Link Layer (MAC)

Unit 11 - Hands-On Networking - 2018

Prof. Dr.-Ing. Thorsten Herfet, Andreas Schmidt, Pablo Gil Pereira

Telecommunications Lab, Saarland Informatics Campus, 22nd Feb. 2018

Recap

- **Network Layer**
  - Provides **communication between hosts**
  - **Forwarding** and **Routing**
  - Today mostly **datagram-based**
- **IPv4**
  - Addresses
  - NAT
- **ICMP**(v4)
  - Error control
  - Debugging


Application


Presentation


Session


Transport


Network


Link


Physical

Link Layer | Principal Functions

- Frame **Delimiting and Recognition**
- Provide **transparent data transfer** for the network layer, e.g.
  - **Unacknowledged connectionless** service (e.g. Ethernet)
  - **Acknowledged connectionless** service (e.g. Wireless LAN, HomePlugAV)
  - **Acknowledged connection-oriented** service (Virtual Circuit Networks)
- Deal with **transmission errors**
  - Physical layer only responsible to **transmit raw bits**
  - Number of **bits received may be less, equal or more** than were transmitted
  - Bit **values may be different**
- **Media Access** and **Data Flow Control**

Link Layer Functions **tightly coupled** with the physical layer

IEEE 802: Local and Metropolitan Area Network Standards

- IEEE Standard(s) for variable-sized-packet networks
- Defines **Data Link Layer** and **Physical Layer** components
- Link layer split up in **Logical Link Control** and **Media Access Control**

LLC (IEEE 802.2)

- Abstracts PHY-specific MAC features
- Mandatory in IEEE 802 protocols (except for 802.3)

MAC

- Provides datagram-style data transfer to upper layer

IEEE 802 | MAC Addresses

- Based on the early Ethernet addressing scheme
- Assigned to **each interface** by manufacturer
- Technically a **globally unique** interface serial number (🕵)
- **48 bit** address
  - **24 bit** vendor prefix **(OUI)**
  - **24 bit** interface identifier
- Notation: **6 bytes** separated by "**-**" or "**:**"

70:54:D2:7B:7A:DB

OUI Interface ID

Special Addresses:	**00:00:00:00:00:00** Unspecified	**FF:FF:FF:FF:FF:FF** Broadcast

Organizationally Unique Identifier (OUI)

A **24-bit** number that **uniquely identifies a vendor/manufacturer/organization**

- Sold by the IEEE
- Used as a portion of **derivative identifiers to uniquely identify** a piece of equipment (e.g. MAC addresses, but there are others)

The **two least-significant bits** of the **most-significant byte** convey special meaning

- Least-significant bit **(I/G bit):**
  - **0:** Address for an individual device
  - **1:** Address for a group of devices
- Second-least-significant bit **(U/L bit):**
  - **0:** Globally unique address
  - **1:** Locally administered address (OUI does not identify the vendor anymore!)

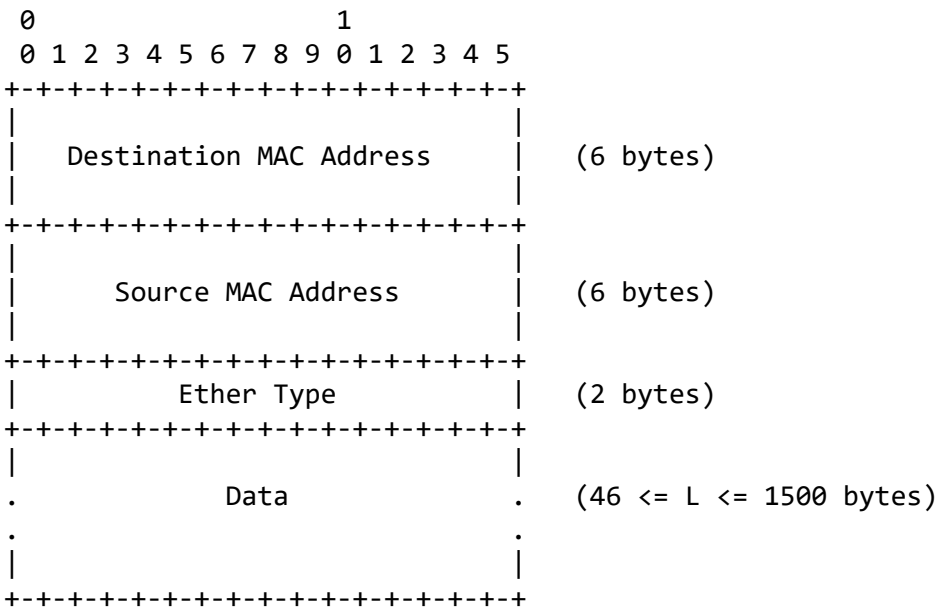
Ethernet
(IEEE 802.3)

Ethernet

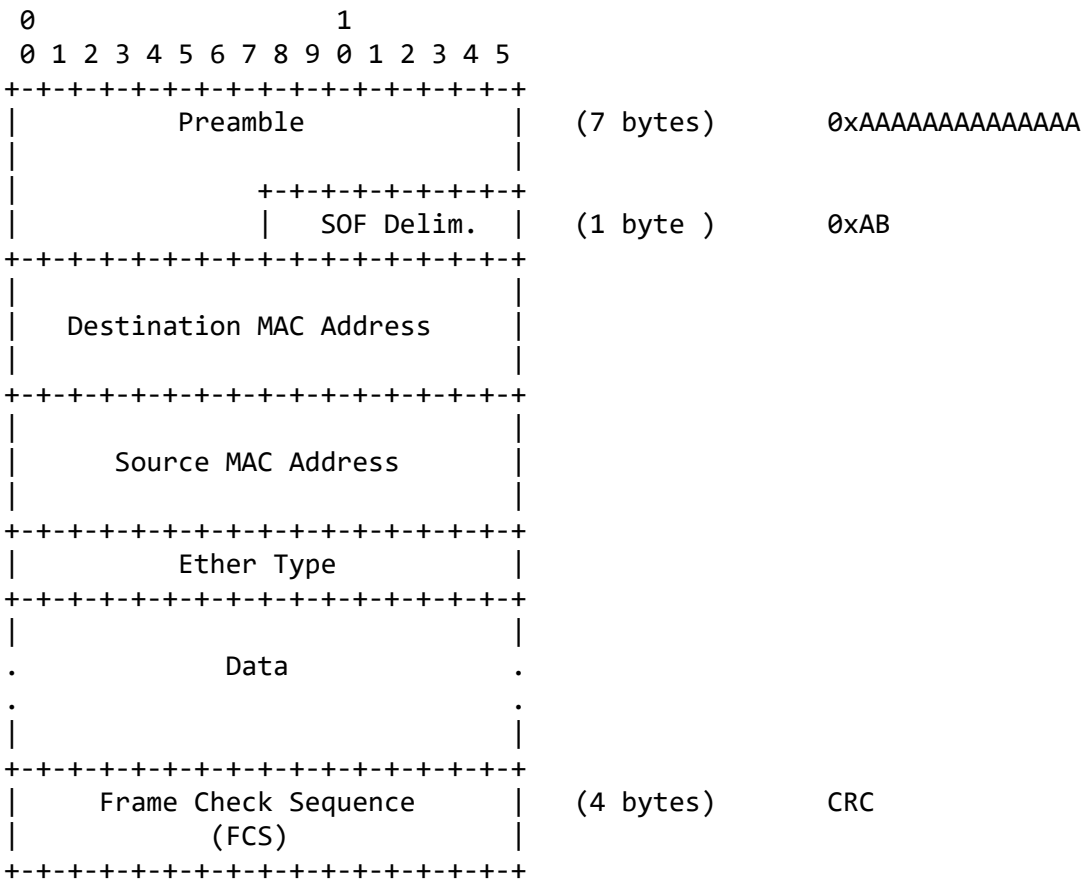
- Networking technology usually used in **LAN** and **MAN** networks
- Uses **guided transmission** media (a.k.a. "wires")
- Provides **unacknowledged connectionless datagram** service
- Ethernet datagrams are called **frames**
- **MTU** of **1500 bytes**

More on that later ...

Frame Format



Wire Frame Format

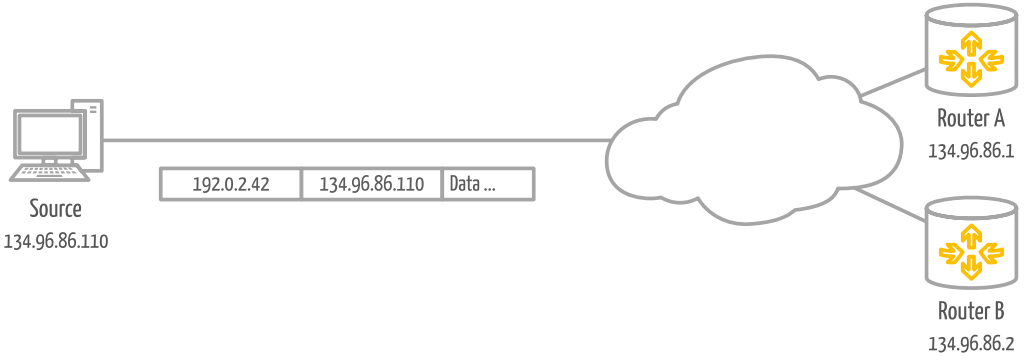


Most Ethernet hardware filters the preamble and FCS.

Quiz

Consider the following network and the abbreviated datagram below the link.

Forwarding Table	
Destination	Via
131.246.0.0/16	134.96.86.1
134.96.1.0/28	134.96.86.1
78.47.42.112/29	134.96.86.2
0.0.0.0	134.96.86.2



Which router handles the datagram?

✔ Router B

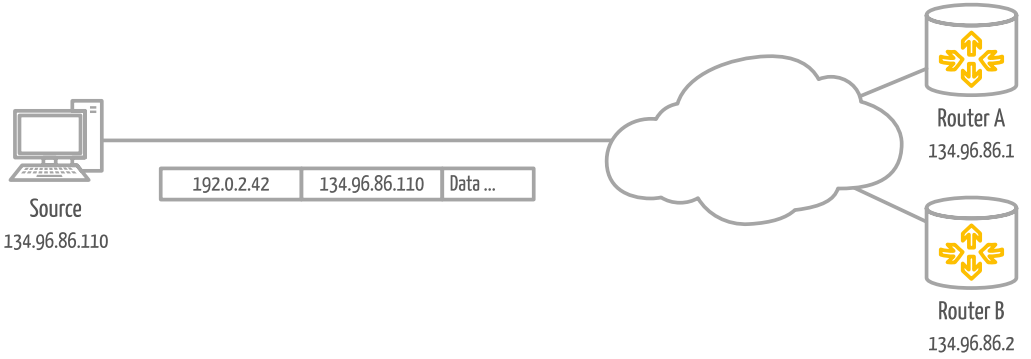
But how does the network know to deliver the packet to router B?

✔ The link layer takes care of that...

Quiz

Consider the following network and the abbreviated datagram below the link.

Forwarding Table	
Destination	Via
131.246.0.0/16	134.96.86.1
134.96.1.0/28	134.96.86.1
78.47.42.112/29	134.96.86.2
0.0.0.0	134.96.86.2



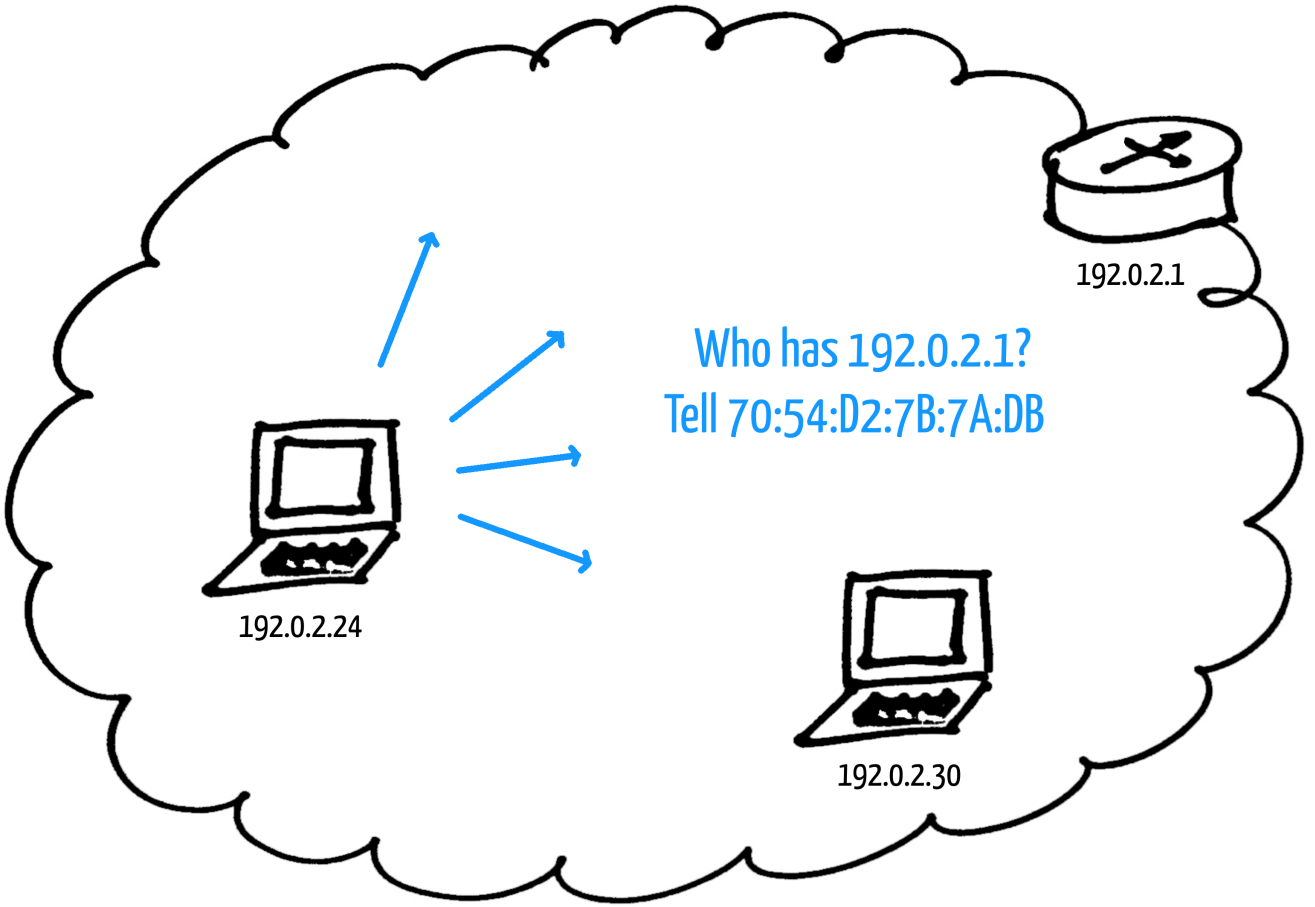
Meanwhile on the link layer ...

```
+===================+===================+========+------------+---------------+----------+
| 00:23:5d:ff:e4:00 | 70:54:d2:7b:7a:db | 0x0800 | 192.0.2.42 | 134.96.86.110 | Data ... |
+===================+===================+========+------------+---------------+----------+
```

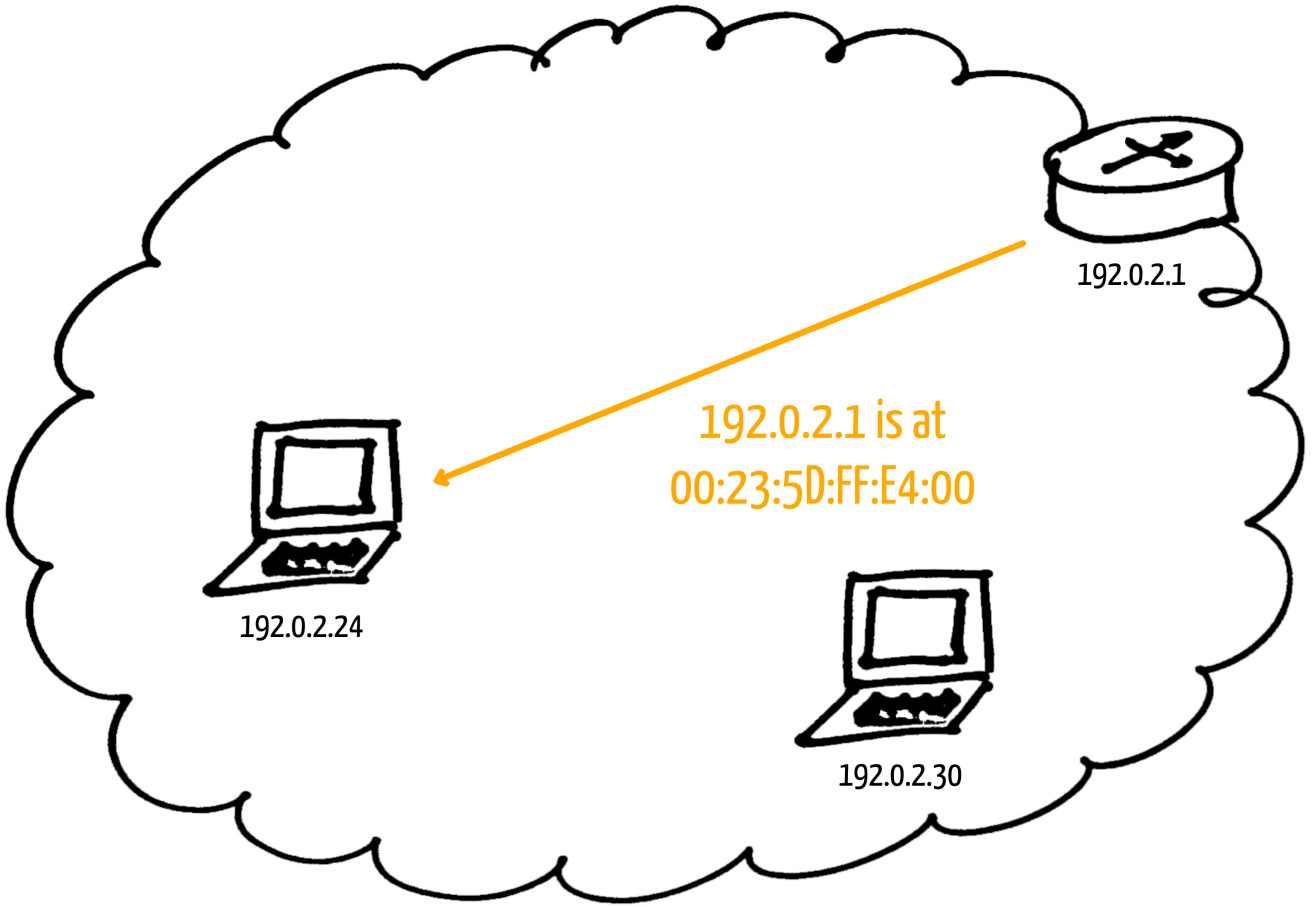
ARP

- **Address Resolution Protocol**
- Defined in RFC826
- Translates **network-layer addresses** to **link-layer addresses**
- **Used in many IEEE802 standards** (notably 802.3 and 802.11)
- **Request-Response** Protocol
- As a link-layer protocol it is **never forwarded** by routers

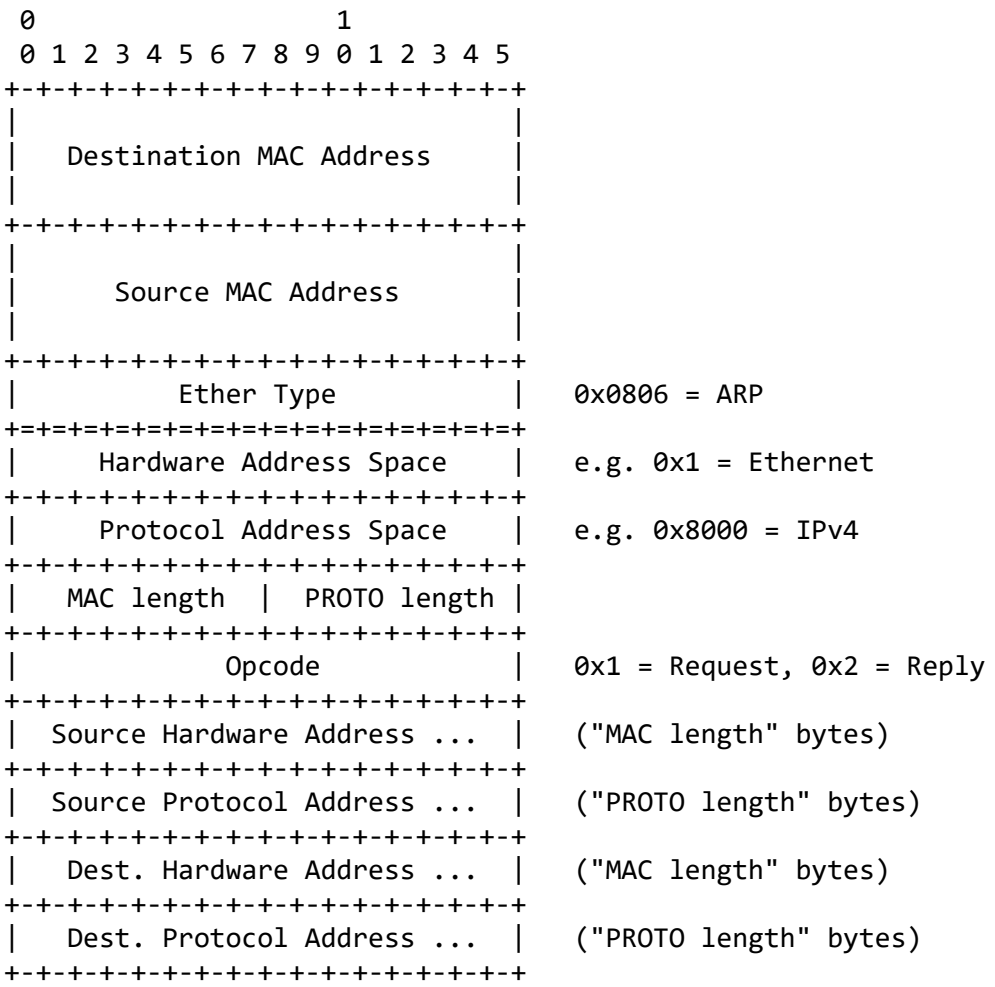
ARP



ARP

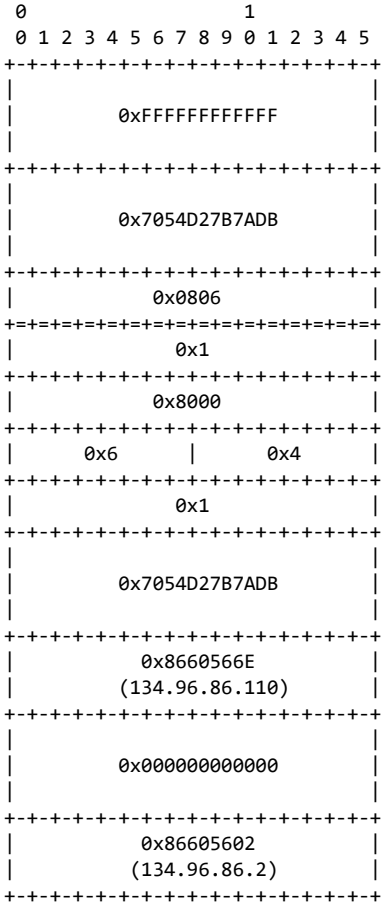


ARP | Frame Format

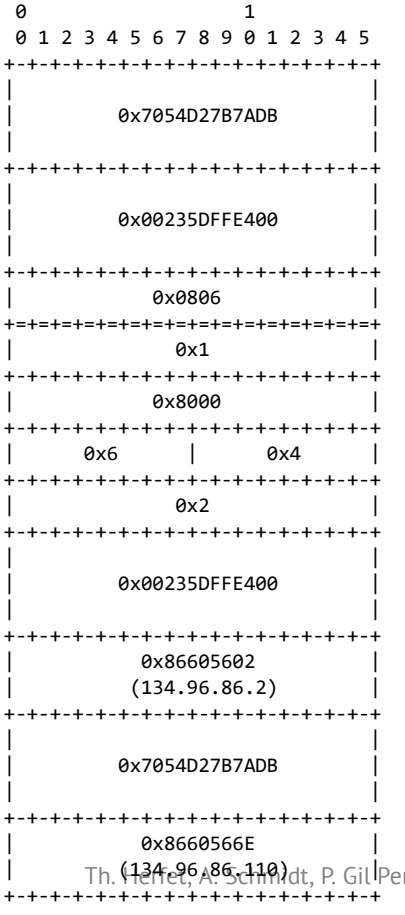


ARP | Frame Format Example

Request



Reply



ARP

Does every device send an ARP Request for every datagram transmitted to the network?

✔ Of course not!

Nearly all network devices cache ARP responses for a while

ARP / MAC Table

Protocol Address	HW Address	Interface
192.0.2.2	44:8A:5B:27:88:0B	1
10.14.0.10	00:19:99:92:E7:06	2
10.14.0.11	00:23:5D:FF:E4:00	2

Gratuitious ARP

ARP request or reply which is not normally needed according to RFC826

Often sent by the OS after an **IP address change** to refresh ARP tables of neighboring devices or detect IP address conflicts (RFC5227)

- A **request** sent to the MAC layer broadcast address where **both source and destination IP are set to the IP address of the sending station**
- A **reply** for which there was **no preceeding request**

Sometimes also sent by someone trying to impersonate somebody else.

ARP | Tools

- Display / modify ARP table:
  - iproute2 (Linux)
  - arp (Linux, macOS, Windows)
- Manually send ARP Requests / Replies.
  - arping (Linux, macOS)

```
$ ip neighbour show
134.96.86.108 dev eth0 lladdr 44:8a:5b:27:88:0b STALE
134.96.86.83 dev eth0 lladdr e8:40:f2:3e:95:2c STALE
134.96.86.35 dev eth0 lladdr 00:19:99:92:e7:06 STALE
134.96.86.34 dev eth0 lladdr 00:19:99:92:e7:08 STALE
134.96.86.1 dev eth0 lladdr 00:23:5d:ff:e4:00 REACHABLE
```

```
$ arp -n
Address                  HWtype  HWaddress           Flags Mask            Iface
134.96.86.108            ether   44:8a:5b:27:88:0b   C                     eth0
134.96.86.83             ether   e8:40:f2:3e:95:2c   C                     eth0
134.96.86.35             ether   00:19:99:92:e7:06   C                     eth0
134.96.86.34             ether   00:19:99:92:e7:08   C                     eth0
134.96.86.1              ether   00:23:5d:ff:e4:00   C                     eth0
```

ARP | Exploits 

- MAC addresses are not a security feature
  Even though a unique address is assigned to an interface, it can easily be changed. Authorization based only on a device's MAC address suggests security where there is none, e.g. in Wireless LANs

- ARP is not authenticated
  Everyone can reply to an ARP request regardless of whether they hold the requested protocol address enabling MitM attacks (ARP Spoofing).

- ARP requests are optional
  Any device can send gratuitious ARP messages. Because of a lack of authentication this opens the door to all kinds of device impersonation or flooding attacks and enables eavesdropping. (ARP Cache Poisoning)

These attacks can be mitigated or at least detected using properly configured professional network equipment.


Application


Presentation


Session


Transport


Network


Link


Physical

Ethernet

Ethernet

- Networking technology usually used in **LAN** and **MAN** networks
- Uses **guided transmission** media (a.k.a. "wires")
- First standardized in 1983
- Provides **unacknowledged connectionless datagram** service
- Ethernet datagrams are called **frames**

- Ethernet supports **different media types**
- Depending on the media, Ethernet supports **varying data rates** and **transmission distance**

Some Ethernet Media

PHY Layer Standard	Data Rate	Medium	Max. Distance
10BASE5	10 Mbit/s	RG-8 Coaxial Cable	500m
10BASE2	10 Mbit/s	RG-58 Coaxial Cable	200m
10BASE-T	10 Mbit/s	Twisted Pair	100m
100BASE-T	100 Mbit/s	Twisted Pair	100m
100BASE-FX	100 Mbit/s	Fiber	412m
1000BASE-T	1 Gbit/s	Twisted Pair	100m
10GBASE-T	10 Gbit/s	Twisted Pair	100m
10GBASE-SR	10 Gbit/s	Fiber	300m

Format: (Data Rate)(Signaling Type)-(Media Type/Encoding)

A little History

Ethernet over shared Coax

- **One** Cable running by **each node**

All nodes **on the same subnet** share the **same link**

- Cables were **tapped** using **"Vampire Taps"**

Source: Wikimedia Commons

Source: Wikimedia Commons

A little History

Ethernet over shared Coax

- Good **EMI protection**

- Complex / **error-prone** wiring
- **Every transmission** is electrically speaking a **broadcast**
- Only **one node** can transmit at a time
- What if two nodes send **at the same time**?

Media Access Control to the Rescue

CSMA/CD (Carrier Sense Multiple Access with Collision Detection)

Check for **interfering signals while transmitting** a frame. A special **jam signal** is used to notify other stations about collisions. Stations receiving the jam signal **stop sending immediately** and wait for a **random time interval** (a few µs) before retrying.

Simplified Algorithm

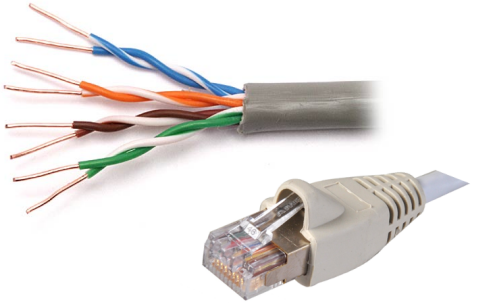
- Is the medium idle? If not wait until it is.
- Start transmitting and monitor for collision during transmission.
- Did a collision occur? If so:
  - Transmit the jam signal.
  - Increment restransmission counter.
  - Was the maximum number of transmissions reached. If yes, abort.
  - Wait for a random time interval and start over.

Carrier sensing and random backoff **lower the probability of collisions**

Modern Ethernet

Ethernet over Twisted Pair Cables

- **Four** twisted **wire pairs**
- **One pair for transmitting** and **one pair for receiving** (<= 100 MBit/s)
- **One** cable **per node**



- Twisted pair cables were **cheaper** and **already available** for phones
- Using **point-to-point links** instead of a shared bus greatly **simplifies troubleshooting**

- Slightly less EMI protection
- **Separate links** need to be **interconnected** somehow

Twisted Pair

Common Types

Name	Bandwidth	Applications
Cat 3	16MHz	10BASE-T, 100BASE-T4
Cat 5(e)	100MHz	100BASE-TX, 1000BASE-T
Cat 6	250MHz	10GBASE-T
Cat 6A	500MHz	10GBASE-T
Cat 7	600MHz	10GBASE-T, CATV
Cat 7A	1000MHz	10GBASE-T, CATV

Twisted Pair

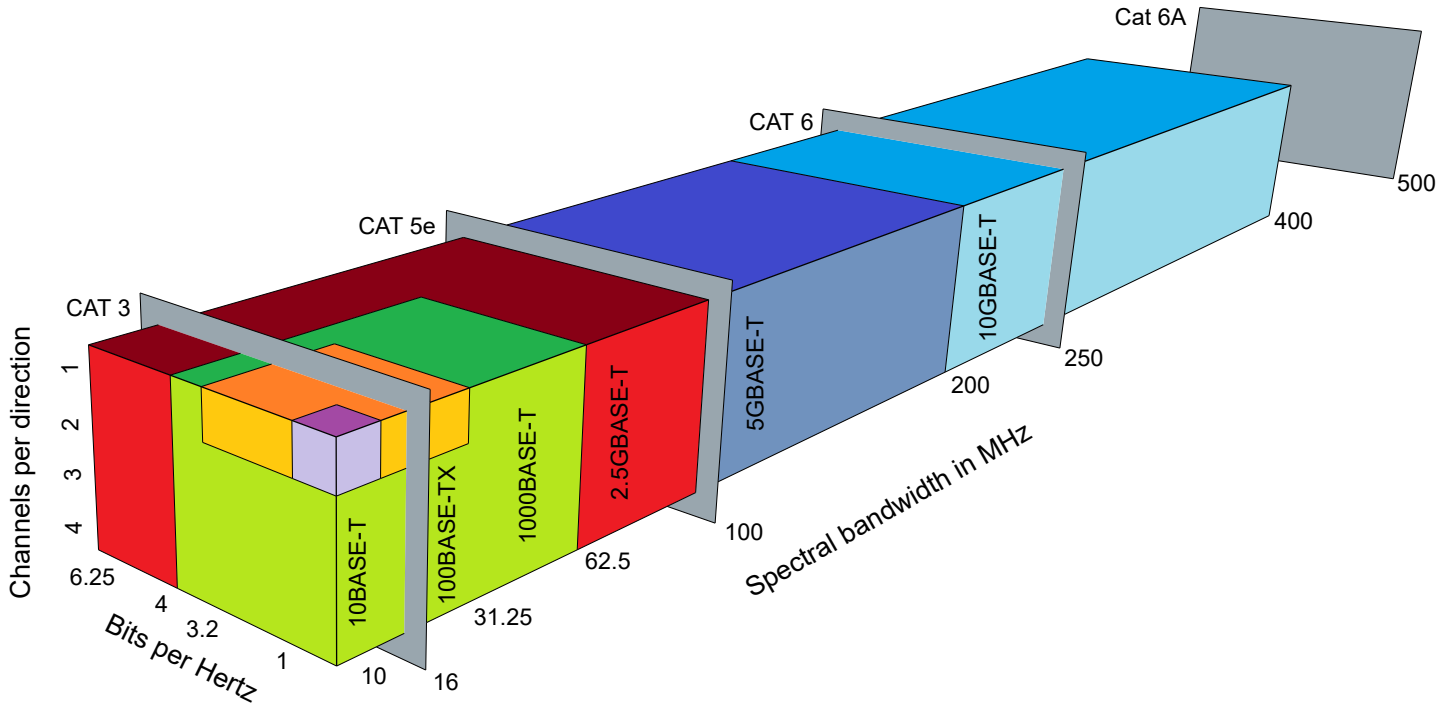


Image by Per Mejdal Rasmussen via Wikimedia Commons

⚠ The lower axis should actually be: $\frac{bps}{Hz}$

Connecting Link Segments

Repeater / Hub

A hub **simply repeats** every frame it receives on **all** its ports **but the receiving one**.

- **Cheap**
- Problems can (sometimes) be **localized more easily**

- **Physical star** topology is **logically still a bus**
- **Collisions**

A hub operates only on the **physical layer**.
It does **not inspect or alter data** in any way.

True hubs are **basically extinct**.

Connecting Link Segments

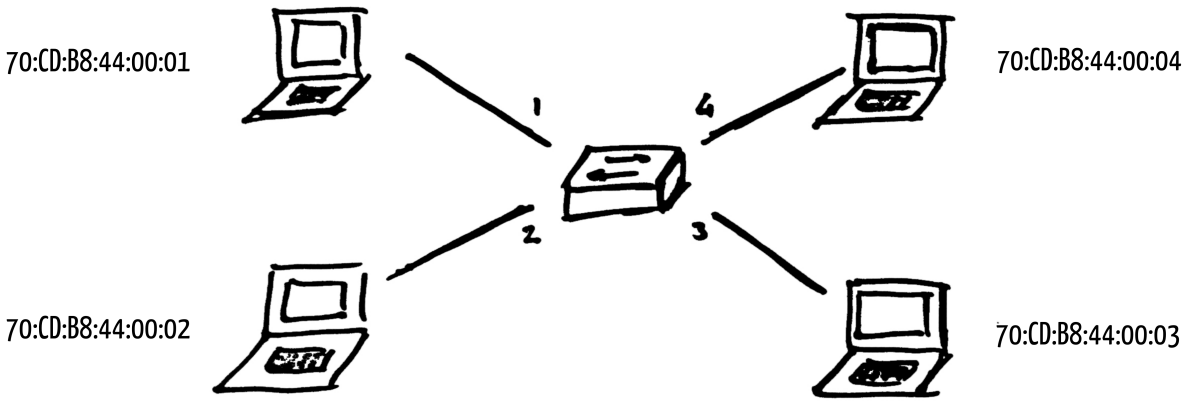
Switch / Multiport Bridge

A switch **only forwards frames** it receives **to stations that need to receive it**.

- Switch needs to know on **which port a station is connected**.
  It **learns MAC addresses** from **ARP** messages.
- **Store-and-Forward** Switching
  The switch **buffers frames** it receives **until the outgoing link becomes available**. All devices attached to a switch can **transmit at the same time** without producing collisions.
- **Full-Duplex Transmission**
  Because of the **point-to-point link** and because transmissions can now be **buffered** the stations are enabled to **send and receive at the same time** (**Full-Duplex transmission**).

Learning Switch

Initial situation (devices are connected and powered up).

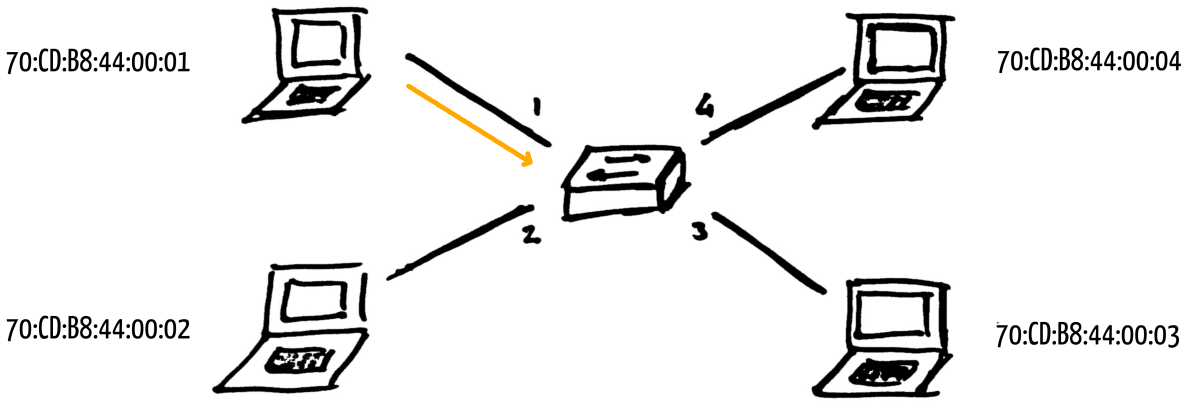


Switch MAC Table

Address	Port

Learning Switch

Node 1 sends an ARP Request.

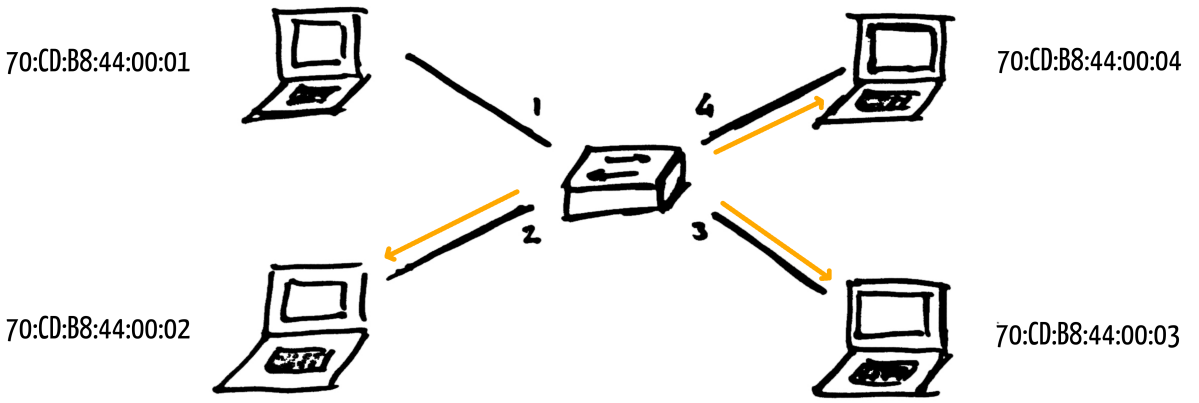


Switch MAC Table

Address	Port
70:cd:b8:44:00:01	1

Learning Switch

Switch learns the source address and forwards it to all other nodes.

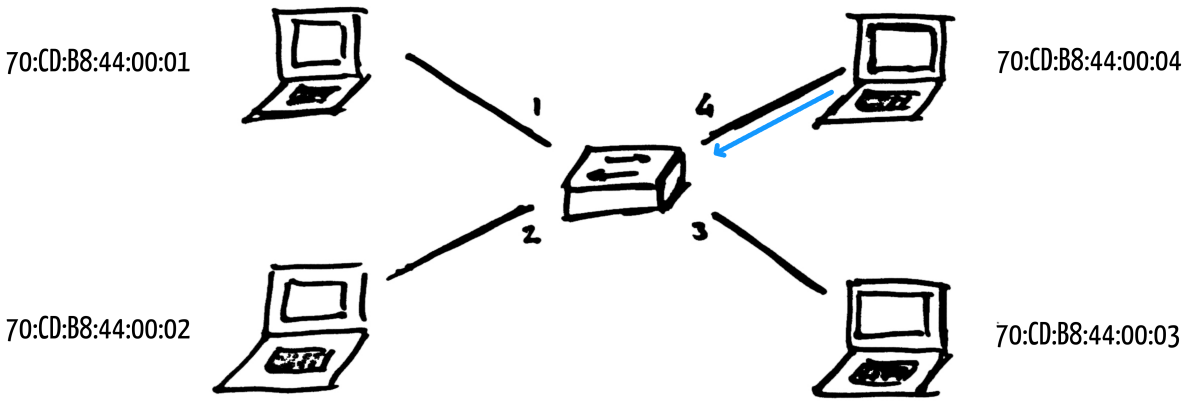


Switch MAC Table

Address	Port
70:cd:b8:44:00:01	1

Learning Switch

Node 4 replies. The switch learns address from the reply.

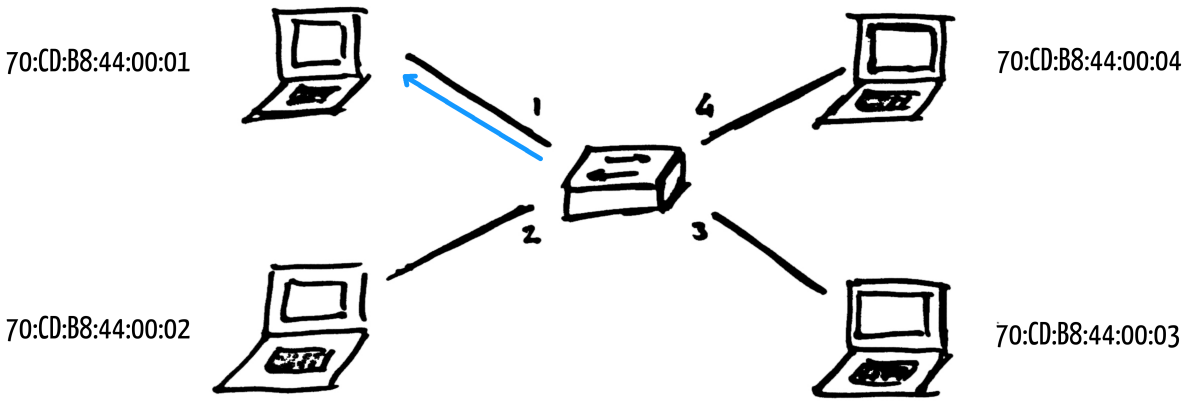


Switch MAC Table

Address	Port
70:cd:b8:44:00:01	1
70:cd:b8:44:00:04	**4**

Learning Switch

The recipient of the reply is in the MAC table, so it is only sent to node 1.



Switch MAC Table

Address	Port
70:cd:b8:44:00:01	**1**
70:cd:b8:44:00:04	4

Connecting Link Segments

Switch / Multiport Bridge

- **Exclusive Media Access** per station
- No collisions, the switch can buffer and coordinate
- **1000BASE-T** and later:
  - Uses **all four wire pairs**
  - **Full-Duplex** transmission

- Costs (not anymore)

Broadcasts will still be sent to **all stations** on the link.

HON

Segmenting the Network

It is good practice to separate a network into multiple segments.

Why?

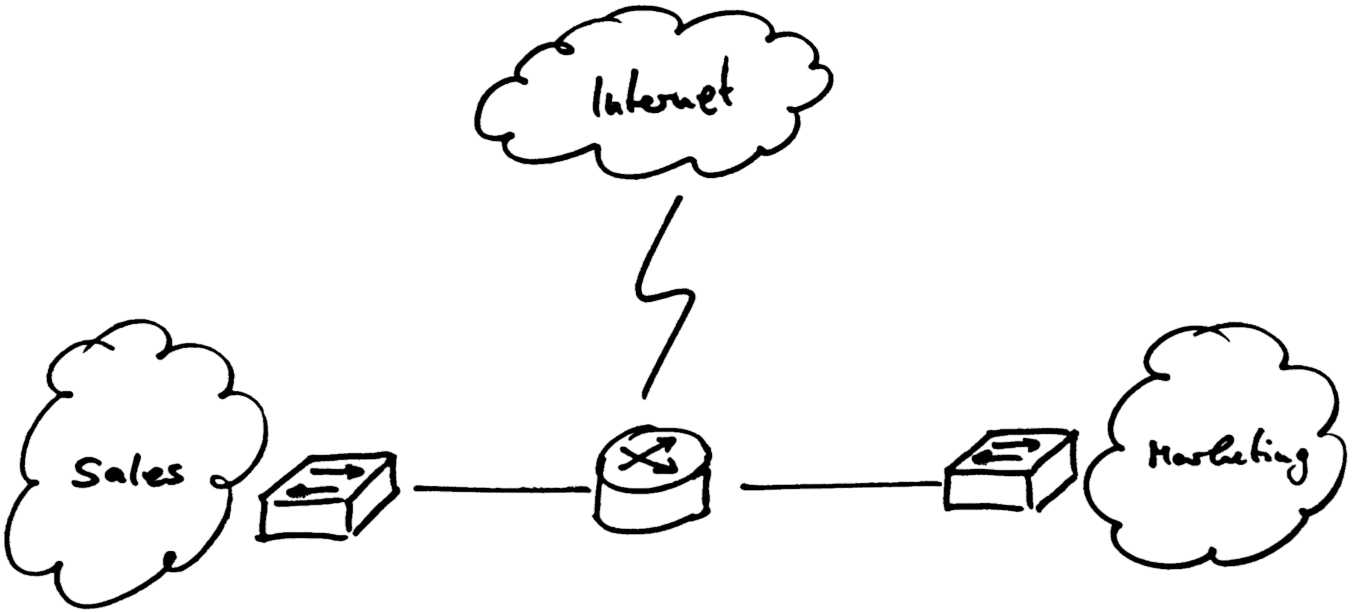
- **Smaller Broadcast Domains**

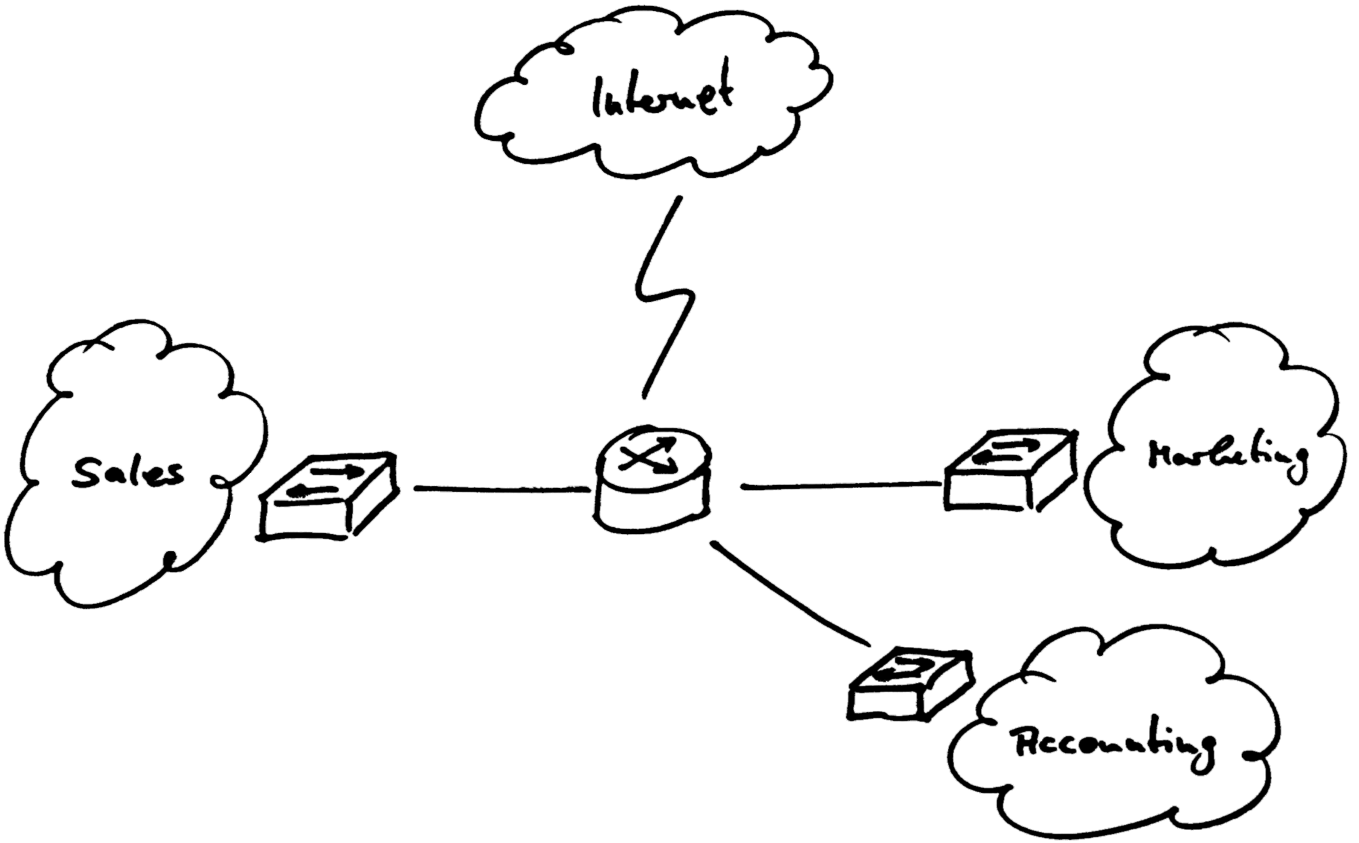
  Too many hosts sending broadcasts can have a negative impact on a segment's performance.

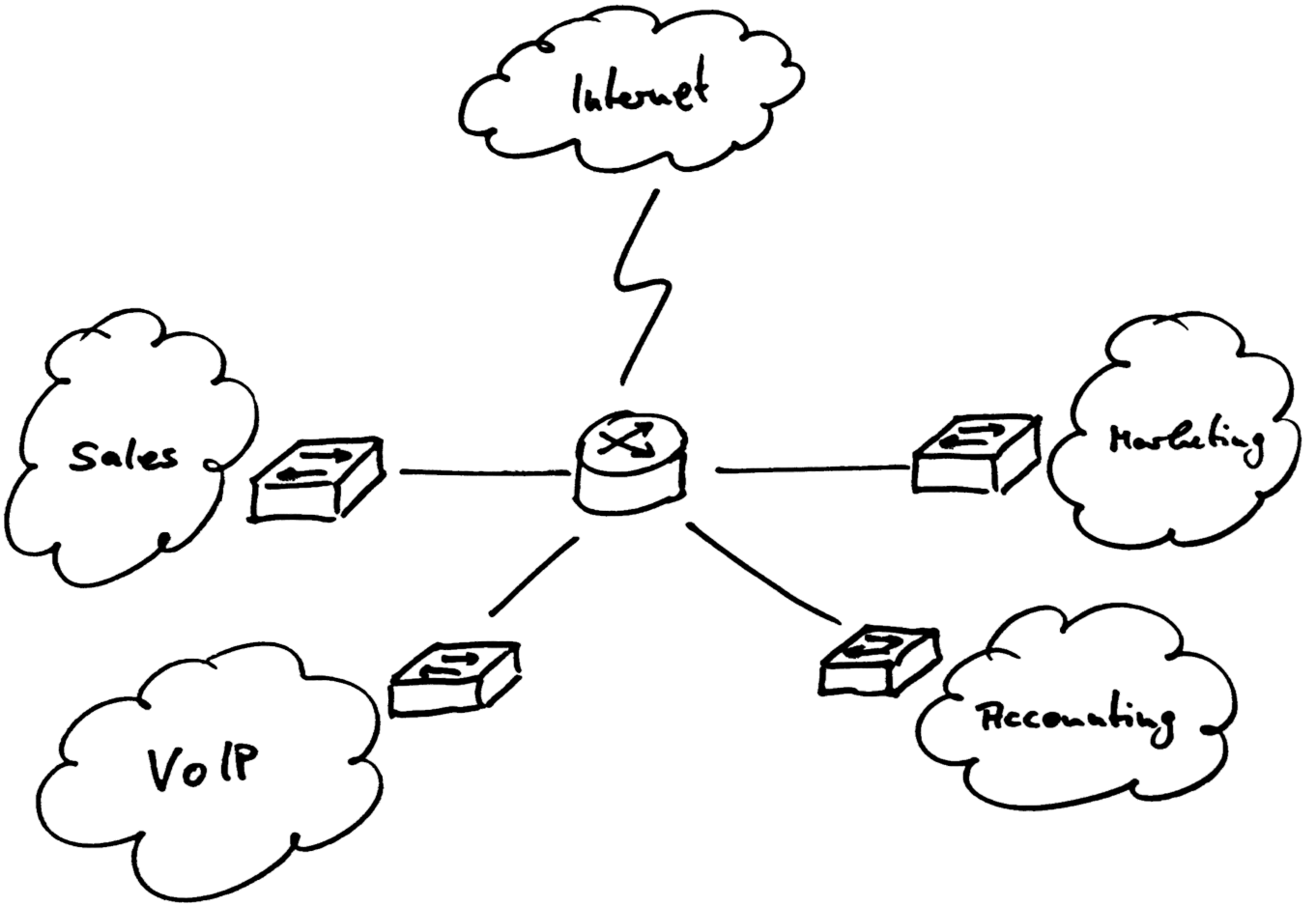
  A **broadcast domain** is a logical part of a network, in which all nodes can reach each other using **link-layer broadcasts**. Routers mark the boundaries of a broadcast domain, since they will not forward broadcasts from one link to another.

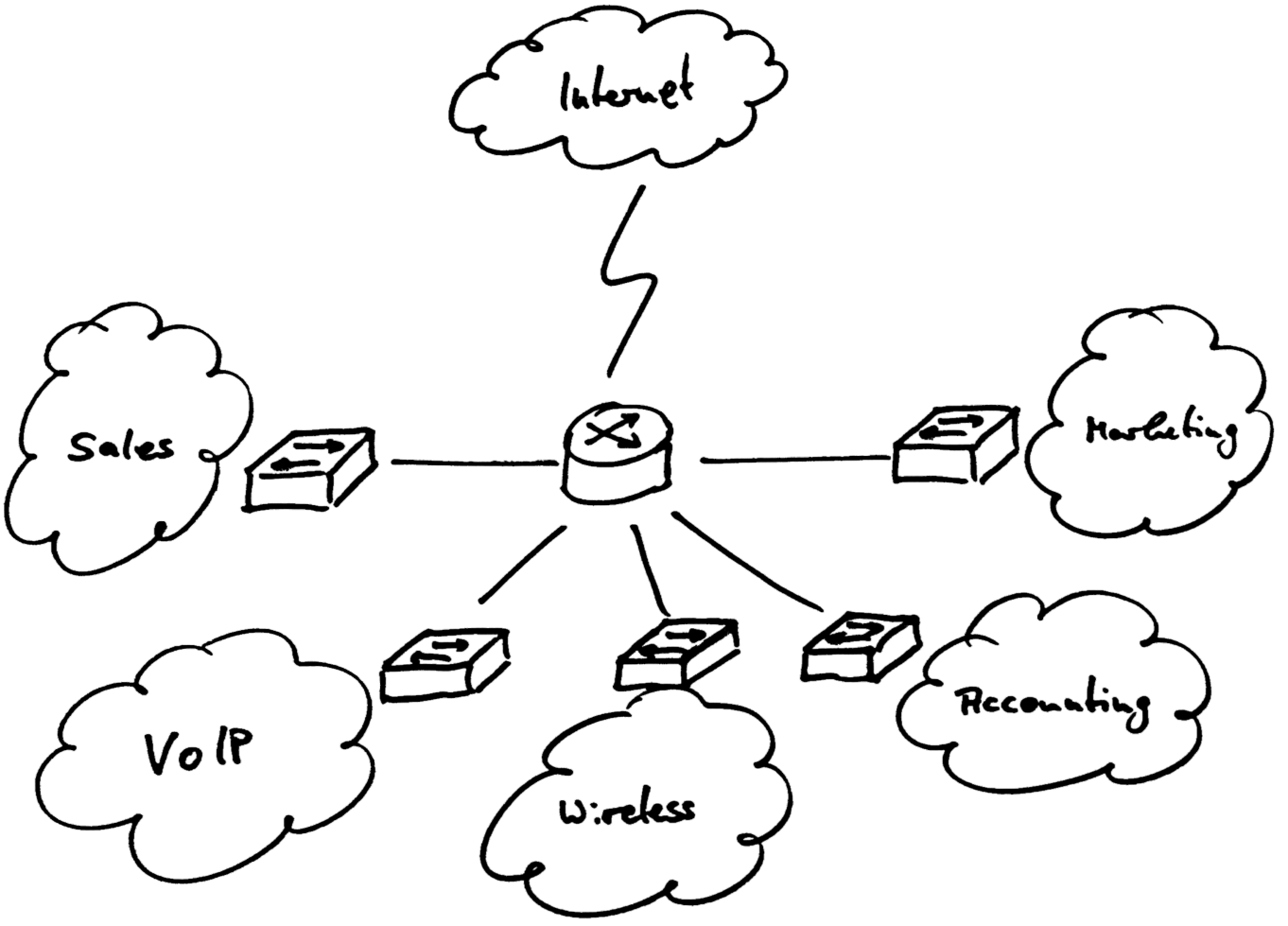
- **Separation of Concerns**

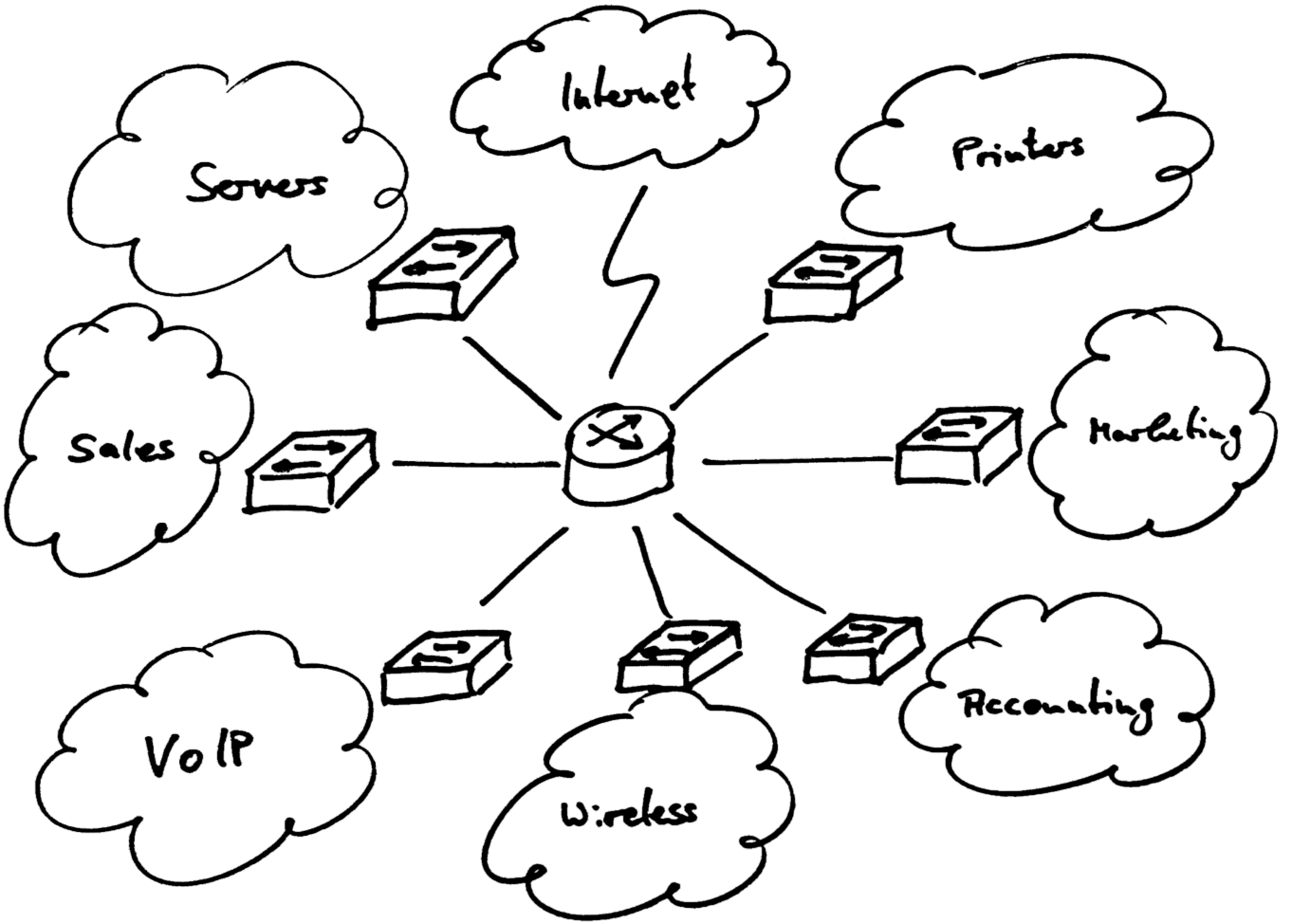
  - **Security Policies**
  - Avoid **interference** (e.g. file transfers vs. voice data)
  - **Optimize segment parameters** (e.g. for storage or VoIP networks)


Internet
Sales
Marketing


Internet
Sales
Marketing
Accounting


Internet
Sales
Marketing
VoIP
Accounting


Internet
Sales
Marketing
VoIP
Wireless
Accounting

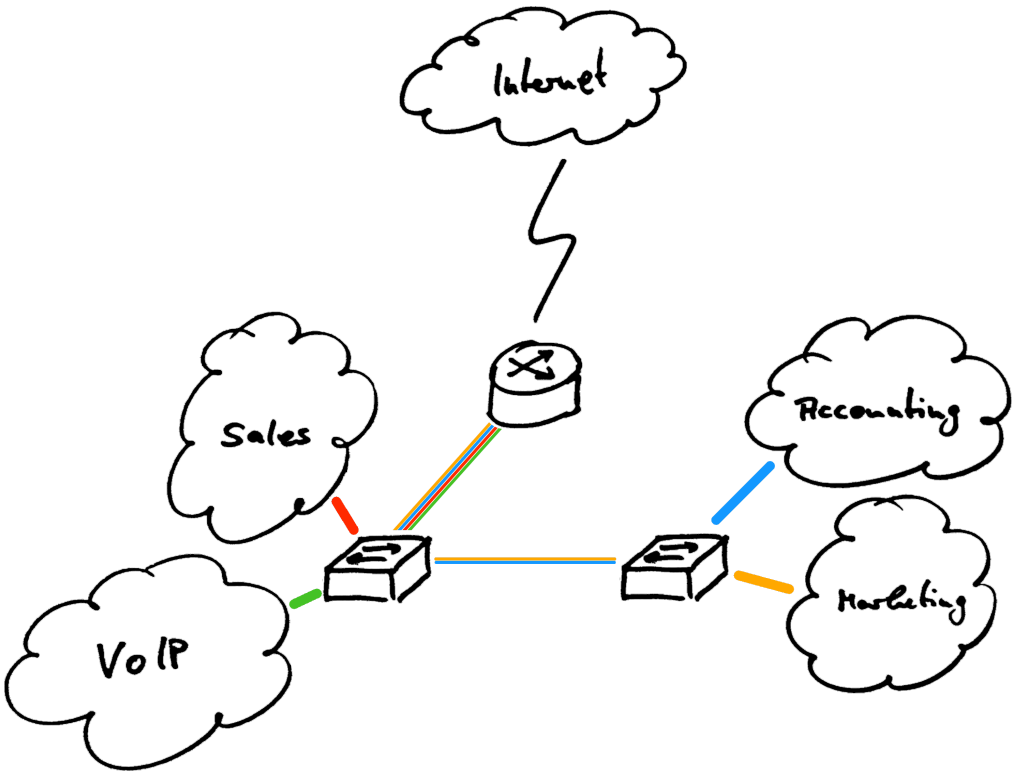

Internet
Servers
Printers
Sales
Marketing
VoIP
Wireless
Accounting

VLAN

(IEEE 802.1q)

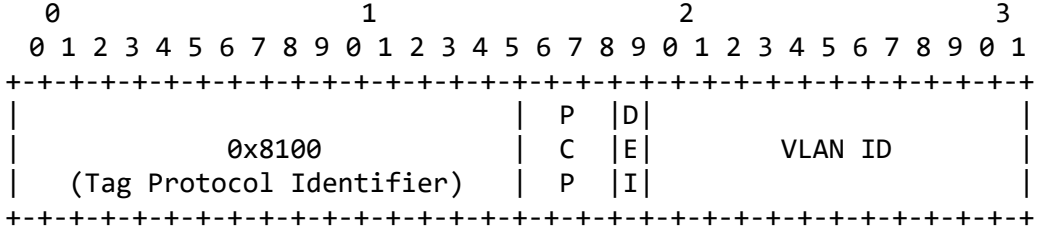
Sharing is Caring ...

Modern switches can use the **same physical link** for **multiple independent LANs**



VLAN (IEEE 802.1q)

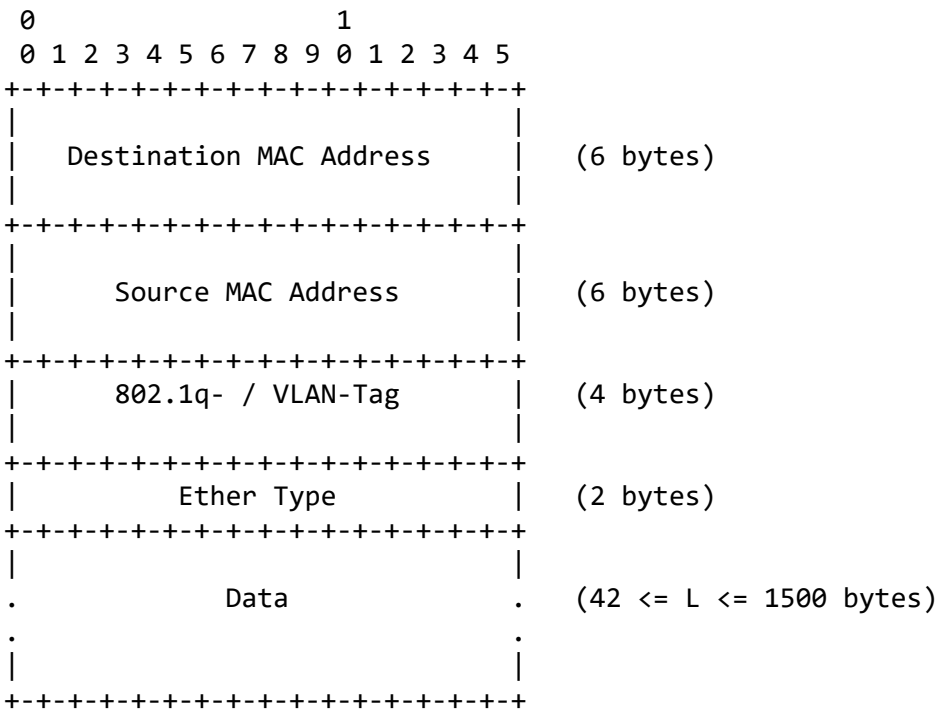
- Open standard to enable the use of virtual LANs between different vendors
- A 4 byte VLAN tag is inserted into the Ethernet frame
- Alternatives exist (e.g. Cisco ISL, QinQ/IEEE 802.1ad)



- **PCP (Priority Code Point)**
- **DEI (Drop Eligible Indicator)**
  If set, the packet may be dropped by the switch in case of congestion.
- **VLAN ID**
  Numeric ID of the VLAN.

PCP	Priority	Traffic Types
1	(lowest)	Background
0		Best Effort
2		Excellent Effort
3		Critical Applications
4		Video, < 100 ms latency and jitter
5		Voice, < 10 ms latency and jitter
6		Internetwork Control
7	(highest)	Network Control

Frame Format with 802.1q



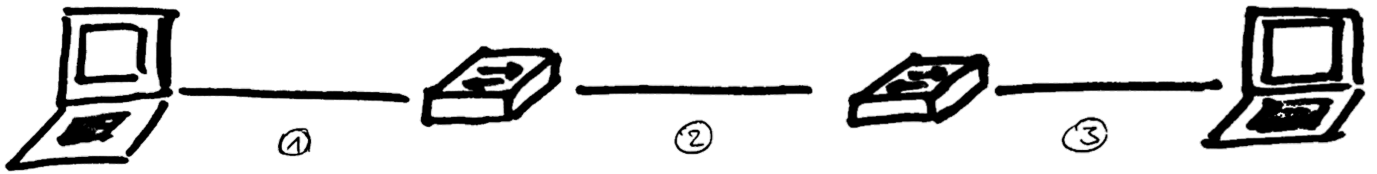
VLAN | Terminology

- **Access Port / Untagged Mode**
  Access Ports are a member of **one fixed VLAN**. Frames **entering** the switch on an access port are forwarded **only to ports belonging to the same VLAN**. Frames **exiting** the port **do not include a VLAN tag**. Workstations are typically connected to access ports.

- **Trunk Port / Tagged Mode**
  Trunk Ports are a member of **one or more VLANs**. Frames **entering** the switch include a VLAN tag and are forwarded only to ports belonging to **the VLAN ID specified in the tag**. Frames **without a VLAN tag** are assigned the **default/physical VLAN ID** of the switch/port. Trunk ports are typically used to connect to other switchs, servers or routers.

- **Default VLAN / Physical VLAN**
  A **default VLAN ID** used for **untagged frames received** by the switch **on a trunk port**. Depending on the switch model or vendor this can be a global default or a per-port setting. Often the default VLAN is used for network management traffic.

VLAN | Example



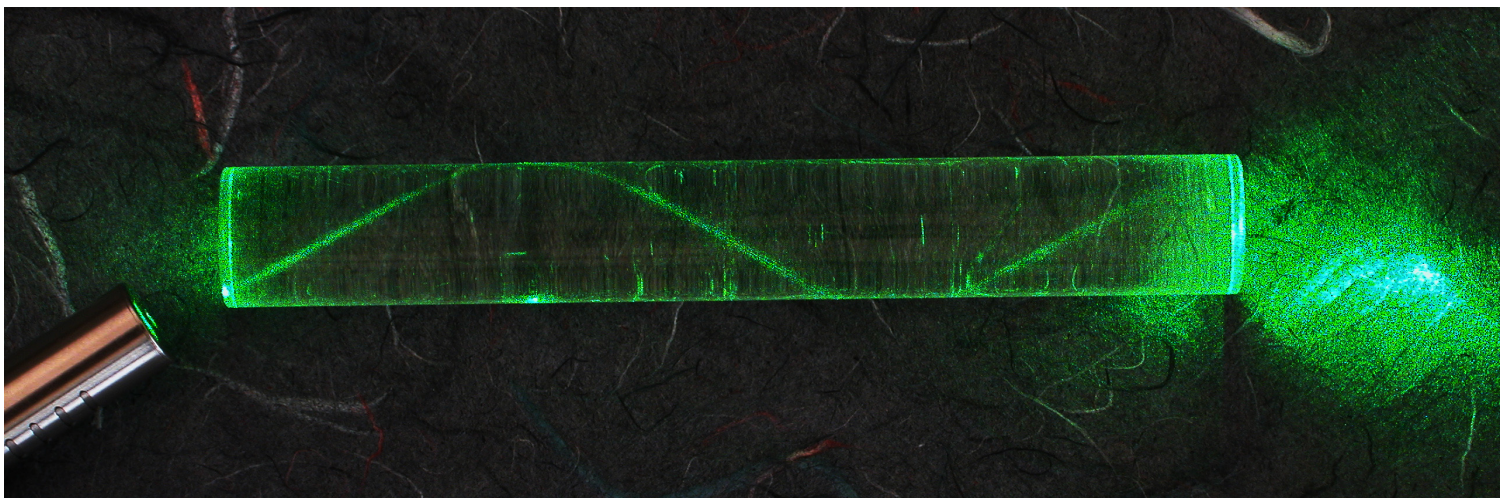
Frames on the Wire

```
    +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--...--+--+
(1) |     DST MAC     |     SRC MAC     |      Payload      |              (untagged)
    +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--...--+--+

    +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--...--+--+
(2) |     DST MAC     |     SRC MAC     |   802.1q  |      Payload      |   (tagged)
    +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--...--+--+

    +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--...--+--+
(3) |     DST MAC     |     SRC MAC     |      Payload      |              (untagged)
    +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--...--+--+
```

Optical Fiber

Optical Fiber

- Data Transmission using **light pulses** through glass or plastic fibers
- Typically **2 - 144 fibres** per cable
- Used in **Industry automation**, **sensor networks**, **MANs** and **WANs**
- Light wave **guided inside the medium** using **reflection principles**



Source: Wikimedia Commons

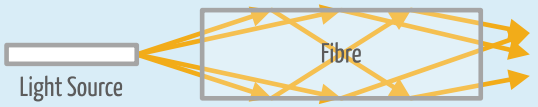
Optical Fiber

- Supports **higher bandwidths** than traditional electrical wires
- **Less attenuation loss** (allows transmission over **longer distances**)
- **Immune** to electromagnetic interference
- Takes up **less space** than electrical wiring
- **Each fiber** can carry **many independent channels** using **Wavelength Division Multiplexing**

- **Complex** and **costly** installation
- Requires **more expensive hardware**

Optical Fiber | Types

Multi mode



- **Core Diameter:** 50µm - 62.5µm
- **Distance:** <550m
- Light can travel using **multiple paths** (aka modes)

Single mode

- **Core Diameter:** 9µm
- **Distance:** several kilometers
- Light can only travel in **one confined path**

For an overview of the different fiber types see here.

Practical Fiber

The **multitude of fiber types and applications** makes it more practical to use **modular transceivers** for fiber optical transmission.

Practical Fiber

Optical Fiber | Wavelengths

Never check a fibre optic transceiver or cable by looking inside!

Light sources used in fibre-optic transceivers are typically **laser diodes**. You will not only **not see any light**, but maybe **damage your eyes** in the process.

The typical wavelengths for optical fibers lie between **850nm** and **1600nm**, which is **outside of the visible spectrum**.

Link-Layer Discovery Protocol

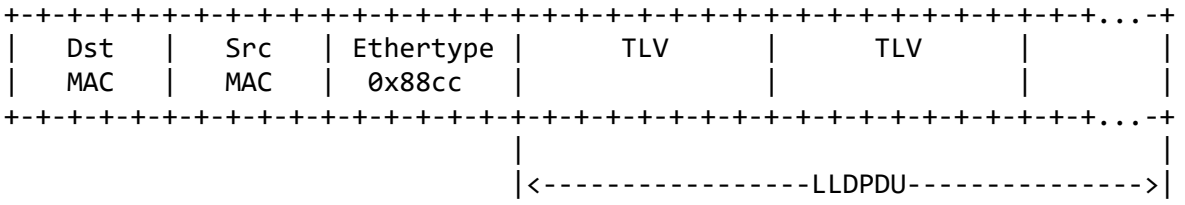
(IEEE 802.1ab)

LLDP

- Used by devices to **advertise their presence, identity and capabilities**, e.g.
  - VLAN name
  - Management IP Address
  - Power over Ethernet capabilities
- **Link-layer only**
- **"One-way"**
- **Vendor Neutral** (unlike CDP, NDP and LLTD)
- Supports **extensions**

- Used for:
  - Endpoint / Device Discovery
  - Basic device monitoring
  - Automated Network Mapping

Frame Structure

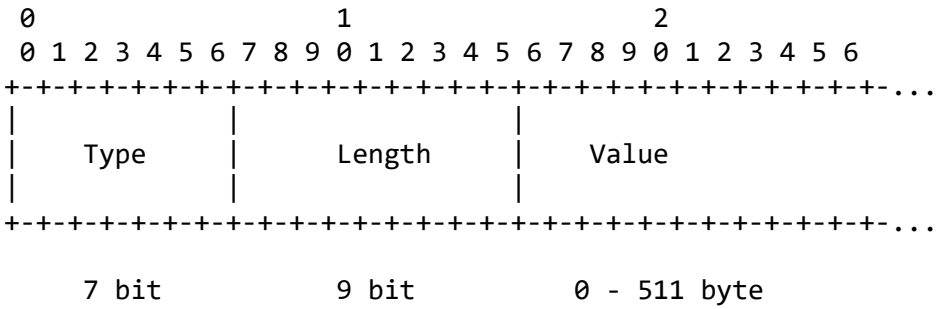


- Sent to well-known multicast addresses
- Information packaged in modules called **TLVs**
- The sequence of TLVs is called the **LLDP Data Unit (LLDPDU)**

Destination Addresses
01:80:c2:00:00:0e
01:80:c2:00:00:03
01:80:c2:00:00:00

Source Address
The MAC address of the sending interface.

Type-Length-Value (TLV)



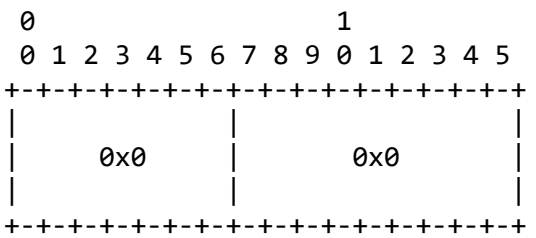
- Space conserving encoding for multi-type data
- Literally consists of a **type**, a **length** and a **value**
- **Type** specifies how to **interpret the value**
- The length is given in **byte**

LLDP TLVs

Type	Name	LLDP Usage
0	End of LLDPDU	Optional
1	Chassis ID	**Mandatory**
2	Port ID	**Mandatory**
3	Time to Live	**Mandatory**
4	Port Description	Optional
5	System Name	Optional
6	System Description	Optional
7	System Capabilities	Optional
8	Management Address	Optional
127	Organizationally Specific	Optional

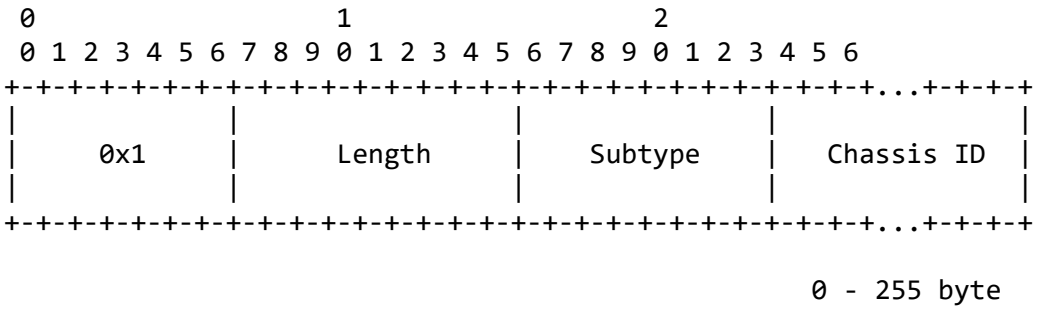
TLVs | EoLLDPDU

- Marks the **end of the TLV sequence**
- Two byte, **all-zeroes** TLV
- The EoLLDPU TLV is **always the last TLV** in the sequence



TLVs | Chassis ID

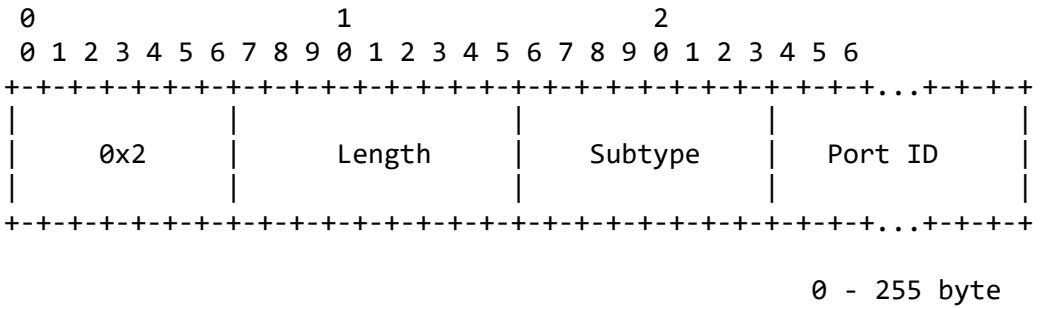
- ID of the **chassis** (i.e. device) associated with the transmitting LLDP agent
- Uses a subtype to further distinguish the type of ID



Subtype	ID Basis	Example
0	Reserved	
1	Chassis Component	cl-SJ17-3-006:rack1:rtr-U3
2	Interface Alias	office net
3	Port Component	backplane1
4	MAC Address	02:04:df:88:a2:b4
5	Network Address	134.96.86.110
6	Interface Name	eth0
7	Locally Assigned	Frank's Computer

TLVs | Port ID

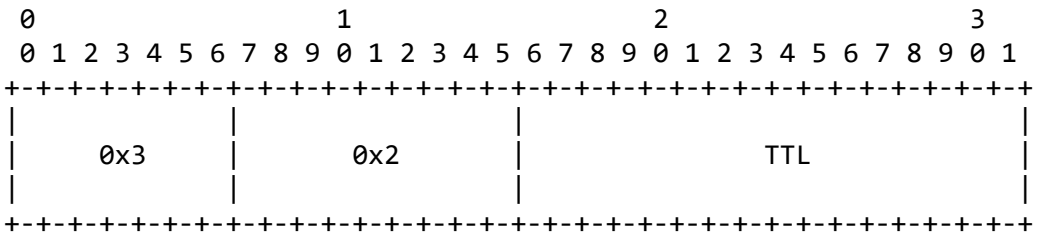
- ID of the **port/interface** associated with the transmitting LLDP agent
- Uses a subtype to further distinguish the type of ID



Subtype	ID Basis	Example
0	Reserved	
1	Interface Alias	office net
2	Port Component	backplane1
3	MAC Address	02:04:df:88:a2:b4
4	Network Address	134.96.86.110
5	Interface Name	GigabitEthernet12
6	Agent Circuit ID	
7	Locally Assigned	USB network interface

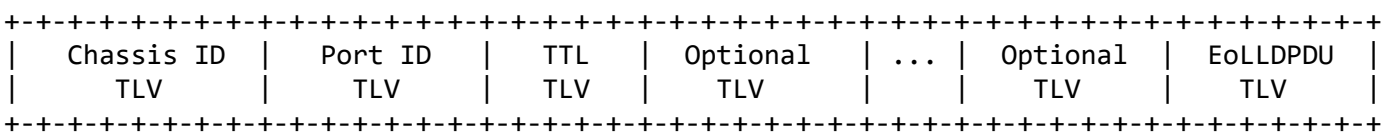
TLVs | TTL

- Number of seconds the receiver of an LLDPDU may regard the transmitted information as valid
- After the TTL has expired the information may be dropped from a receiving station's database



The LLDP Data Unit

- Every LLDPDU contain the following **ordered sequence** of mandatory TLVs
  - Chassis ID
  - Port ID
  - TTL
- The mandatory TLVs be followed by **zero or more optional TLVs**
- **Optional** TLVs be contained **in any order**
- The LLDPDU be **terminated with the EoLLDPDU TLV**



LLDP at the receiver

```
# show lldp info remote-device

 LLDP Remote Devices Information

  LocalPort | ChassisId                 PortId PortDescr SysName
  --------- + ----------------------- ------ -------- --------------------
  15        | 6c c2 17 c9 b0 d0       10     10       sw-wh-og
  24        | 70 54 d2 7b 7a db       eth0   eth0     fw
  25        | f0 92 1c 68 0f e0       26     26       sw-pb
  26        | f0 92 1c 68 7f c0       26     26       sw-nb
```

- Receiving LLDP agent keeps track of the transmitted data
- Data about neighbours can be queried by the user / administrator
- Entries will be purged once the TTL has exceeded and no further LLDPDUs have been received

Wrap-Up

Take-Home Messages

- Link Layer is responsible for:
  Frame Transport. **Error Handling**. **Media Access** and **Data Flow Control**.
- Ethernet defines a protocol for **guided transmission** (e.g. over coax, twisted pair, fiber).
- Defines **media access strategies (CSMA/CD)**.
- **ARP** is used to resolve **network-layer addresses into link-layer addresses**.
- Network segments are **connected using hubs and switches**.
- **LLDP** allows hosts to **announce themselves and learn about others**.

Further Reading

- Kurose-Ross "Computer Networking" (Sections 5.1 - 5.4)

- IEEE 802
- IEEE 802.3 (Ethernet)
- IEEE 802.1AB (LLDP)
- RFC826 (ARP)