# Ethernet

Chapter 12 in Stallings 10th Edition



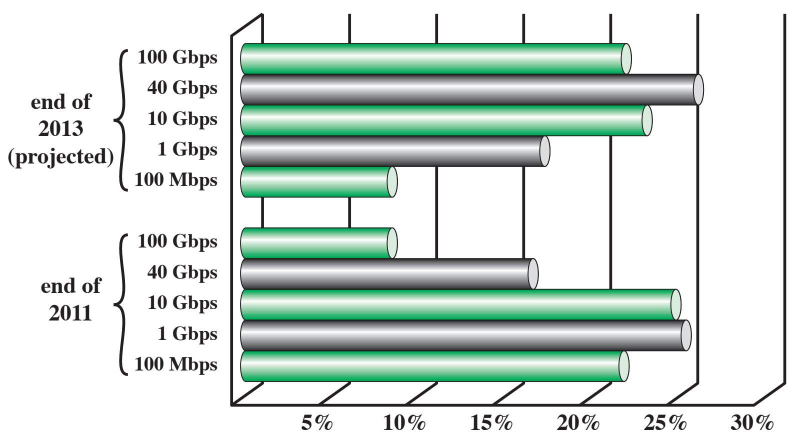



**Figure 12.1 Data Center Study—Percentage of Ethernet Links by Speed**

## Early on...

### Earliest was ALOHA

- •Developed for packet radio networks
- •Station may transmit a frame at any time
- •If frame is determined invalid, it is ignored
- •Maximum utilization of channel about 18%

### Next came slotted ALOHA

- •Organized slots equal to transmission time
- •Increased utilization to about 37%



## ALOHA

- Packet Radio
- When station has frame, it sends it!
- Station listens (for max round trip time) plus small increment
- If ACK, fine. If not, retransmit
- If no ACK after repeated transmissions, give up.
- Frame check sequence (as in HDLC)
- If frame OK and address matches receiver, send ACK
- Frame may be damaged by noise or by another station transmitting at the same time (collision)
- Any overlap of frames causes collision
- Max utilization 18%

## Slotted ALOHA

- Time in uniform slots equal to frame transmission time
- Need central clock (or other sync mechanism)
- Transmission begins at slot boundary
- Frames either miss or overlap totally
- Max utilization 37%



## CSMA/CD Precursors

- Carrier Sense Multiple Access (CSMA)
  - Station listens to determine if there is another transmission in progress
  - If idle, station transmits
  - Waits for acknowledgment
  - If no acknowledgment, collision is assumed and station retransmits
  - Utilization far exceeds ALOHA

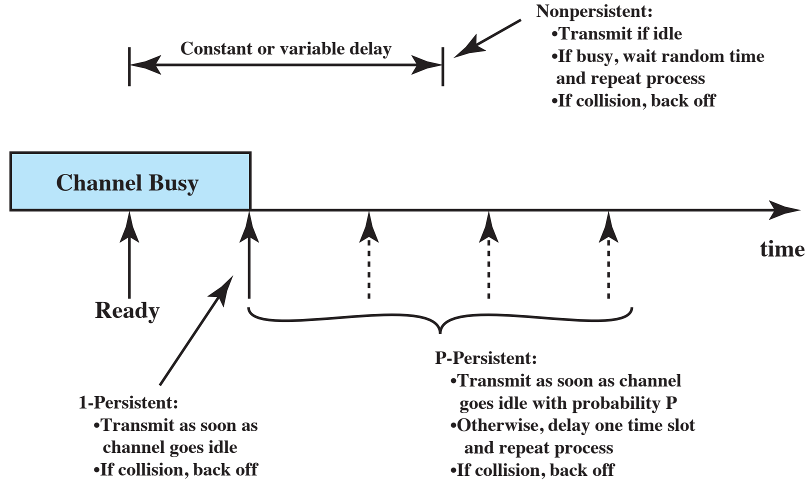



**Figure 12.2 CSMA Persistence and Backoff**



# Nonpersistent CSMA

If the medium is idle, transmit; otherwise, go to step 2

If the medium is busy, wait an amount of time drawn from a probability distribution and repeat step 1

**Disadvantage:**
Capacity is wasted because the medium will generally remain idle following the end of a transmission even if there are one or more stations waiting to transmit

## 1-Persistent CSMA

- Avoids idle channel time
- Rules:
    1. If medium is idle, transmit
    2. If medium is busy, listen until idle; then transmit immediately
- Stations are selfish
- If two or more stations are waiting, a collision is guaranteed



## P-Persistent CSMA

- A compromise to try and reduce collisions and idle time
- P-persistent CSMA rules:
    1. If medium is idle, transmit with probability *p*, and delay one time unit with probability (1–*p*)
    2. If medium is busy, listen until idle and repeat step 1
    3. If transmission is delayed one time unit, repeat step 1
- Issue of choosing effective value of *p* to avoid instability under heavy load

# Value of *p*?

- Have *n* stations waiting to send
- At end of transmission, expected number of stations is *np*
  - If *np*>1 on average there will be a collision
- Repeated transmission attempts mean collisions are likely
- Eventually all stations will be trying to send, causing continuous collisions, with throughput dropping to zero
- To avoid catastrophe *np*<1 for expected peaks of *n*
  - If heavy load expected, *p* must be small
  - Smaller *p* means stations wait longer



# Description of CSMA/CD

1. If the medium is idle, transmit; otherwise, go to step 2
2. If the medium is busy, continue to listen until the channel is idle, then transmit immediately
3. If a collision is detected, transmit a brief jamming signal to assure that all stations know that there has been a collision and cease transmission
4. After transmitting the jamming signal, wait a random amount of time, referred to as the ***backoff***, then attempt to transmit again

# CSMA/CD Operation

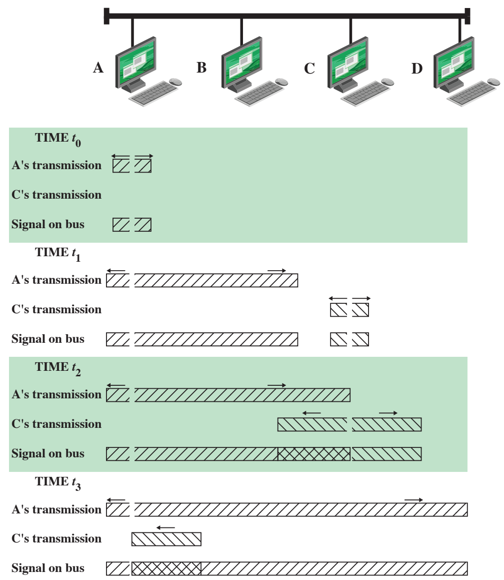



Figure 12.3 CSMA/CD Operation



# Which Persistence Algorithm?

- IEEE 802.3 uses 1-persistent
- Both nonpersistent and p-persistent have performance problems

**1-persistent seems more unstable than p-persistent**

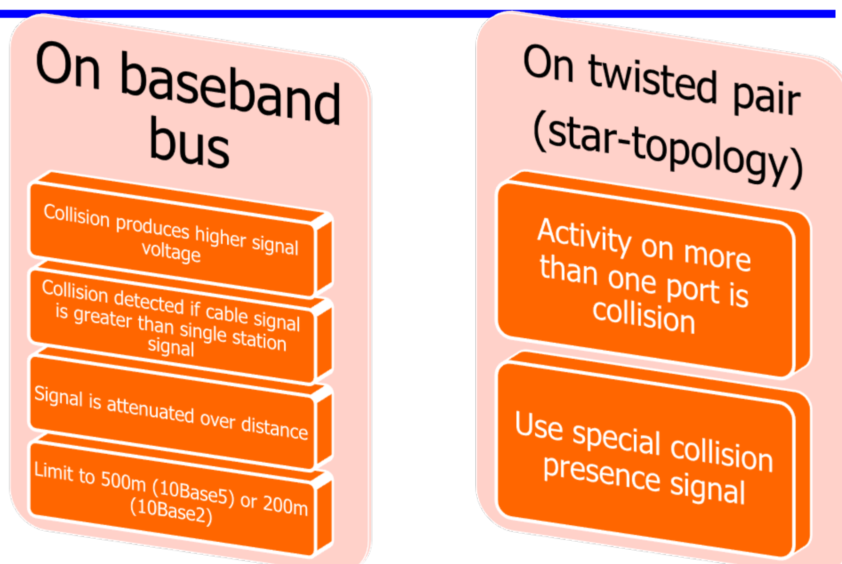
- **Because of greed of the stations**
- **Wasted time due to collisions is short**
- **With random backoff unlikely to collide on next attempt to send**

## Binary Exponential Backoff

- IEEE 802.3 and Ethernet both use binary exponential backoff
- A station will attempt to transmit repeatedly in the face of repeated collisions
  - —On first 10 attempts, mean random delay doubled
  - —Value then remains the same for 6 further attempts
  - —After 16 unsuccessful attempts, station gives up and reports error
- 1-persistent algorithm with binary exponential backoff is efficient over wide range of loads
- Backoff algorithm has last-in, first-out effect



## Collision Detection



# IEEE 802.3 Frame Format

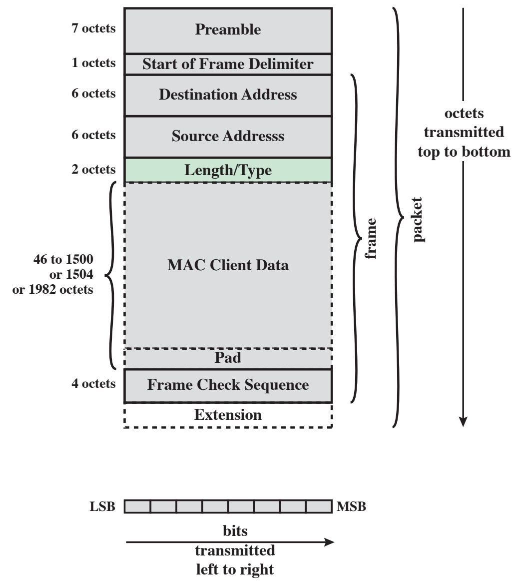



Figure 12.4 IEEE 802.3 MAC Frame Format



# Table 12.1

# IEEE 802.3 10-Mbps Physical Layer Medium Alternatives

| | 10BASE5 | 10BASE2 | 10BASE-T | 10BASE-FP |
|---|---|---|---|---|
| **Transmission medium** | Coaxial cable (50 ohm) | Coaxial cable (50 ohm) | Unshielded twisted pair | 850-nm optical fiber pair |
| **Signaling technique** | Baseband (Manchester) | Baseband (Manchester) | Baseband (Manchester) | Manchester/on-off |
| **Topology** | Bus | Bus | Star | Star |
| **Maximum segment length (m)** | 500 | 185 | 100 | 500 |
| **Nodes per segment** | 100 | 30 | — | 33 |
| **Cable diameter (mm)** | 10 | 5 | 0.4 to 0.6 | 62.5/125 $\mu$m |

# Table 12.2

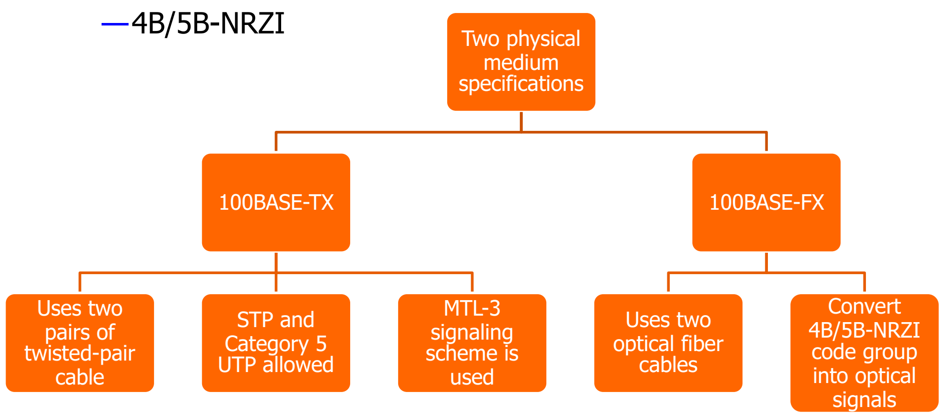
## IEEE 802.3 100BASE-T Physical Layer Medium Alternatives

| | 100BASE-TX | | 100BASE-FX | 100BASE-T4 |
|---|---|---|---|---|
| **Transmission medium** | 2 pair, STP | 2 pair, Category 5 UTP | 2 optical fibers | 4 pair, Category 3, 4, or 5 UTP |
| **Signaling technique** | MLT-3 | MLT-3 | 4B5B, NRZI | 8B6T, NRZ |
| **Data rate** | 100 Mbps | 100 Mbps | 100 Mbps | 100 Mbps |
| **Maximum segment length** | 100 m | 100 m | 100 m | 100 m |
| **Network span** | 200 m | 200 m | 400 m | 200 m |



# 100BASE-X

- Uses a unidirectional data rate 100 Mbps over single twisted pair or optical fiber link
- Encoding scheme same as FDDI
  - 4B/5B-NRZI



# 100BASE-X Data Rate and Encoding

- Unidirectional data rate 100 Mbps over single link
  - —Single twisted pair, single optical fiber
- Encoding scheme same as FDDI
  - —4B/5B-NRZI
  - —Modified for each option



# 100BASE-X Media

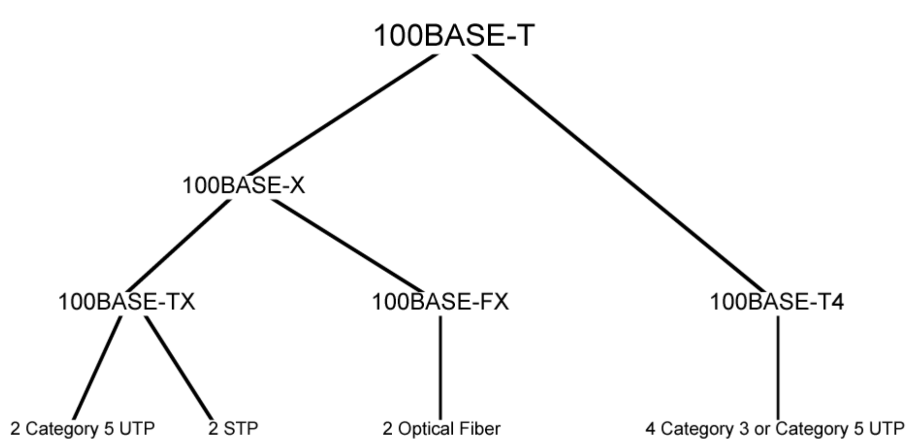
- Two physical medium specifications
- 100BASE-TX
  - —Two pairs of twisted-pair cable
  - —One pair for transmission and one for reception
  - —STP and Category 5 UTP allowed
  - —The MTL-3 signaling scheme is used
- 100BASE-FX
  - —Two optical fiber cables
  - —One for transmission and one for reception
  - —Intensity modulation used to convert 4B/5B-NRZI code group stream into optical signals
  - —1 represented by pulse of light
  - —0 by either absence of pulse or very low intensity pulse

## 100BASE-T4

- 100-Mbps over lower-quality Cat 3 UTP
  - —Takes advantage of large installed base
  - —Does not transmit continuous signal between packets
  - —Useful in battery-powered applications
- Can not get 100 Mbps on single twisted pair
  - —So data stream split into three separate streams
  - —Four twisted pairs used
  - —Data transmitted and received using three pairs
  - —Two pairs configured for bidirectional transmission
- Use ternary signaling scheme (8B6T)



## 100BASE-T Options



# 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



# Mixed Configurations

- Fast Ethernet supports mixture of existing 10-Mbps LANs and newer 100-Mbps LANs
- E.g. 100-Mbps backbone LAN to support 10-Mbps hubs
  - —Stations attach to 10-Mbps hubs using 10BASE-T
  - —Hubs connected to switching hubs using 100BASE-T
    - Support 10-Mbps and 100-Mbps
  - —High-capacity workstations and servers attach directly to 10/100 switches
  - —Switches connected to 100-Mbps hubs using 100-Mbps links
  - —100-Mbps hubs provide building backbone
    - Connected to router providing connection to WAN

## Gigabit Ethernet - Differences

- Carrier extension
  - —At least 4096 bit-times long (512 for 10/100)
- Frame bursting
  - —allow multiple frames to be transmitted consecutively, i.e. "together"
- Not needed if using a switched hub to provide dedicated media access



## Gigabit Ethernet – Physical

- 1000Base-SX
  - —Short wavelength, multimode fiber
- 1000Base-LX
  - —Long wavelength, Multi or single mode fiber
- 1000Base-CX
  - —Copper jumpers <25m, shielded twisted pair
- 1000Base-T
  - —4 pairs, cat 5 UTP

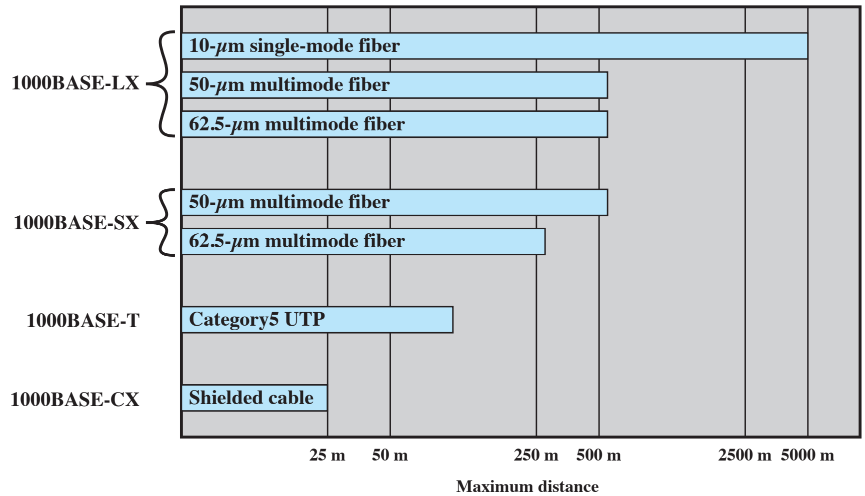



**Figure 12.5 Gigabit Ethernet Medium Options (log scale)**



# 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

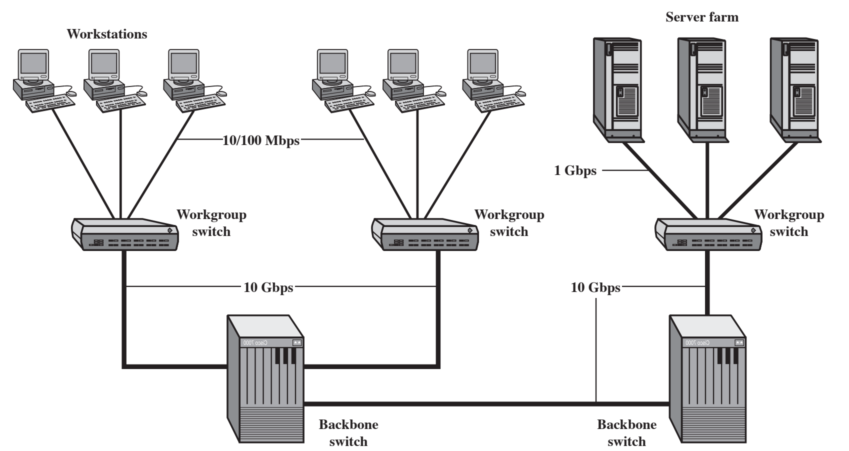
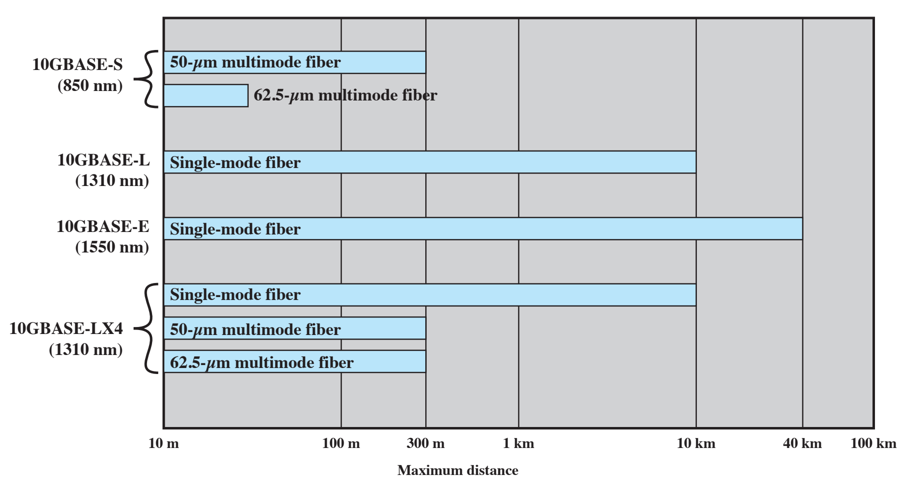



**Figure 12.6 Example 10 Gigabit Ethernet Configuration**



# 10Gbps Ethernet Distance Options (log scale)



## 100-Gbps Ethernet (100GbE)

- 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



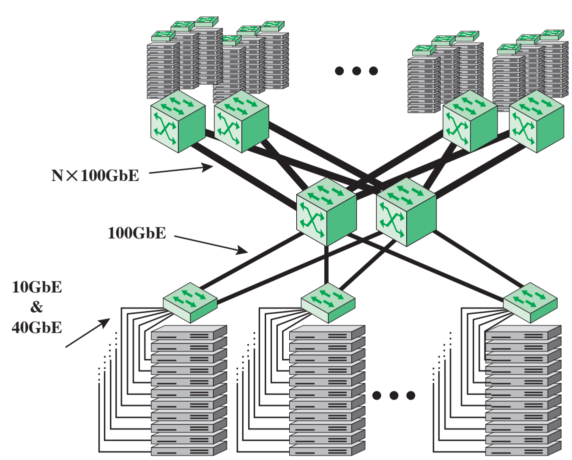



Figure 12.8 Example 100-Gbps Ethernet Configuration for Massive Blade Server Site

# Multilane Distribution

used to achieve the required data rates

- Multilane distribution:
  - Switches implemented as multiple parallel channels
    - Separate physical wires
- Virtual lanes:
  - —If a different number of lanes are actually in use, virtual lanes are distributed into physical lanes in the PMD (physical medium dependent) sublayer
  - —Form of inverse multiplexing



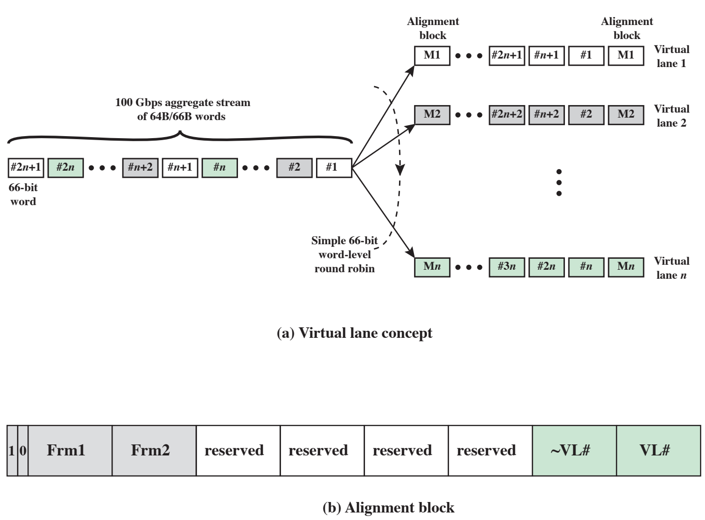



**Figure 12.9 Multilane Distribution for 100-Gbps Ethernet**

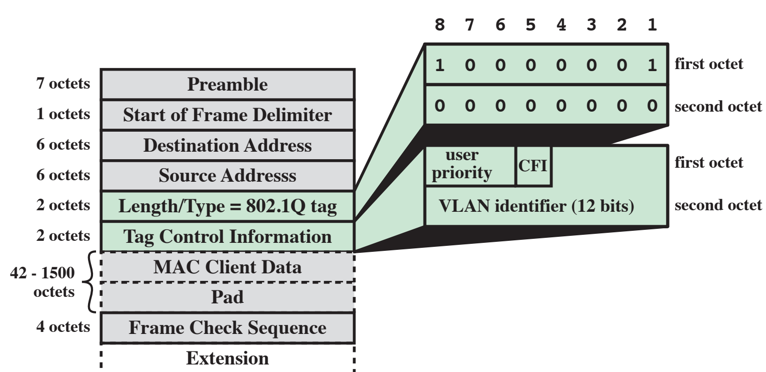
# Media Options for 40-Gbps and 100-Gbps Ethernet

Table 12.3

| | 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)






CFI = Canonical Format Indicator
VLAN = virtual local area network

**Figure 12.10 Tagged IEEE 802.3 MAC Frame Format**

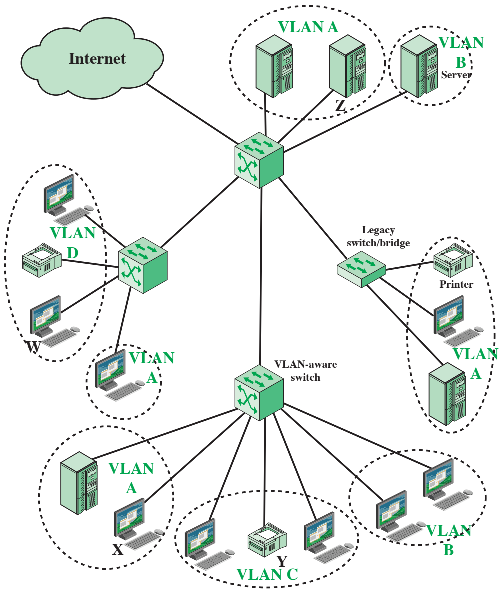



Figure 12.11 A VLAN Configuration



# Summary

- Traditional Ethernet
  - —IEEE 802.3 medium access control
  - —IEEE 802.3 10-Mbps specifications (Ethernet)
- IEEE 802.1Q VLAN standard
- High-speed Ethernet
  - —IEEE 802.3 100-Mbps specifications (Fast Ethernet)
  - —Gigabit Ethernet
  - —10-Gbps Ethernet
  - —100-Gbps Ethernet