Switches

- **Switches** are link-level devices; they switch link-layer frames rather than network-layer datagrams.

- They don’t recognize network-layer datagram or addresses and don’t use routing algorithms like RIP or OSPF.
Link-layer addresses

- A LAN node’s link-layer address (variously physical, Ethernet, or MAC* address) properly belongs to its adapter.
- Generally these are 48 bits long.
- Although MAC addresses were designed to be permanent, (originally burnt into the adapter’s ROM) it is now possible to set an adapter’s MAC via software.

*Media Access Control

Expressed in dashed hexadecimal notation

1A-23-F9-CD-06-9B
5C-66-AB-90-75-B1
49-BD-D2-C7-56-2A
88-B2-2F-54-1A-01

No two are ever alike

- How is this possible? MAC address space is managed by IEEE.
- When a company wants to manufacture adapters, it purchases a chunk of $2^{24}$ addresses for a nominal fee.
- The first 24 bits of a MAC address are fixed by IEEE, and companies get to create unique combinations of the last 24 bits.
Flat tax?

- An adapter's MAC address has a flat structure, as opposed to IP's hierarchical structure.

- A slaptop has the same MAC address, no matter where so ever it goes*; meanwhile its IP address is bouncing all over the place**.

*Kinda like your social security number.
**Kinda like your postal address (at least when you are young).

An adaptor sends a frame into the LAN

- If it is a broadcast LAN*, it is received and processed by all other adapters on the LAN.
- On the receiving end,
  - If a destination address matches the node’s MAC address, it extracts the network-layer datagram and passes it up the protocol stack.
  - If the destination address doesn’t match, the node discards the frame. The host is never disturbed.

*Such as 802.11 or the old fashion Ethernet LANs.
MAC Broadcast address

- Sometimes the sending adaptor wants all other adapters on the LAN to receive and process the frame.
- The sending adapter inserts a MAC broadcast address into the destination field of the frame.*

*For an Ethernet and token passing LAN, the broadcast address is FF-FF-FF-FF-FF.

ARP: address resolution protocol

**Question:** how to determine interface's MAC address, knowing its IP address?

**ARP table:** each IP node (host, router) on LAN has table

- IP/MAC address mappings for some LAN nodes:
  <IP address; MAC address; TTL>
- **TTL (Time To Live):** time after which address mapping will be forgotten (typically 20 min)
ARP protocol: same LAN

- A wants to send datagram to B
  - B’s MAC address not in A’s ARP table.
- A broadcasts ARP query packet, containing B’s IP address
  - dest MAC address = FF-FF-FF
  - all nodes on LAN receive ARP query
- B receives ARP packet, replies to A with its (B’s) MAC address
  - frame sent to A’s MAC address (unicast)
- A caches (saves) IP-to-MAC address pair in its ARP table until information becomes old (times out)
  - soft state: information that times out (goes away) unless refreshed
- ARP is “plug-and-play”:
  - nodes create their ARP tables without intervention from net administrator

Addressing: routing to another LAN

- Send datagram from A to B via R
  - focus on addressing – at IP (datagram) and MAC layer (frame)
  - assume A knows B’s IP address
  - assume A knows IP address of first hop router, R (how?)
  - assume A knows R’s MAC address (how?)
Addressing: routing to another LAN

- A creates IP datagram with IP source A, destination B
- A creates link-layer frame with R’s MAC address as dest, frame contains A-to-B IP datagram

Frame sent from A to R
Frame received at R, datagram removed, passed up to IP
Addressing: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B’s MAC address as dest, frame contains A-to-B IP datagram
Addressing: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as dest, frame contains A-to-B IP datagram

```
IP src: 111.111.111.111
IP dest: 222.222.222.222
MAC src: 1A-23-F9-CD-06-9B
MAC dest: 49-BD-D2-C7-56-2A
```

Ethernet
Ethernet

- Since the mid 1990s Ethernet has dominated the LAN market. Partially because,
  - First to be widely deployed.
  - Generally cheaper and simpler than its competitors.
  - Always managed to maintain comparable data rates with emerging technologies.

Original bus topology for Ethernet LANs*

- Metcalfe and Boggs's original design used a coaxial bus to interconnect the nodes.

- A topology that lingered on into the mid-nineties.
By the late 1990s ...

- the bus LANs had been replaced with Ethernet installations using a hub-based star topology.

- A hub is a physical-layer device that acts on individual bits rather than frames.

- When a bit arrives from one interface, the hub recreates the bit, boosts its energy, and transmits the bit onto all other interfaces.

Hubs also provide network management features

- If an adapter is misbehaving (e.g., hogging the network), the hub can detect this and disconnect the offending one.

- Most hubs also gather information and report the information back to a host that connects directly to the hub.
In the early 2000s ...

- ... the hub at the center of the star topology was replaced with a switch.

- A switch is a collision-less, store-and-forward packet switching device which operates at the physical and link layers.

---

Ethernet frame structure

- Sending adapter encapsulates IP datagram (or other network layer protocol packet) in

**Ethernet frame**

```
| preamble | dest. address | source address | data (payload) | CRC |
```

- Preamble:
  - 7 bytes with pattern 10101010 followed by one byte with pattern 10101011
  - used to synchronize receiver, sender clock rates
Ethernet frame structure (more)

- Addresses: 6 byte source, destination MAC addresses
  - if adapter receives frame with matching destination address, or with broadcast address (e.g. ARP packet), it passes data in frame to network layer protocol
  - otherwise, adapter discards frame

- Type: indicates higher layer protocol (mostly IP but others possible, e.g., Novell IPX, AppleTalk)

- CRC: cyclic redundancy check at receiver
  - error detected: frame is dropped

Demultiplexing using the type field

- Hosts can use multiple network-layer protocols.*
- When Ethernet frame arrives at adapter B, B needs to know to which network-layer protocol it should pass its contents.
- This field is analogous to the protocol field in the network-layer datagram and the port number in the transport-layer segment.

*Each protocol has its own type number (e.g., a frame containing an ARP packet has type field hex 0806).
What could possible go wrong?

• Causes of errors include attenuation in signal strength and ambient electromagnetic energy that leaks into the Ethernet cables and interface cards.

• CRC calculated on everything except preamble.

Unreliable connectionless service

• Ethernet provides connectionless service;
  – no handshaking
  – the sending adapter simply encapsulates the datagram and drops it into the LAN

• Similarly, the service is unreliable.
  – Receiver checks for errors, but sends no acknowledgements one way or the other
Many flavors

• We’ve been talking about Ethernet as if it were one protocol standard when fact it is many.

• Many of these have been standardized by the IEEE 802.3 CSMA/CS (Ethernet) working group.
10BASE-2 and 10BASE-5 Ethernet

- Legacy technology from the 90s, 10Base used thin coaxial cable to connect nodes in a bus topology.
- The 10 stands for 10 Mbps, BASE refers to baseband Ethernet and the 2 (5) stands for 200 meters (500 meters), the length limit of a coaxial cable.

- Longer runs could be obtained by using repeaters, physical-layer devices to boost signals.

Generations pass

- Today's Ethernet is very different from the original coaxial bus technology.
- Nodes are connected to a switch via point-to-point segments made of twisted-pair copper wires or fiber-optic cable.
100BASE-T, 100-BASE-FX, 100BASE-SX, 100BASE-BX

- In the mid-nineties, Ethernet was standardized at 100Mbps.
- The original MAC protocol and frame format were preserved, 100 Mbps Ethernet is limited to a 100 meter distance over twisted pair, and several kilometers over fiber.

*The T stands for twisted pair, the FX, SX, and BX for fiber.

Gigabit and 10 Gbps Ethernet

- Gigabit Ethernet (IEEE 802.3z) is backwardly compatible with 10BASE-T/100BASE-T technologies and uses the same frame format.
- Allows for point-to-point links (switches) as well as shared broadcast channels (hubs).
- Uses CSMA/CD for shared broadcast channels. Max distances restricted for acceptable efficiency.
Which brings us back to switches

- The switch is transparent to the nodes; a node addresses and sends a frame into the LAN unaware that a switch will be doing the dirty work.

LANs with switches are good

1. Switches permit interdepartmental communication while preserving isolated collision domains.
2. Switches can interconnect different LAN technologies, including 10 Mbps, 100 Mbps, Cat 5, and fiber Ethernets.
3. There is no limit to how large a LAN can grow when switches are used.
Switches filtering and forwarding

**Filtering**
The ability to determine whether a frame should be forwarded to some interface or simply dropped.

**Forwarding**
The ability to determine the interface to which a frame should be forwarded, and then to get it there.

---

Switch tables

<table>
<thead>
<tr>
<th>Address</th>
<th>Interface</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>7C:6A:29:84:71:10</td>
<td>3</td>
<td>9:36</td>
</tr>
</tbody>
</table>

---
Filtering & forwarding with switch tables

Frame arrives with destination address DD-DD-DD-DD-DD-DD.

*But there is no entry in the table for DD-DD-DD-DD-DD-DD.

No entry in the switches ARP table

Switch forwards copies to all interfaces except the one on which it arrived (2)*

*In other words, it broadcasts it.
Filtering & forwarding with switch tables

Frame arrives with destination address DD-DD-DD-DD-DD-

*And this time there is an entry in the table for DD-DD-DD-DD-DD (4).

Off you go then
Filtering & forwarding with switch tables

Frame arrives with destination address DD-DD-DD-DD-DD-DD*.  
*And this time there is an entry in the table for DD-DD-DD-DD-DD-DD, but the entry is 4.

*Hold on, how realistic is that?
But that address isn’t in the switch table

- Switches are self-learning;
  - their tables are built automatically.

- For each frame received, the switch stores in its table
  1. The LAN address in the frame’s source field;
  2. The interface from which the frame arrived;
  3. The current time.

Frame arrives with source address 10-23-34-45-56 on interface 2
Address 10-23-34-45-56 is entered with interface 2

<table>
<thead>
<tr>
<th>Address</th>
<th>Interface</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>62:FE:71:3AA3</td>
<td>1</td>
<td>9:32</td>
</tr>
<tr>
<td>7C:8A:22:91:10</td>
<td>3</td>
<td>9:56</td>
</tr>
<tr>
<td>01-12-23-34-45-56</td>
<td>2</td>
<td>9:39</td>
</tr>
</tbody>
</table>

Summary comparison

<table>
<thead>
<tr>
<th>Feature</th>
<th>Hubs</th>
<th>Switches</th>
<th>Routers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic isolation</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Plug &amp; Play</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Optimal Routing</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
Virtual Networks

Problems with the institutional network

Group traffic is localized to a single switch, but broadcast traffic is not. Limiting the scope of such broadcast traffic would improve LAN performance*. Not to mention security/privacy reasons.

*Not to mention security/privacy reasons.
Inefficient use of switches and managing users

More, smaller groups would require more (under-utilized) switches.

Employees moving between groups would require physical cabling changes.*

*And what do we do about an employee that belongs to two groups?

Virtual Local Area Networks (VLANs)

- Switches that support VLANs allow multiple virtual local area networks over a single physical LAN.
- Switch’s ports are divided into VLAN groups, each forming its own broadcast domain.*
- VLANs solve all three of the previously mentioned problems.
- But now we have a new problem: How can traffic from EE be sent to CS?

*That is, broadcast traffic from one port can only reach other ports in the group.
Hybrid router/switches

- One way to connect EE and CS is to connect a VLAN switch port to an external router and configure that port to belong to both EE and CS.

- To make life easier, vendors provide single devices that contain both a VLAN switch and a router.

Okay, here's another problem

Suppose now some EE and CS faculty are housed in separate buildings... and they still would like to be part of the department's VLAN.
Well, we could ...

Define a port belonging to the CS VLAN on each switch and connect these ports to each other.

And do the same for ports belonging to the EE VLAN.

*But this approach doesn't scale.

VLAN trunking

Special ports on each switch are configured as a trunk port to interconnect the two VLAN switches.

*Fair enough, but how does a switch know that a frame arriving on a trunk port belongs to a particular VLAN?