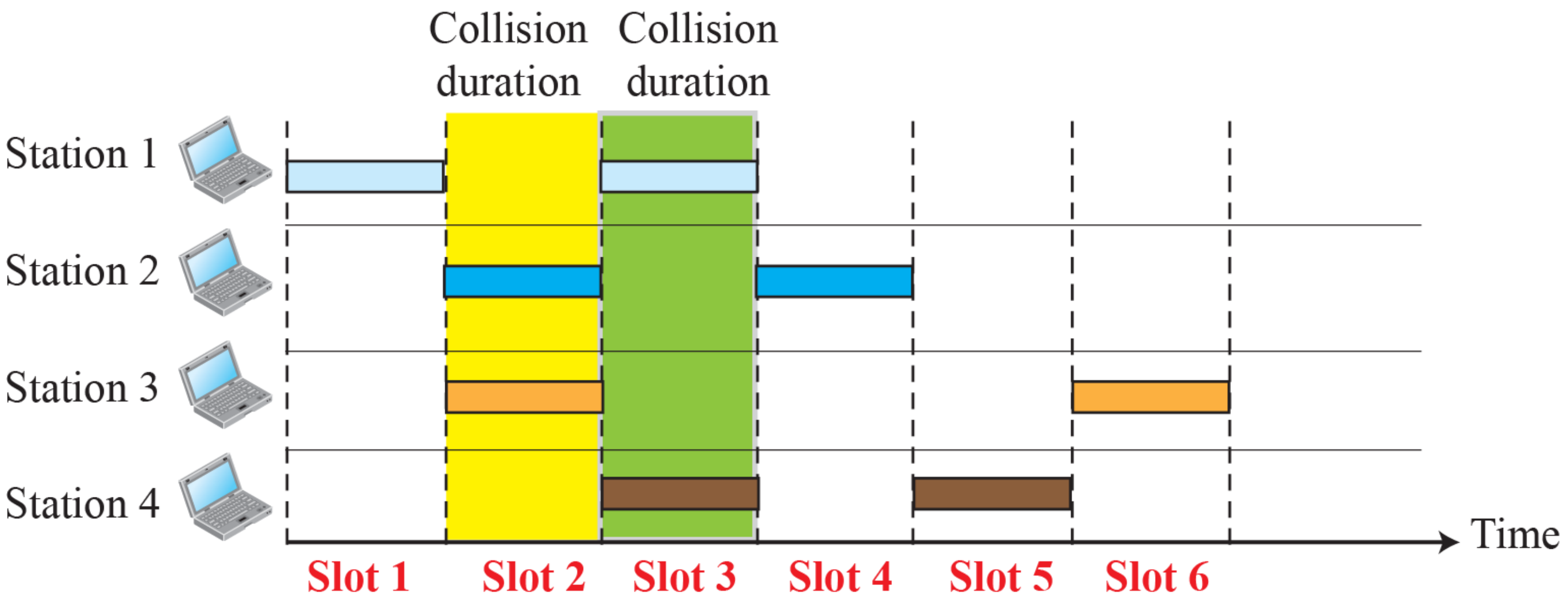
Vulnerable time for pure ALOHA protocol


A's end collides with B's beginning
B's end collides with C's beginning
A
B
C
Time
t – T_fr
t
t + T_fr
Vulnerable time = 2 × T_fr

Slotted ALOHA network

- Slotted ALOHA was invented to improve the efficiency of pure ALOHA.
- In **slotted ALOHA** we divide the time into slots of T_{fr} seconds and force the station to send only at the beginning of the time slot
- Because a station is allowed to send only at the beginning of the synchronized time slot, if a station misses this moment, it must wait until the beginning of the next time slot.
- The Slotted ALOHA protocol relies on acknowledgments from the receiver.
- A collision involves two or more stations. If all these stations try to resend their frames after the time-out, the frames will collide again.
- The vulnerable time is now reduced to one-half, equal to T_{fr}

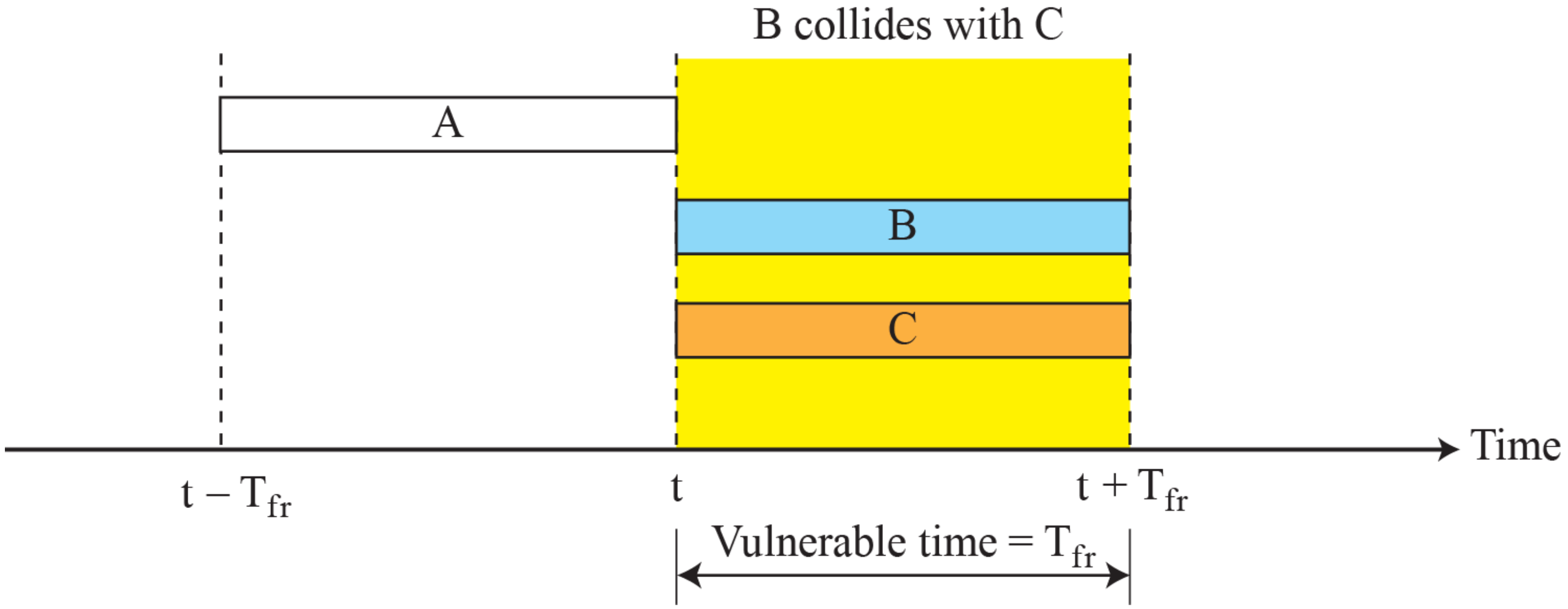
Frames in a slotted ALOHA network



Maximum Throughput $S = G \times e^{-G}$
when G=1 → $S_{max} = 0.36$

Vulnerable time = T_{fr}

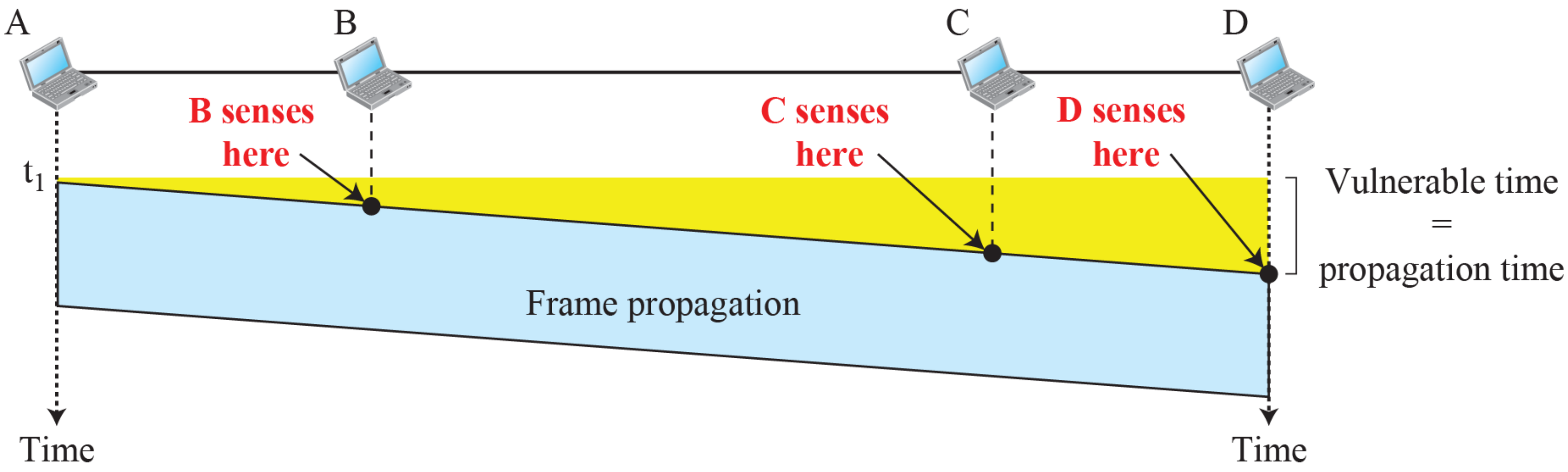
Figure 5.33: *Vulnerable time for slotted ALOHA protocol*



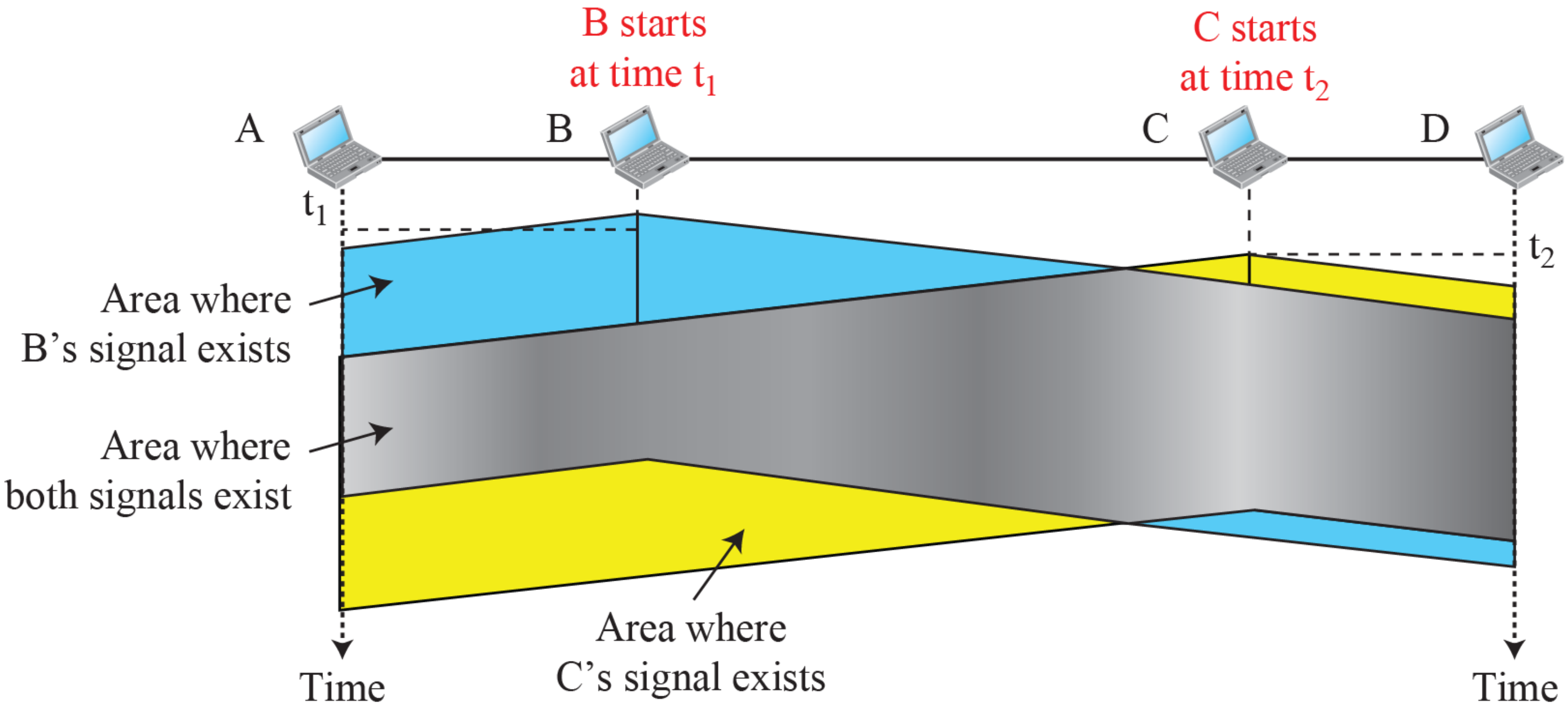
Carrier sense multiple access(*CSMA)*

- **Carrier sense multiple access (CSMA)** requires that each station first listen to the medium (or check the state of the medium) before sending.
- CSMA can reduce the possibility of collision, but it cannot eliminate it.
- The possibility of collision still exists because of propagation delay.
- The vulnerable time for CSMA is the ***propagation time*** T_p.
- A collision involves two or more stations. If all these stations try to resend their frames after the time-out, the frames will collide again.
- The vulnerable time is now reduced to one-half, equal to T_{fr}

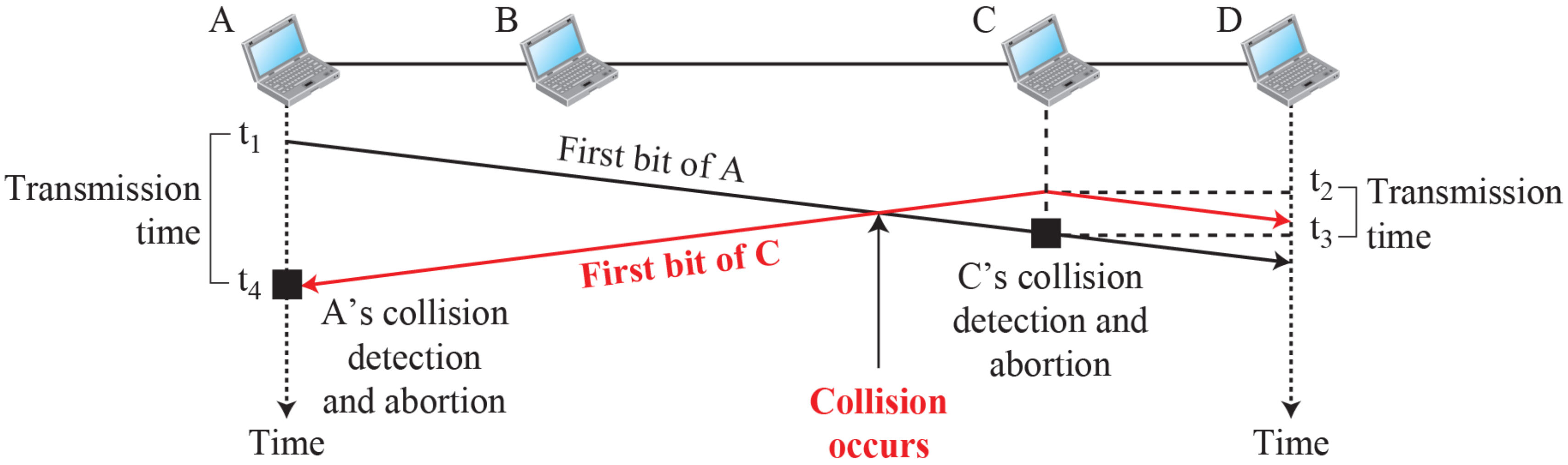
Vulnerable time in CSMA



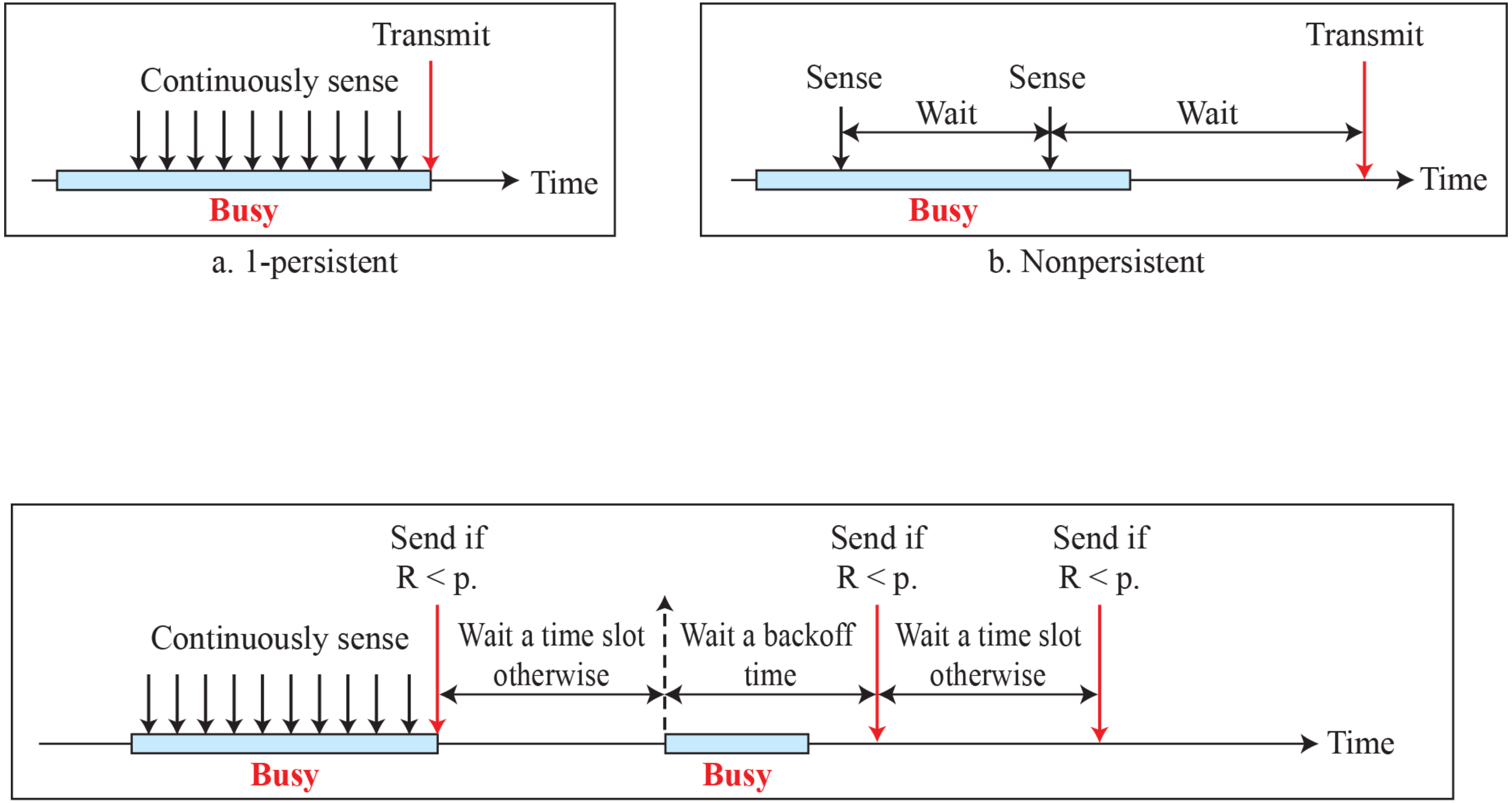
Space/time model of a collision in CSMA



Collision of the first bits in CSMA/CD



Behavior of three persistence methods

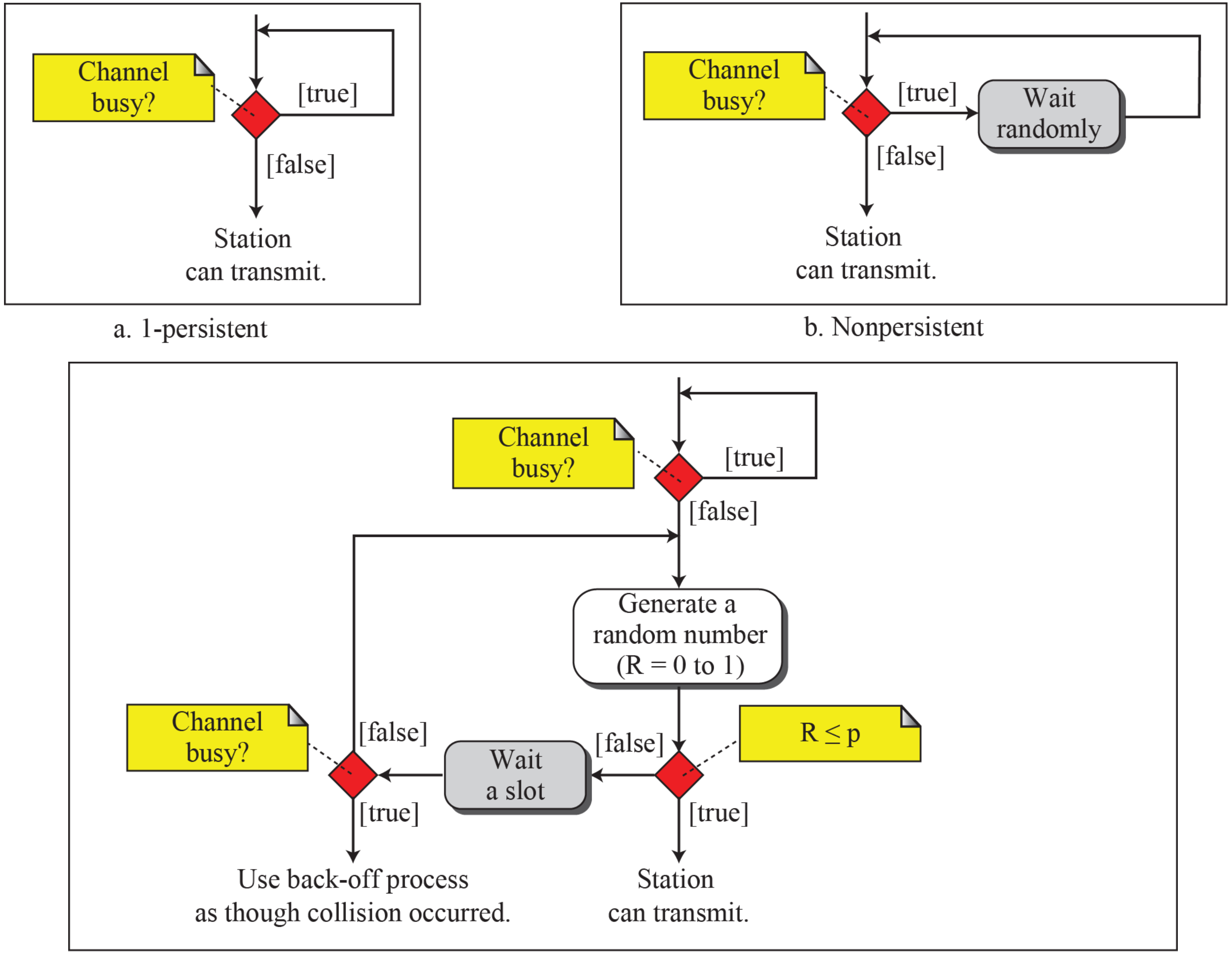


a. 1-persistent

b. Nonpersistent

c. *p*-persistent

Flow diagram for three persistence methods



a. 1-persistent

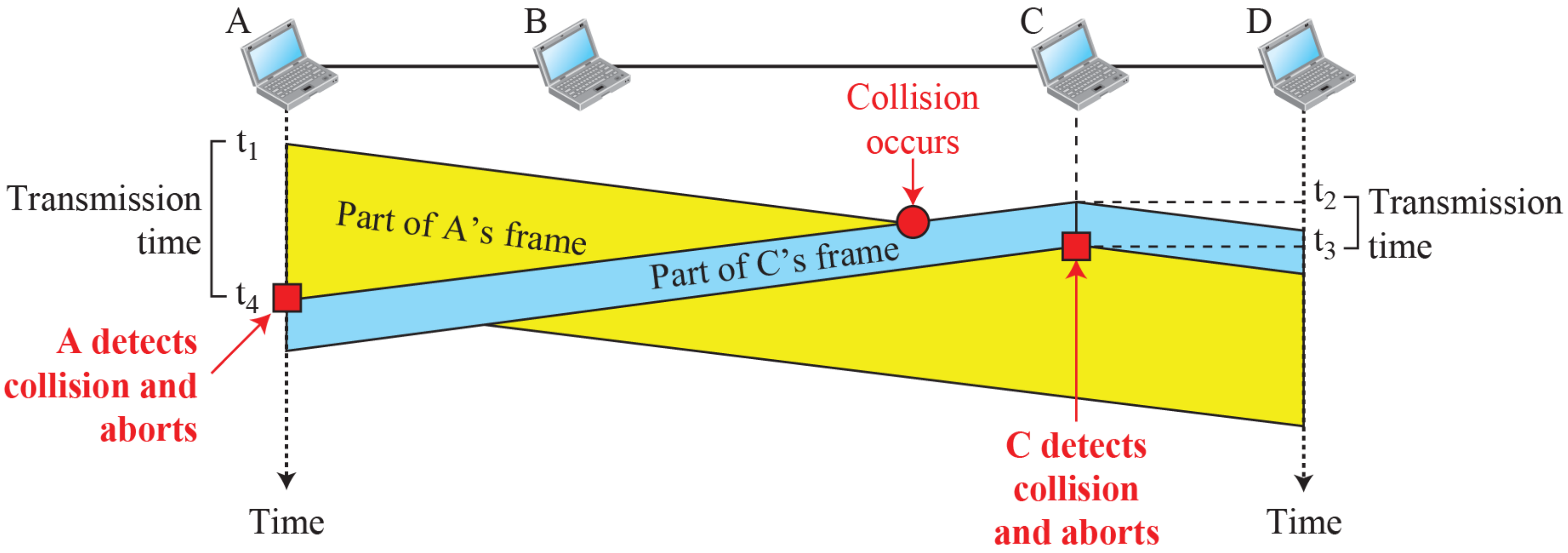
b. Nonpersistent

c. *p*-persistent

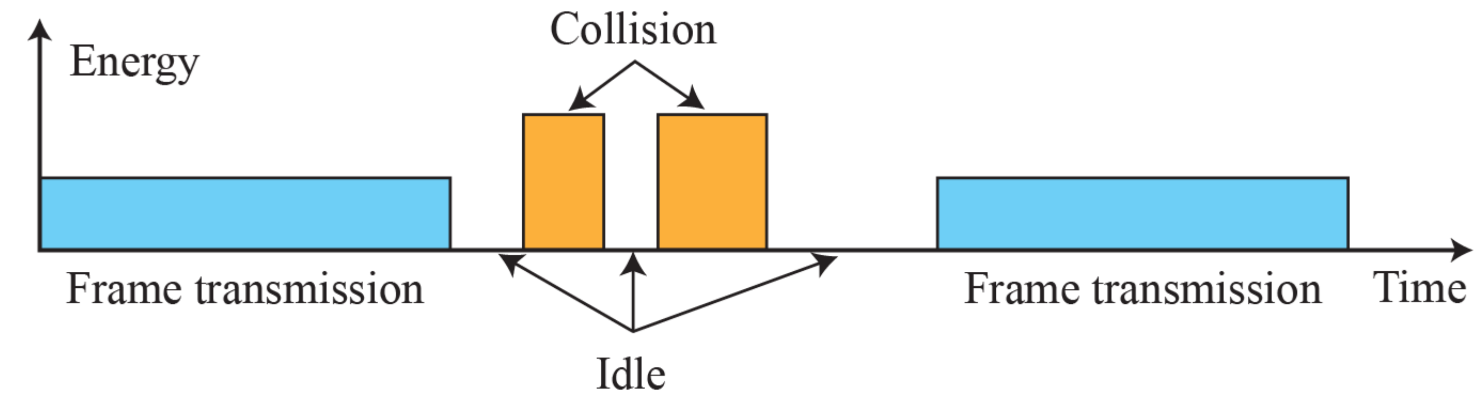
Carrier Sense Multiple Access/ Collision Detection (*CSMA/CD)*

- The CSMA method does not specify the procedure following a collision.
- CSMA can reduce the possibility of collision, but it cannot eliminate it.
- For CSMA/CD to work, we need a restriction on the frame size. Before sending the last bit of the frame, the sending station must detect a collision, if any, and abort the transmission.
- This is so because the station, once the entire frame is sent, does not keep a copy of the frame and does not monitor the line for collision detection. Therefore, the frame transmission time T_{fr} must be at least two times the maximum propagation time T_p.
- One of the LAN protocols that used CSMA/CD is the traditional Ethernet with the data rate of 10 Mbps.

Collision and abortion in CSMA/CD



Energy level during transmission, idleness, or collision



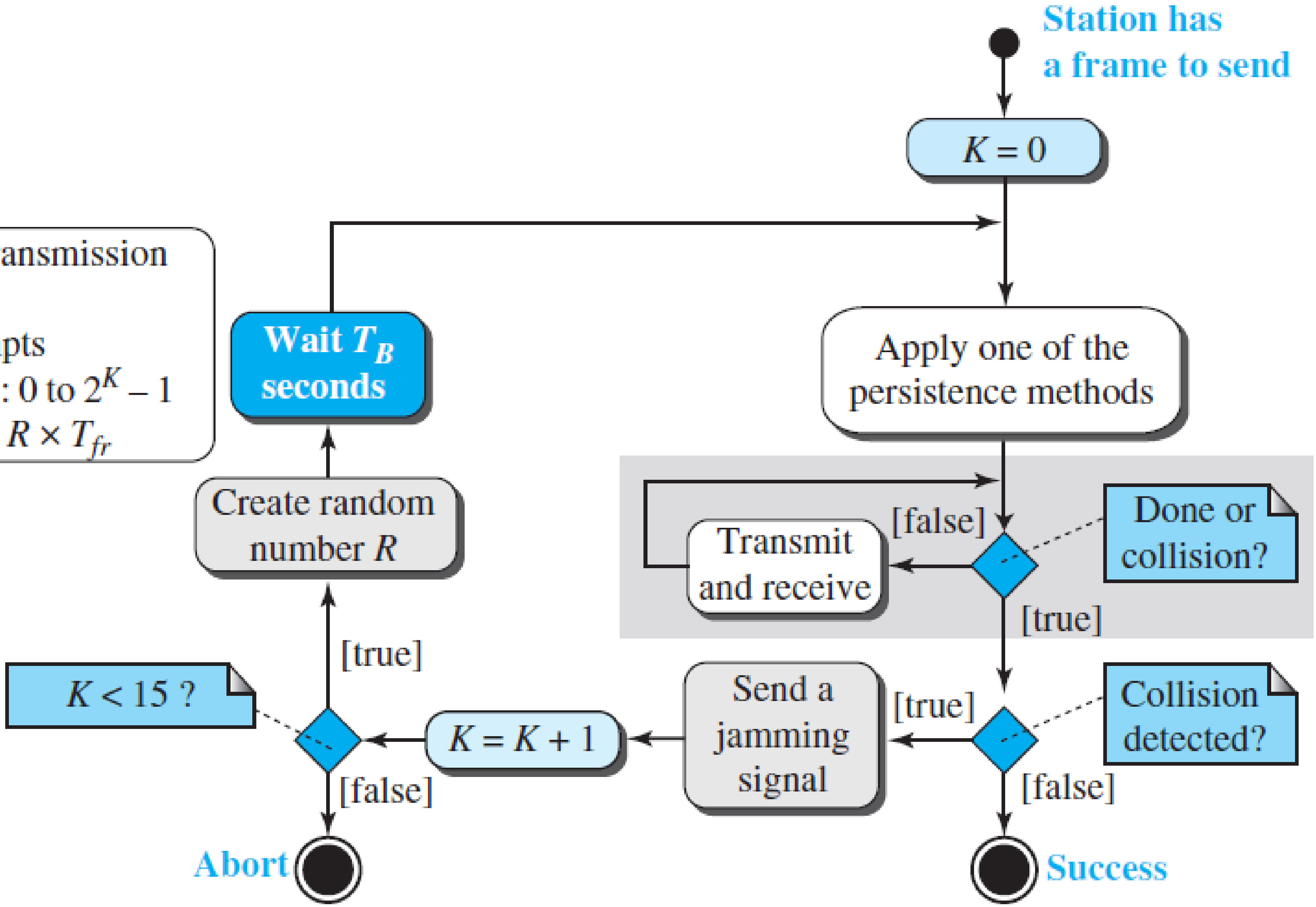
Procedure

- It is similar to the one for the ALOHA protocol, but there are differences.
  - The first difference is the addition of the persistence process
  - The second difference is the frame transmission and acknowledgement
    - CSMA/CD → No ack. i.e. either transmission is finished or a collision is detected. Either event stops transmission.
- The third difference is the sending of a short jamming signal to make sure that all other stations become aware of the collision.

Flow diagram for the CSMA/CD

Legend

T_{fr}: Frame average transmission time
K : Number of attempts
R : (random number): 0 to $2^K - 1$
T_B: (Backoff time) = $R \times T_{fr}$



Carrier Sense Multiple Access/ Collision Avoidance (*CSMA/CA)*

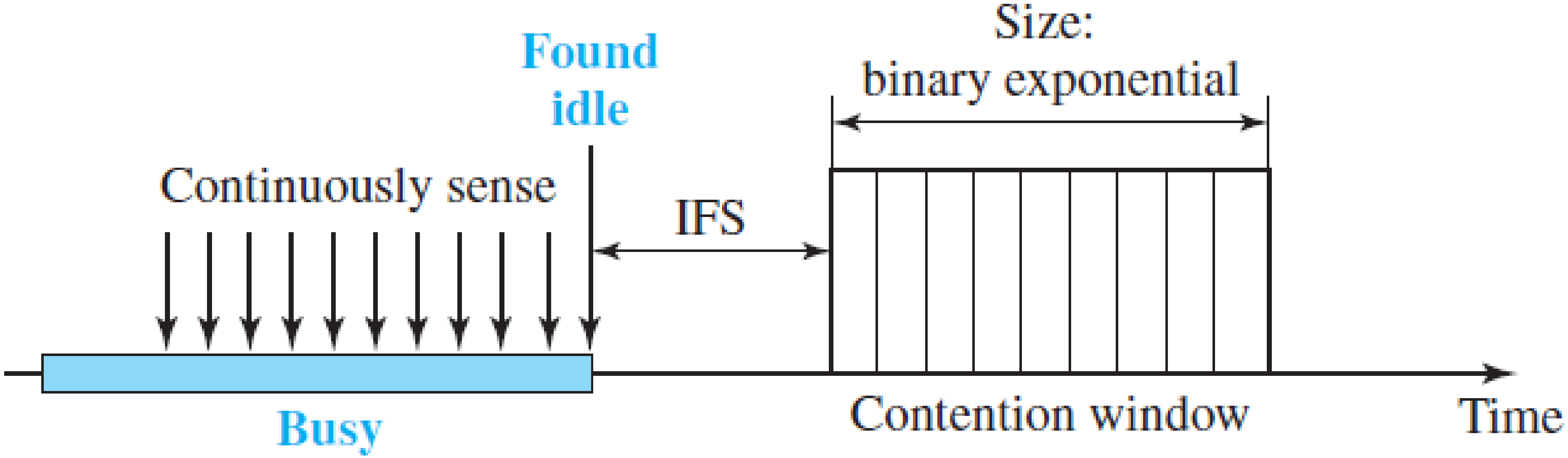
- **CSMA/CA** was invented for wireless networks.
- Collisions are avoided through the use of CSMA/CA's three strategies:
  - Interframe space,
  - Contention window
  - Acknowledgments

Interframe space

First, collisions are avoided by deferring transmission even if the channel is found idle. When an idle channel is found, the station does not send immediately. It waits for a period of time called the ***interframe space*** or ***IFS.*** The IFS time allows the front of the transmitted signal by the distant station to reach this station. After waiting an IFS time, if the channel is still idle, the station can send, but it still needs to wait a time equal to the contention window.

Contention window

The **contention window** is an amount of time divided into slots. A station that is ready to send chooses a random number of slots as its wait time. The number of slots in the window changes according to the binary exponential backoff strategy. One interesting point about the contention window is that the station needs to sense the channel after each time slot. However, if the station finds the channel busy, it does not restart the process; it just stops the timer and restarts it when the channel is sensed as idle. This gives priority to the station with the longest waiting time.



Acknowledgments

With all these precautions, there still may be a collision resulting in destroyed data. In addition, the data may be corrupted during the transmission. The positive acknowledgment and the time-out timer can help guarantee that the receiver has received the frame.

Frame Exchange Time Line

1. Before sending a frame, the source station senses the medium by checking the energy level at the carrier frequency.

a. The channel uses a persistence strategy with backoff until the channel is idle.

b. After the station is found to be idle, the station waits for a period of time called the ***DCF interframe space (DIFS);*** then the station sends a control frame called the ***request to send (RTS)*****.**

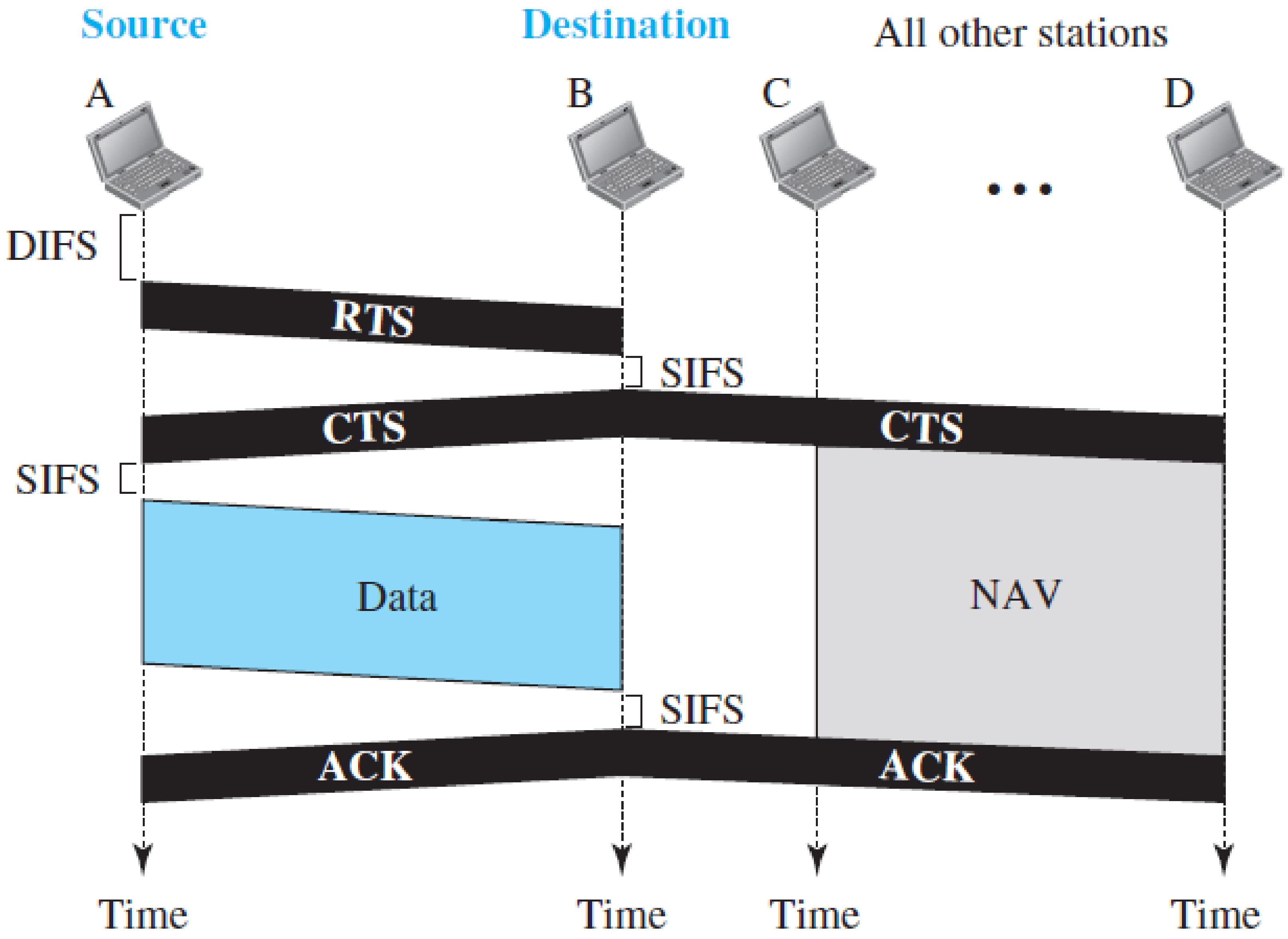
2. After receiving the RTS and waiting a period of time called the ***short interframe space (SIFS),*** the destination station sends a control frame, called the ***clear to send (CTS)****,* to the source station. This control frame indicates that the destination station is ready to receive data.

3. The source station sends data after waiting an amount of time equal to SIFS.

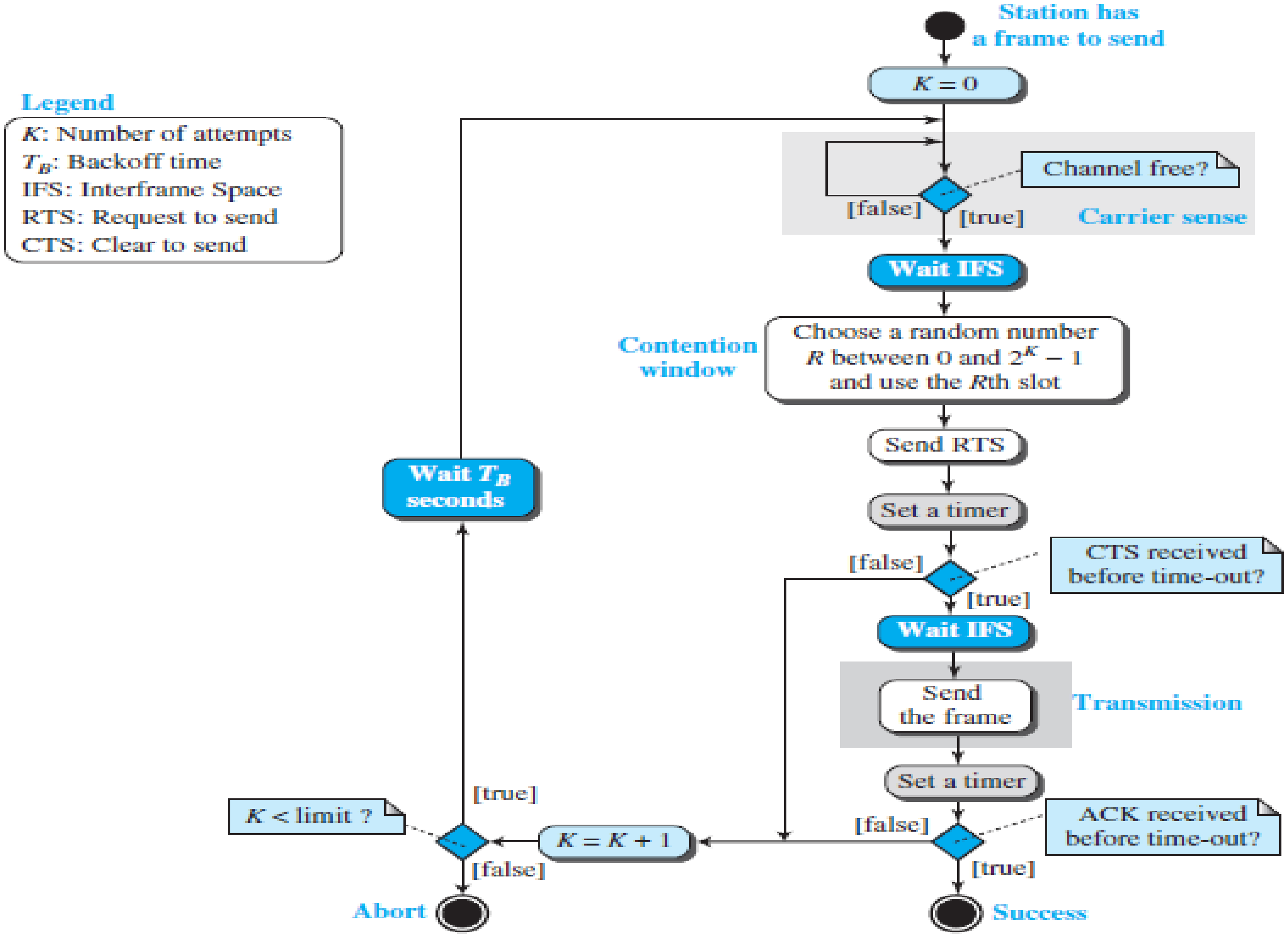
4. The destination station, after waiting an amount of time equal to SIFS, sends an acknowledgment to show that the frame has been received. Acknowledgment is needed in this protocol because the station does not have any means to check for the successful arrival of its data at the destination. On the other hand, the lack of collision in CSMA/CD is a kind of indication to the source that data have arrived.

Network Allocation Vector

When a station sends an RTS frame, it includes the duration of time that it needs to occupy the channel. The stations that are affected by this transmission create a timer called a **network allocation vector (NAV)** that shows how much time must pass before these stations are allowed to check the channel for idleness. Each time a station accesses the system and sends an RTS frame, other stations start their NAV. In other words, each station, before sensing the physical medium to see if it is idle, first checks its NAV to see if it has expired..


Source
Destination
All other stations
A
B
C
D
DIFS
RTS
SIFS
CTS
CTS
SIFS
Data
NAV
SIFS
ACK
ACK
Time
Time
Time
Time

Flow diagram for the CSMA/CA



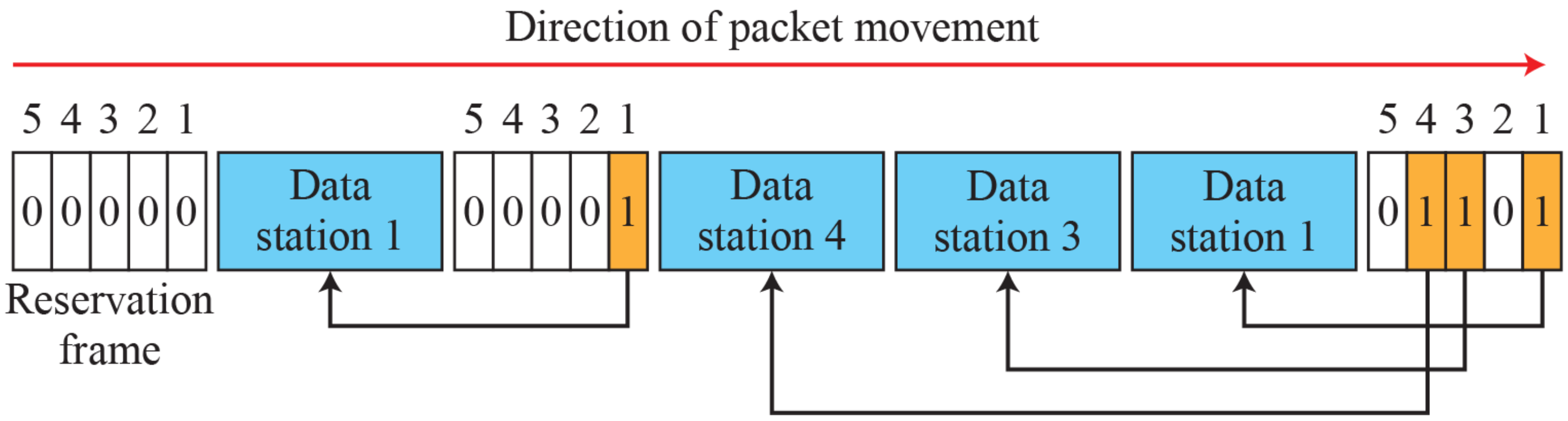
Controlled Access

In controlled access, the stations consult one another to find which station has the right to send. A station cannot send unless it has been authorized by other stations.

- Reservation
- Polling
  - Select
  - Poll
- Token Passing
  - Logical Ring

Reservation access method

- A station needs to make a reservation before sending data.
- Time is divided into intervals. In each interval, a reservation frame precedes the data frames sent in that interval.
- If there are *N* stations in the system, there are exactly *N* reservation minislots in the reservation frame. Each minislot belongs to a station. When a station needs to send a data frame, it makes a reservation in its own minislot. The stations that have made reservations can send their data frames after the reservation frame.



Polling-access method

- **Polling** works with topologies in which one device is designated as a ***primary station*** and the other devices are ***secondary stations*****.**
- All data exchanges must be made through the primary device even when the ultimate destination is a secondary device. The primary device controls the link; the secondary devices follow its instructions.
- It is up to the primary device to determine which device is allowed to use the channel at a given time
- The drawback is if the primary station fails, the system goes down.

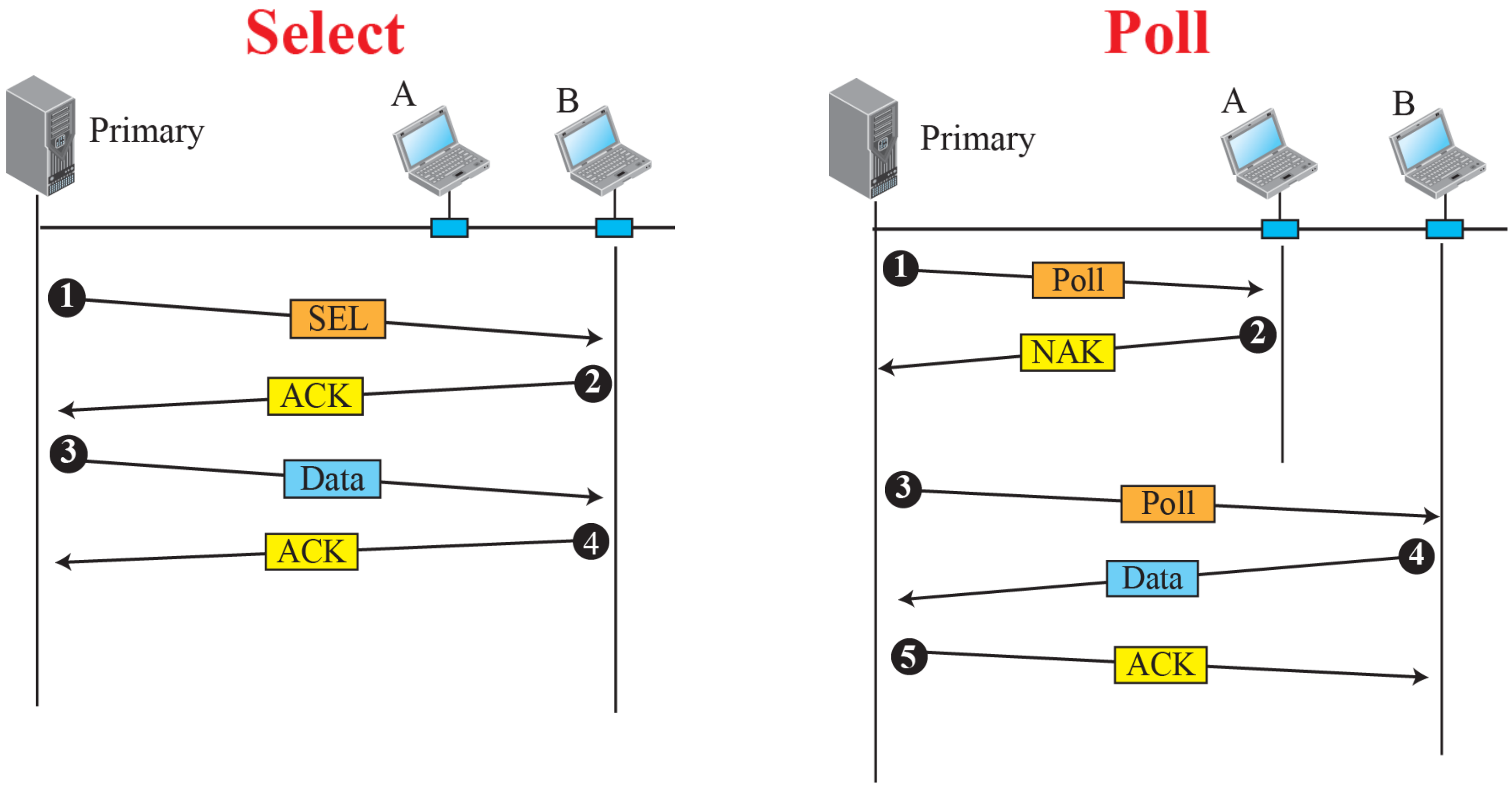
Select

The *select* function is used whenever the primary device has something to send. Remember that the primary controls the link. If the primary is neither sending nor receiving data, it knows the link is available. If it has something to send, the primary device sends it.

Poll

The *poll* function is used by the primary device to solicit transmissions from the secondary devices. When the primary is ready to receive data, it must ask (poll) each device in turn if it has anything to send. When the first secondary is approached, it responds either with a NAK frame if it has nothing to send or with data.

Select and poll functions in polling-access method



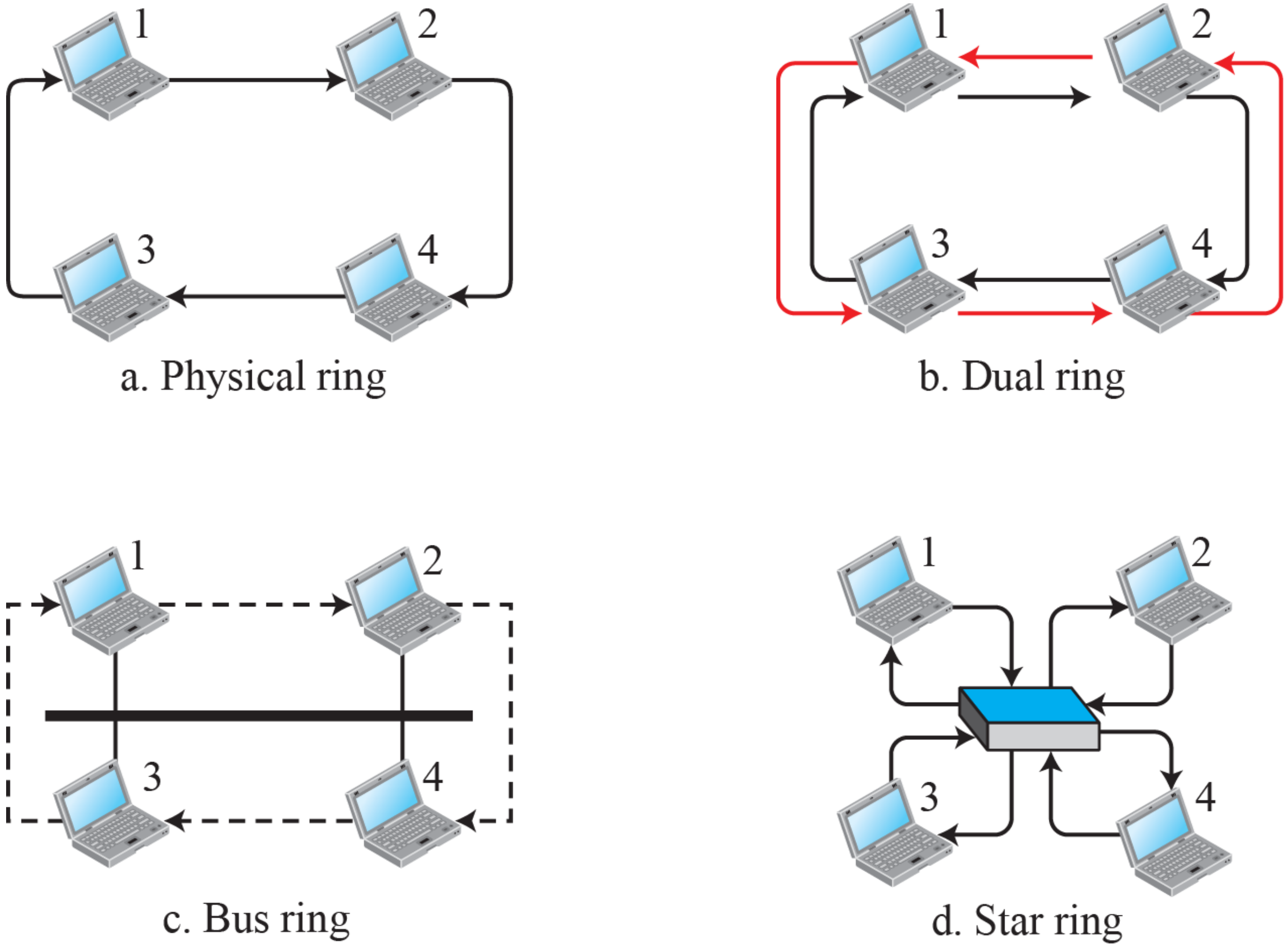
Select

Poll

Token-passing access method

- The **token-passing** method, the stations in a network are organized in a logical ring.
- For each station, there is a *predecessor* and a *successor*. The predecessor is the station which is logically before the station in the ring; the successor is the station which is after the station in the ring. The current station is the one that is accessing the channel now.
- In this method, a special packet called a ***token*** circulates through the ring.
- The possession of the token gives the station the right to access the channel and send its data.
- When a station has some data to send, it waits until it receives the token from its predecessor. It then holds the token and sends its data. When the station has no more data to send, it releases the token, passing it to the next logical station in the ring. The station cannot send data until it receives the token again in the next round.
- In a token-passing network, stations do not have to be physically connected in a ring; the ring can be a logical one.

Logical ring and physical topology in token-passing access method



a. Physical ring

b. Dual ring

c. Bus ring

d. Star ring