LAN devices

Fulvio Risso
Politecnico di Torino
Copyright notice

- This set of transparencies, hereinafter referred to as slides, is protected by copyright laws and provisions of International Treaties. The title and copyright regarding the slides (including, but not limited to, each and every image, photography, animation, video, audio, music and text) are property of the authors specified on page 1.

- The slides may be reproduced and used freely by research institutes, schools and Universities for non-profit, institutional purposes. In such cases, no authorization is requested.

- Any total or partial use or reproduction (including, but not limited to, reproduction on magnetic media, computer networks, and printed reproduction) is forbidden, unless explicitly authorized by the authors by means of written license.

- Information included in these slides is deemed as accurate at the date of publication. Such information is supplied for merely educational purposes and may not be used in designing systems, products, networks, etc. In any case, these slides are subject to changes without any previous notice. The authors do not assume any responsibility for the contents of these slides (including, but not limited to, accuracy, completeness, enforceability, updated-ness of information hereinafter provided).

- In any case, accordance with information hereinafter included must not be declared.

- In any case, this copyright notice must never be removed and must be reported even in partial uses.
LAN devices

- Repeater
  - Hub
- Bridge
  - Switch
Repeater (1)

- Receives and propagates a sequence of bits
- Interconnection at the physical layer
  - E.g. fiber to copper
  - Same MAC
  - Recover signal degradation (long cables) and allows larger distances
Repeater (2)

- Functions
  - Signal Amplification
  - Signal Symmetry
  - Signal Retiming
  - Carrier Sense and Data Repeat
  - Fragment Extension (min 96 bit)
  - Collision Detection and Jam Generation
  - Test functions

- Please note that all the ports of a repeater have the same speed
  - If this does not happen, the device is not a pure repeater

- No longer used (at least on Ethernet LANs)
Repeater and Hubs

- Classical repeater: 2 port devices
- Multiport repeater: repeater with more than 2 ports
  - Known as “hub”
  - Common device when the adoption of structured cabling took place
  - More flexible (and robust) than the old coax cable
- Please note: in Ethernet, same collision domain (max diameter unchanged)
Bridge (1)

- Receives and re-transmit a frame
- Interconnection at the data-link layer
  - E.g. Ethernet to WiFi
  - Different MACs
    - Medium access mechanism, framing, update FCS
Bridge (2)

- Original objective
  - Interconnection between different LANs
    - Usually not possible anyway due to MTU issues (data-link does not have fragmentation)
  - LAN extension (total diameter)
    - Especially useful for FastEthernet and upper speed (200m)
- Bridges decouple broadcast domain from collision domain
  - Different LANs, same broadcast domain
  - Different collision domain (in Ethernet)
Bridge (3)

- By-product: if the bridge has a smart forwarding process, it can implement traffic segregation
  - Increase the aggregate bandwidth of the network
  - Right now, the most important reason for using these devices
  - Forwarding technique based on MAC Destination address
Transparent bridge

- IEEE standardized the bridging function in 802.1D
- 802.1D defines Transparent Bridges
  - Other (non transparent) bridges have been proposed in the past (e.g. Token Ring networks)
  - Transparent bridges have been proposed in Ethernet
- Transparency
  - End systems must operate in the same way (same packets, some format, etc) with or without bridges
  - Performance (throughput, max distances) may vary, but functionalities are the same
  - Requires a local forwarding table
Transparent bridge (2)

- Requires additional components
  - Stations auto-learning (*backward learning*)
  - Loop detection (*spanning tree algorithm*)
Smart forwarding process (1)
Smart forwarding process (2)

Begin

Received frame on port X

Errors (collision, CRC)?

Y

MAC destination in DB?

N

N

Y

Destination port == X?

N

Discard frame

Y

Forward on selected port

End

Forward on all ports
Smart forwarding process (3)

- A bridge can implement the “store and forward” technology
  - Receives a packet on one interface
  - Stores the packet into a local buffer
  - Analyzes the destination address
    - Why is the MAC Destination before MAC Source in Ethernet frames?
  - Forward it on the right port (if needed)

- Ports belong to different collision domain, therefore a bridge can send/receive at the same time
  - Bridges have buffers in order to absorb bursts and to wait for the proper transmission slot
Smart forwarding process (4)

- Smarter forwarding rules
  - Unicast: only on the port toward we can reach the destination (Destination MAC-based forwarding)
  - Multicast, Broadcast: flooding
    - All ports except the port on which the frame has been received

- Transient
  - If the MAC address is not present in the MAC forwarding table
  - Bridge = hub
    - Frame duplicated on all ports except the one on which it was received

- A MAC forwarding table must be available locally
Changes in sent/received frames

- Changes in sent frames
  - Nothing at all

- Changes in received frames at the NIC level
  - No receiving all frames anymore
  - Only broadcast/multicast
  - MAC filtering on the NIC becomes useless
  - NIC in promiscuous mode becomes useless

- Changes in received frames at the OS level
  - Nothing at all
  - Bridges filters frames that were previously filtered by the NIC
    - The result at the OS level is the same
The Filtering Database

- Table with the “location” of any MAC source address
  - MAC address
  - Destination port
  - Ageing time (default expire after 300 s)
  - Port status (depending by spanning tree protocol)

- Entry types
  - Dynamic
    - Populated and updated by the backward learning process
    - Max entries: 2 ÷ 65 K
  - Static
    - Not updated by the learning process
    - Usually < 1K entries
Filtering database: example

Switch-1> show cam dynamic

* = Static Entry. + = Permanent Entry.
# = System Entry X = Port Security Entry

<table>
<thead>
<tr>
<th>Dest MAC Address</th>
<th>Ports</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>00-00-86-1a-a6-44</td>
<td>1/1</td>
<td>1</td>
</tr>
<tr>
<td>00-00-c9-10-b3-0f</td>
<td>1/1</td>
<td>0</td>
</tr>
<tr>
<td>00-00-f8-31-1c-3b</td>
<td>1/2</td>
<td>4</td>
</tr>
<tr>
<td>00-00-f8-31-f7-a0</td>
<td>1/1</td>
<td>2</td>
</tr>
<tr>
<td>00-01-e7-00-e3-80</td>
<td>2/2</td>
<td>0</td>
</tr>
<tr>
<td>00-02-a5-84-a7-a6</td>
<td>2/1</td>
<td>1</td>
</tr>
<tr>
<td>00-02-b3-1e-b4-aa</td>
<td>2/1</td>
<td>5</td>
</tr>
<tr>
<td>00-02-b3-1e-da-da</td>
<td>2/5</td>
<td>1</td>
</tr>
<tr>
<td>00-02-b3-1e-dc-fd</td>
<td>2/4</td>
<td>2</td>
</tr>
</tbody>
</table>
Backward learning (1)
Begin

Received frame on port X

MAC address found in the DB?

Update port and ageing time

Y

End

Add new entry in DB

N
Backward learning (3)

- Bridges learn the topology by inspecting received frames
  - Analysis of MAC source address
- Works also in presence of multiple bridges
  - Remote bridges learn the position anyway, even if the end-system is not connected locally
- The network topology cannot include loops
  - Network topology is transformed into a tree by the spanning tree algorithm
Moving end systems (1)

- End-system generates broadcast frame immediately
  - No problems
Moving end systems (2)

- End-system generates unicast traffic immediately
  - We may have forwarding errors
    - D → B is correctly delivered
    - C → B is lost
Moving end systems (3)

- End-system does not generate traffic
  - We may have forwarding errors
    - D → B is correctly delivered
      - The frame is forwarded also to the original destination
  - C → B is lost
Moving end systems (4)

- The aging time
  - Usually enough in order to cope with manual movements
    - A laptop moved from office to lab
  - Represents the worst-case black-out time for an end system

- Please note that...
  - An end-system whose MAC address is not in the DB is *always* reachable
  - An end-system whose MAC address is in the DB may be unreachable
Moving end systems (5)

- Some problems may appear in specific environments
  - E.g. fault-tolerant NICs
    - The speed of the movement is << 5min
  - NIC driver has to generate an additional broadcast frame
Bridge architecture

Diagram showing the forward and backward learning processes, filtering database, and forwarding process in a bridge architecture.
Bridges and switches (1)

- Bridge
  - Originally 2 ports, then more
  - Software-based architecture
  - No longer used in real networks
  - Still some PC-based implementations
    - For research or some special purpose
  - WiFi access points are bridges
Bridges and switches (2)

Switch

- Same device, different technology
- Hardware based forwarding and learning
- Lookup through CAMs (Content Addressable Memories)
- Spanning Tree in software
  - Convergence time in several seconds, hence hardware implementation is useless
- Can implement a “cut-through” forwarding technology
  - Faster than “store and forward”
  - Requires all ports operating at the same speed
Switch internals

- Shared bus or switching matrix
  - speed
  - complexity

- Central CPU and memory
  - intelligence
  - complexity

- Filtering Database
  - efficient lookup

- Queuing system (often on the output link)
  - decoupling of different physical speed
  - absorption of bursts
  - can drop packets

- Filtering Database
  - useful only on some links (e.g., intra-switch)

- Table may become full
- transient

No CSMA/CD (Full Duplex)
- speed

IN

OUT
Switched LANs (1)

- Progressive replacement of shared segments with switches
Switched LANs (2)

Data center (CED)

Internet
Switched LANs (3)

-Currently, end systems directly connected to switches
  - More aggregated bandwidth
  - No need to replace NIC on clients when moving from hubs to switches
  - Switches may be 10/100/1000 and support different speed on the client side
    - Possibility to smooth upgrade of the network (NICs, hubs/switches), mixing different Ethernet technologies
    - Hub did not support multiple speed
Switched LANs (4)

- No longer used
  - CSMA/CD
    - Only one station can be attached to a physical link (no need to arbitrate the channel)
  - Frame bursting
  - Carrier Extension

- What remains
  - Framing

- Maximum diameter of an Ethernet (Fast/Giga/…) network
  - Max diameter (for collision domain) is no longer a problem
  - Max cable length (due to signal attenuation) is still a problem
    - E.g., 100m from end-system to a switch (twisted pair) is still a valid limit
**Full-duplex**

- Introduced with Fast Ethernet
  - Allows to send/receive at the same time
    - No CSMA/CD, no collision
  - Requires a direct connection to a switch
    - End-system – switch, switch – switch
- Most LANs currently use Full Duplex
- 2x in aggregated bandwidth
  - Useful in switch – switch connections (often symmetric traffic)
  - Not very useful in end system – switch connection
    - Servers tend to saturate uplinks
    - Clients tend to saturate downlink
- Traditional mode: Half Duplex
**Switched networks and throughput**

- Aggregate bandwidth increases
- Throughput may not!
- Uplink speed is a critical factor
  - Uplinks must sustain the traffic of all the attached station
  - Links toward servers must be fast enough
  - Is it a good choice to have clients connected at 1Gbps?
- Buffers play an important role in switches
  - Classical Ethernet implements a “reliable” transmission
    - Why can CIFS and NFS use UDP for data transfers?
  - Switches may drop packets due to congestions (limited buffer size)
  - TCP timeouts come into play
    - Dramatic decline in throughput
Bridges and meshes: the loop problem

Loop!

Where is B?
Where is B?

Filtering Database

<table>
<thead>
<tr>
<th>MAC</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>5</td>
</tr>
</tbody>
</table>

A → B

Where is B?
Bridges and meshes

Two problems
  - Looping frames
  - Backward learning no longer able to operate

Looping frames when
  - Multicast/broadcast packet
  - Packet to a non-existing station
    - MAC address unknown in the filtering DB

Backward learning
  - Switches may have inconsistent filtering database
  - An entry in the filtering database may change the port indefinitely
Spanning Tree

Where is B?
Spanning Tree

- 802.1D
  - Original idea from Radia Perlman, PhD @DEC
- Prunes the network
- Meshes are disabled and the network becomes a tree
- Operate periodically (every second)
  - Decide which port set to forwarding state and which port set to blocking state