521150A
Introduction to Internet

Lecture 5 – Data link layer, part III: Wired LANs & Ethernet

Erkki Harjula 2023

Schedule of the course

PART I: Basics of networking and Internet, data link layer

Mon 13.3. 10-12 L6/Zoom — Lecture 1: Introduction & motivation

Tue 14.3. 10-12 L6/Zoom — Lecture 2: Architecture & design principles

Wed 15.3. 10-12 L5/Zoom — Lecture 3: Data link layer – basics part I

Thu 16.3. 10-12 L4/Zoom — Exercise session 1A

Mon 20.3. 10-12 L6/Zoom — Lecture 4: Data link layer – basics part II

Tue 21.3. 10-12 L6/Zoom — Exercise session 1B

Tue 21.3. 14-18 AT122/Zoom — Lab exercise 1 – group 3

Wed 22.3. 10-12 L5/Zoom — Lecture 5: Data link layer – Wired networks

Wed 22.3. 14-18 AT122/Zoom — Lab exercise 1 – group 4

Thu 23.3. 14-18 AT122/Zoom — Lab exercise 1 – group 2

Fri 24.3. 12-16 AT122/Zoom — Lab exercise 1 – group 1

Mon 27.3. 14-16 L5/Zoom — Exercise session 1C

Tue 28.3. 10-12 L6/Zoom — Lecture 6: Data link layer – Wireless networks

Wed 29.3. 10-12 Moodle — Theory exam 1

PART II: Network and transport layers

Thu 30.3. 8-10 L5/Zoom — Lecture 7: Network layer part I

Mon 3.4. 14-16 L5/Zoom — Exercise session 2A

Tue 4.4. 10-12 L6/Zoom — Lecture 8: Network layer part II

Wed 5.4. 10-12 L5/Zoom — Lecture 9: Transport layer part I

Course work intro

Tue 11.4. 10-12 L6/Zoom — Lecture 10: Transport layer part II

Tue 11.4. 14-18 AT122/Zoom — Lab exercise 2 – group 3

Wed 12.4. 10-12 Moodle — Theory exam 2

Wed 12.4. 14-18 AT122/Zoom — Lab exercise 2 – group 4

Thu 13.4. 8-10 L5/Zoom — Exercise session 2B

Thu 13.4. 14-18 AT122/Zoom — Lab exercise 2 – group 2

Fri 14.4. 12-16 AT122/Zoom — Lab exercise 2 – group 1

Course work (independent work)

PART III: Application layer, network security and multimedia

Fri 14.4. 10-12 L5/Zoom — Lecture 11: Networking applications

Mon 17.4. 14-16 L5/Zoom — Lecture 12: Network security

Tue 18.4. 10-12 L6/Zoom — Exercise session 3A

Wed 19.4. 10-12 L5/Zoom — Lecture 13: Multimedia and QoS

Mon 24.4. 14-16 L5/Zoom — Exercise session 3B

Tue 25.4. 14-18 AT122/Zoom — Lab exercise 3 – group 3

Wed 26.4. 10-12 L5/Zoom — Lecture 14: Challenges&Future Internet trends

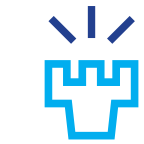
Wed 26.4. 14-18 AT122/Zoom — Lab exercise 3 – group 4

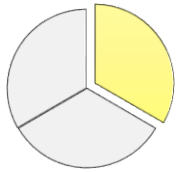
Thu 27.4. 14-18 AT122 /Zoom — Lab exercise 3 – group 2

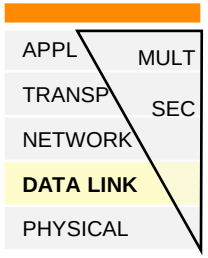
Fri 28.4. 12-16 AT122 /Zoom — Lab exercise 3 – group 1

Wed 3.5. 10-12 Moodle — Theory exam 3

Thu 25.5. 16-19 L1 — Final exam

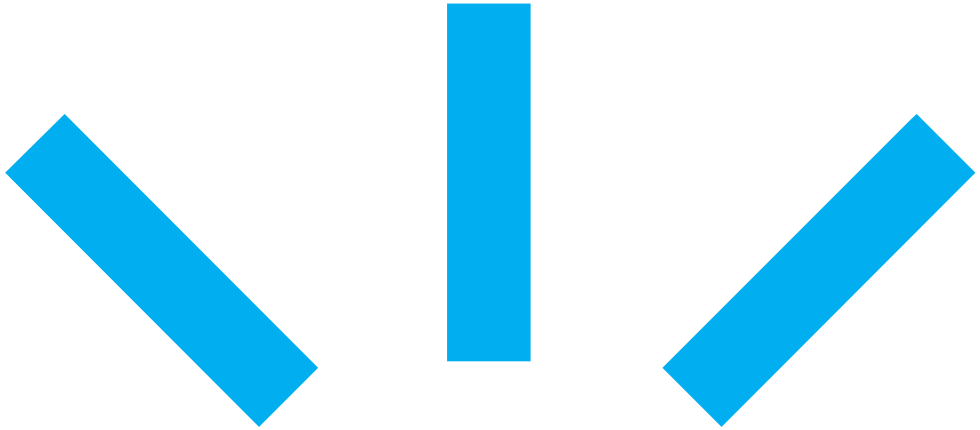




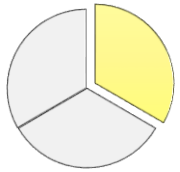


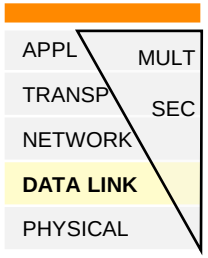
Main learning objectives of this lecture

1. **Know the basics of Local Area Networks (LAN)**
   - **Topologies**
   - **Addressing**
   - **Interconnection devices (hubs, bridges, switches)**
     - Self learning
     - Link-layer switching
2. **Know the basics of Ethernet**
   - **Medium access control (CSMA/CD)**
   - **Frame structure**
   - **Switched Ethernet**

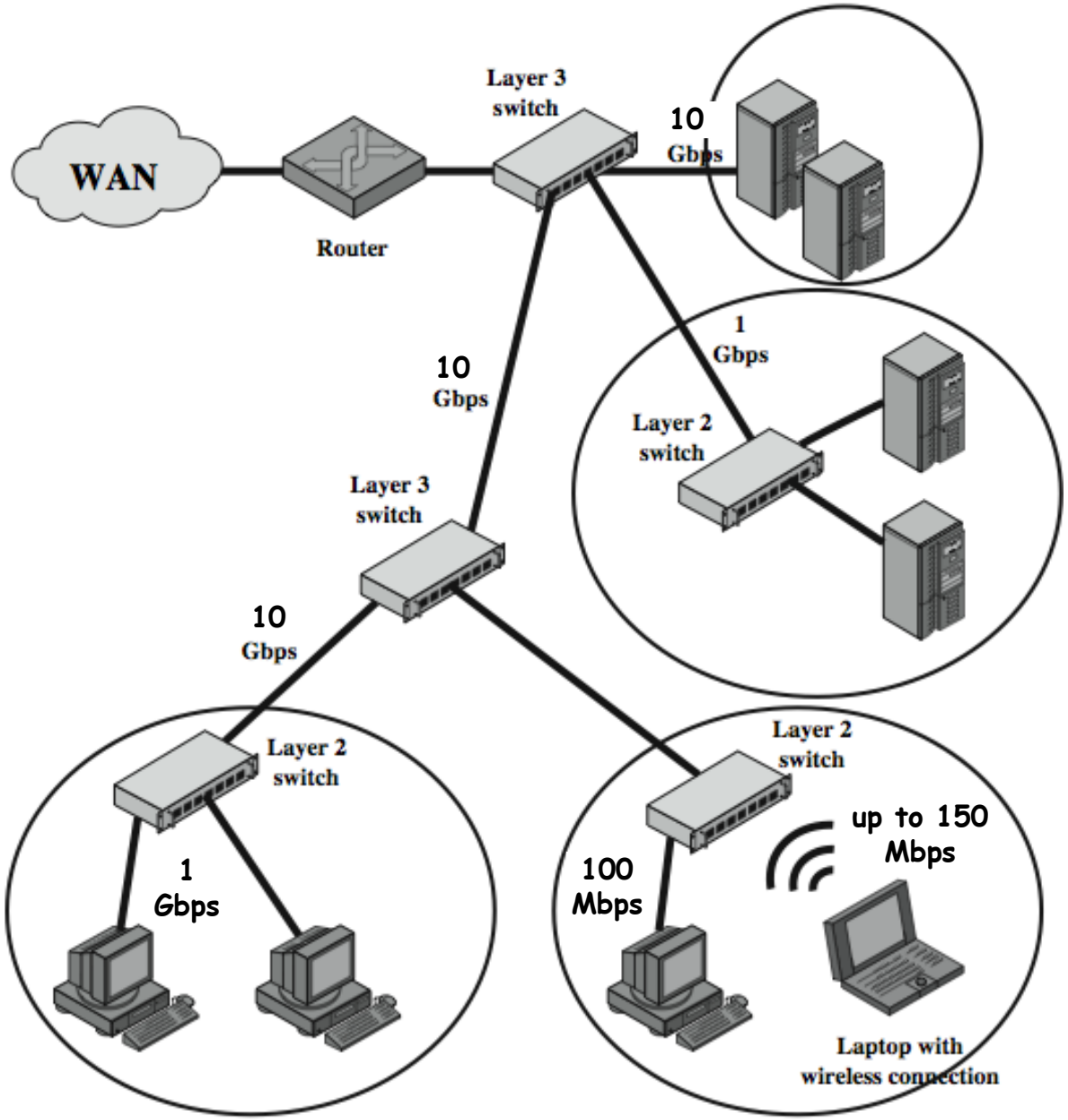


Logical Link Control (LLC): Principles, topologies and reference model

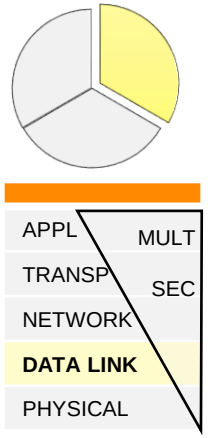




Typical large organization LAN

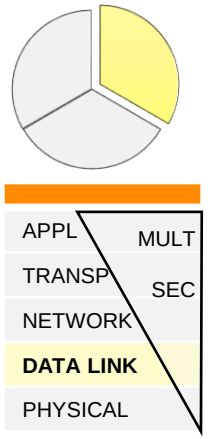


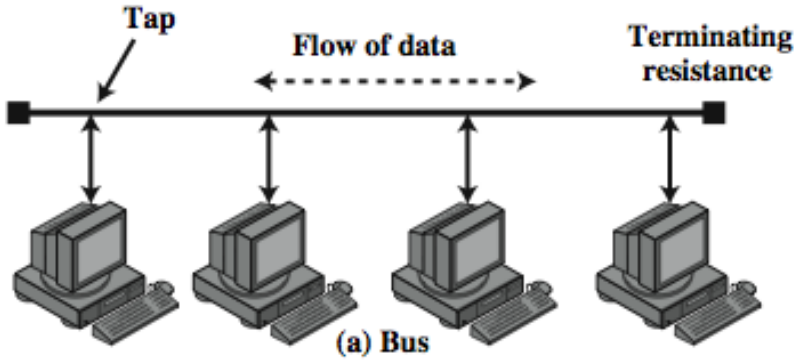
Common LAN applications

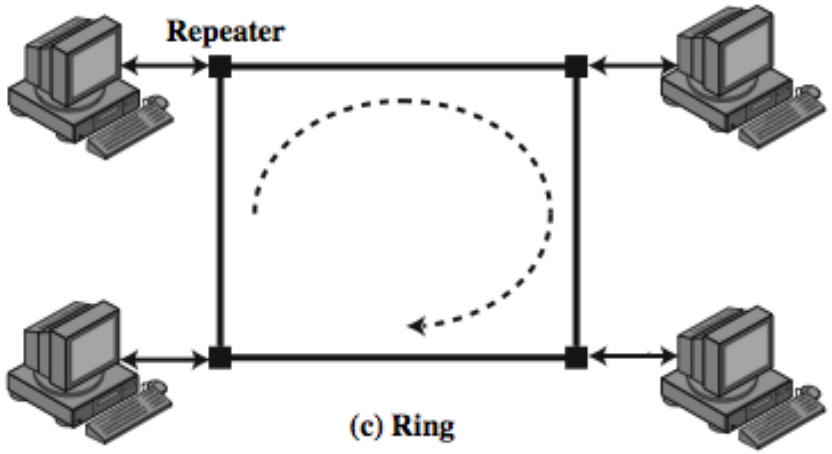


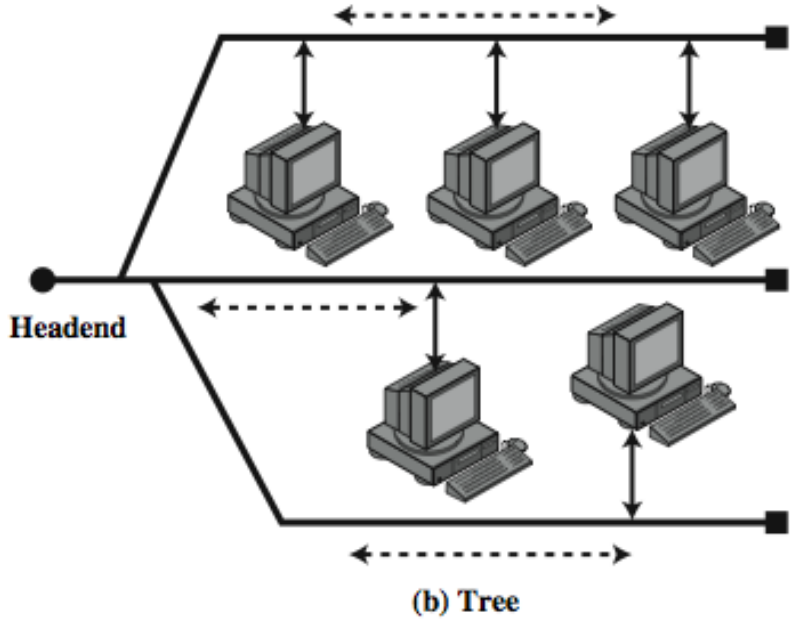
- **Personal computer LANs**
  - Low cost
  - Limited distance
  - Limited number of devices
- **Local backend networks, storage area networks (SAN), data center networks**
  - Interconnecting large systems (mainframes and large storage devices)
  - High data rate
  - High-speed interfaces
  - Distributed access
  - Limited distance
  - Limited number of devices
- **High speed office networks**
  - Desktop image processing
  - High capacity local storage
- **Backbone LANs**
  - Interconnect low speed local LANs
  - Reliability
  - Capacity
  - Cost

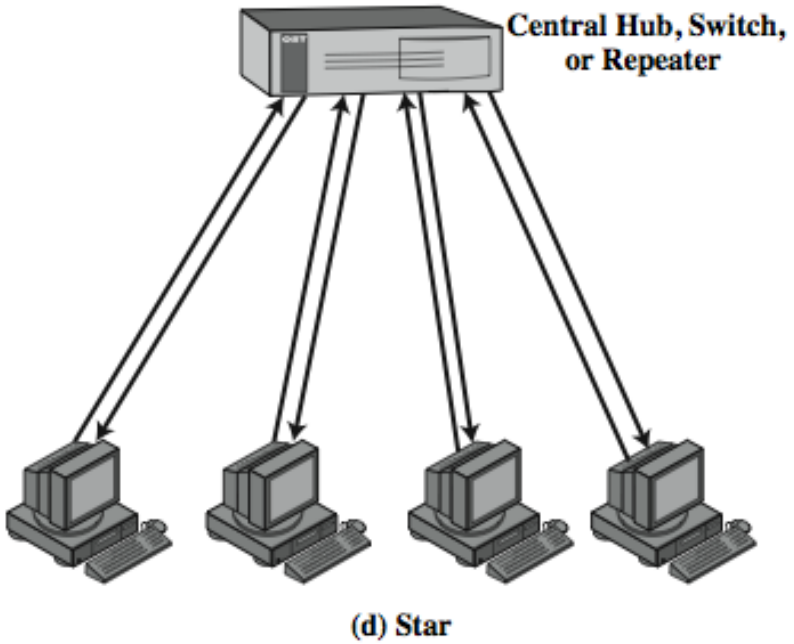
Basic LAN topologies



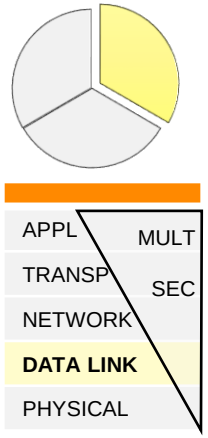




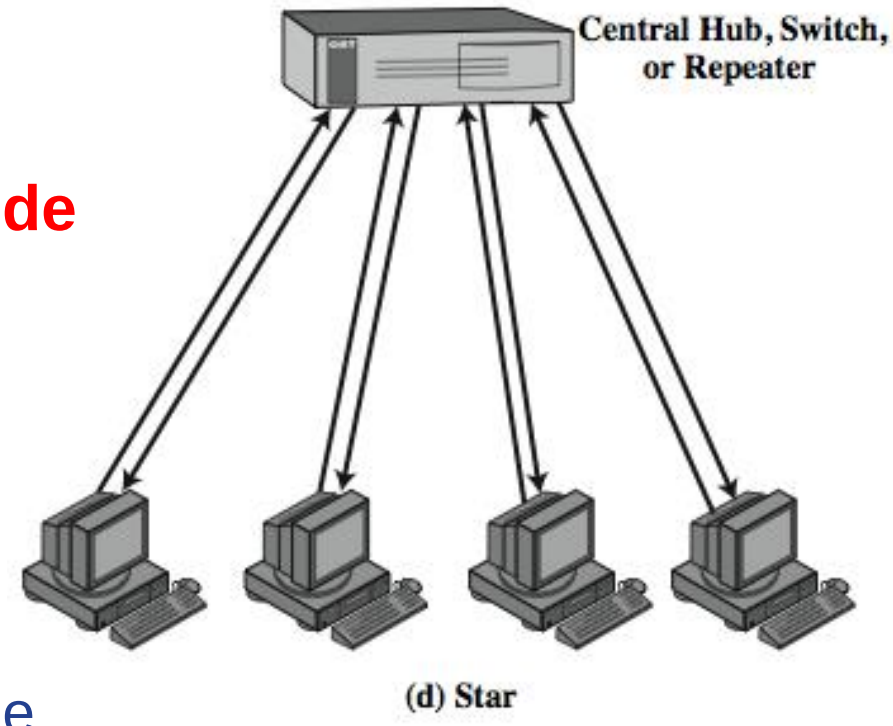




Star topology

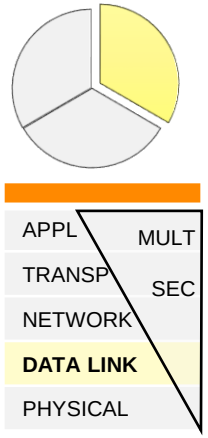


- **Each station connects to central node**
  - **Usually via two point-to-point links for full-duplex communication**
- **Central node types**
  - **Hub and repeater**
    - Physical star, logical bus
    - Only one station can transmit at a time
  - **Switch and bridge**
    - Store&forward, i.e. isolates nodes
    - Stations can transmit simultaneously
- **Most modern LANs use star topology with frame switching**

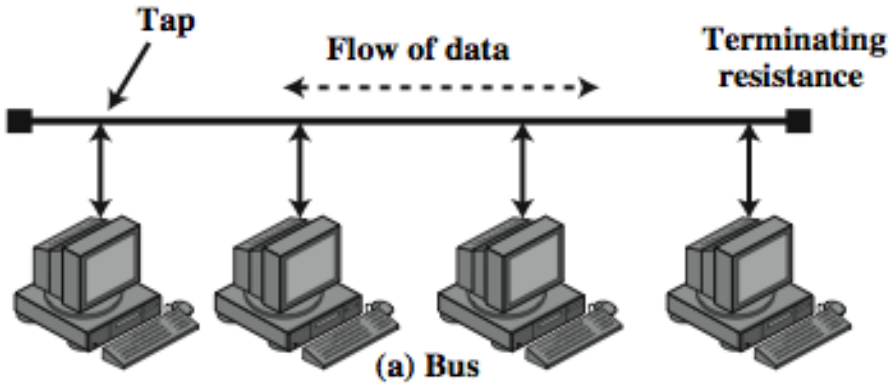


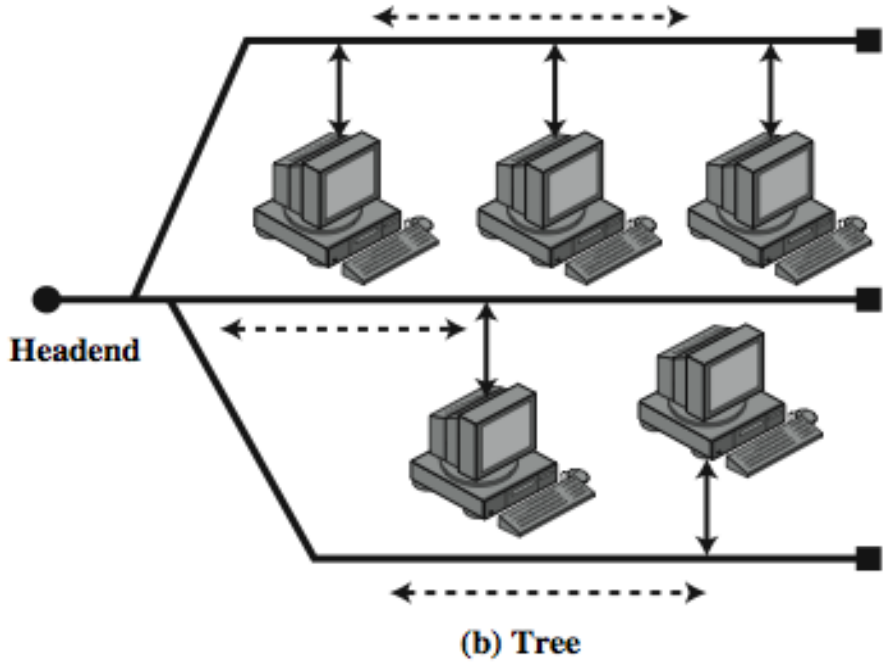
(d) Star

Bus and tree topology

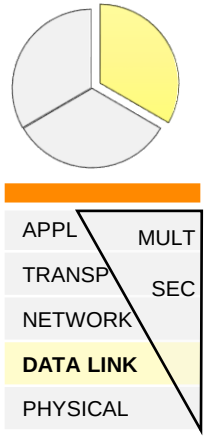


- Tree is a generalization of bus
- Used with multipoint medium
- Transmission propagates throughout medium
- Heard by all stations
- Full duplex connection between station and tap
  - Allows for transmission and reception
- Need to regulate transmission
  - To avoid collisions and hogging
- Terminator absorbs frames at end of medium
- Headend connected to branching cables

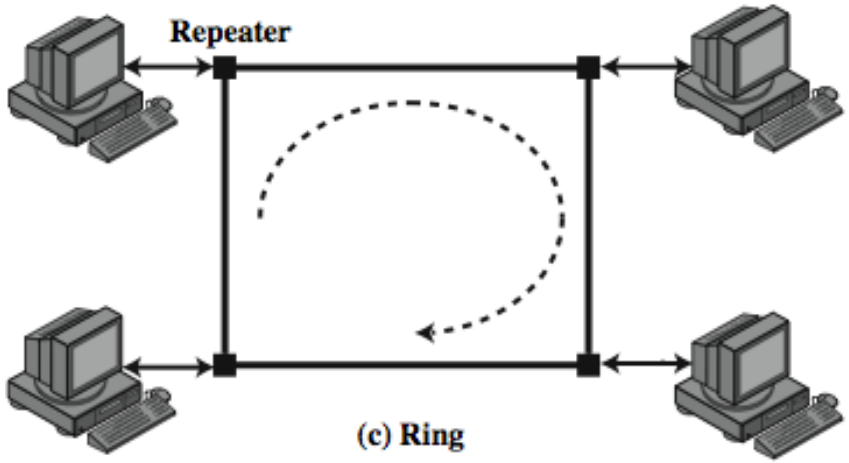


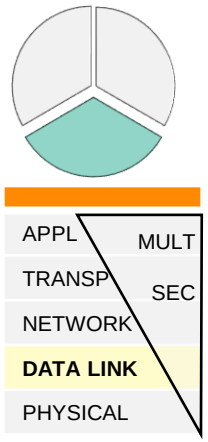


Ring topology

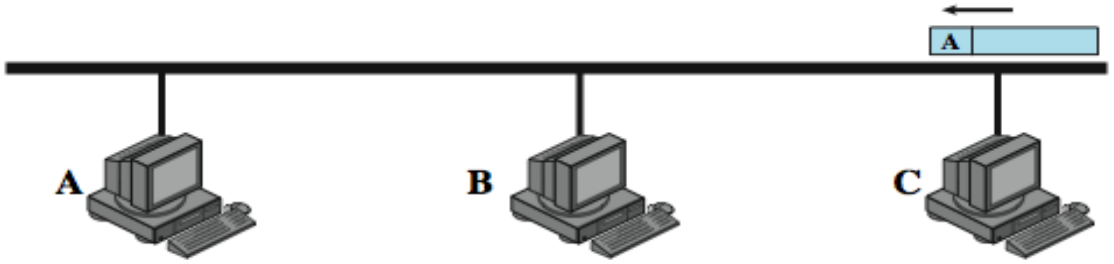


- A **closed loop of repeaters** joined by point to point links
- Receive data on one link retransmit on another
  - **Links unidirectional**
  - Stations attach to repeaters
- Data in frames
  - **Circulate past all stations**
  - Destination recognizes address and copies frame
  - Frame circulates back to source where it is removed
- Medium access control determines when a station can insert frame

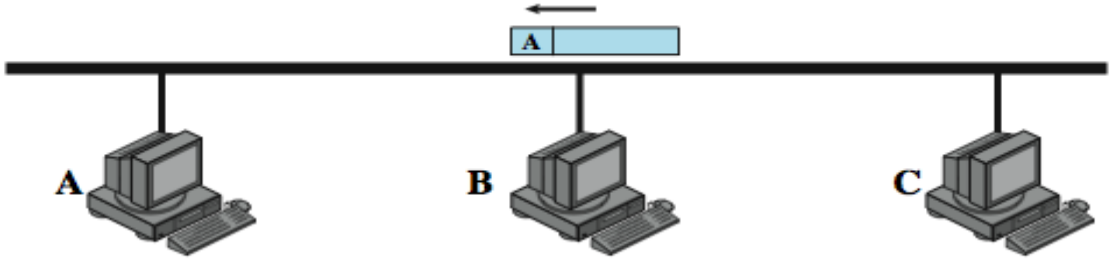




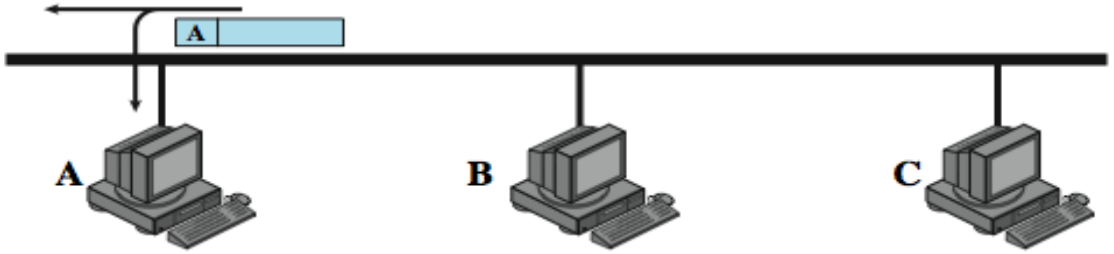
Frame transmission in bus LAN



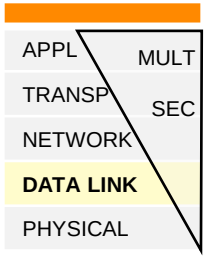
C transmits frame addressed to A



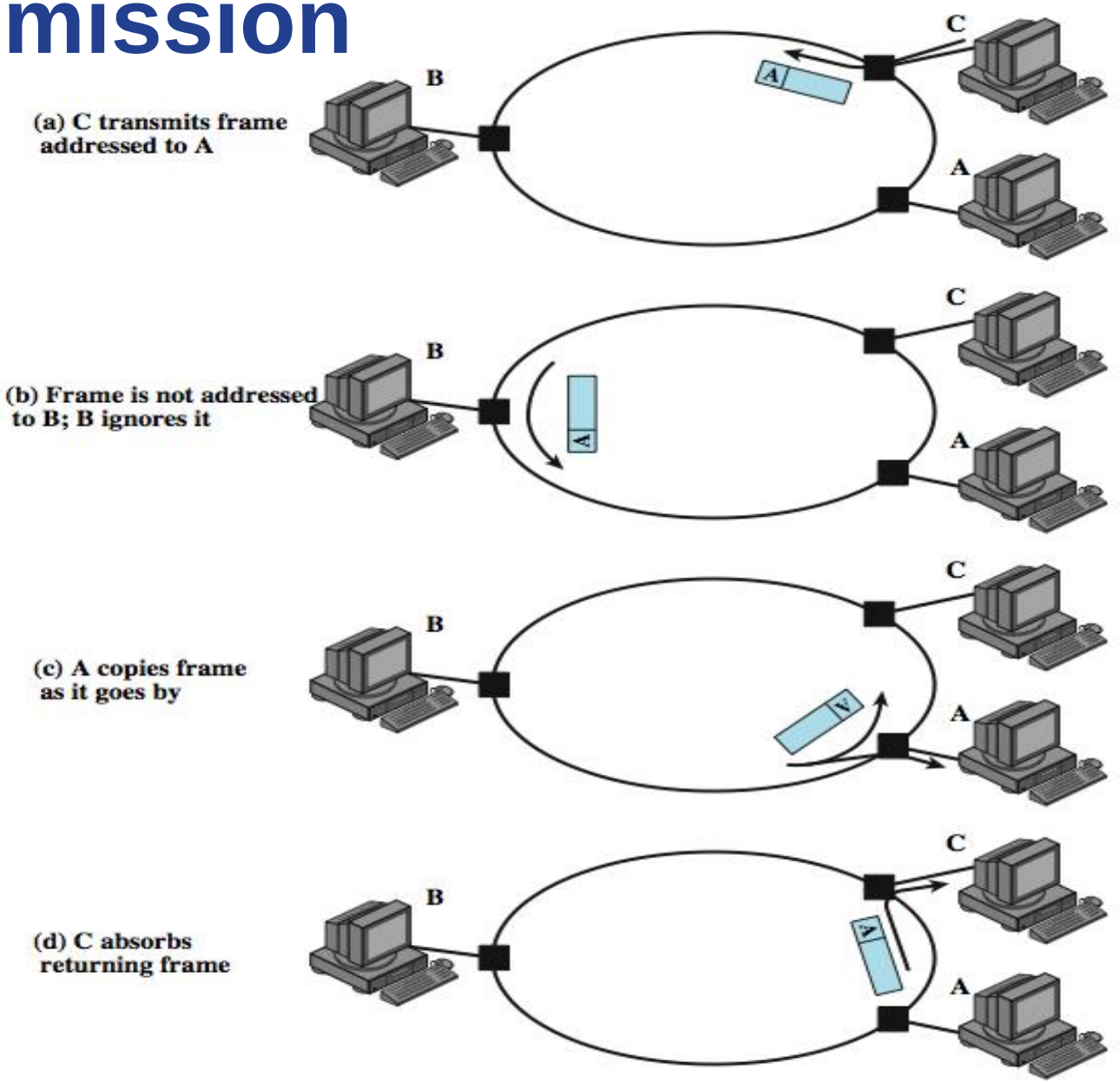
Frame is not addressed to B; B ignores it

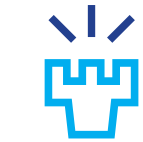


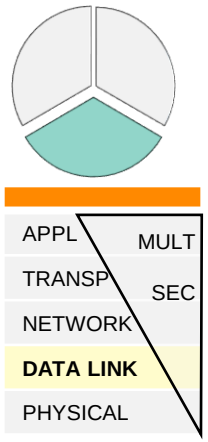
A copies frame as it goes by



Frame transmission in ring LAN



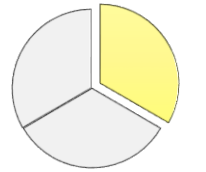


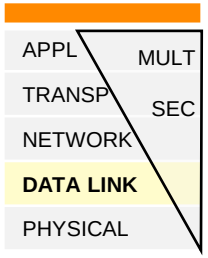


Choice of LAN topology

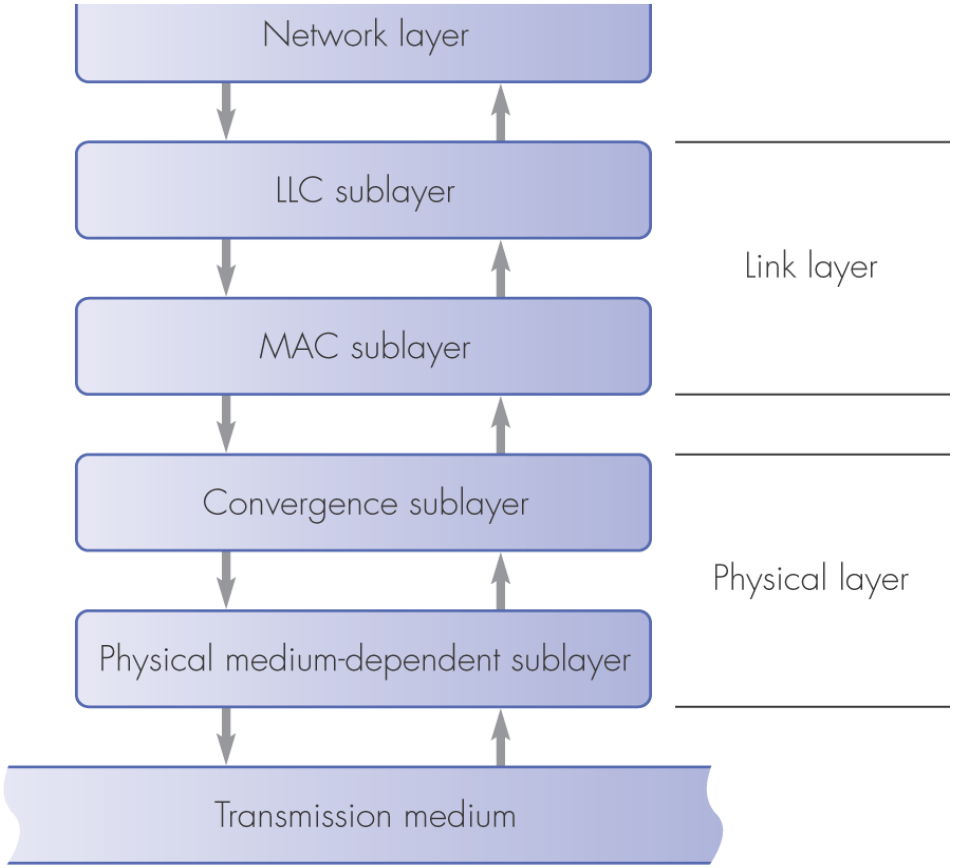
- Reliability
- Expandability
- Performance
- Needs considerations in context of
  - Medium
  - Wiring layout
  - Access control

IEEE 802 LAN/MAN reference model



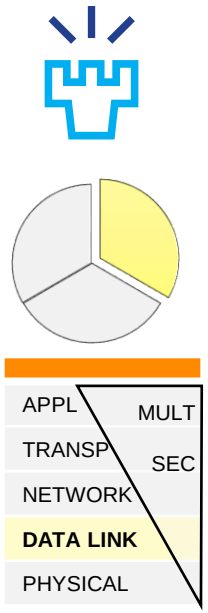


Protocol stack



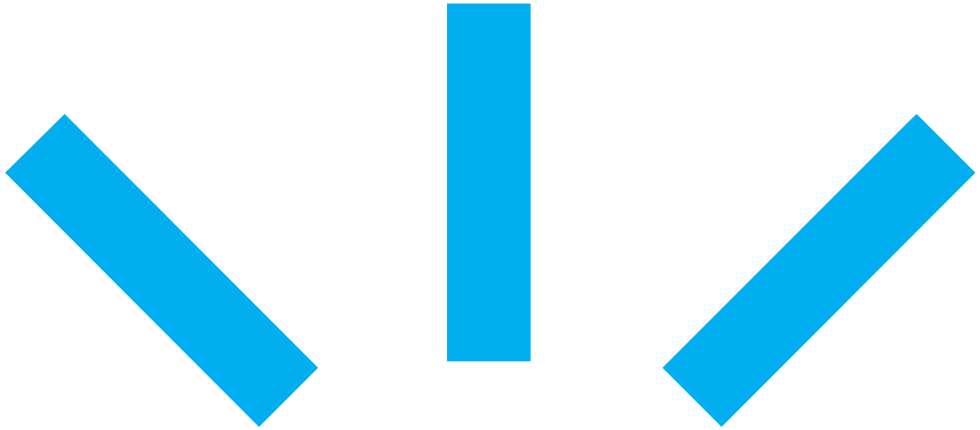
Examples

IEEE	802.1	Station management
	802.1d	Transparent bridges
	802.1Q	Virtual LANs
	802.2	Logical link control (LLC)
	802.3	CSMA/CD (Ethernet) bus
	802.3u	Fast Ethernet
	802.3x	Hop-by-hop switch flow control
	802.3z	Gigabit Ethernet
	802.3ae	10 Gigabit Ethernet

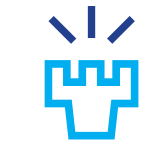


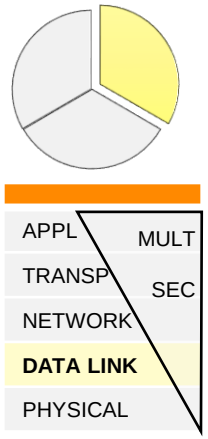
IEEE 802: Layers

- **Logical Link Control (LLC)**
  - Interface to higher levels
  - Flow and error control
- **Medium Access Control (MAC)**
  - On transmit assemble data into frame
  - On receive disassemble frame
  - Govern access to transmission medium
  - For same LLC, may have several MAC options
- **Physical (PHY)**
  - Encoding/decoding of signals
  - Preamble generation/removal
  - Bit transmission/reception
  - Transmission medium and topology



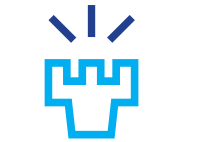
LAN addressing

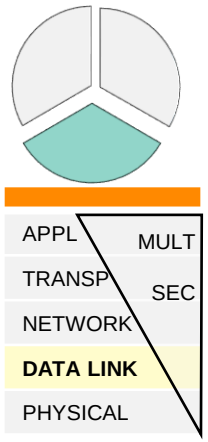




IEEE 802.2: Logical Link Control (LLC)

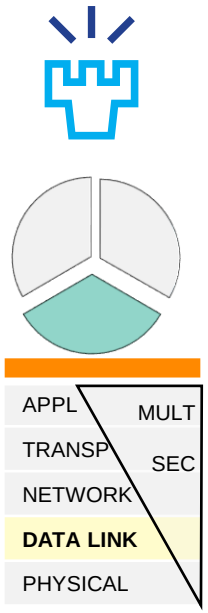
- **Transmission of link level PDUs between stations**
- **Must support multiple access, shared medium**
- **However, MAC layer handles link access details**
- **Hides differences of various MAC (Ethernet, WLAN etc.)**
  - Provides single format and interface to network layer





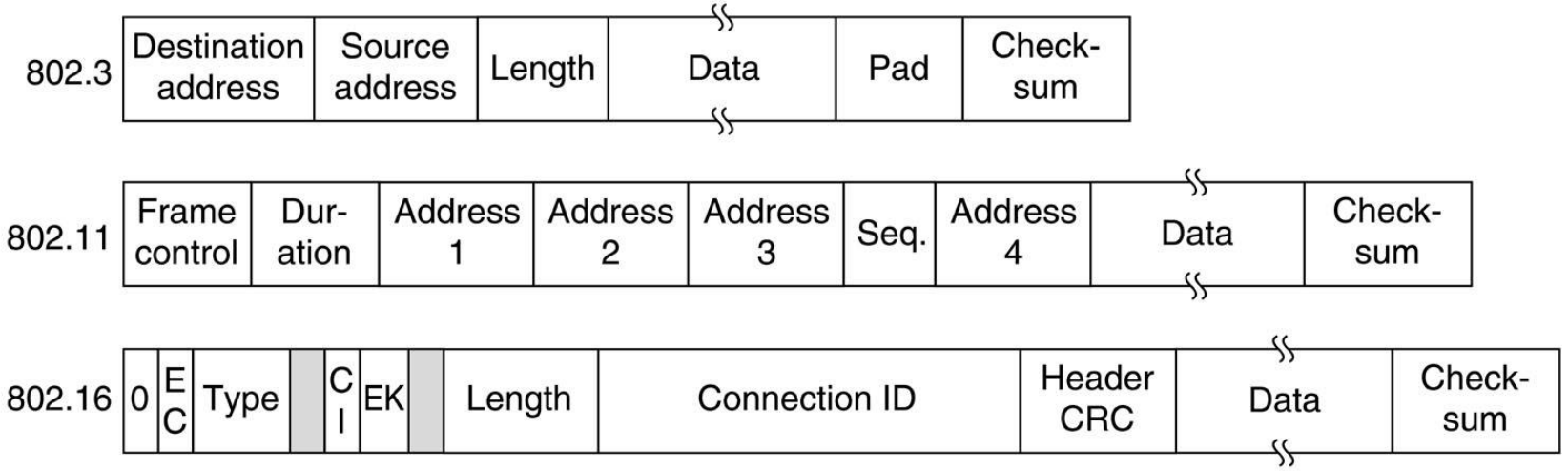
IEEE 802.2: LLC service options

1. **Unacknowledged connectionless service**
   - Unreliable datagram service
   - No connection setup
   - No error control
   - No flow control
2. **Acknowledged connectionless service**
   - Acknowledged datagram service
   - No connection setup
3. **Connection-mode service**
   - Reliable connection-oriented service
   - HDLC asynchronous balanced mode
   - For simple devices

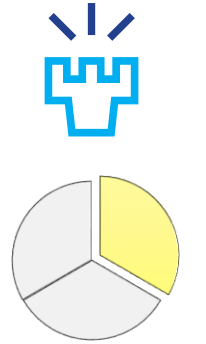


IEEE 802.2: Bridging different LAN types

- **Different LAN's use different frame formats**

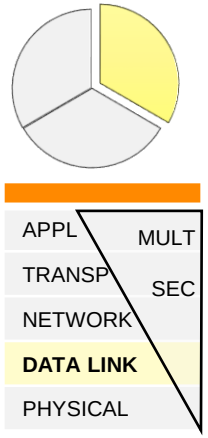


- **Different LAN's may run at different data rates**
  - **Buffering problems**
- **Different LAN's have different maximum frame lengths**
- **Security**
  - **802.11 and 802.16 support encryption in the link layer, 802.3 not**
- **Quality of Service**



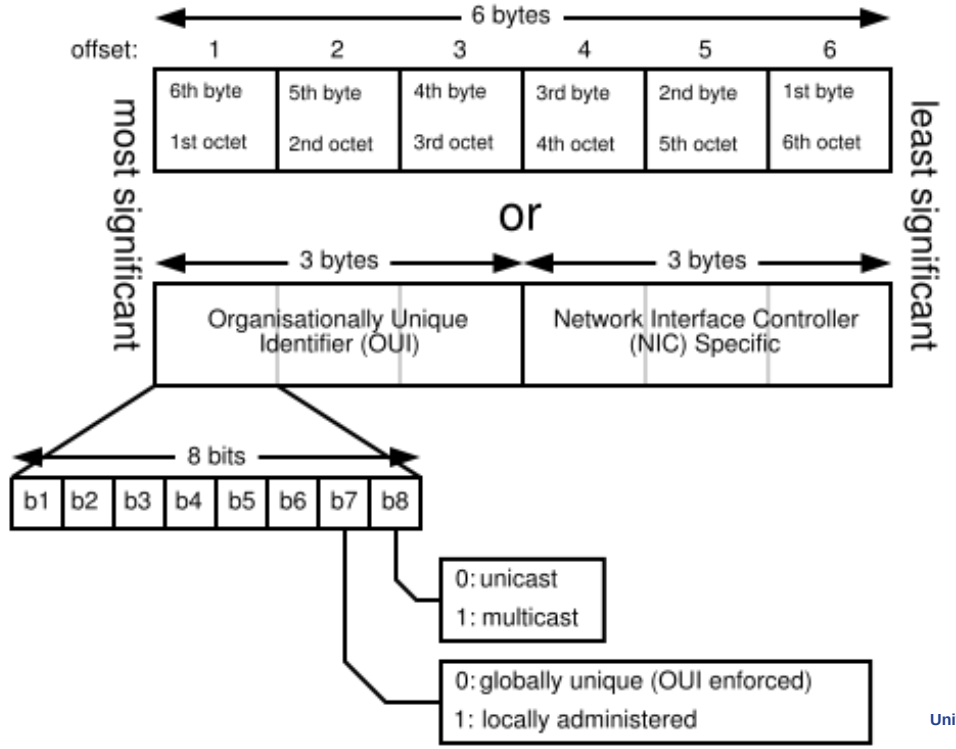
IEEE 802.2: Example of bridging from 802.11 to 802.3

Example: WLAN (WiFi)
AP (access point)

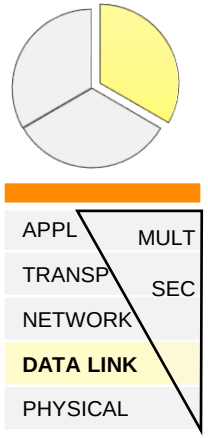


LAN addressing (1)

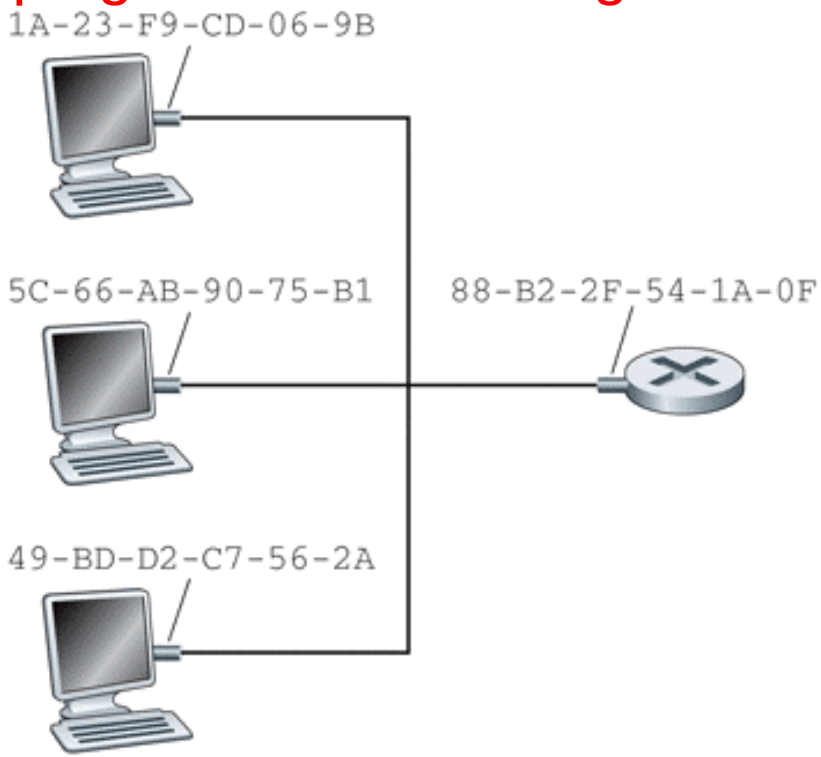
- **In 802 LANs and in some other LANs, each LAN adapter has fixed and universal 48-bit address**
- **Called LAN/MAC/Ethernet/physical address**
- **Address expressed in hexadecimal AA-BB-CC-DD-EE-FF**
  - 3-byte manufacturer ID (OUI) + 3-byte adapter ID
- **Address burned into adapter ROM (sometimes software settable)**
- **Address allocation governed by IEEE**
  - Manufacturer purchases portion of address space (given OUI) and assures uniqueness of addresses
  - E.g. OUI C8-97-9F has been allocated for Nokia



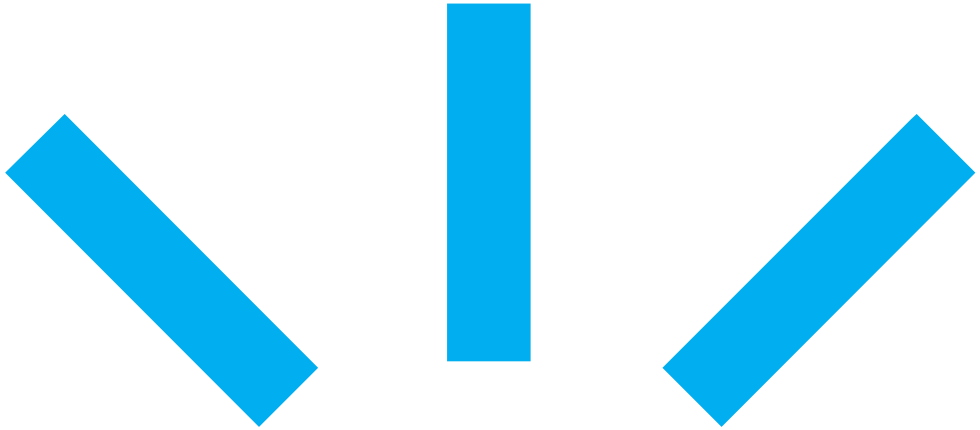
LAN addressing (2)



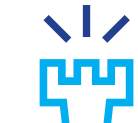
- **Fixed and universal MAC address → portability**
  - Can move LAN card from one LAN to another
  - Note: IP address is NOT portable, but depends on the IP subnet to which node is attached
  - MAC address ↔ IP address mapping is resolved using ARP (Address Resolution Protocol)

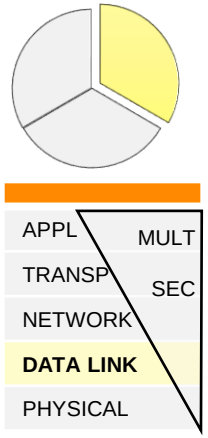


- **Frame destination address FF-FF-FF-FF-FF-FF (all 1's) denotes LAN broadcast**
  - Each adapter in LAN receives the frame



LAN interconnection





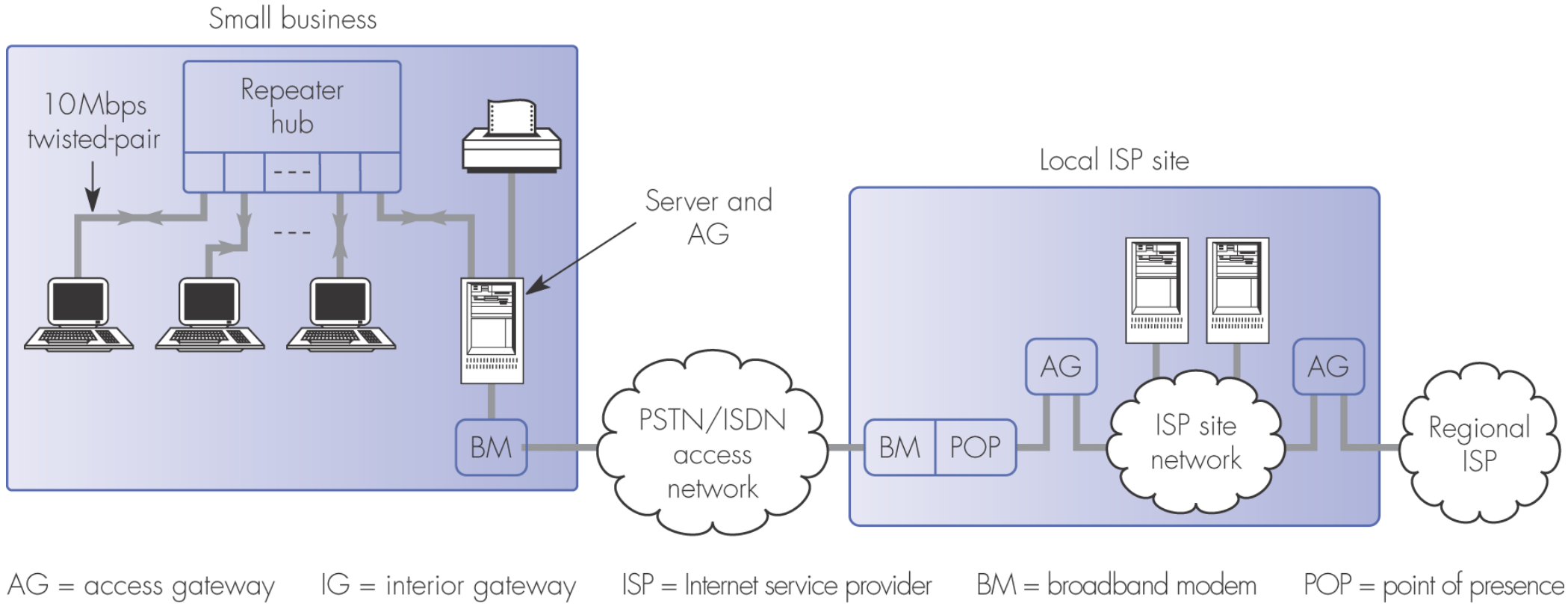
LAN interconnection technologies

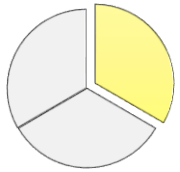
Hubs

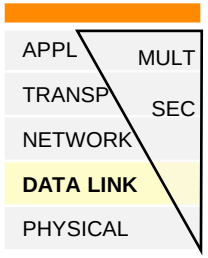
Bridges

(Obsolete technology)

Switches

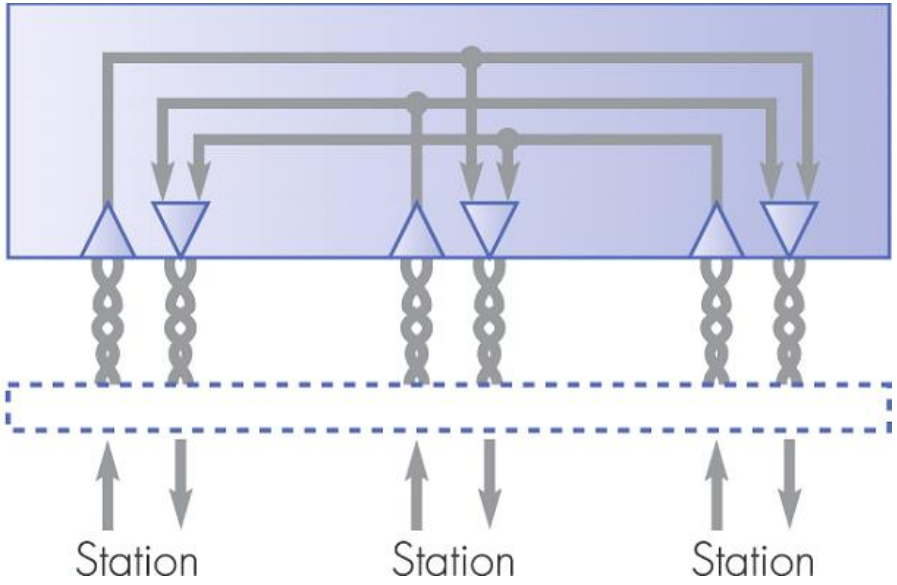




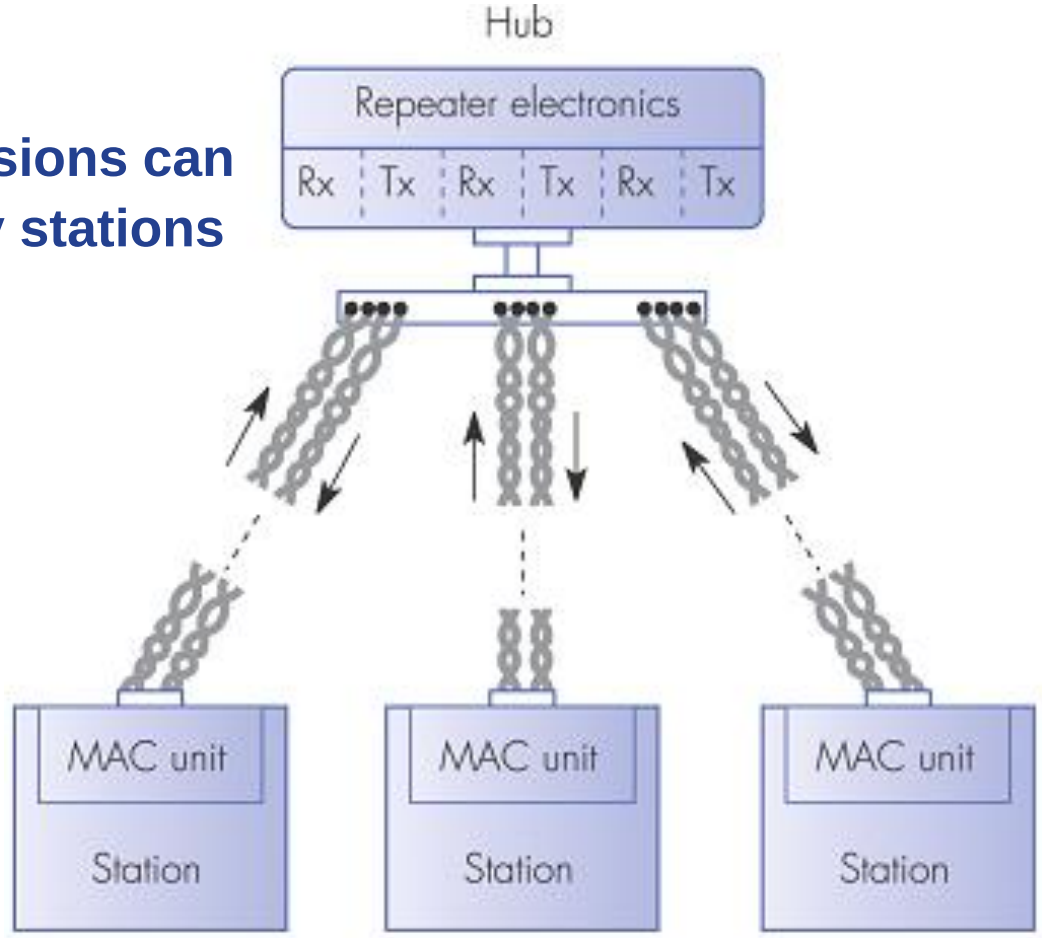


Hub

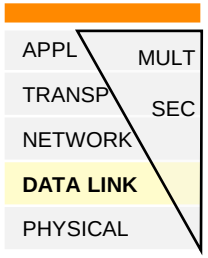
- Essentially **physical-layer (L1) repeater**
  - Bits coming from one link go out all other links
  - At the same rate
  - No frame buffering
  - No CSMA/CD: collisions can be detected only by stations



repeater schematic

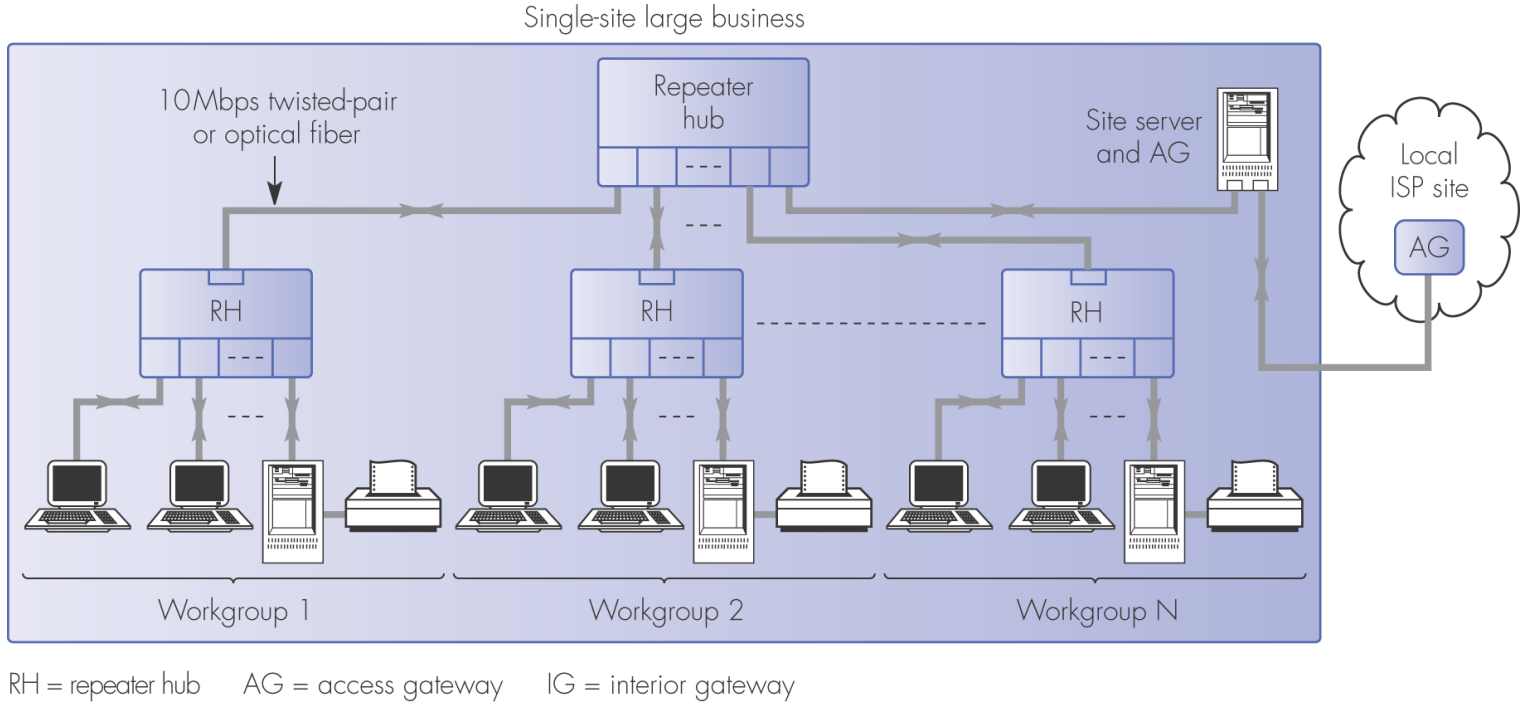


topology



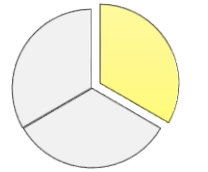
Interconnecting with hubs

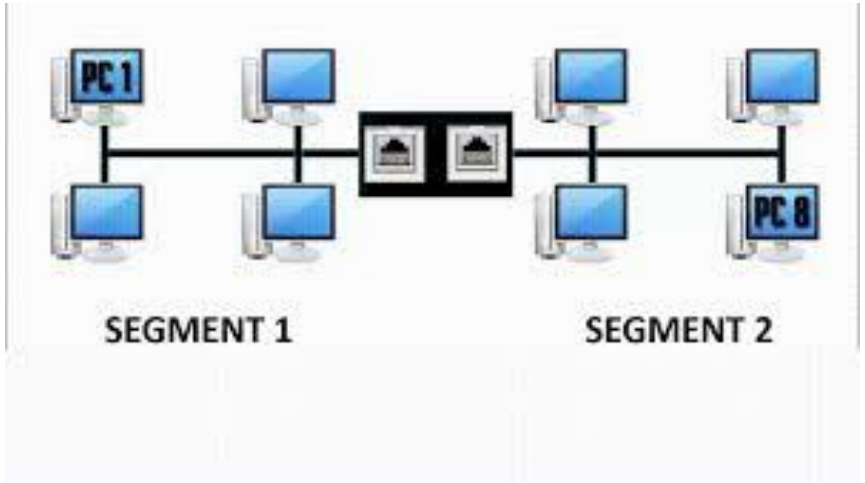
- Backbone hub interconnects LAN segments
- Extends max distance between nodes
- But: individual segment collision domains become one large collision domain
- Can NOT interconnect different LAN segments (technologies)



Large single-site network implemented with hubs

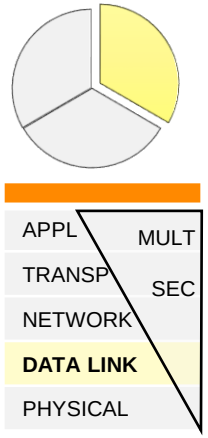
Bridge

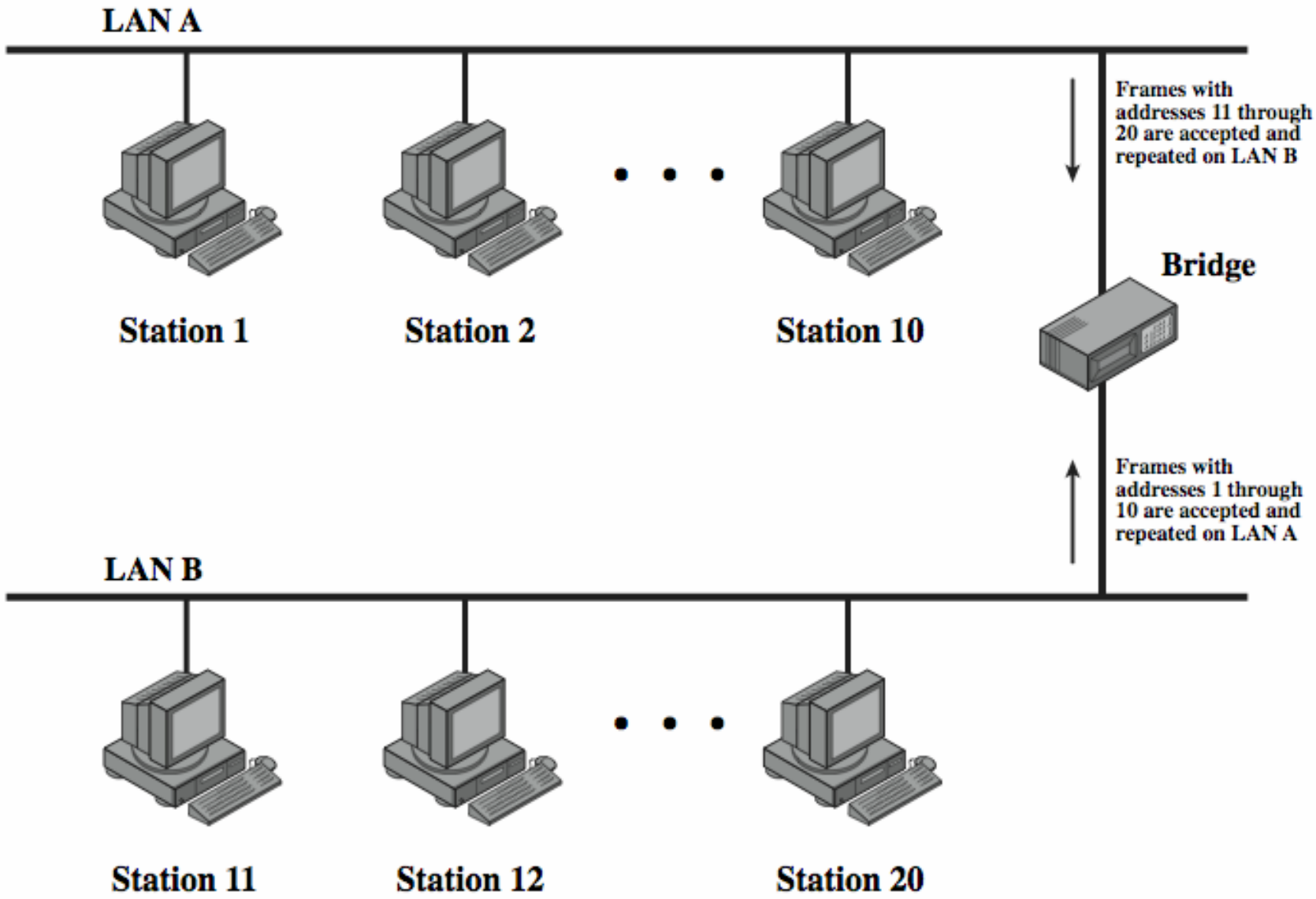


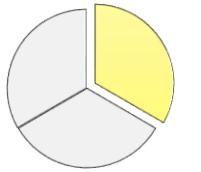


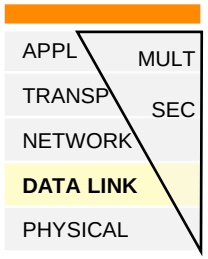
- **Link layer (L2) device**
  - Stores and forwards frames
  - Examines frame header and **selectively forwards** frame based on MAC destination address
  - When frame is to be forwarded on a LAN segment, uses CSMA/CD to access segment
  - **Promiscuous** operation mode
    - Receives and buffers frames received on each of its ports
- **Transparent**
  - Hosts and hubs are unaware of presence of bridges
- **Plug-and-play, self-learning**
  - Bridges do not need to be configured
- **Can interconnect different LAN segments**

Bridge: Role


APPL
MULT
TRANSP
SEC
NETWORK
DATA LINK
PHYSICAL

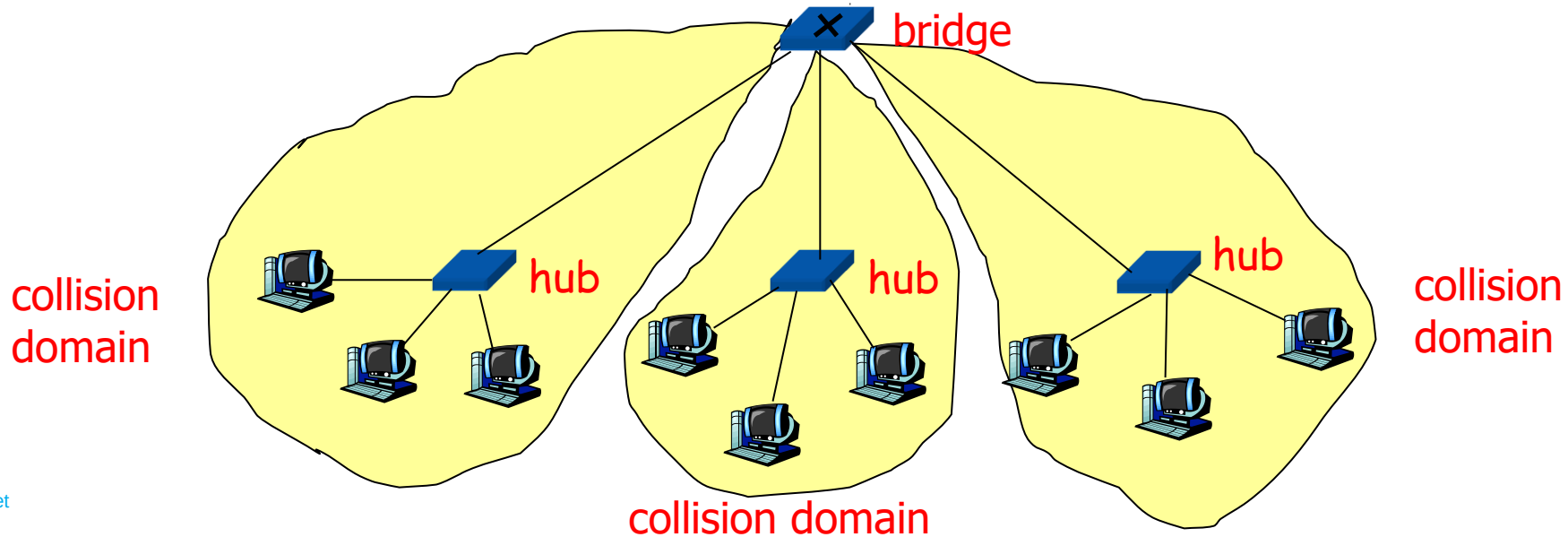

LAN A
Station 1
Station 2
Station 10
Frames with addresses 11 through 20 are accepted and repeated on LAN B
Bridge
Frames with addresses 1 through 10 are accepted and repeated on LAN A
LAN B
Station 11
Station 12
Station 20

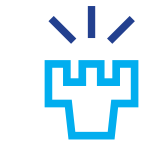


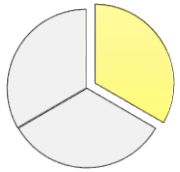


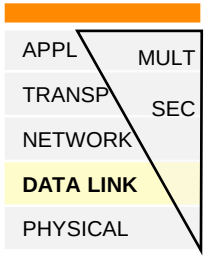
Bridge: Traffic isolation

- **Bridge breaks subnet into LAN segments and filters frames between segments**
  - Erroneous frames are not forwarded
  - Same-LAN-segment frames not usually forwarded onto other LAN segments (do not load rest of the network)
  - LAN segments become separate collision domains
    - **Allows simultaneous transmission on attached segments!**





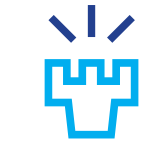


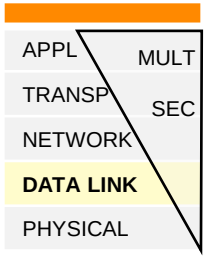


Bridge: Frame forwarding

- Maintain forwarding database for each port
  - Lists station addresses reached through each port

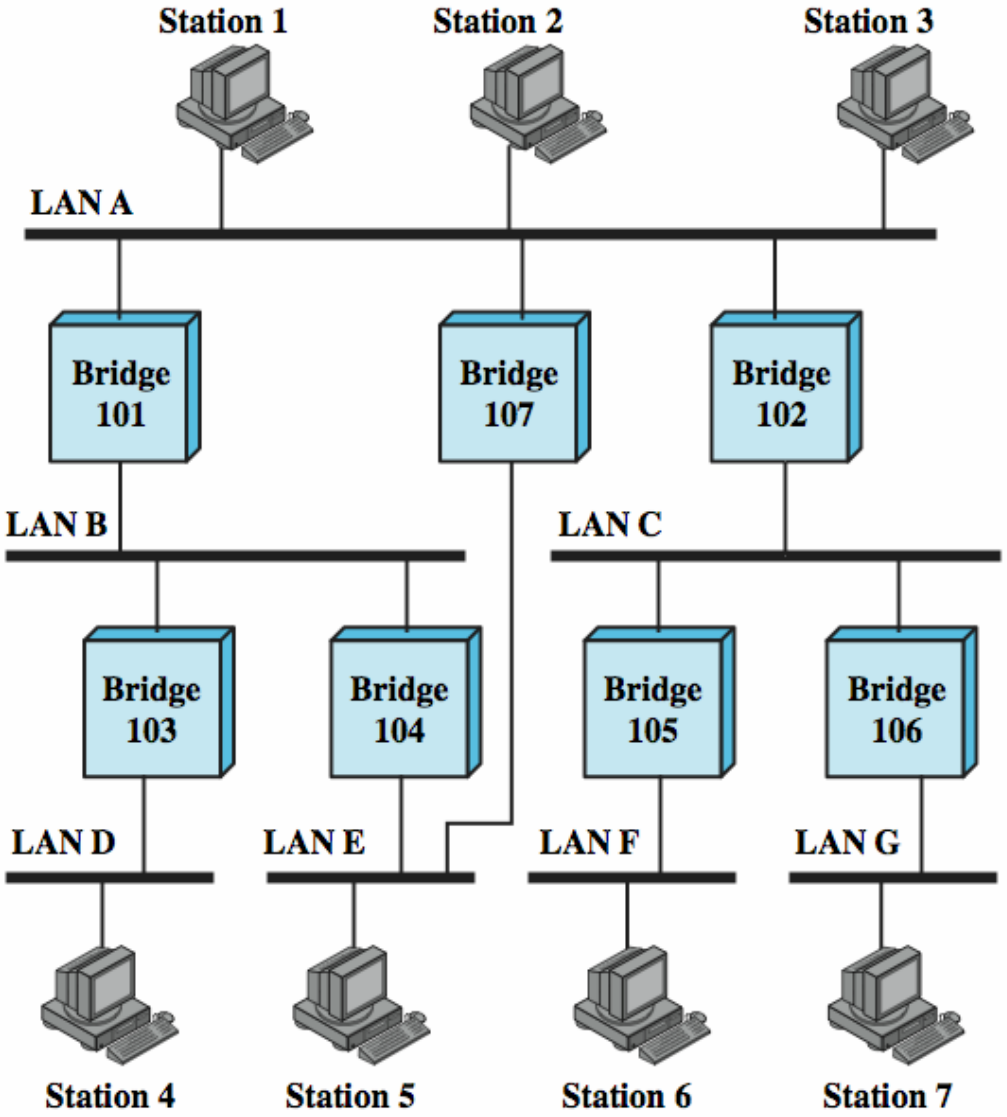
- For a frame arriving on port X
  - Search forwarding database to see if MAC address is listed for any port except X
  - If address not found, forward to all ports except X
  - If address listed for port Y, check port Y for blocking or forwarding state
  - If not blocked, transmit frame through port Y

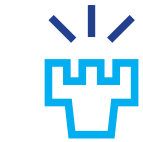


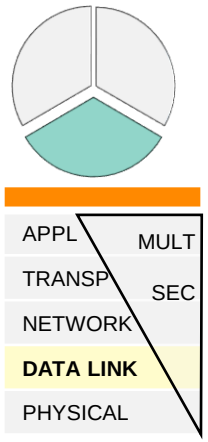


Bridge: Alternative routes

- Complex large LANs need alternative routes
  - Load balancing and fault tolerance
- Bridge must decide whether to forward frame
- Bridge must decide LAN to forward frame to

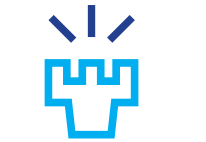


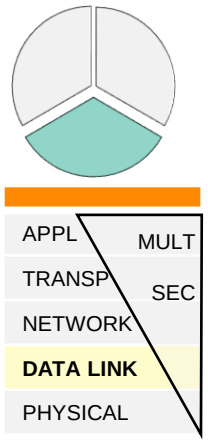




Bridge: Fixed routing

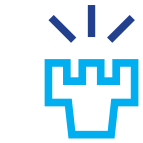
- For each source-destination pair of LANs
- Done in configuration
- Usually least hop route
- Only changed when topology changes
- Was widely used but provided limited flexibility

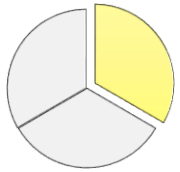


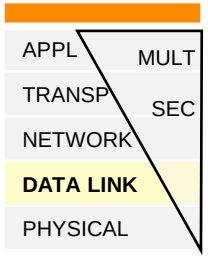


Bridge: Spanning tree routing

- **Bridge automatically develops routing table (spanning tree)**
- **Automatically updates routing table in response to changes in the network**
- **Three mechanisms**
  - Frame forwarding
  - Address learning
  - Loop resolution



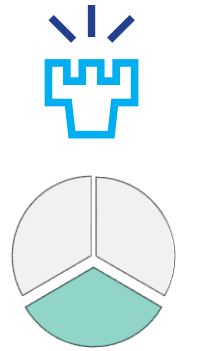




Bridge: Address learning (self learning)

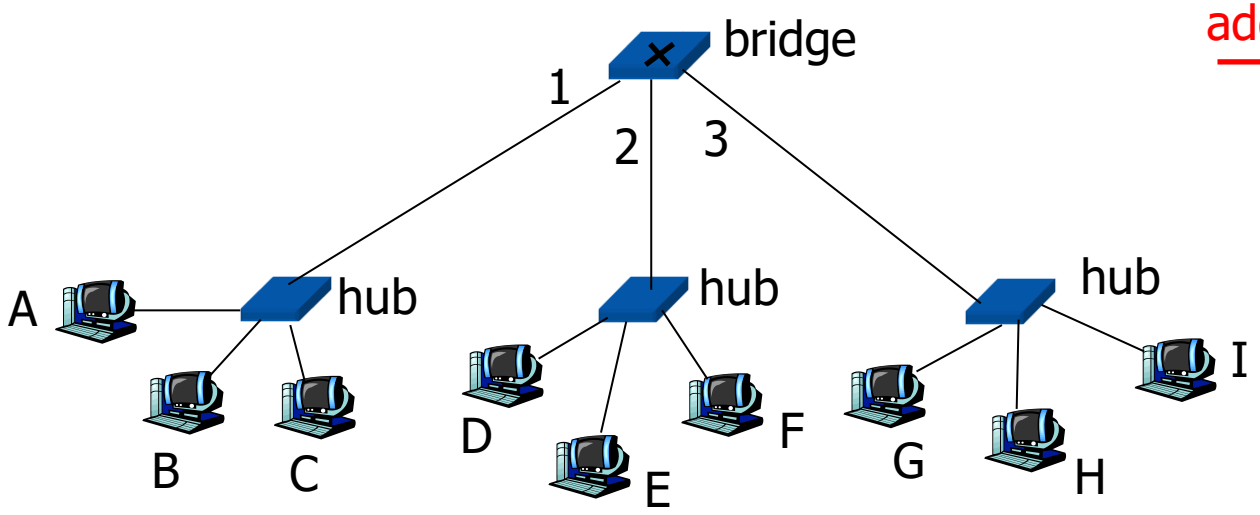
- Bridge has a **forwarding table**
  - Entry in forwarding table:
    <MAC address, Interface, Time-to-Live timestamp>
  - Stale entries in table dropped (TTL can be 60 min)

- Bridge ***learns*** which hosts can be reached through which interfaces
  - When frame received, bridge “learns” location of sender (incoming LAN segment)
  - Records sender/location pair in bridge table



Bridge: Learning example (1)

Suppose C sends frame to D

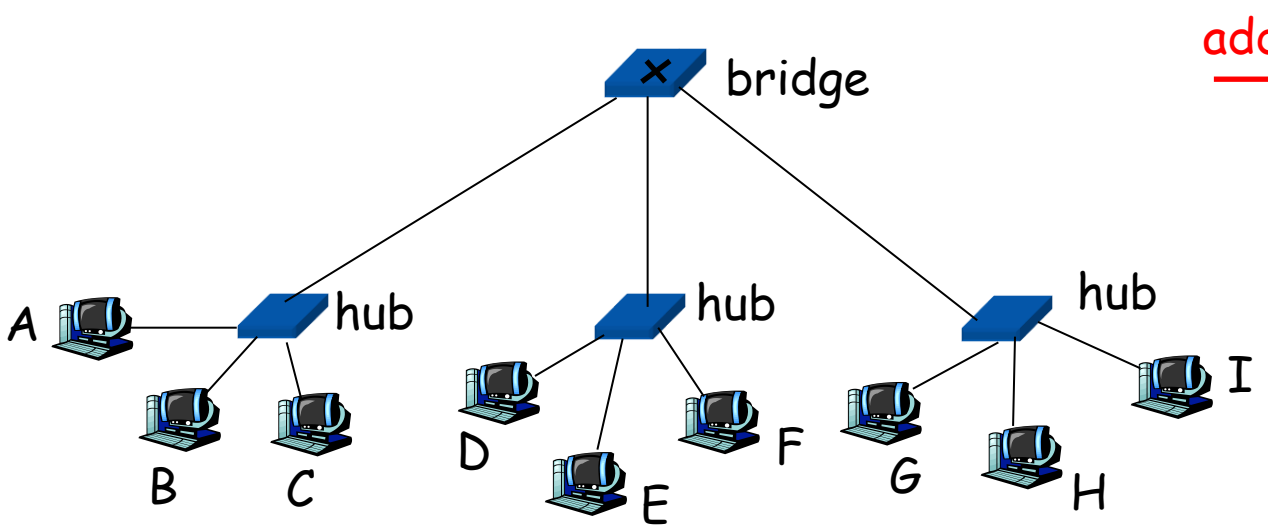


address	interface
A	1
B	1
E	2
G	3

1. Bridge receives a frame addressed to D from C
   - Bridge checks its forwarding table and notices that C is on interface 1
   - Because D is not in forwarding table, bridge forwards frame into interfaces 2 and 3 (flooding)
2. D receives the flooded frame

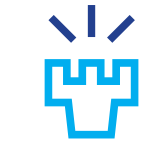
Bridge: Learning example (2)

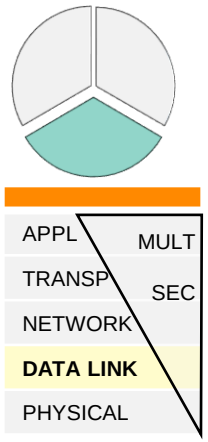
Suppose D replies back with frame to C



address	interface
A	1
B	1
E	2
G	3
C	1

- Bridge receives frame from D
  - Notes in forwarding table that D is on interface 2
  - Because C is in table, bridge forwards frame only to interface 1
- C receives the frame

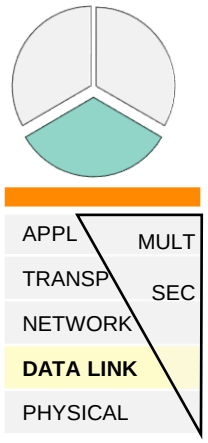


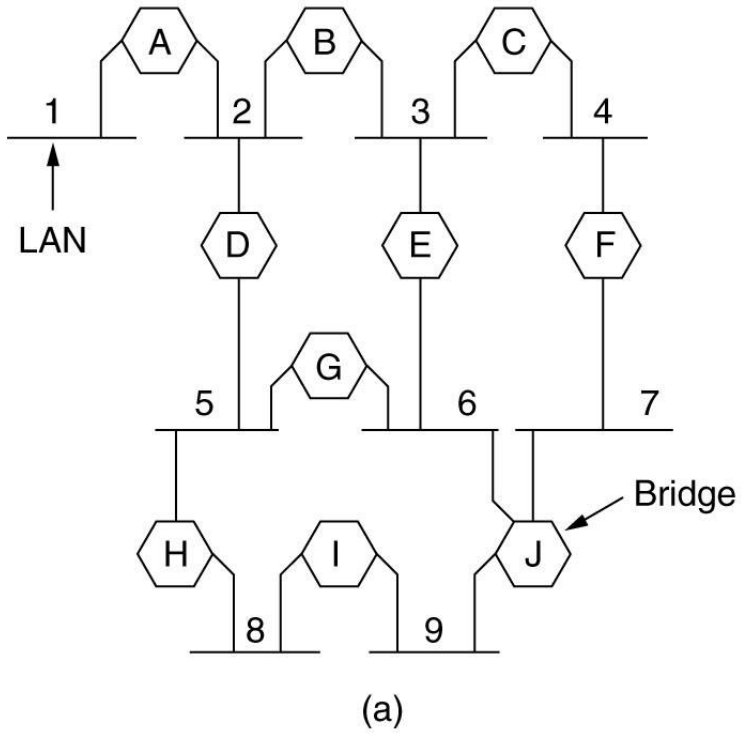


Bridge: Loop resolution

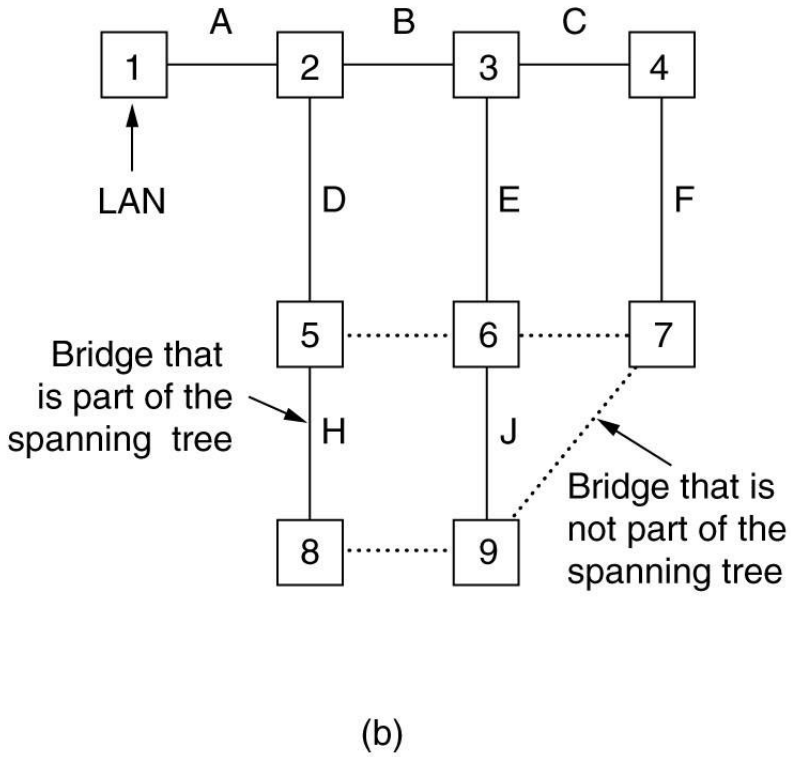
- **Address learning works for tree layout**
- **In general graphs (LANs) have loops**
- **For any connected graph there is a spanning tree maintaining connectivity with no closed loops**
  - A subset of interfaces is disabled
- **IEEE 802.1D spanning tree protocol finds this**
  - Each bridge assigned unique identifier
  - Exchange info between bridges to find spanning tree
  - Automatically updated whenever topology changes

Bridge: Spanning tree

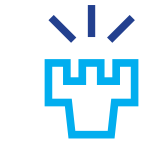


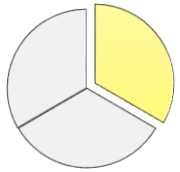


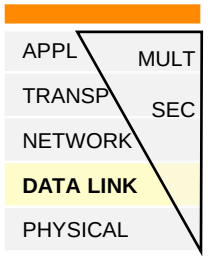
Interconnected LANs



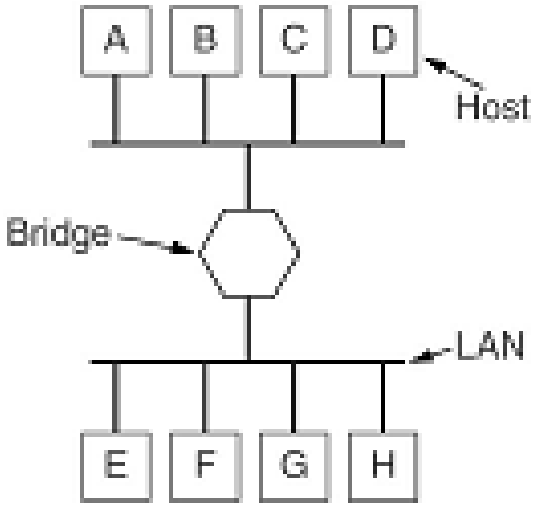
A spanning tree covering the LANs. The dotted lines are not part of the spanning tree.

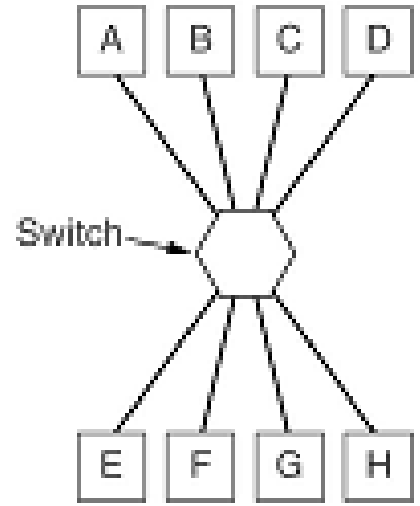






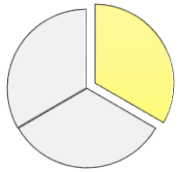
Switch



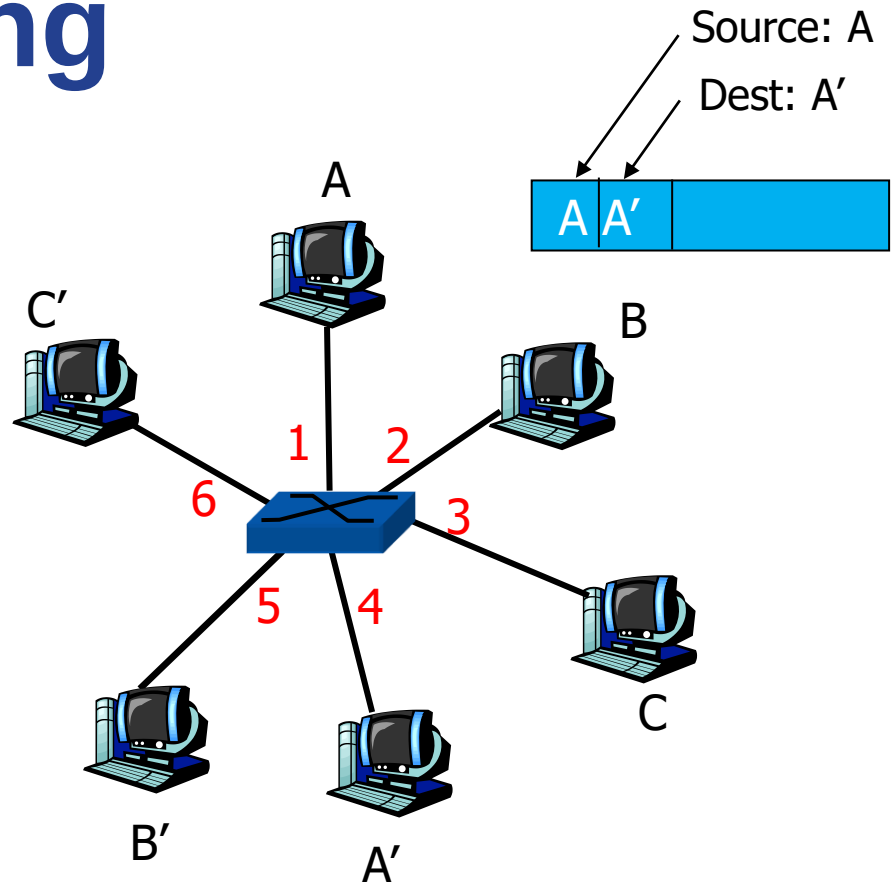


- (Just like bridge, but ...)
- Buffers/switches/forwards all frames it receives
- Full-duplex connections between switch and stations/bridges
  - All stations can transmit and receive frames continuosly
  - No collisions, no CSMA/CD needed
  - Multiplies LAN capacity
- Combinations of shared/dedicated, 10/100/1000/10000 Mbps interfaces

Switch: Self learning

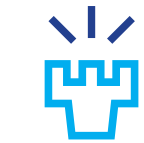


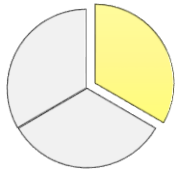
- Switch learns which hosts can be reached through which interfaces
  - When frame received, switch “learns” location of sender (incoming interface)
  - Records sender/location pair in switch table

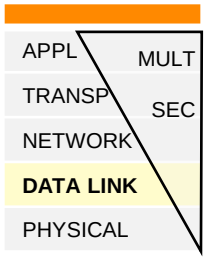


MAC addr	interface	TTL
A	1	60

Switch table (initially empty)





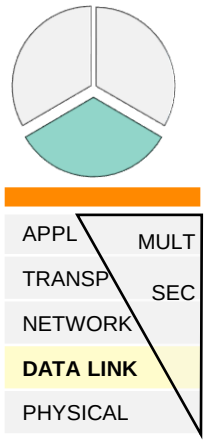


Switch: Frame filtering/forwarding

When frame received:

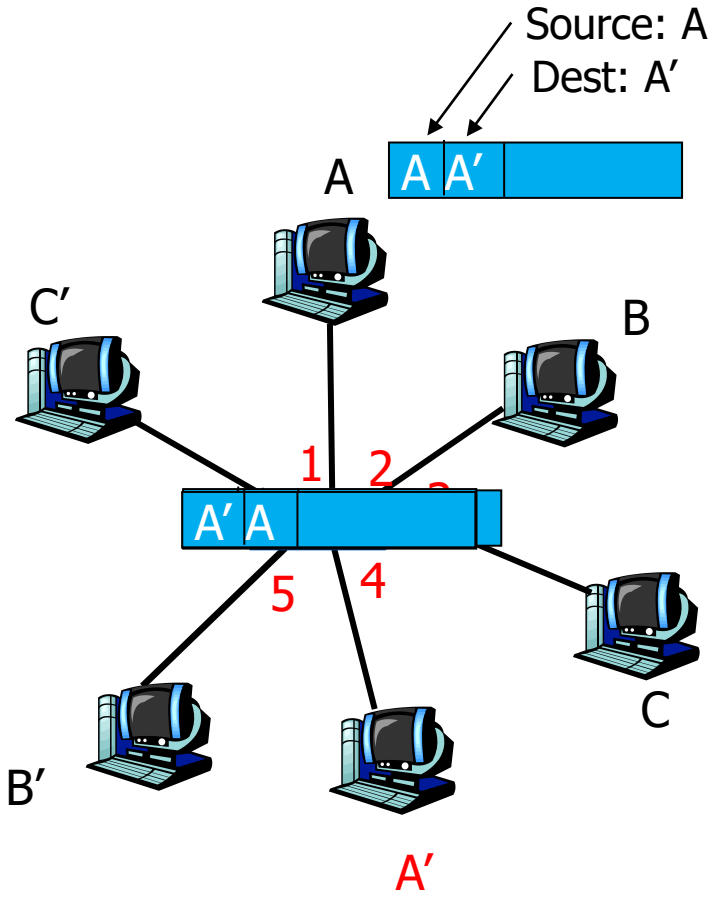
1. Record link associated with sending host
2. Index switch table using MAC destination address
3. if entry found for destination
   then {
   if destination on segment from which frame arrived
   then drop the frame
   else forward the frame on interface indicated
   }
   else flood

forward on all but the interface on which the frame arrived



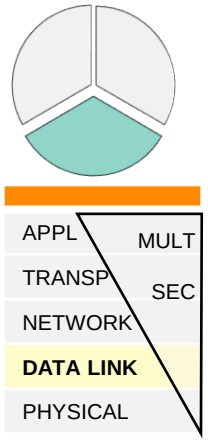
Switch: Self learning & forwarding example

- Frame destination unknown: flood
- Destination A location known: selective forwarding



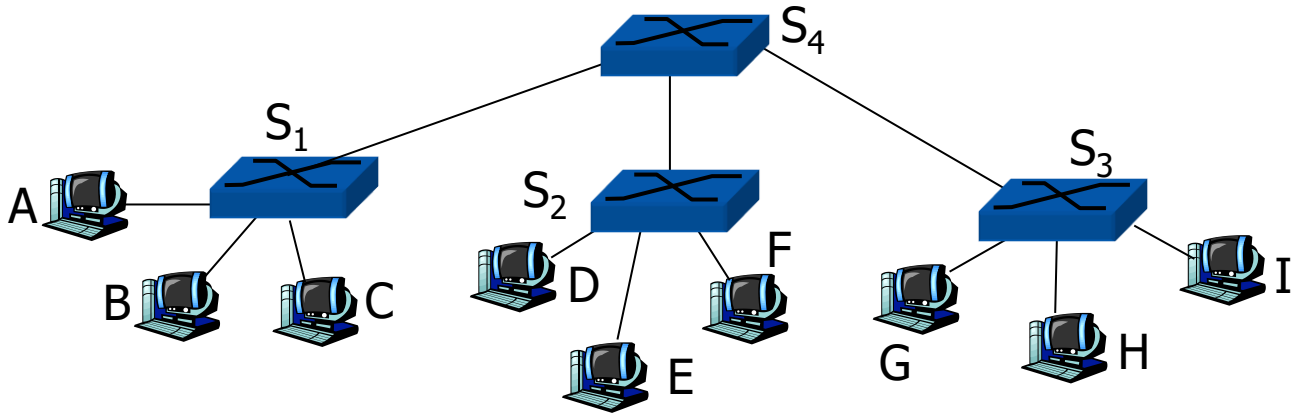
MAC addr	interface	TTL
A	1	60
A′	4	60

Switch table (initially empty)

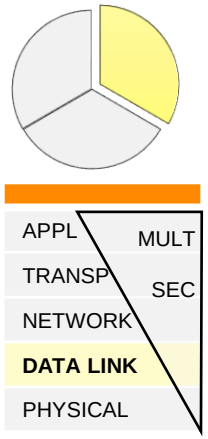


Switch: Interconnecting switches

Switches can be connected together



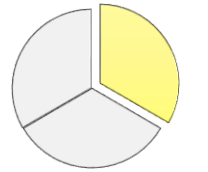
- Q: Sending from A to G: how does S_1 know to forward frame destined to F via S_4 and S_3?
- A: Self learning! (Works exactly the same as in single-switch case!)

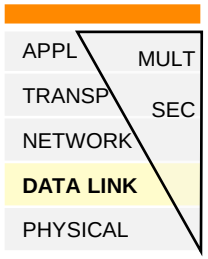


Switch: Store-and-forward vs cut-through

- **Store-and-forward** switch
  - Accepts frame on input line, buffers briefly, routes to destination port
  - Consider delay between sender and receiver
  - Better integrity

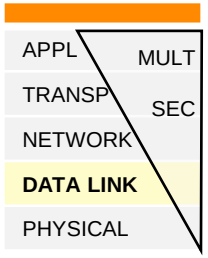
- **Cut-through** switch
  - Use destination address at beginning of frame
  - Switch begins repeating frame onto output line as soon as destination address recognized
  - Highest possible throughput
  - Risk of propagating bad frames





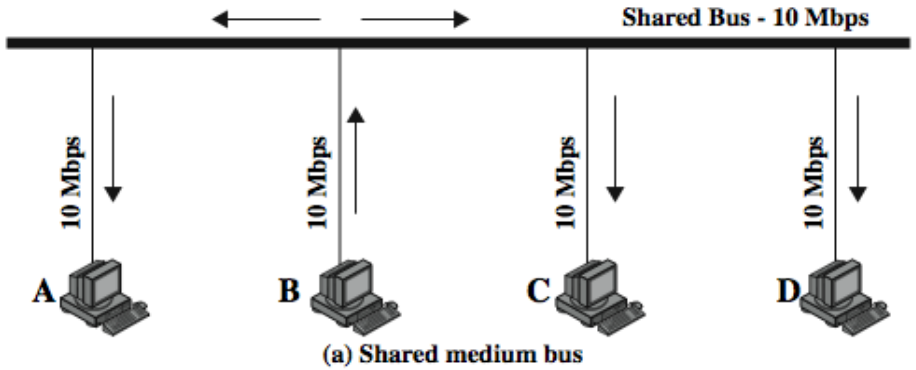
Switch: Benefits

- **Heterogeneous links**
  - Since links are isolated from each other, they can be different
  - No changes to attached devices required to convert bus LAN or hub LAN to switched LAN
    - E.g. Ethernet LANs use Ethernet MAC protocol
- **Capacity**
  - Elimination of collisions with full-duplex point-to-point links
  - Dedicated capacity equal to original LAN
    - Assuming switch has sufficient capacity to keep up with all devices
- **Scales up easily**
  - Additional devices attached to switch by increasing capacity of layer 2
- **Management and security**
  - Transparent, plug and play (self learning)
  - Automatic disconnection of malfunctioning nodes
  - Enhanced security against sniffing
    - Nodes receive only frames that have explicitly sent to them

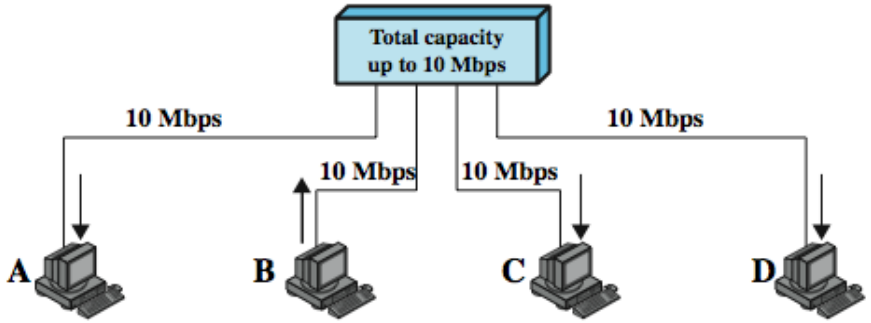


Switch: Multiplying LAN capacity

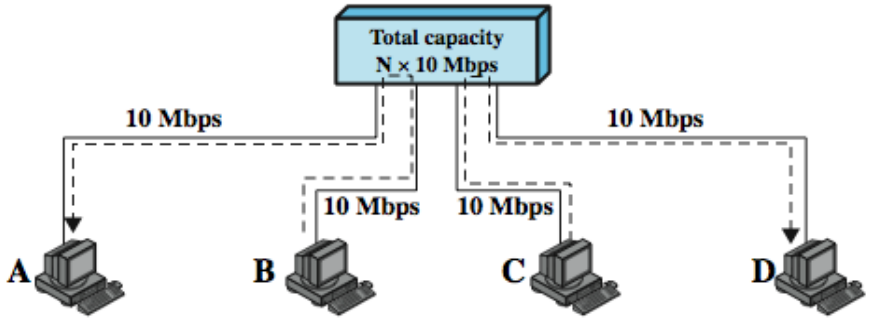
- Requires
  - Full-duplex point-to-point connections
  - Sufficient switching capacity from the switch



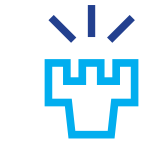
(a) Shared medium bus

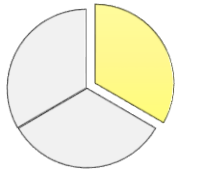


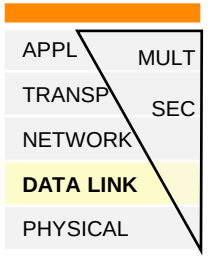
(b) Shared medium hub



(c) Layer 2 switch

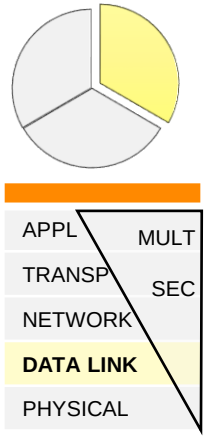






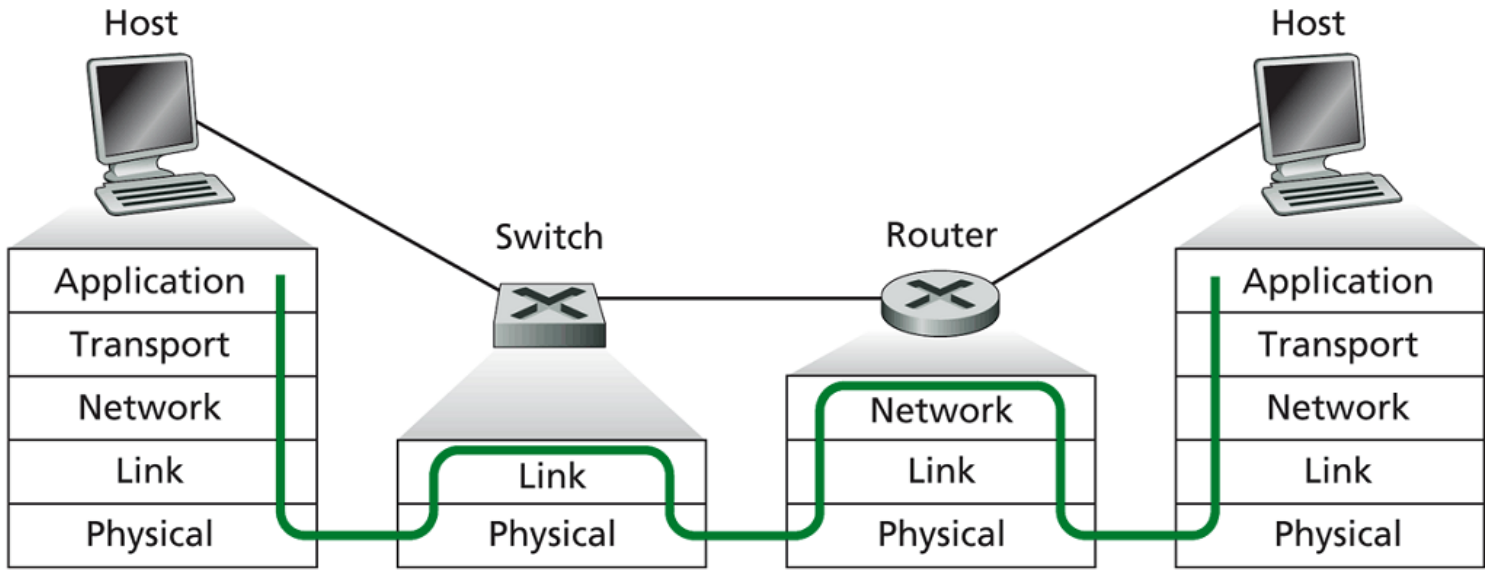
Switch vs bridge

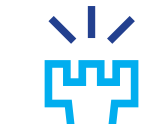
- **Differences between switches and bridges**
  - Bridge frame handling done in software
  - Switch performs frame forwarding in hardware
  - Bridge analyzes and forwards one frame at a time
  - Switch can handle multiple frames at a time
  - Bridge uses only store-and-forward operation
  - Switch can use also cut-through operation
- **Thus, bridge has suffered commercially**
  - Bridges are no longer sold as such
  - New installations typically use switches with bridge functionality than bridges

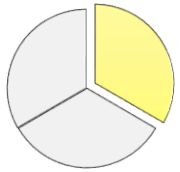


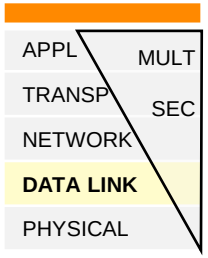
Switch vs router

- Routers are network layer (L3) devices (examine datagram headers)
- Switches are link layer (L2) devices (examine frame headers)
- Routers maintain routing tables, implement routing algorithms
- Switches maintain switch tables, implement filtering, learning algorithms



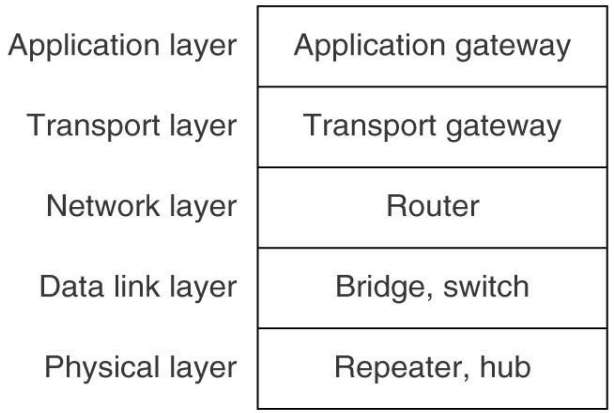




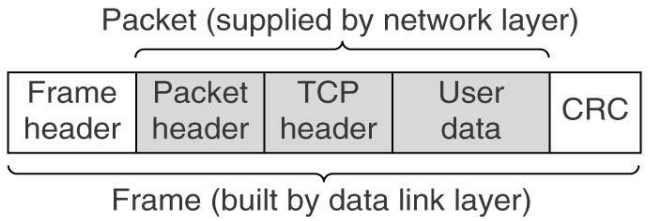


Summary comparison of popular LAN interconnection devices

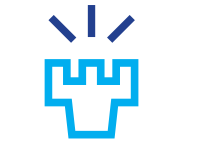
	Hubs	Switches	Routers
Traffic isolation	No	Yes	Yes
Plug and play	Yes	Yes	No
Optimal routing	No	No	Yes
Cut-through	Yes	Yes	No
	(layer 1)	(layer 2)	(layer 3)

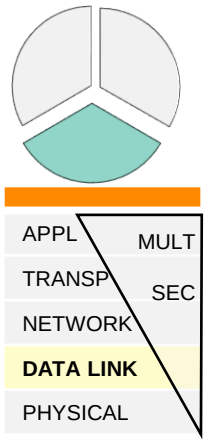


(a)



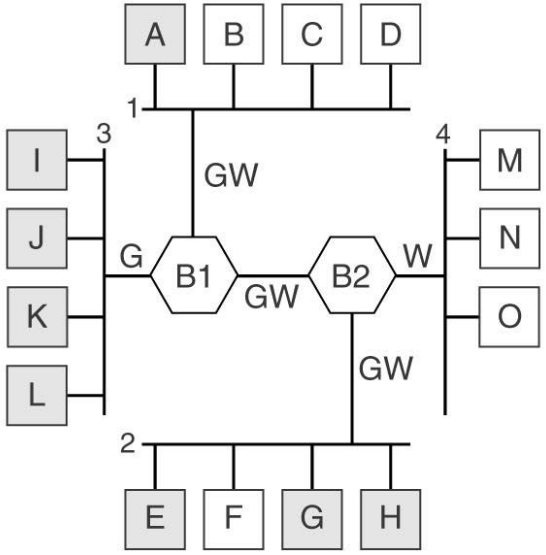
(b)





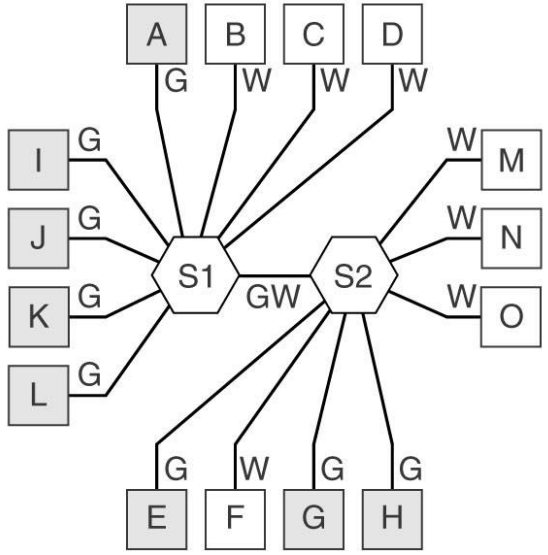
VLAN (Virtual LAN)

- Logical subnetwork that can group together a collection of devices from different physical LANs
- Flexibility and traffic isolation by LAN virtualization
- Based on VLAN-aware switches (earlier also with bridges)
- Different VLAN's are often named by colors/numbers



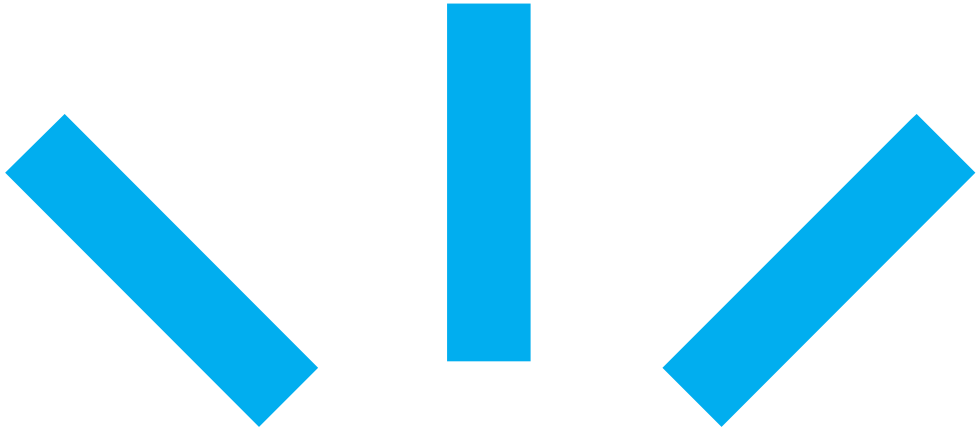
(a)

(a) Four physical LANs organized into two VLANs, gray and white, by two bridges

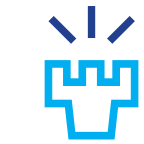


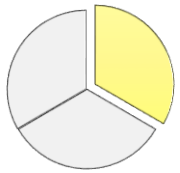
(b)

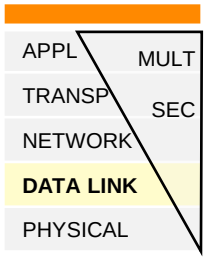
(b) The same 15 machines organized into two VLANs by two switches



Ethernet (IEEE 802.3)

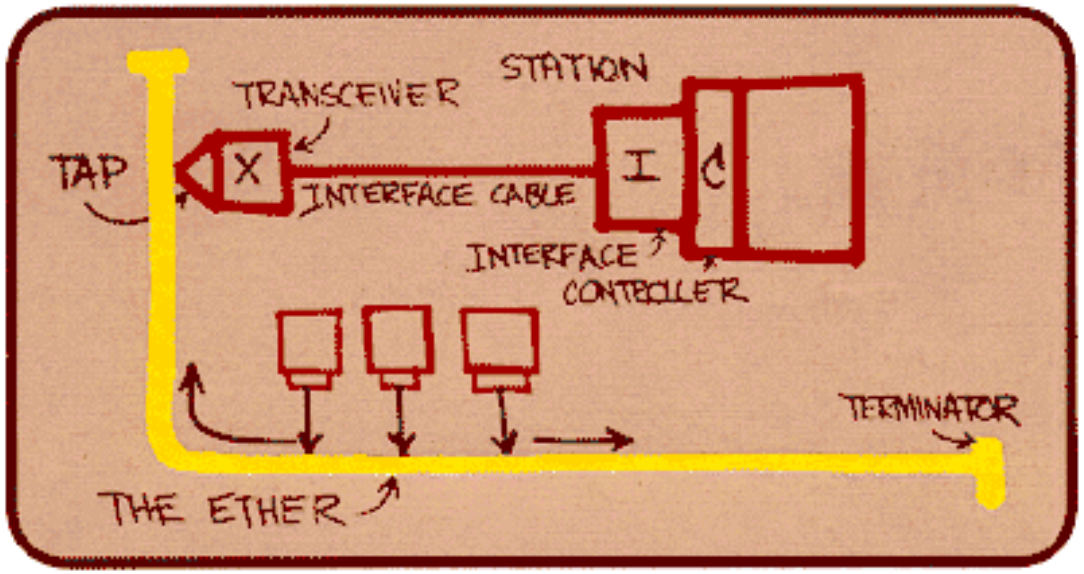




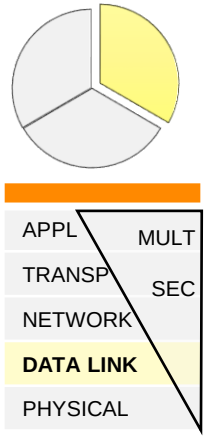


Ethernet (IEEE 802.3)

- De-facto (dominant) wired LAN technology
- First widely used LAN technology
- Simpler, cheaper than token LANs and ATM
- Kept up with speed race: 10 Mbps – 100 Gbps



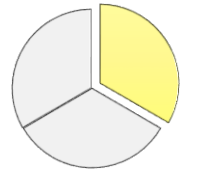
Metcalfe's original Ethernet sketch from 1973



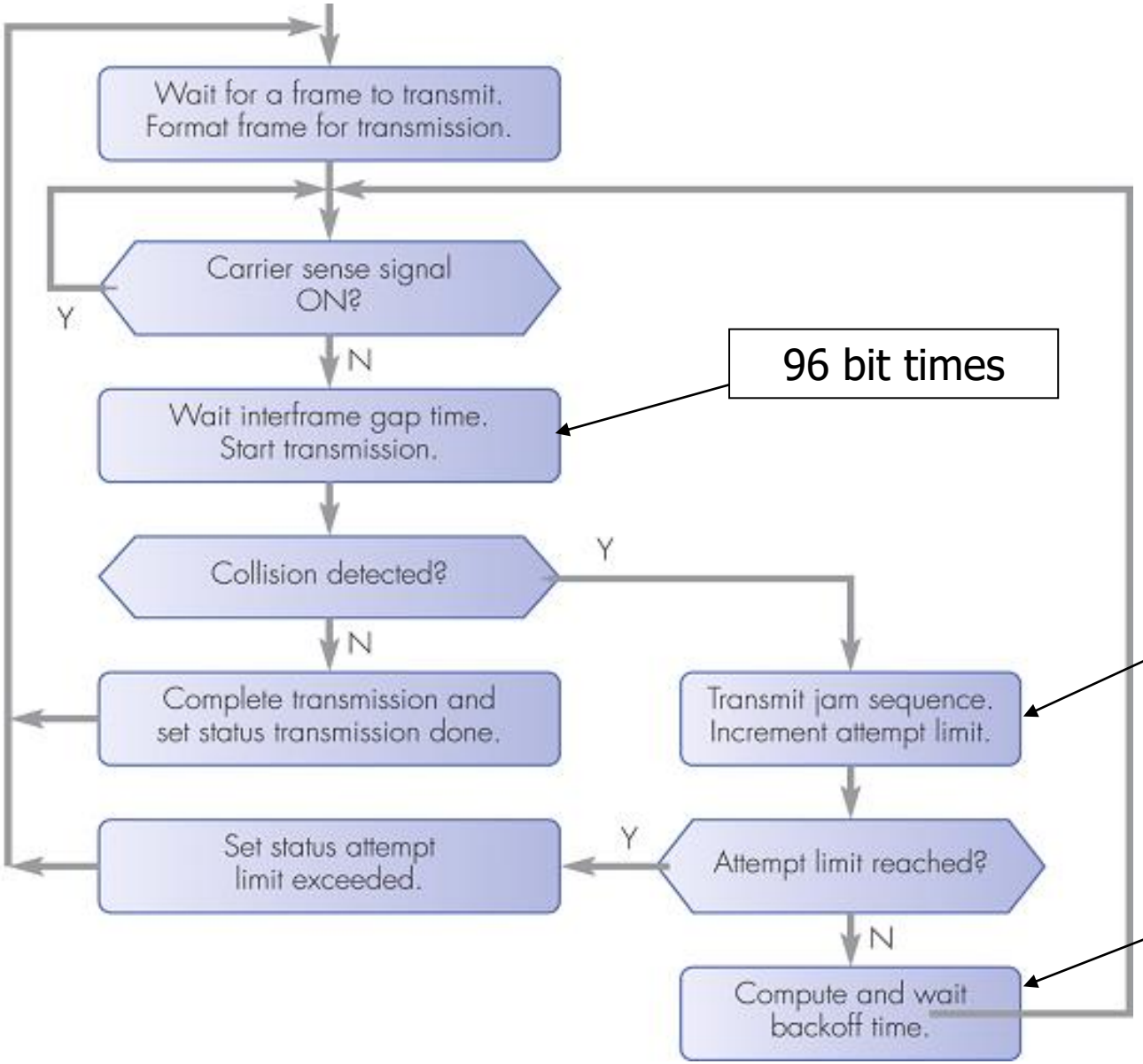
Ethernet: CSMA/CD

- No slots
- Adapter does not transmit if it senses that some other adapter is transmitting (carrier sense)
- Transmitting adapter aborts when it senses that another adapter is transmitting (collision detection)
- Before attempting a retransmission, adapter waits a random time (random access)

Ethernet: CSMA/CD: Transmit



APPL
MULT
TRANSP
SEC
NETWORK
DATA LINK
PHYSICAL



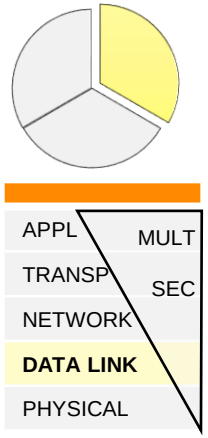
96 bit times

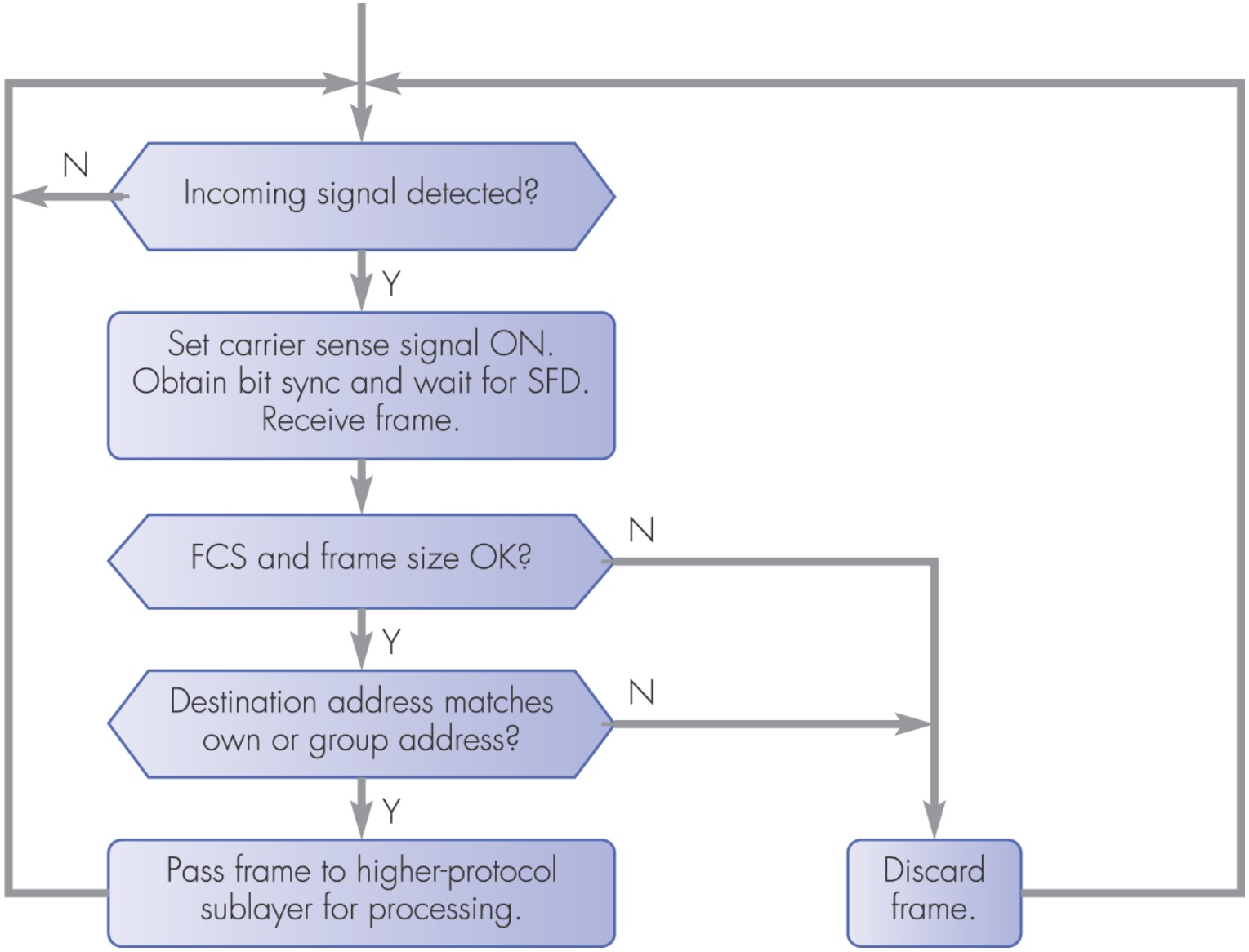
Jam signal
(32-48 random bits)

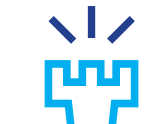
make sure all other stations are aware of collision

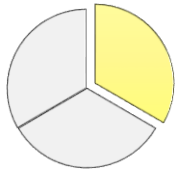
Exponential backoff algorithm

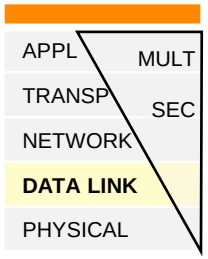
Ethernet: CSMA/CD: Receive





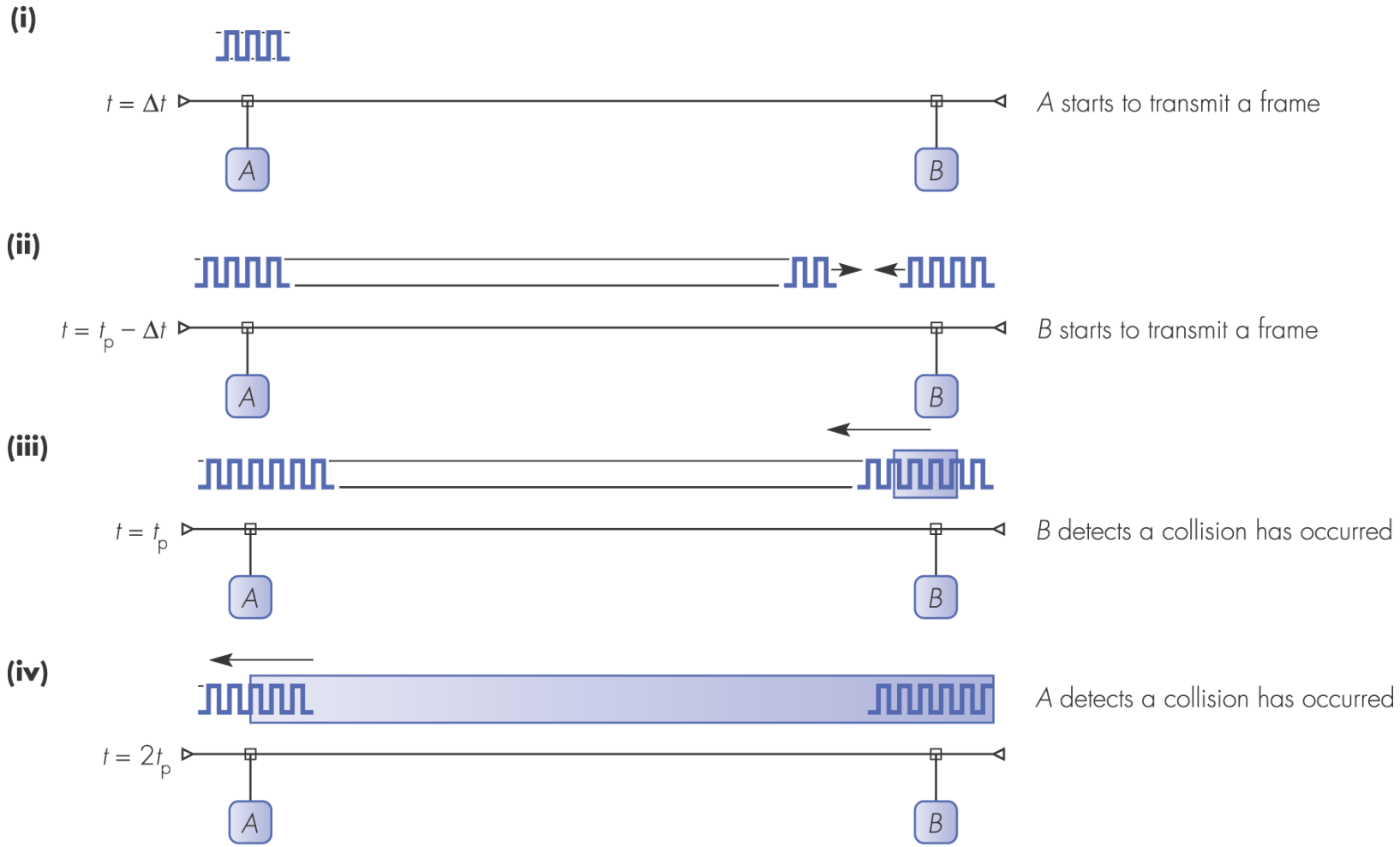




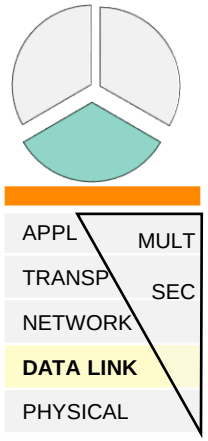


Ethernet: CSMA/CD: Collision detection

Collision detection can take $2t_p$ (round-trip propagation delay)



t_p = (worst-case) transmission propagation (path) delay



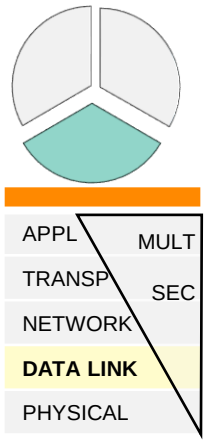
Ethernet: CSMA/CD:

Collision detection vs Segment length vs Frame size

- 10 Mbps Ethernet LAN
  - Maximum segment length is 2500 m (5x500 m) with 4 repeaters (amplify analog signals between cable segments)
- Assume signal propagation speed of $2x10^8$ m/s

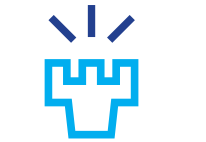
  →The round-trip propagation delay is 25 μs
- Each repeater adds few μs to delay
  - The worst-case total delay is estimated at 50 μs
  - At 10 Mbps 500 bits needs to be transmitted to cover 50 μs
  - Add 12 bits as safety margin

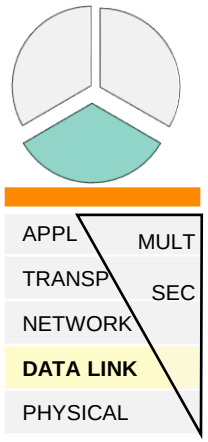
  → 512 bit minimum frame size for reliable collision detection
- Also, upon sensing a collision station sends jam signal (32-48 random bits) to make sure all other stations detect the collision.
- Longer segments and higher bit rates → bigger frames!



Ethernet: CSMA/CD:
Exponential backoff algorithm

- Adapt retransmission attempts to estimated current load (heavy load: random wait will be longer)
- After the m^{th} collision, adapter chooses a K at random from $\{0,1,2,\ldots,2^{\min(m,10)}-1\}$ and waits for K· 512 bit times
  - First collision: choose K from {0,1}; delay is K· 512 bit times
  - After second collision (m=2): choose K from {0,1,2,3}…
  - After ten collisions (m=10), choose K from {0,1,2,3,4,…,1023}
- After 16 collisions, give up transmitting
- **Bit time:** 0.1 μs for 10 Mbps Ethernet; for K=1023 wait time is about 50 ms





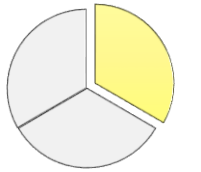
Ethernet: CSMA/CD: Efficiency

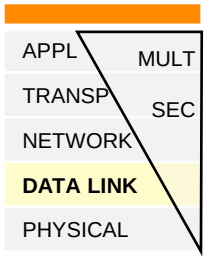
- T_p = max propagation delay between 2 nodes in LAN
- T_t = time to transmit maximum-size frame

- It can be shown (Lam 1980, Bertsekas 1991) that

$$\text{efficiency} = \frac{1}{1 + 5T_p / T_t}$$

- $T_p \rightarrow 0$, efficiency $\rightarrow 1$
- $T_t \rightarrow$ infinity, efficiency $\rightarrow 1$
- In practice efficiency 80-95%
- Much better than ALOHA, but still decentralized, simple, and cheap



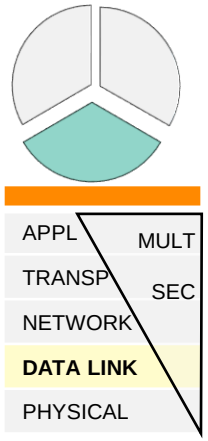


Ethernet: Connectionless, unreliable service

- **Connectionless:** No handshaking between sending and receiving adapter

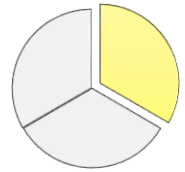
- **Unreliable:** Receiving adapter does not send ACKs or NAKs to sending adapter (for efficiency reasons)
  - Stream of datagrams passed to network layer can have gaps
  - Gaps will be filled if application is using reliable transport layer protocol (e.g. TCP)
  - Otherwise, application will see the gaps

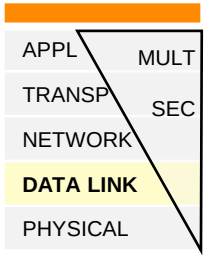
Ethernet: Technology (PHY) evolution (1)



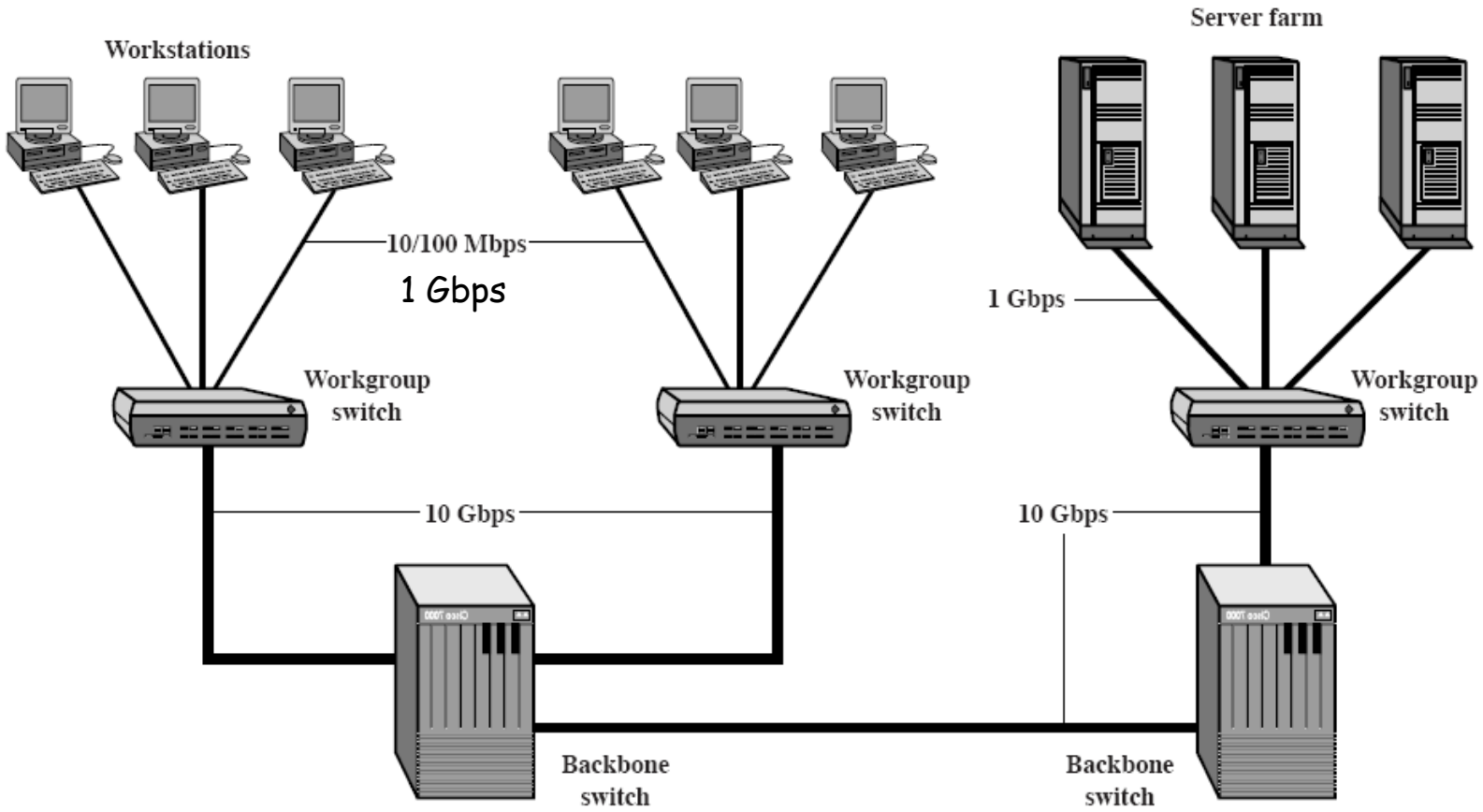
	Ethernet Type	Bandwidth	Cable Type	Maximum Distance
1990	10Base-T	10Mbps	Cat 3/Cat 5 UTP	100m
1995	100Base-TX	100Mbps	Cat 5 UTP	100m
	100Base-TX	200Mbps	Cat 5 UTP	100m
	100Base-FX	100Mbps	Multi-mode fiber	400m
	100Base-FX	200Mbps	Multi-mode fiber	2Km
1998	1000Base-T	1Gbps	Cat 5e UTP	100m
	1000Base-TX	1Gbps	Cat 6 UTP	100m
	1000Base-SX	1Gbps	Multi-mode fiber	550m
	1000Base-LX	1Gbps	Single-mode fiber	2Km
2002	10GBase-T	10Gbps	Cat 6a/Cat 7 UTP	100m
	10GBase-LX	10Gbps	Multi-mode fiber	100m
	10GBase-LX	10Gbp	Single-mode fiber	10Km

2006	40 Gigabit Ethernet	100 Gigabit Ethernet
At least 1m backplane	40GBASE-KR4	
At least 7m copper cable	40GBASE-CR4	100GBASE-CR10
At least 100m OM3 MMF	40GBASE-SR4	100GBASE-SR10
At least 150m OM4 MMF	40GBASE-SR4	100GBASE-SR10
At least 10km SMF	40GBASE-LR4	100GBASE-LR4
At least 40km SMF		100GBASE-ER4

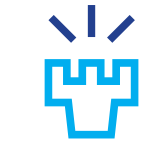


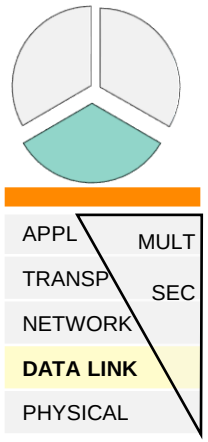


Modern LAN deployment with switched Ethernet



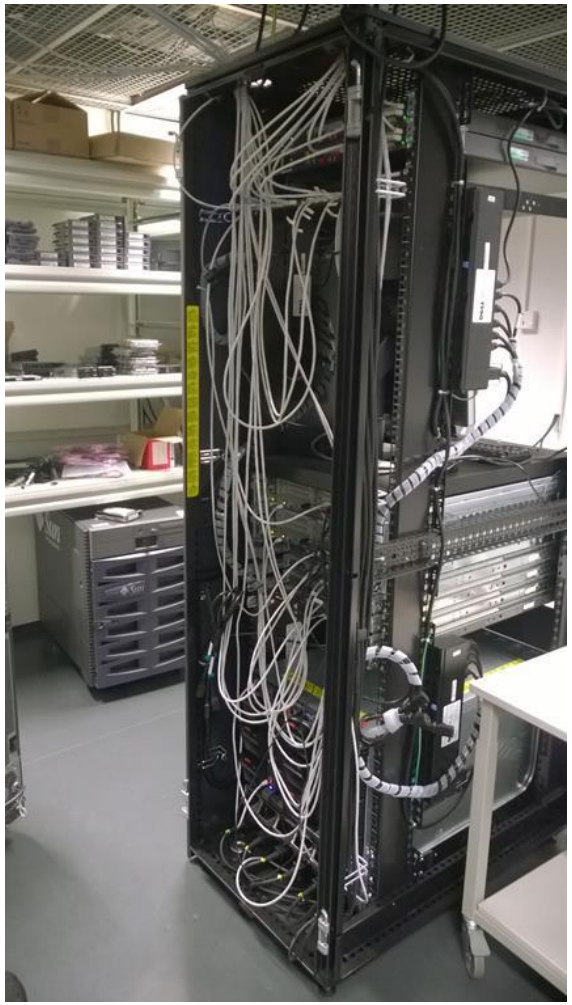
- Full-duplex links between nodes and Ethernet switches
- No collisions, no CSMA/CD, all lines are buffered
- Autoconfiguration
- Signal strength determines cable length





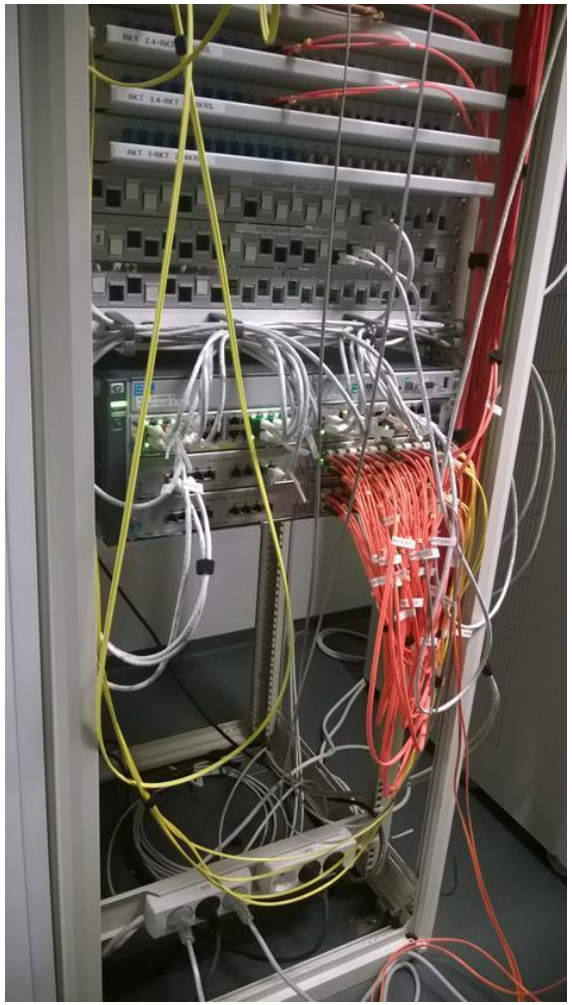
ITEE LAN: Central server room

File servers

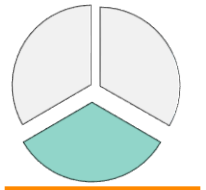


Local switch

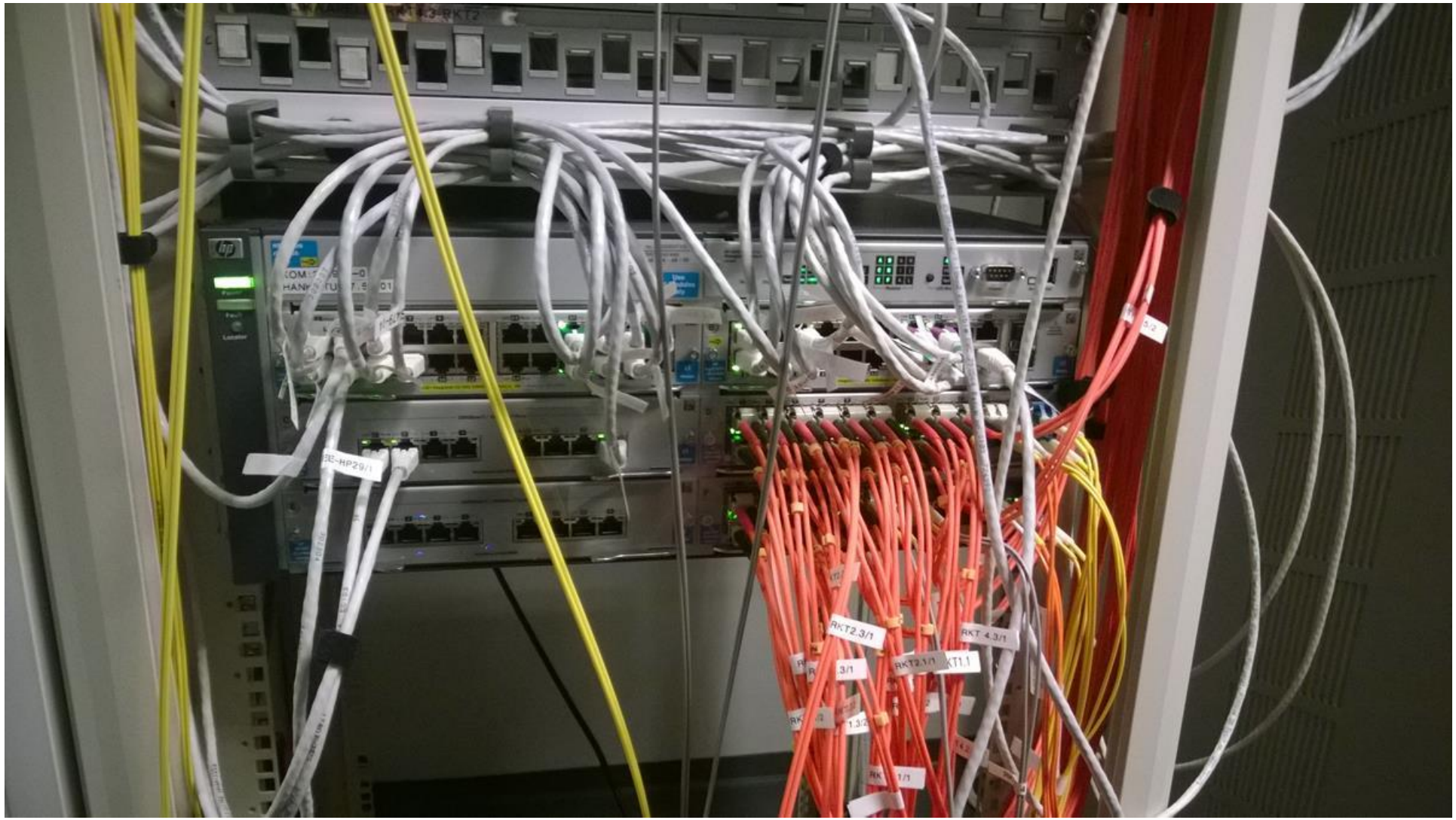
Computing servers

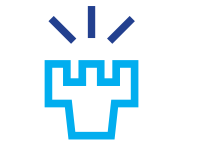


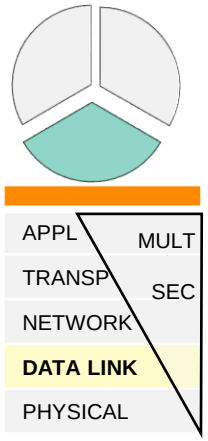
Central switch



ITEE LAN: Central Ethernet switch HP 5406 zl (1)



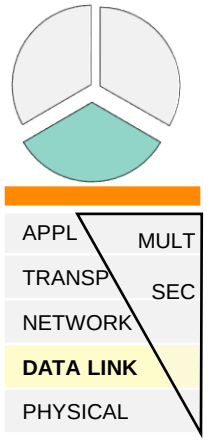




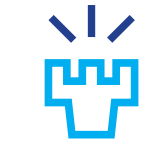
ITEE LAN: Central Ethernet switch HP 5406 zl (2)

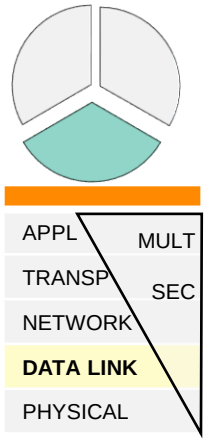
- L2-L4 device
  - Can manage packet traffic using L4 protocol headers!
- Port configuration
  - 16 x 10 Gbps UTP
  - 2 x 10 Gbps optic fiber
  - 44 x 1 Gbps UTP
  - 48 x 1 Gbps optic fiber
- Switch fabric capacity: 379.2 Gbps
- Throughput (forwarding rate): up to 282.1 Mbps
- Routing table size: 10000 entries (IPv4), 5000 entries (IPv6)
- MAC address table size: 64000 entries
- Price: ~16000 EUR (particular configuration)

ITEE LAN: Workgroup switches Dell N2048



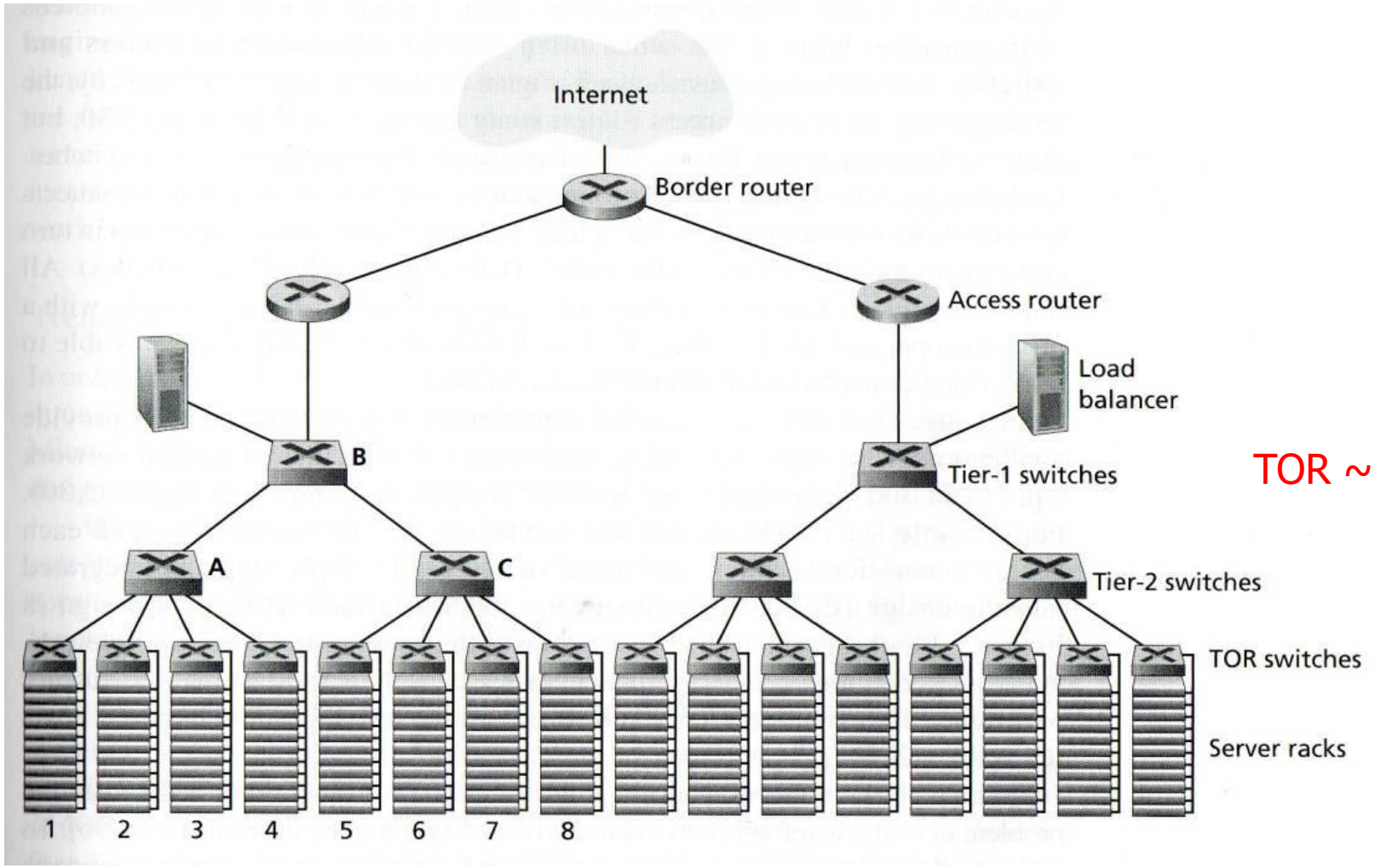
- L2-L3 device
- Port configuration
  - 48 x 10/100/1000 Mbps UTP
  - 2 x 10 Gbps optic fiber
- Switch fabric capacity: 220 Gbps
- Throughput (forwarding rate): up to 164 Mpps
- Routing table size: 256 entries (IPv4),
  256 entries (IPv6)
- MAC address table size: 8192 entries
- Price: ~1300 EUR (base unit)



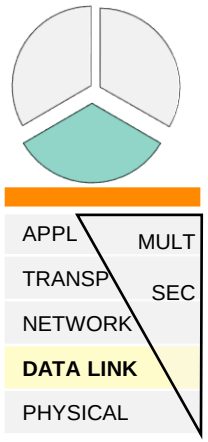


Data center networking (1)

Data center network with a hierarchical topology

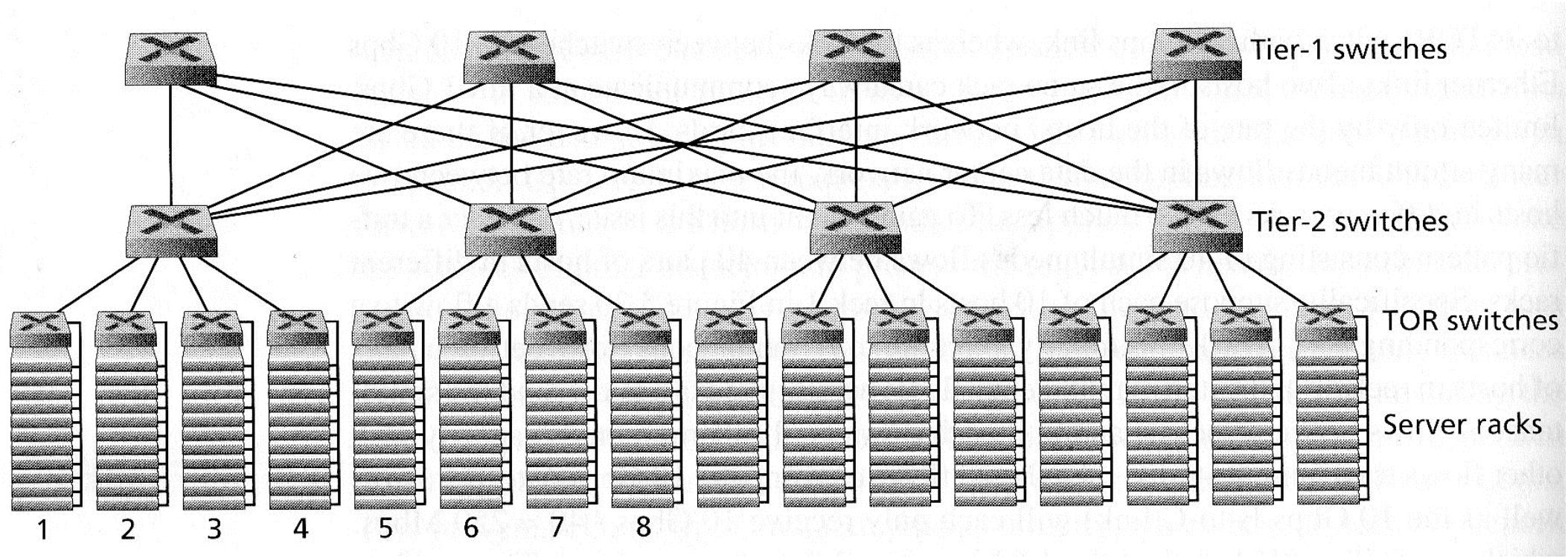


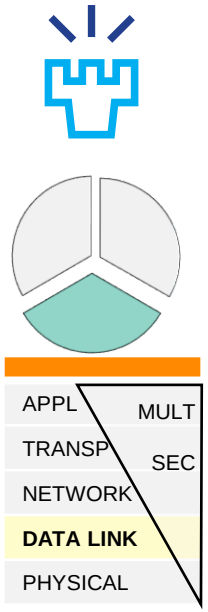
TOR ~ ”Top Of Rack”



Data center networking (2)

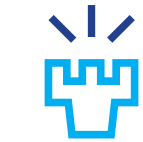
Highly connected data center network with a hierarchical topology

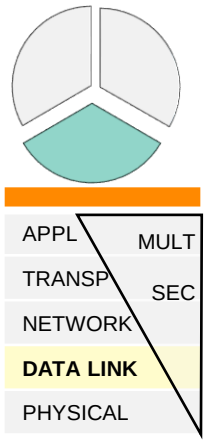




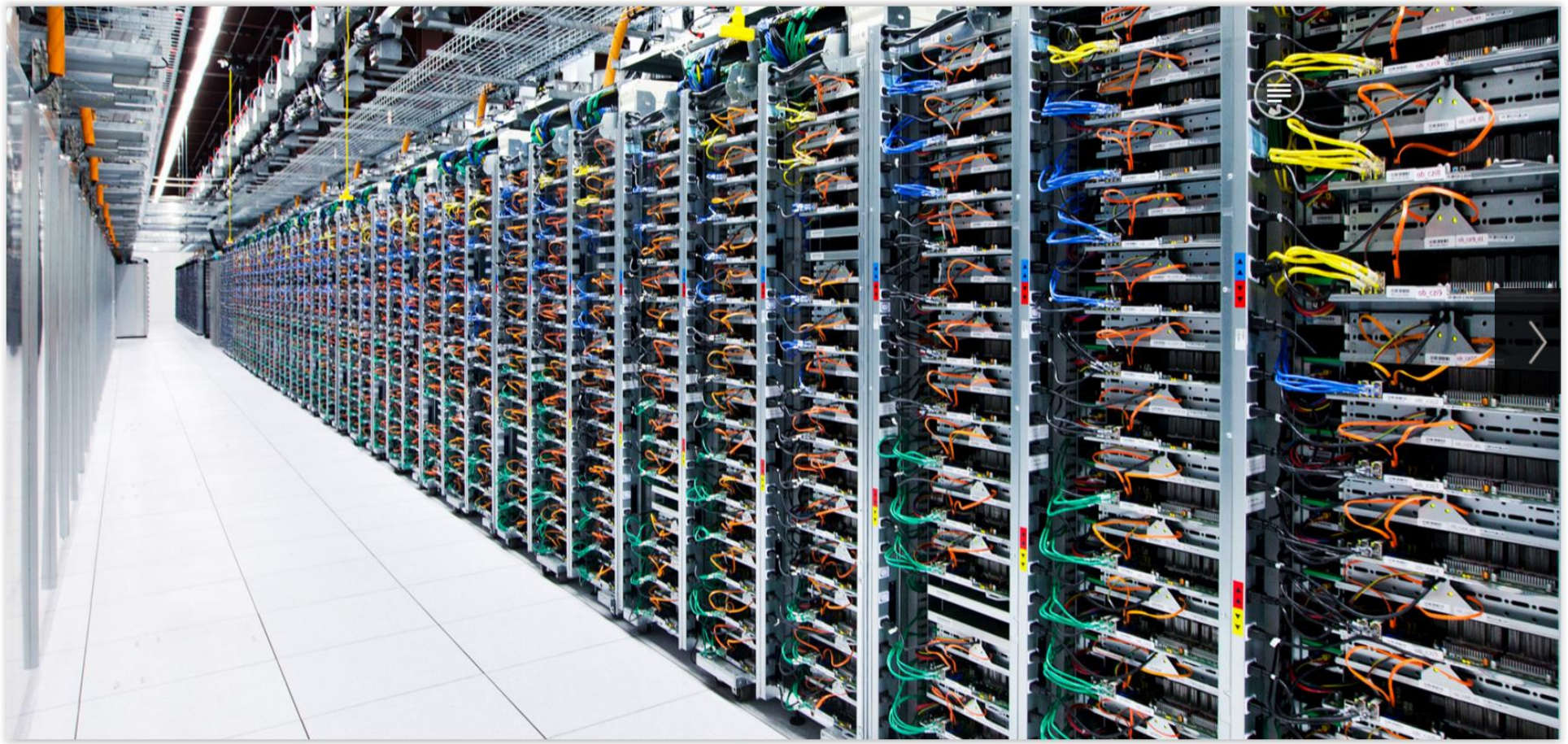
Data center networking (3)

- Data center networks are characterized by large bandwidths and very small delays
- Challenges
  - Load balancing between computing nodes
  - Providing high bandwidth for each flow in a highly connected data center network
- IEEE 802 enhancements relevant for data centers
  - IEEE 802.1Qbb: Priority-based Flow Control
  - IEEE 802.1Qau: Congestion Notification
  - IEEE 802.1aq: Shortest Path Bridging
    - Similar to IETF TRILL (Transparent Interconnect of Lots of Links)
  - IEEE 802.1Qaz: Enhanced Transmission Selection

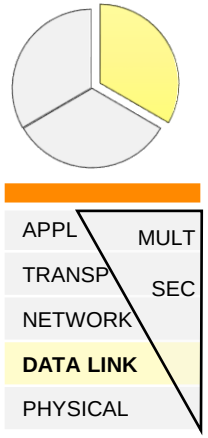




Data center networking (4)



Server racks (4 switches per rack) in Google's data center in Mayes County



Key points to remember

1. **Local Area Networks (LAN)**
   - **Topologies**
   - **Addressing**
   - **Interconnection devices (hubs, bridges, switches)**
     - Self learning
     - Link-layer switching
2. **Ethernet**
   - **Medium access control (CSMA/CD)**
   - **Frame structure**
   - **Switched Ethernet**

Thank you!