Local area networks

Katia Jaffrès-Runser and Gentian Jakllari

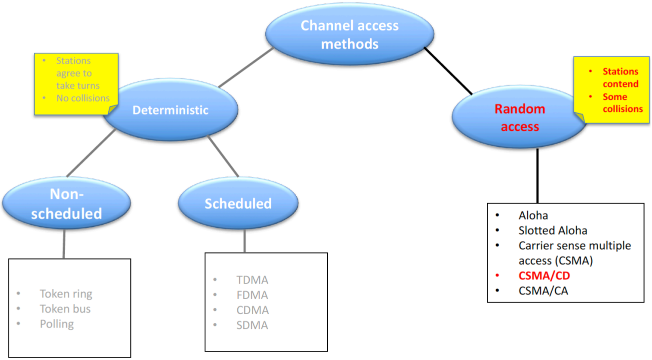
{kjr,jakllari}-at-n7.fr

Toulouse INP - ENSEEIHT

Département Sciences du Numérique
1ère annèe

2017-2018

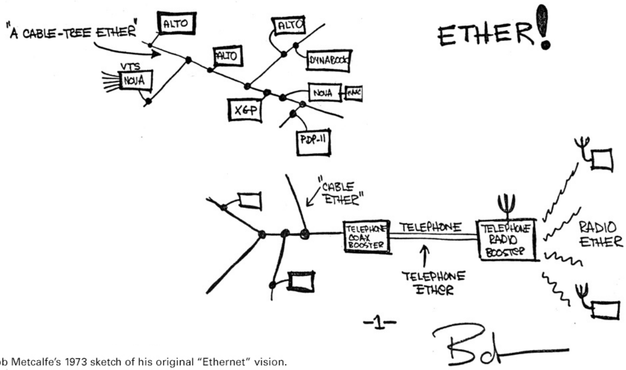
Ethernet and Switched Ethernet



Ether[1]+net

History

- Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s



Bob Metcalfe's 1973 sketch of his original "Ethernet" vision.
Image courtesy of PARC, a Xerox company

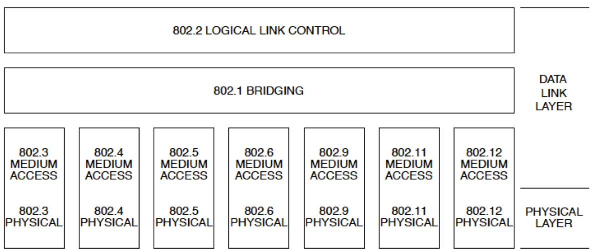
[1] In the 19th century, luminiferous aether or ether, was the postulated medium for the propagation of light

Ethernet

History

- Rooted in Aloha packet radio network
- Introduced in 1973, first products in 1980, became an IEEE standard in 1983
- Originally 2.94 Mbps, the latest reaches 100 Gbps
- Standardized, open to multiple vendors, became quickly fast and cheap
- CSMA/CD: Ethernet's medium access control (MAC) policy
  - CS = Carrier Sensing
    Send only if medium is idle
  - MA = Multiple Access
  - CD = Collision Detection
    Stop sending immediately if collision is detected

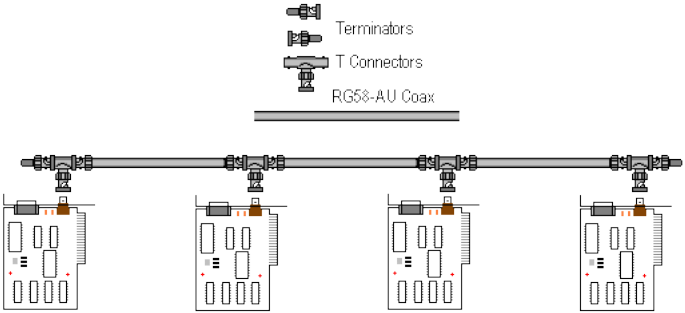
IEEE LAN Standards



- The goal of the standardization is to create vendor-neutral solutions
  - In part in response to IBM's dominance in the 70's
- IEEE LAN standards define the MAC and physical layers
- IEEE 802.3 CSMA/CD - Ethernet standard, 10 Mbps (originally 2Mbps)
- IEEE 802.3u standard for 100Mbps Ethernet
- IEEE 802.3z standard for 1Gbps Ethernet

10Base2

- "10" means 10Mbps; "2" means under 200 meters (actually 185m)
- Thin coaxial cable in a bus topology

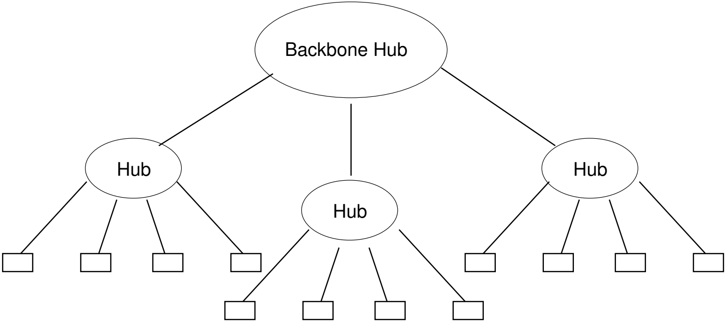


- Repeaters used to connect different segments
- Repeater repeats bits it hears on one interface to its other interface
  A physical layer device only!

Ethernet technologies

10BaseT / 100BaseT

- "10" for 10Mbps ("100" for 100Mbps); "T" for Twisted pair
- Hub(s) connected by twisted pair facilitate star topology
  - Distance of any node to the hub must be $<$ 100 meters
  - Hub repeats the bits it hears on one interface on all other interfaces
    Physical layer device as well !



Ethernet

Overview

Most popular packet-switched LAN technology

- Maximum bus length : 2500 meters divided into 500 meter segments with 4 repeaters
- Bus and start topologies are possible
  - Hosts attach to network via Ethernet transceiver or hub (or switch) detects line state (idle/busy) and sends/receives signals
  - Hubs are used to facilitate shared connections
- Network layer packets are transmitted over Ethernet after encapsulation
- A broadcast protocol by nature
  - Any signal can be received by all
  - All hosts on Ethernet are competing for access to medium

Problem: Distributed algorithm that provides a fair access

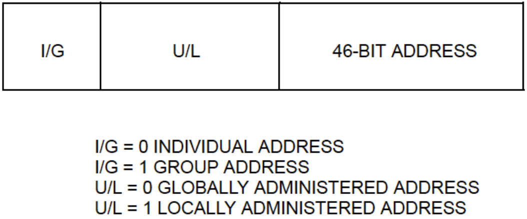
Frame format

There are two MAC frame format versions: for Ethernet II (legacy DIX commercial version) and for the IEEE802.3 standard.

Preamble	Dest Address	Src Address	Type / Length	DATA / LLC+DATA	CRC
64 bits	48 bits	48bits	16 bits	Up to 1500 bytes	32 bits

- Preamble is a sequence of 7 bytes, each set to “10101010”
  Used to synchronize the receiver before actual data is sent and delimit the start of a frame
- In Ethernet II: “Type” field is a demultiplexing key used to determine which higher layer protocol the frame should be delivered to.
- In the IEEE802.3 standard, this field holds the length of DATA in bytes. LLC header stores the equivalent of Type field then.

Addresses



- Globally unique, 48-bit unicast address assigned to each adapter
  Example: f8:1e:df:e4:9b:9b
  Each manufacturer gets their own address range
- Broadcast : all 1s
- Contrast with IP addresses ...

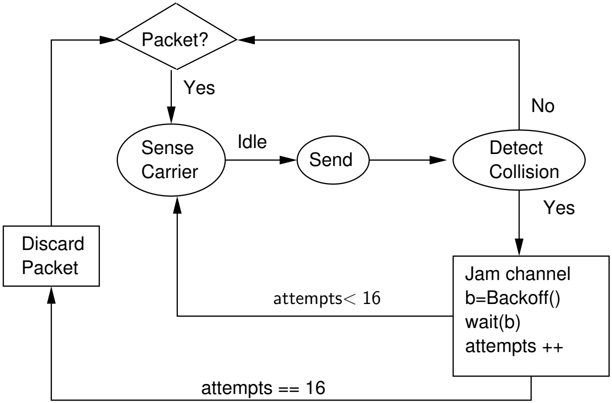
Medium Access Control (MAC) Protocol

CSMA/CD

- In ALOHA, decisions to transmit are made without paying attention to what other nodes might be doing
- In CSMA/CD, the device listens to the line before (CSMA) and **during sending (CD)**
- Perform 1-persistent CSMA:
  - If channel is idle, send message immediately.
  - If channel is busy, wait until idle and transmit packet after waiting a short Inter-Frame space (IFS) of 9.6 μs.
- AND collision detection (CD) : while sending, detect collision.
  - If collision, stop sending and jam signal: $\rightarrow$ **increases throughput compared to 1-persistent CSMA**
  - Try again later: **Back-off mechanism**

Medium Access Control (MAC) Protocol

CSMA/CD State Diagram



Medium Access Control (MAC) Protocol

Remember collisions

Collisions are caused when two stations transmit at the same time if:

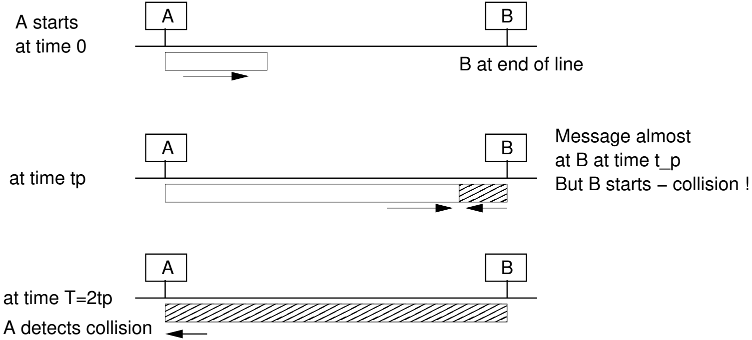
- Both stations found line to be idle
- Both had been waiting for a busy line to become idle

How can we be sure that both stations will eventually know that there was a collision?

Medium Access Control (MAC) Protocol

Vulnerability period

How can station A know that a collision took place?



- A's message collides with B's message at time t_p
- B's message reaches A at time $2t_p$.

Medium Access Control (MAC) Protocol

Ensure minimum message size

So, station A must still be transmitting (and thus listening) at time $2t_p$ to detect a collision

- IEEE 802.3 specifies a max value of $2t_p$ to be 51.2μs.
  - Relates to the maximum distance of 2500m between hosts
  - At 10Mbps it takes 0.1μs to transmit one bit, so 512 bits (64bytes) take 51.2μs to send
  - So Ethernet frames must be at least 64 bytes long
    14 bytes of header, 46 bytes of data and 4 bytes of CRC
  - Padding is used if data less 46 bytes
- Send jamming signal after collision to insure all hosts see collision
  - 48 bits signal

Truncated Exponential Backoff algorithm

Basic idea:

- If a collision is detected, each station waits until its slot of 51.2μs is finished.
- Then it backs off for a random amount of time and tries again if channel is idle.
- Backoff Time = Random() x SlotTime (51.2μs)
  - First collision: choose Random() from $\{0,1\}$
  - Second collision: choose Random() from $\{0,1,2,3\}$
  - n^{th} time: choose Random() from $\{0,\ldots,2^n-1\}$
  - Max value for Random()=1023 (i.e. $n=10$) – Truncated !
  - Give up and drop packet after several tries (usually 16)

Ethernet MAC Protocol

MAC algorithm from the receiver side

- Sender handles all access control
- Receiver simply reads frames with acceptable address:
  - Address to host
  - Broadcast address
  - Address to multicast to which host belongs to
  - All frames if host is in promiscuous mode

Ethernet MAC Protocol

Exercise

Two hosts A and B are connected to a 10Mbps Ethernet LAN.

- At $t = 0$, A sends a 520 byte frame.
- At $t = 5 \times 10^{-6}$s, B sends a 64 byte frame.

The propagation delay from A to B is 9×10^{-6}s.

1. Show on a temporal diagram the time the collision is detected by each host
2. Assume that the binary exponential backoff provides a backoff value of '1' to host A and a '0' to host B.
   Give the sequence of frame exchanges on a temporal diagram.

Fast and Gigabit Ethernet

Fast Ethernet 100Mbps

Has technology very similar to 10Mbps Ethernet

- Uses different PHY layer encoding
- Many NIC's (Network Interface Controllers) are 10/100 capable

Gigabit Ethernet 1000Mbps

- Compatible with lower speeds
- Uses standard framing and CSMA/CD
- Distances are very limited
- Typically used for backbone and inter-router connectivity

Experiences with Ethernet

Ethernet works best under light loads

Utilization over 30 % is considered heavy
Network capacity is wasted by collisions

- Most networks are limited to about 200 hosts
  Specification allows for up to 1024
- Most networks are much shorters
- Transport layer flow control helps reduce load (number of back to back packets)
- Ethernet is fast, inexpensive and easy to administer

IEEE 802.3 MAC Parameters

Parameters	MAC data rate			
	Up to and including 100 Mb/s	1 Gb/s	10 Gb/s	40 Gb/s and 100 Gb/s
slotTime	512 bit times	4096 bit times	not applicable	not applicable
interPacketGap[a]	96 bits	96 bits	96 bits	96 bits
attemptLimit	16	16	not applicable	not applicable
backoffLimit	10	10	not applicable	not applicable
jamSize	32 bits	32 bits	not applicable	not applicable
maxBasicFrameSize	1518 octets	1518 octets	1518 octets	1518 octets
maxEnvelopeFrameSize	2000 octets	2000 octets	2000 octets	2000 octets
minFrameSize	512 bits (64 octets)	512 bits (64 octets)	512 bits (64 octets)	512 bits (64 octets)
burstLimit	not applicable	65 536 bits	not applicable	not applicable
ipgStretchRatio	not applicable	not applicable	104 bits	not applicable

Limitations of Ethernet

Ethernet Problems

- Pretty low peak utilization
- Peak throughput worst with:
  - More hosts
    More collisions needed to identify single sender
  - Longer links
    Collisions take longer to observe, more wasted bandwidth
  - Efficiency is improved by avoiding these conditions of course.

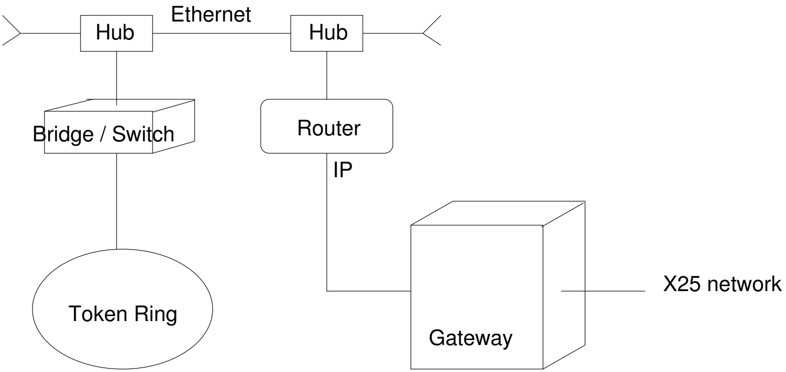
But we need larger networks, with more hosts ....

Legacy 802.3/Ethernet not used anymore !

Switched Ethernet

Interconnection devices

There are several types of devices to interconnect networks



Interconnection devices

Ethernet Hub

- Interconnects Ethernet segments
- PHY layer interconnection device:
  just copies the bits received on the output port
- Collisions are propagated - just extends the broadcast domain

Bridge / LAN switch

- Interconnects two or more LANs together
- DLC layer interconnection device:
  Interconnects identical or dissimilar networks
- Switch is often used in the context of Ethernet interconnection.

Interconnection devices

Routers

- NETwork layer interconnection device:
  Typically interconnects IP networks
- Decides on routes for packets by implementing the IP routing protocol

Gateway

- A more generic term for routers
- NETwork layer device as well
- Often designates devices that interconnect different layer 3 protocols

Switched Ethernet

Legacy Ethernet suffers from low peak utilization because it has to handle collisions on the shared medium

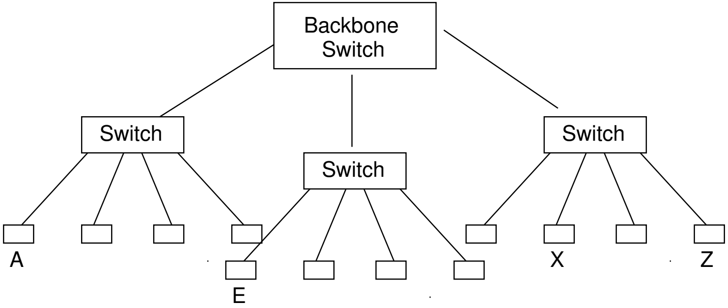
To improve Ethernet efficiency, current LANs use a 'Switched Ethernet' technology (legacy Ethernet is rarely used nowadays)

- It replaces the shared medium with a *dedicated segment* for each station
  i.e. 1 segment per station
- Segments connect to a *switch*, which connects many of these segments in a star topology
- Each segment is *full duplex*

No collisions anymore on a segment !
$\rightarrow$ up to 100% utilization per segment

Switched Ethernet

Architecture example



Switched Ethernet

Ethernet switch

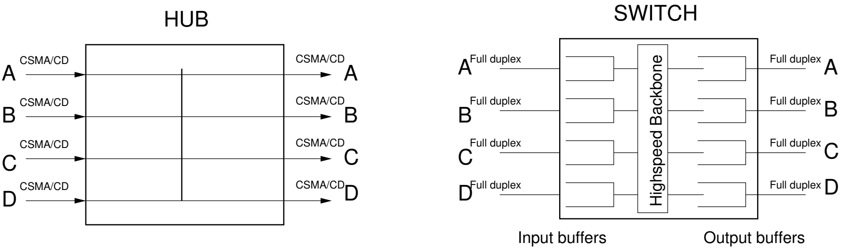
- Looks similar to a hub but no shared medium
- Processes and sends data in a more sophisticated way than a hub
- The switch reads the destination address of the packet and sends the packet only to the port the destination is reachable at (NO broadcast as for a hub)
- Thus, for each output port, there is a *buffer* that stores packets waiting for service.

Switches can support today hundreds of dedicated segments

Very scalable !

Ethernet switch

Each port is isolated and builds its own collision domain



Switched Ethernet

Ethernet switch

Overall design goal: *Complete Transparency*

"Plug and play"

Should be self-configuring, without hardware or software changes
Should not impact operation of existing LANs

Main parts for understanding bridges:

1. Forwarding of Frames
2. Learning of Addresses

Switched Ethernet

(1) Forwarding of frames

A switch relies on a *MAC forwarding table* (or switching table)

- It stores entries of the form:

{ MAC address - output port - age}

MAC address:	host name or group address
port:	outgoing port number of bridge
age:	age timing of entry (in seconds)

- As a packet comes in the switch, it reads the destination address and looks in the table for its output port.
- If no entry exists in the table?
    - It floods the packet on all output ports (except its incoming port).

(2) Learning of MAC addresses

Forwarding tables are set automatically with a simple heuristic:

- The source field of frames that arrive on a port tells which hosts are reachable from this port
- Learning algorithm
  - For each frame received, the source stores the source field in the forwarding database together with the port where the frame was received
  - All entries are deleted after some time (default is $\sim$15s)