# *LAN, Legacy Ethernet, Bridging*

LAN Introduction, IEEE Standards,
Logical Link Control (LLC), Ethernet Fundamentals,
Bridging Fundamentals, Bridging vs. Routing







## Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**

## LAN History

- **Local Area Network (LAN), invented late 70s**
  - Initially designed for a common transmission medium
    - Shared media
  - High speed
    - 4 Mbit/s, 10 Mbit/s, 16 Mbit/s, 100 Mbit/s
    - nowadays up to 10 Gbit/s
  - Limited distance
    - Up to some km -> hence local
  - High speed
    - Did not allow any network elements with store and forward behavior
  - Therefore simple topologies
    - Bus, ring, star
  - Base for distributed computing

Originally there was no packet switching or routing in the LAN world because of the high speed wanted to be achieved. Note: Ethernet bridging / Ethernet switching was invented as L2 packet switching technology in the late 80s.

The LAN technologies of these days allowed a high speed extension of internal computer bus for the first time in data communication history. Hence LANs became the basis for any kind of distributed computing and client-server computing.

Some basic LAN characteristics:

All network stations share the same media and all stations have equal rights. A station can directly communicate with all other stations of the same LAN. In HDLC we have seen already a shared media in the multipoint topology using MSD. But here the duties and rights of HDLC primary and secondary station were unbalanced. Remember only secondary stations have a L2 address but primary station - as one and only master - needs no address.

LANs have a natural broadcast behavior: A message sent out by one station reaches all other stations on same LAN. A LAN is therefore a multipoint line which needs addressing and access control.

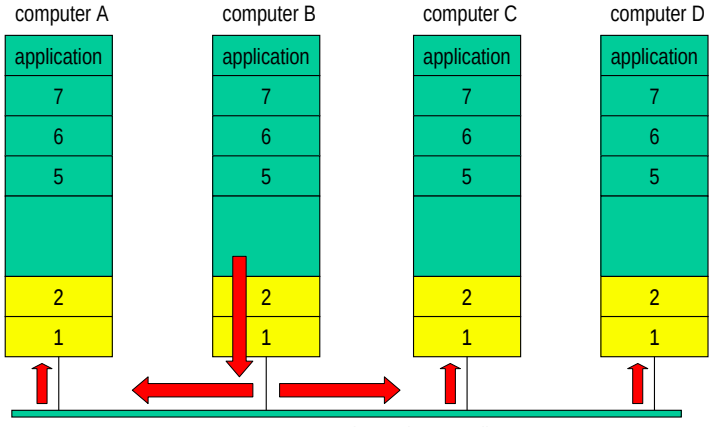
Therefore the terms MAC (media access control) and MAC address as unique but unstructured address are used in the LAN world. Initially there were no routing requirements and hence no need for structured addresses because store and forward (packet switching) done by CPUs was too slow. Layer 1 and layer 2 of the OSI model are sufficient to allow communication between systems connected to a single LAN. Have a look to the next slide to see the resulting protocol stack for LAN systems.





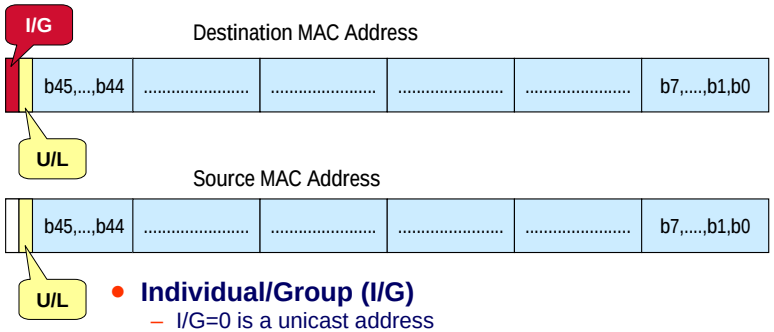


## LAN Communication



## 6 Byte MAC Addresses




- **Individual/Group (I/G)**
  - I/G=0 is a unicast address
  - I/G=1 is a group (broadcast) address
- **Universal/Local (U/L)**
  - U/L=0 is a global, IEEE administered address
  - U/L=1 is a local administered address

 LAN and Ethernet Fundamentals, v5.2 5

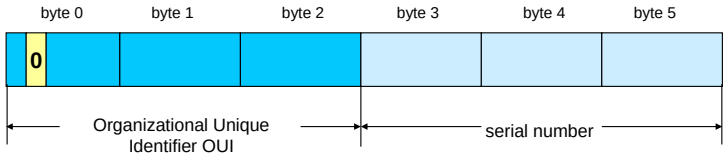
Every station on a LAN is identified by a unique MAC-address which may be used either as source or destination MAC-address in LAN frames.

A MAC address is 6 bytes or 48 bits long and is typically written in hexadecimal notation. The first two bits of a MAC address on the have a special meaning. The first bit (I/G) specifies whether the MAC address is a unicast address (0) or a broadcast/multicast address (1). A multicast is a broadcast for a group whereas broadcast addresses all stations on a single LAN. The broadcast-address is an address with all bits set to 1 (hex FFFF FFFF FFFF, U/L is set to 1). Please recognize that bit 47 (I/G) has no meaning in the source address of a LAN frame. The second bit (U/L) specifies whether it's a global and unique MAC address administered by the IEEE, or a local administered address.





## MAC Address Structure




- **Each vendor of networking component can apply for an unique vendor code**
- **Administered by IEEE**
- **Called "Burnt In" Address (BIA)**

 LAN and Ethernet Fundamentals, v5.2 6

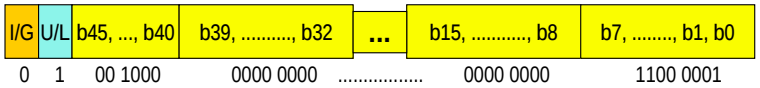
The MAC addresses are globally administered by the International Electrical and Electronic Engineering (IEEE) standardization organization.

Each vendor of networking components can apply for an globally unique vendor code. The vendor code costs 1000$ and occupies the first three bytes of the MAC address.

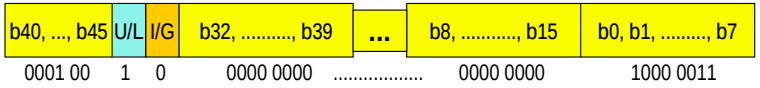
The remaining three bytes of the MAC address may be used by the vendor to address its components. Every Ethernet network card has one burnt in MAC address (BIA). Network cards of some vendors even support the use of programmable local administered MAC addresses.

## Storage Format of 802.3 MAC-Address

- **Basic rule:**
  - I/G bit must be the first bit on the medium, so the transmitted address must have the following format:




  - 802.3 sends the least significant bit of each octet at first
    - So 802.3 must store each octet in memory in reverse order:
    - also called **"Canonical" Format**




 LAN and Ethernet Fundamentals, v5.2 7

Ethernet is using a canonical address format, which defines the order how bits from the transmission buffer are put onto the medium. In Ethernet systems the least significant bit of each byte is put on the medium first followed by the more significant bits.

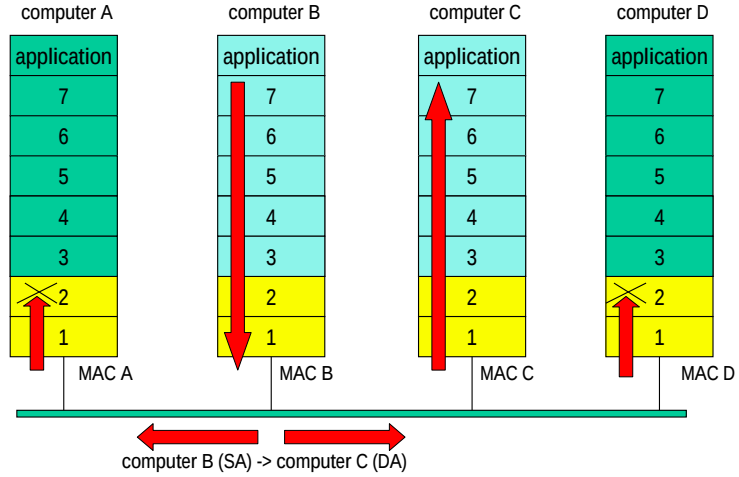
Remember that the first two bits of a MAC address on the **wire** have a special meaning (I/G and U/L) per definition. Hence if MAC addresses - as parameters for LAN communication - are stored in the memory of a computer system, they have to be stored in the right format.







## Direct Communication




 LAN and Ethernet Fundamentals, v5.2 8

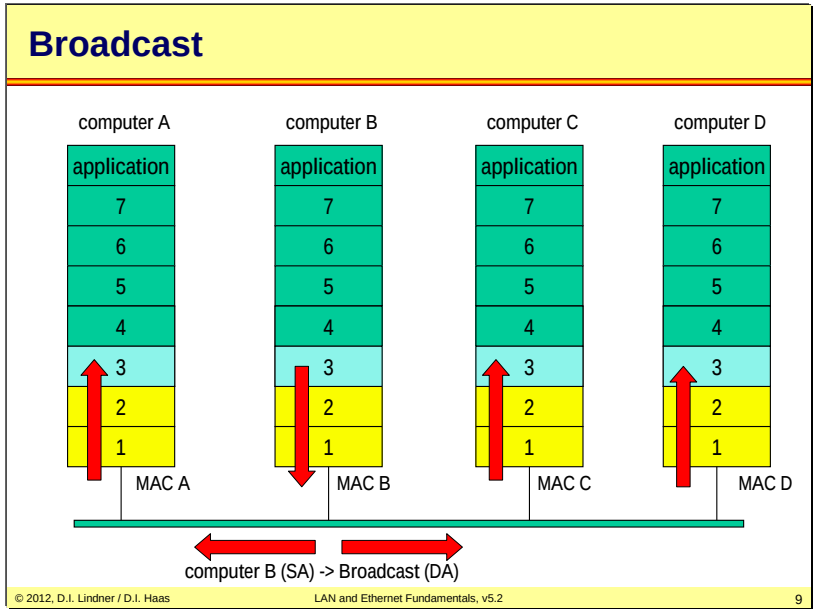
Receipt of frames:

Because of the inherent broadcast behavior of a LAN every frame is received by the Network Interface Card (NIC) of a station. The NIC decides if a frame should be forwarded to the higher layers (3-7) of a station depending on its own BIA and the destination address of the received frame. Usually NIC interrupts the CPU of the station if frame is to be forwarded. Otherwise the received frame is silently discarded by the NIC.

Frame are only forwarded to the higher layers (3-7) :

1. The destination address of the frame is equal with own BIA address
2. The destination address was a broadcast address
3. The destination address was a multicast (group) address and the given station is member of such a group.

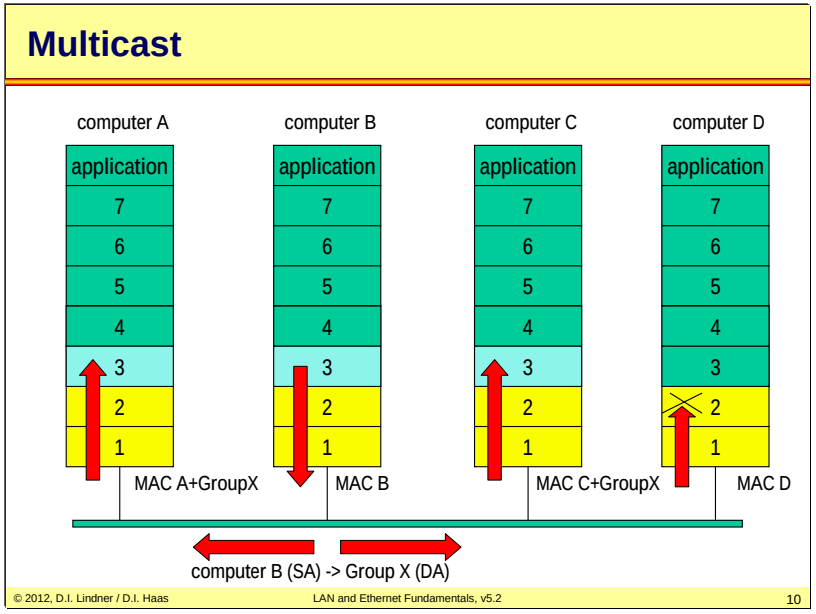
The later two are seen on the next two slides.




Keep in mind that for normal operation frames should be destined to a station specific MAC-addresses (direct communication), because broadcast frames will interrupt all stations for further handling by the higher layers. Even if it turns out that a station needs not to act on such a broadcast frame the CPU time of this stations will be wasted. Broadcast should be used in initialization phases of a network only!








In this example computer A and C are programmed to listen to group address X. Computer D will not be disturbed by this frame because it is not programmed to listen to X.

# Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**



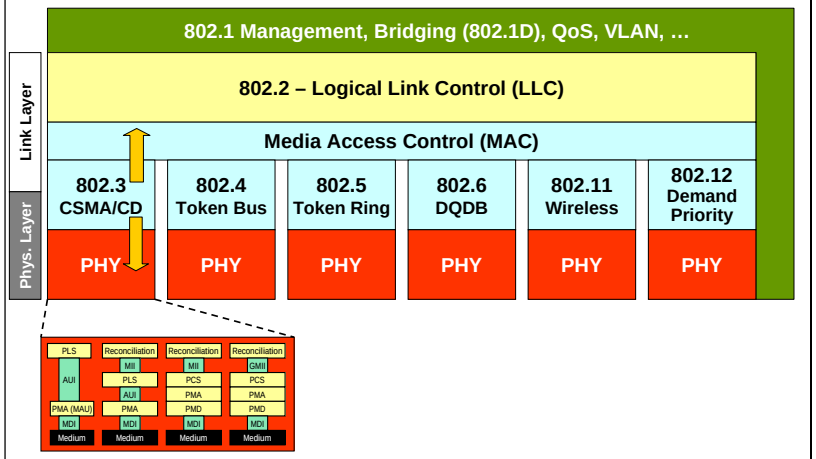




# IEEE 802

- **LAN Standardization is done**
  - By IEEE (Institute of Electrical and Electronics Engineers)
  - Workgroup 802 (Start: February 1980)
- **OSI Data Link Layer (Layer 2)**
  - Was originally designed for point-to-point line
  - But LAN = multipoint line, shared media
- **Therefore OSI Layer 2 was split into two sublayers**
  - Logical Link Control
  - Media Access Control

## IEEE 802 Layer Model






802.3 standard specifies “Ethernet”. The standard covers part of OSI layer 2 (MAC layer) and OSI layer 1 (physical Layer).

The physical layer of Ethernet is responsible for the speed of the transmission (currently there are four different speeds available, 10, 100, 1000, 10000 Mbit/s). In the graphic the physical interface structure of the 10, 100, 1000 Mbit/s systems is shown.

The interface function between the physical layer and the Ethernet data-link layer is performed by the CSMA/CD algorithm.

The Medium Access Control layer is responsible for addressing and it controls whether a data frame is picked up from the wire and is loaded into the buffer of the Ethernet card or not.

The Logical Link Control layer is responsible for the interface function between the Ethernet layer and higher layers on top of Ethernet plus the support of connection-less or connection-oriented mode.

The 802.1 Management cannot be seen as an separate Ethernet layer but it describes additional optional Ethernet functions like bridging, QoS, flow control, SPT, etc.





## The IEEE Working Groups

- 802.1 Higher Layer LAN Protocols
- 802.2 Logical Link Control
- 802.3 Ethernet
- 802.4 Token Bus
- 802.5 Token Ring
- 802.6 Metropolitan Area Network (DQDB, MAN)
- 802.7 Broadband TAG
- 802.8 Fiber Optic TAG
- 802.9 Isochronous LAN
- 802.10 Security
- 802.11 Wireless LAN (WLAN)
- 802.12 Demand Priority (VGAnyLAN)
- 802.13 Not Used — **Superstition?**
- 802.14 Cable Modem
- 802.15 Wireless Personal Area Network (Bluetooth)
- 802.16 Broadband Wireless Access
- 802.17 Resilient Packet Ring



On this slide you can see a summary of the most important IEEE standards so far. An Ethernet system is covered by the standards 802.1, 802.2 and 802.3.

The 802.1 describes management and optional functions inside the Ethernet technology like the Spanning-tree (STP) process, Ethernet bridging, VLAN systems, etc.

The 802.2 standards describes the Logical Link Control (LLC) function, which is only used in 802.3 Ethernet systems. With LLC it is possible to provide connection-oriented or connection-less mode service to the upper layers.

The 802.3 standard describes the physical layer of the Ethernet system plus the media access that is controlled by the CSMA/CD procedure.

Note: At the early days of Ethernet there were two competing organizations: the IEEE committee responsible for the 802.X standards and the companies Digital, Intel and Xerox (DIX) which where responsible for the Ethernetv2 aka DIX standard. In the year 1984 the DIX committee disappeared and the IEEE took over the responsibility to maintain and adapt the DIX standard for new upcoming Ethernet technologies.

Today all Ethernet interface cards support both frame types the 802.3 and the ETH 2 frame.

# Tasks of LAN Layers (1)

- Physical layer (PHY) specifies actual transmission technique
- Provides
  - Electrical/optical specification
  - Mechanical interface
  - Encoding
  - Bit synchronisation
- Consists of
  - MAU (Medium Attachment Unit)
  - AUI (Attachment Unit Interface)
  - PLS (Physical Layer Signalling)
  - Later expanded by PCS, PMA, PMD







# Tasks of LAN Layers (2)

- MAC (Media Access Control) takes care for medium access algorithms, framing, addressing and error detection
  - Avoid collisions
  - Grant fairness
  - Handle priority frames
- LLC (Logical Link Control) provides original services of data link layer
  - "HDLC on LAN"
  - Connection-oriented services with error-recovery
  - Connection-less service (best-effort)
  - SAPs (Service Access Points) for the higher layers

# IEEE 802.1 Standards

- **Specifies a common framework for all 802.x LANs**
  - Addressing rules, relations to the OSI model
  - Subnet addressing, Bridging Ethernetv2 to 802.2 LANs
  - Management (802.1B)
  - Bridging (802.1D-1998) including STP (Spanning Tree Protocol)
    - Single STP in case of VLANs
  - System Load Protocol (802.1E)
  - Virtual (V) LANs (802.1Q)
    - Tagging
  - STP Rapid Configuration (802.1w or 802.1D-2004)
  - Multiple STP (802.1Q-2003)
    - Multiple STP instances in case of VLANs
  - EAP Authentication (802.1x)
    - Extensible Authentication Protocol







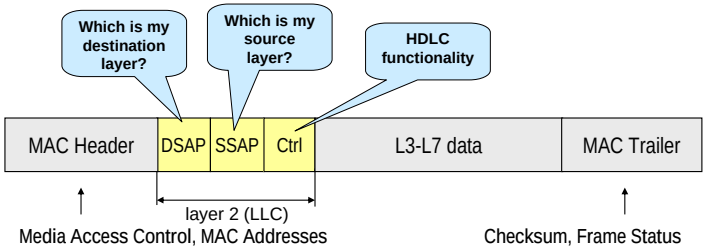
# Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**

## IEEE LAN Framing

- **Every IEEE LAN/MAN protocol carries the Logical Link Control header**
  - DSAP (Destination Service Access Point),
  - SSAP (Source Service Access Point)
  - Control Field = HDLC heritage

**Basic L2 frame format of every IEEE protocol**



The LLC (802.2) is part of every basic frame format that is specified by the IEEE e.g. Token ring, Token bus, Ethernet, etc.

The DSAP and SSAP field are both eight bit in length and are used to address layer 3 processes. With the SSAP the layer 2-3 interface used at the source is specified, while the DSAP specifies the layer 2-3 interface at the destination. But typically it is very unlikely to use a SSAP value different from the DSAP value, because only layer 3 processes of the same kind are able to communicate with each other. For IP to IP communication a SSAP and DSAP value of 0 x 06 is defined by IEEE.

The Control field inside the LLC can be used for connection-oriented or connection-less communication and the way it works is basically the same what HDLC does.

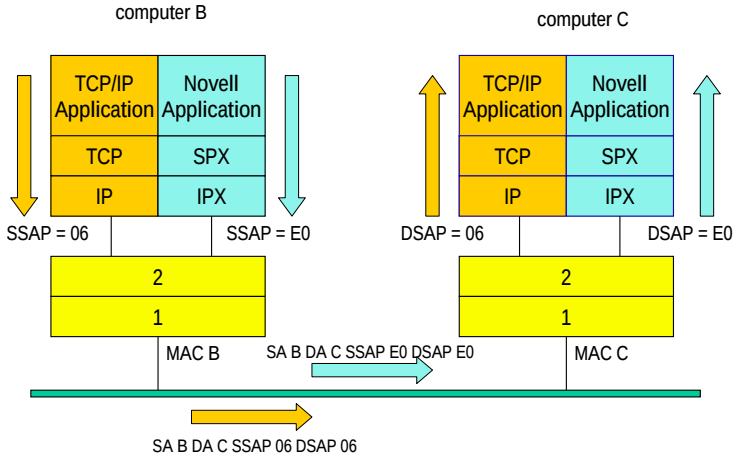
The connection-less mode of LLC is used by IP, IPX, AppleTalk. The connection-oriented mode of LLC is used by SNA over LLC Type 2 and NetBIOS over LLC Type 2 (NetBeui -> old style Microsoft Network –> already obsolete).







## Protocol Stack Distinction



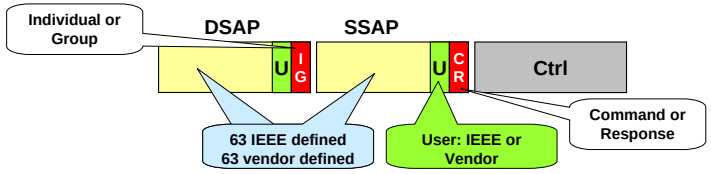

 

An IEEE 802 LAN can be used by different protocol families sharing the same communication media e.g. TCP/IP parallel to Novell IPX, IBM SNA, NetBeui, AppleTalk

DSAP and SSAP identify the higher level protocol family, which is the destination and the source of the given frame. They can be seen as protocol type or protocol stack identifier.

## SAP Identifiers

- **128 possible values for protocol identifiers**
- **Examples:**
  - 0x42 … Spanning Tree Protocol 802.1d
  - 0xAA… SNAP
  - 0xF0… NetBIOS
  - 0xE0… Novell
  - 0x06... IP (but not used because IP rides on top of Ethernetv2)






The DSAP and the SSAP are both 8 bit in length. The least significant bit in the DSAP is reserved to indicate whether it's a individual or group access point. In the SSAP this bit is the command/response bit which works together with LLC connection-oriented protocol mode. The U bit is used to specify whether its an IEEE or vendor specific access point.

Some examples:

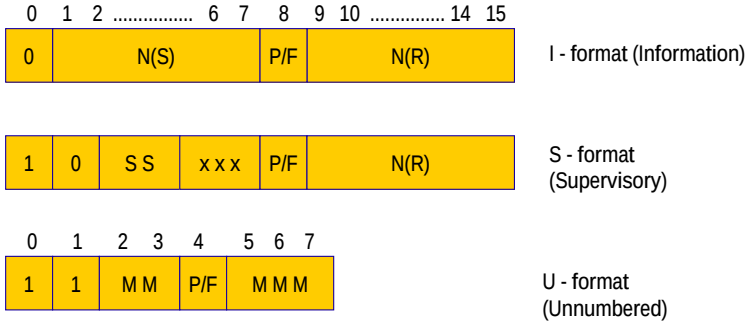
| | |
|---|---|
| Hex E0 ..... | Novell (U=0) |
| Hex Fy ..... | reserved for IBM (U=0) |
| Hex F0 ..... | NetBIOS (U=0) |
| Hex F4 ..... | IBM LAN manager individual (U=0) |
| Hex F5 ..... | IBM LAN manager group (U=0, I/G =1) |
| Hex F8 ..... | remote program load (U=0) |
| Hex 04 ..... | SNA path control individual (U=0) |
| Hex 05 ..... | SNA path control group (U=0, I/G =1) |
| Hex 03 ..... | LLC sub-layer management (U=1) |
| Hex 06 ..... | DoD IP (U=1) |
| Hex 42 ..... | 802.1d Spanning Tree Protocol (U=1) |
| Hex AA ..... | TCP/IP SNAP (U=1) |
| Hex FE ..... | ISO Network Layer (U=1) |
| Hex 00 ..... | Null SAP, a station with operating LLC software always responds to a frame destined to the Null SAP -> a kind of LLC Ping can be implemented |

The range Hex 8y to 9C (with U=0) is reserved for free usage except y = xx1x (binary notation); U=1





## LLC Control Field = HDLC Control Field




N(S), N(R) ..... send- and receive - sequence numbers
S S, MMM ..... selection bits for several functions
P / F ..... poll / final bit ( P in commands, F in responses; distinction of commands and responses through a dedicated SSAP bit -> C/R bit)



LLC Control field and protocol procedures are very similar to HDLC

remember: HDLC procedures allow connection-less and connection-oriented services on a layer 2 link

Connection-less mode of LLC is used by

IP, IPX, AppleTalk, etc

Connection-oriented mode of LLC was used by

SNA over LLC Type 2

NetBIOS over LLC Type 2 (NetBeui)

e.g. Microsoft Network (old style – already obsoleted)

## LLC Frame Types and Classes

| | Cmd | Control | Resp | Control | Class 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|
| Type 1 | UI | 1100p000 | | | x | x | x | x |
| | XID | 1111p111 | XID | 1111f111 | x | x | x | x |
| | TEST | 1100p111 | TEST | 1100f111 | x | x | x | x |
| Type 2 | I | 0 n(s) p n(r) | I | 0 n(s) f n(r) | | x | | x |
| | RR | 10000000 p n(r) | RR | 10000000 f n(r) | | x | | x |
| | RNR | 10100000 p n(r) | RNR | 10100000 f n(r) | | x | | x |
| | REJ | 10010000 p n(r) | REJ | 10010000 f n(r) | | x | | x |
| | SABME | 1111p110 | UA | 1100f110 | | x | | x |
| | DISC | 1100p010 | DM | 1111f001 | | x | | x |
| | | | FRMR | 1110f001 | | x | | x |
| Type 3 | AC0 | 1110p110 | AC0 | 1110f110 | | | x | x |
| | AC1 | 1110p111 | AC1 | 1110f111 | | | x | x |

 LAN and Ethernet Fundamentals, v5.2 23

Frames used for connectionless service, datagram service (type1):
- UI (Unnumbered Information) -> Datagram Info -> Best-Effort
- XID (Exchange Identification) -> LLC Ping
- TEST -> Ping plus test data

Frames used for connection-oriented service (type2): “HDLC in extended ABM mode”
- SABME (Set Asynchronous Balanced Mode Ext.)
  connection establishment
- UA (Unnumbered Acknowledgement)
  connection establishment acknowledgement
- DM (Disconnected Mode)
  negative acknowledgement for connection establishment or connection abort
- DISC (Disconnect)
  connection tear down
- I (Information)
  data frame
- RR (Receiver Ready)
  ACK plus station ready
- RNR (Rec. Not Ready)
  ACK plus station not ready
- REJ (Reject)
  NACK with GoBackN
- FRMR (Frame Reject)
  for signalling error situations

Frames additionally used for acknowledged datagram-service
- ACx command with data immediately acknowledged by ACx response, next ACy command allowed only after arrival of ACx
- Idle RQ protocol (stop and wait)

Note: Information what frame is command and what frame is response is taken from the C/R bit in the SSAP field.







## LLC Service Methods:

- Class 1:
  - Connectionless unacknowledged service (datagram)
  - Type 1 - frames: UI,XID,TEST
- Class 2:
  - Connection oriented service plus Class 1
  - Type 2 - frames: I,RR,RNR,REJ, SABME,UA,DM
- Class 3:
  - Class 1 plus connectionless acknowledged service
  - Type 1 -frames plus additional type 3 - frames: AC0, AC1
- Class 4:
  - Class 2 plus connectionless acknowledged service
  - Type 2 - frames plus additional type 3 - frames: AC0, AC1

 LAN and Ethernet Fundamentals, v5.2 24

Class 1 offers best effort service only, while Class 2 works connection-oriented with error recovery and flow control support.

The most important service class is the Class 1 connection-less service, because the tasks of error recovery and flow control are typically performed by higher layer processes e.g. TCP.

Only protocols like Microsoft's NetBeui or IBM's SNA need Class 2 connection-oriented service, because error recovery and flow control is not supported by their protocol stacks.

## Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**







*"Ethernet works in practice but not in theory."*

**Robert Metcalfe**

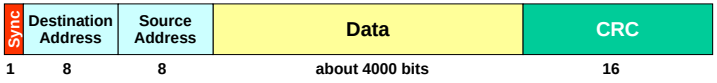


Yeah,...Robert Metcalfe was the inventor of Ethernet.

# History (1)

- Late 1960s: **Aloha protocol University of Hawaii**
- Late 1972: **Robert Metcalfe developed first Ethernet system based on CSMA/CD**
  - Xerox Palo Alto Research Center (PARC)
  - Exponential Backoff Algorithm was key to success (compared with Aloha)
  - 2.94 Mbit/s

**Original Ethernet Frame**






The Aloha protocol was fairly simple: send whenever you like, but wait for an acknowledgement. If there is no acknowledgement then a collision is assumed and the station has to retransmit after a random time. "Pure Aloha" achieved a maximum channel utilization of 18 percent. "Slotted Aloha" used a centralized clock and assigned transmission slots to each sender, hereby increasing the maximum utilization to about 37 percent. Robert Metcalfe perceived the problem: another backoff algorithm was needed but also "listen before talk". Metcalfe created Carrier Sense Multiple Access Collision Detection (CSMA/CD) and a truncated exponential backoff algorithm which allows a 100 percent load.

Robert Metcalfe's first Ethernet system used a transmission rate at 2.94 Mbit/s which was the system clock of the Xerox Alto workstations at that time. Originally, in 1972 Metcalfe called his system Alto Aloha Network, but one year later he renamed it into "Ethernet" in order to emphasize that this networking system could support any computer not just Altos – and of course to clarify the difference to traditional Aloha!

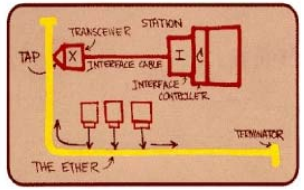




# History (2)

- 1976: **Robert Metcalfe released the famous paper: "Ethernet: Distributed Packet Switching for Local Computer Networks“**
  - In our context it would be better to call it “Asynchronous TDM on a Shared Wire” because there is no thing such an active packet switch

**Original sketch**

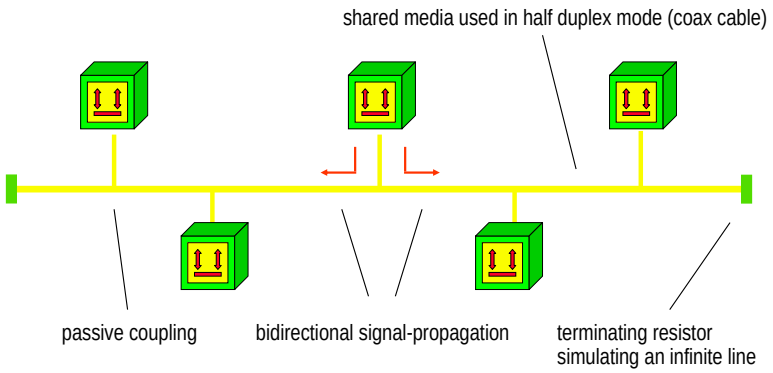





The press has often stated that Ethernet was invented on May 22, 1973, when Robert Metcalfe wrote a memo to his bosses stating the possibilities of Ethernet's potential, but Metcalfe claims Ethernet was actually invented very gradually over a period of several years. In 1976, Robert Metcalfe and David Boggs (Metcalfe's assistant) published a paper titled, "Ethernet: Distributed Packet-Switching For Local Computer Networks."

Metcalfe left Xerox in 1979 to promote the use of personal computers and local area networks (LANs). He successfully convinced Digital Equipment, Intel, and Xerox Corporations to work together to promote Ethernet as a standard. Now an international computer industry standard, Ethernet is the most widely installed LAN protocol.

## Basic Idea of Ethernet Bus System



Basic ideas:

Bus topology based on coax-cables with passive, uninterrupted coupling.

Shared media like the „Ether“ of air.

Bidirectional signal-propagation -> termination resistors avoid signal reflections.

Given transmitting power of a network station limits cable length and number of (receiver-) stations.

10 Mbit/s baseband transmission with Manchester encoding.





Datenkommunikation 384.081 - SS 2012

L06 - LAN and Ethernet Fundamentals (v5.2)

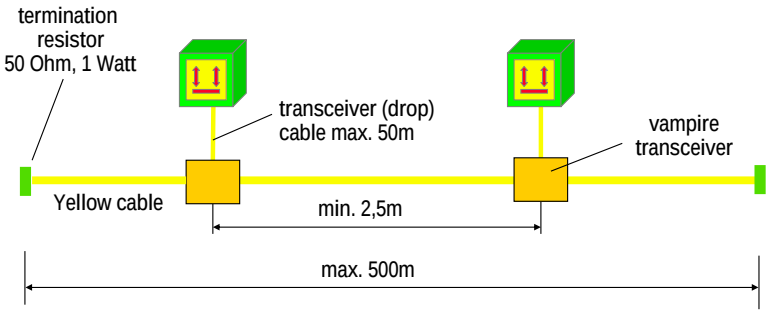
## History (3)

- **1978: Patent for Ethernet-Repeater**
- **1980: DEC, Intel, Xerox (DIX) published the 10 Mbit/s Ethernet standard (Ev1)**
  - "Ethernet II" was latest release (DIX V2.0)
- **Feb 1980: IEEE founded workgroup 802**
- **1985: The LAN standard IEEE 802.3 had been released**
  - Two types with baseband transmission with Manchester encoding, 10 Mbit/s
    - 10Base5 "Yellow Cable“
    - 10Base2 "Cheapernet"
  - One type with broadband transmission
    - 10Broad36 (modulation of carrier) -> like cable TV

First Ethernet standard was entitled "The Ethernet, A Local Area Network: Data Link Layer and Physical Layer Specifications" and focused on thick coaxial cable only.

## 10Base5 Parameter



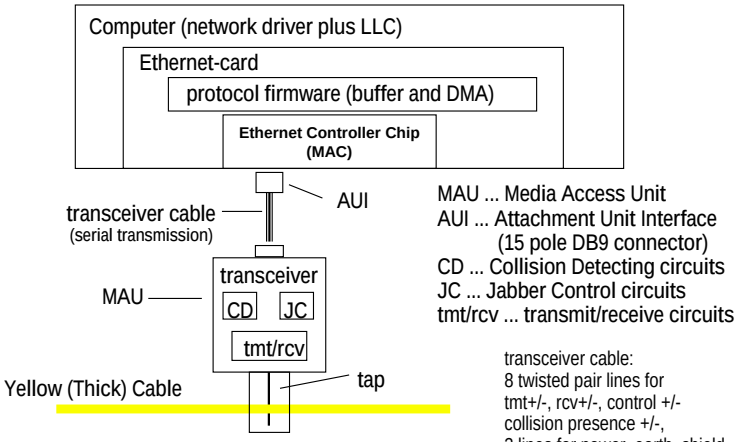

- maximal number of stations: 100
- attachable only at marked points
- cable splitting via coax couplers
- individual cable parts have a length of 23,4m or 70,2m or 117,5m (wave minimum on standing waves due to inhomogeneous media)
- smallest bending radius: 254mm

 LAN and Ethernet Fundamentals, v5.2 31





## AUI-Connection with 10Base5 Transceiver




 LAN and Ethernet Fundamentals, v5.2 32

Different transceiver types for 10Base5, 10Base2, FOIRL (Fiber Optic Inter Repeater Link) and 10BaseT, 10BaseF (these types will be handled later in this presentation).

Transmitter/Receiver = provides electronic circuits for:

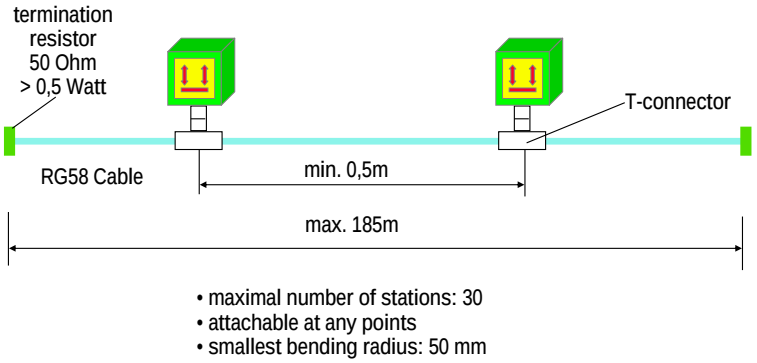
1. Inserting and receiving signal currents
2. Collision detection (measurement of DC level):

10Base5: Level High (1) = 0 mA, Level Low (0) = -80 mA -> DC of Manchester-encoded signal = -40 mA -> two signals at same time: DC Level < -40 mA

3. Heartbeat function
4. SQE Signal Quality Error
5. Jabber control (prevents continuously emitting of frames beyond the maximal frame size caused by an erroneous Ethernet controller)

External transceiver: AUI interface (with or without transceiver cable) connects end system and transceiver; transceiver powered by end system

## 10Base2 Parameter



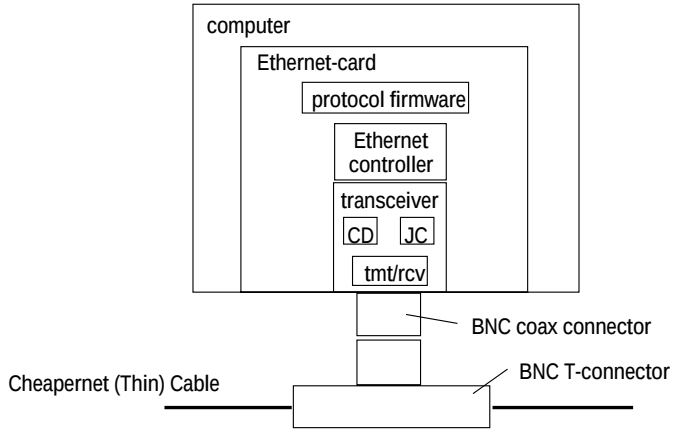

- maximal number of stations: 30
- attachable at any points
- smallest bending radius: 50 mm







## Integrated Transceiver for 10Base2






Integrated transceiver: Transceiver is integrated on network card of end system network card provides necessary physical connector like BNC (10Base2), RJ45 (10BaseT), ST (10BaseF)

## Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**







## CSMA/CD (1)

- **Carrier Sense Multiple Access / Collision Detection**
  - Improvement of ALOHA
  - "Listen before talk" plus
  - "Listen while talk"
- **Fast and low-overhead way to resolve any simultaneous transmissions**

1) Listen if a station is currently sending
2) If wire is empty, send frame
3) Listen during sending if collision occurs
4) Upon collision stop sending
5) Wait a random time before retry



Ethernet is a shared media technology, so a procedure had to be found to control the access onto the physical media. This procedure was called the Carrier Sense Multiple Access Collision Detection (CSMA/CD) circuit.

The way it works is quite simple, every stations that wants to send need to do a Carrier Sense to check if the media is already occupied or not.

If the media is available the station is allowed to perform an Media Access and may start sending data.

In the case that two stations almost at the same time access the media, a collision will happen. To recognize and resolve a collision is the task of the Collision Detect circuit.

Every station listens to its own data while sending. In the case of a collision the currently sending stations recognize the collision by the superimposition of the electrical waves on the wire. A jamming signal will be sent out to make sure all involved stations recognize the occurrence of an collision.

All stations involved in the collision stop sending and start a randomize timer. When the randomize timer expires the station may try to access the media again.

## CSMA/CD (2)

- **Details:**
  - Access control based on contention
  - Network stations listen to the bus before they start a transmission
  - Network stations can detect ongoing transmission (CS) and will not start own transmission before ongoing transmission is over
  - But still simultaneous transmissions (MA) cause collisions
    - Collisions are detected (CD) by observing the DC-level on the medium
  - Bus conflict



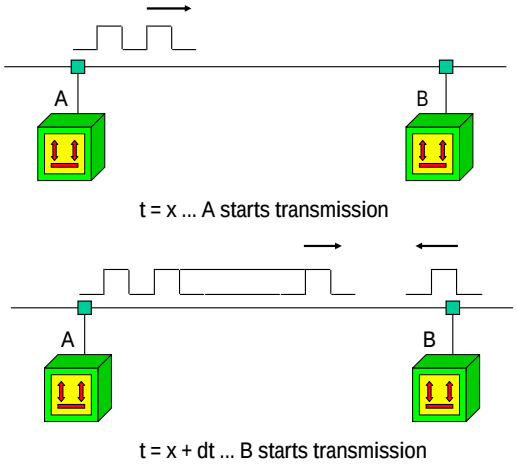




## CSMA/CD (3)

- **Conflict resolution:**
  - Aborting of transmission by all involved stations
  - Sending of a JAM-signal (32 bit)
    - To make sure that every station can recognize the collision
    - Collision is spread to a minimum length
  - Starting a random number generator to create a timeout value
    - Truncated binary exponential backoff algorithm (the more often a collision occurs the larger is the range for the random number)
  - After timeout expired, station attempts a retransmission
  - Number of retransmission-trials is limited to 16
    - After 16 collisions in a sequence a error is signaled to the higher layer

## Collision Window / Slot Time 1

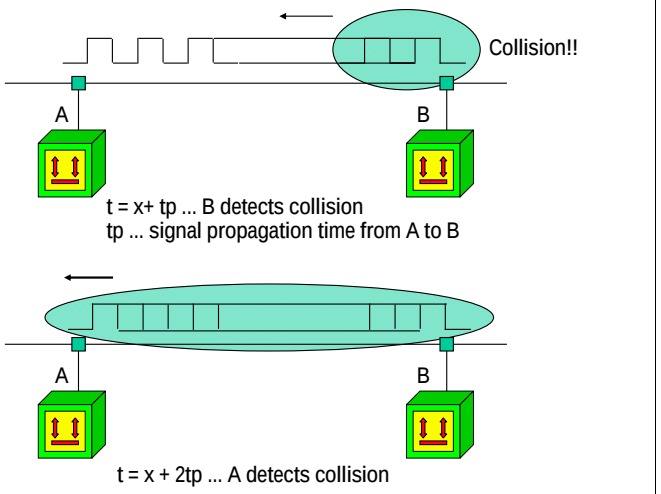









## Collision Window / Slot Time 2






In the worst case stations have to send bits twice the maximum signal propagation time (RTT) for reliable collision detection. Otherwise a collision may not be seen by the transmitting station for the currently transmitted frame.

## Slot Time

- **Minimum frame length has to be defined in order to safely detect collisions**
- **Each frame sent must stay on wire for a RTT duration – at least**
- **This duration is called "slot time" and has been standardized to be 512 bit-times**
  - 51,2 µs for 10 Mbit/s
  - Hence minimal frame length is 64 byte
  - Note:
    - The request for fairness limits the maximum frame size, too
    - 1518 byte is the maximum allowed frame size

 LAN and Ethernet Fundamentals, v5.2 41

There is a very basic Ethernet rule that says a collision must be detected while a station is transmitting data. Therefore a stations needs to keep on sending at least of the duration of the RTT of the Ethernet system. The maximum allowed RTT is standardized and is called the slot time. The slot time for 10Mbit/s Ethernet systems is set to 51,2 µs.

If collisions occur after expiration of the slot time we talk about “late collisions”, which may cause malfunctions in the network.

For example if a station transmits a frame and no collision was detected, the station assumes correct delivery of the frame. Now the station removes the frame from the transmit buffer, leaving no chance to retransmit the frame in the case of a late collision.





## Slot Time Consequences

- **So minimum frame length is 512 bits (64 bytes)**
- **With signal speed of 0.6c and the delay caused by electronic circuits such as interface cards and repeaters the RTT of 512 bit times allows a network diameter of**
  - 2500 meters with 10 Mbit/s
  - 250 meters with 100 Mbit/s
  - 25 meters with 1000 Mbit/s (!)

**NOTE: Only valid on shared media (!)**

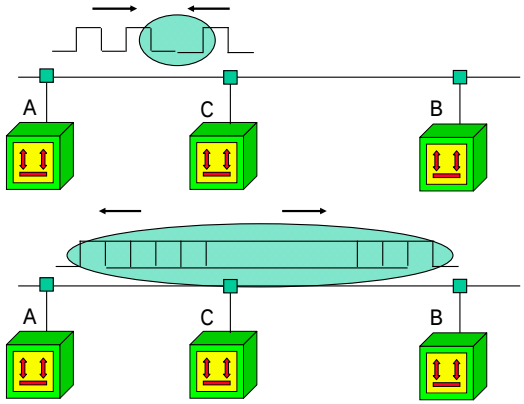
 LAN and Ethernet Fundamentals, v5.2 42

The minimum frame length in Ethernet systems is set to 64 byte or 512 bit. This minimum frame length plus the slot time in combination with the speed of electrical signals on a wire (~ 180.000 km/s) determines the maximum outspread of an Ethernet system.

Therefore we end up at a maximum outspread of 2500 meters for 10Mbit/s Ethernet systems. The maximum outspread of faster Ethernet systems is directly related to their shorter slot times, because of the higher speed.

These distance limitations must only be taken into account in shared media environments like Ethernet Bus and Hub systems. In more modern switched environments using full duplex communication these distance limitations can be neglected.

## Collision Extended by JAM Signal !




 LAN and Ethernet Fundamentals, v5.2 43

The picture shows how a collision between closely located stations is extended by the jam signal to a minimal length in order to be recognized by network station too. Without jam signal the collision would cause A and C to immediately stop their transmission and that short collision (small signal spike on the media) may not be recognized by the receiving circuits of all other stations.





## Exponential Backoff

- **Most important idea of Ethernet !**
- **Provides maximal utilization of bandwidth**
  - After collision, set basic delay = slot time
  - Total delay = basic delay * random
  - 0 <= random < 2^k
    - k = min (number of transmission attempts, 10)
- **Allows channel utilization**
- **After 16 successive collisions**
  - Frame is discarded
  - Error message to higher layer
  - Next frame is processed, if any
- **Truncated Backoff (k<=10)**
  - 1024 potential "slots" for a station
  - Thus maximum 1024 stations allowed on half-duplex Ethernet

 LAN and Ethernet Fundamentals, v5.2 44

The retransmission in case of collisions is controlled by the exponential backoff algorithm.

The retransmission is delayed about a basic delay, which is set to 51,2 microseconds for 10 Mbit/s Ethernet, times a random factor. The range out of which the randomize factor is selected is increasing with the number of retransmission attempts. Repeated collisions indicate a busy medium, therefore the station tries to adjust to the medium load by progressively increasing the time delay between repeated retransmission attempts.

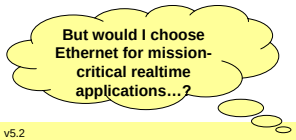
The retransmission of a frame is attempted up to a defined maximum number of retries typically known as the attempt limit. The attempt limit is set to a maximum of 16 retries by the standard.

After 16 retries the frame is discarded and a error message is sent to higher layers. Then the station continuous to process the next frame.

Due to the truncated backup algorithm a maximum of 1024 potential time slots for a station are available. So the maximum number of stations attached to half duplex Ethernet systems should not exceed 1024 stations.

## Channel Capture

- Short-term unfairness on very high network loads
- Stations with lower collision counter tend to continue winning
- 10 times harder to occur on 100 Mbit/s Ethernet
- Rare phenomena, so no solution against it






In the case of very high network loads Ethernet tends to prefer stations with lower collision counters, because they try to access the media in shorter time intervals than stations with a higher collision counter.

This is a phenomena that was never solved in Ethernet systems, but can be disregarded for today's Ethernet networks, because most of them are switched networks where collisions play no or just a minor role.

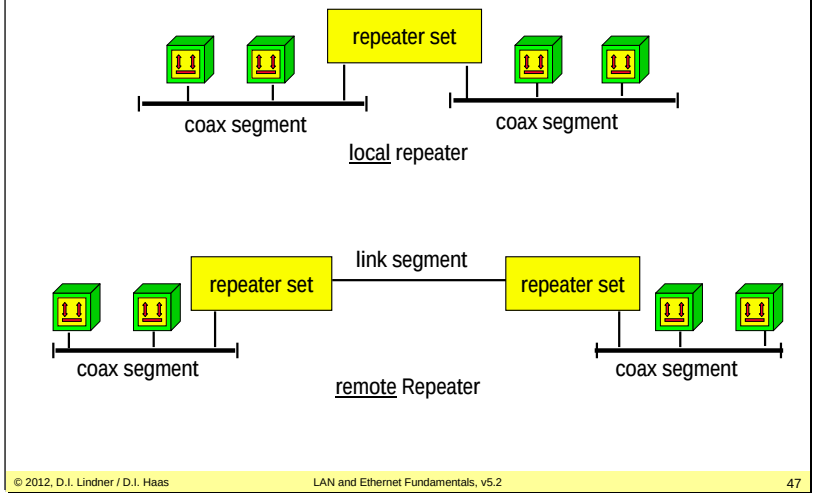




## Agenda

- Introduction to LAN
- IEEE 802
- Logical Link Control
- Legacy Ethernet
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- Transparent Bridging
- Bridging versus Routing

## Local / Remote Repeater




 LAN and Ethernet Fundamentals, v5.2 47

Repeater is an amplifier expanding the maximal distance of an Ethernet-LAN segment. It regenerates signals on the receiving port, amplify them, and send these signals to all other connected net segments (no buffering, just a short delay, which must be taken into account for the collision window / slot-time). In case a collisions is detected all other ports are notified by jam-signal. Optionally auto partition on erroneous ports may be performed by a repeater.

Important: Collision domain is preserved by repeaters.

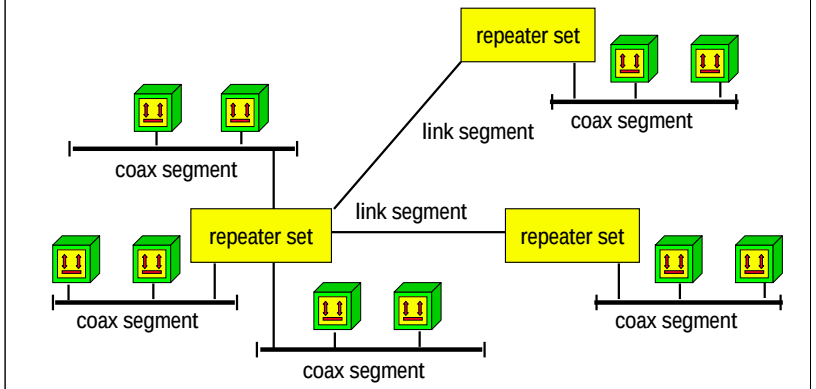
Local repeaters directly connect two (coax) segments.

Remote repeaters are connected by so called link segments.





## Multiport Repeater - One Collision Domain




 LAN and Ethernet Fundamentals, v5.2 48

Link segment: a physical point-to point connection between two devices

First link segments were used for repeater interconnection only. Several types were defined (fiber based, copper based):

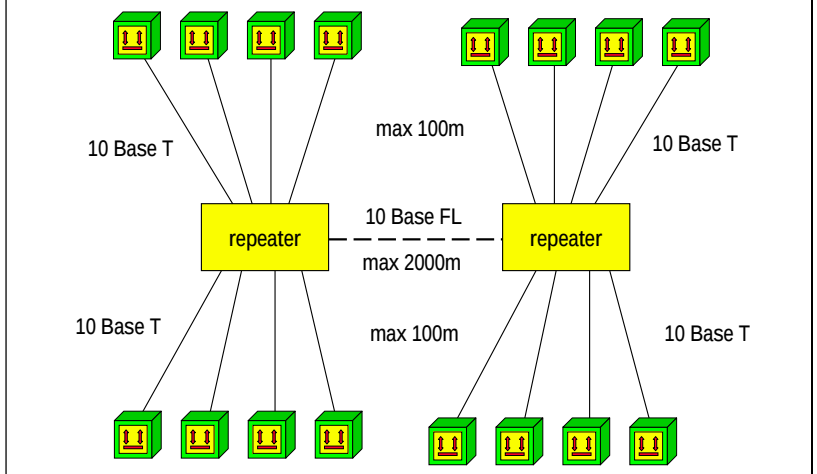
FOIRL (Fibre Optic Inter Repeater Link: maximal length 1000m, for repeater - repeater)

10BaseFL (asynchronous, maximal length 2000m for repeater - repeater, end system - multiport repeater)

10BaseFB (synchronous (idle signals during communication pauses), maximal length 2000m, for repeater - repeater links only

10BaseFP (passive hub, no active repeater function, remark: active means electrically powered)

## Multiport Repeater as „Hub“




 LAN and Ethernet Fundamentals, v5.2 49

The link segment was later also defined for connection of a network station (end system) to a multiport repeater using a dedicated physical point-to-point line.

The reason: Ethernet was originally based on coax cabling and bus topology which was hard to wire in a building. Later an international standard for **structured cabling** of buildings was defined which was star wired to (a) central point(s) based on twisted pair cabling. That excellently fits to Token ring cabling at that time. Ethernet had been adapted to that in order to survive. Otherwise Token ring would have won the LAN battle.

The first implementation of such link segment was 10BaseT (unshielded twisted pair, maximal length 100m, 2 lines Tmt+-, 2 lines Rcv+-, RJ45 connector, Manchester-Code with no DC offset). In such an environment the collisions can not be detected any longer just by measurement of the DC level as done in 10Base5 because tmt and rcv travels on different physical lines . Now a collision is interpreted, if signals are on the tmt and rcv line at the same time. The repeater has to watch out, if two or more signals are received at the same time (-> that means collision in the LAN). Now the hub has to produce a Jam signal on all ports in order to signal a collision to all systems.

We can see: the method of collision detection is different for every physical layer.

Here are the details:

In coaxial Ethernet, transceivers send their Manchester code using the DC offset method. A "high" value is nominally zero current; a "low" value is nominally -80 mA. This results in a DC component to the signal of -40 mA, which creates a voltage of -1 VDC (the transceiver sees a 25 ohm load from the two 50 ohm cables going "left and right" away from the transceiver). When two transceivers send at the same time, their currents add, increasing the DC component of the combined signal to -2 VDC Thus, we can detect collisions by looking for DC signals in excess of the maximum that could possibly be generated by a single transmitter.

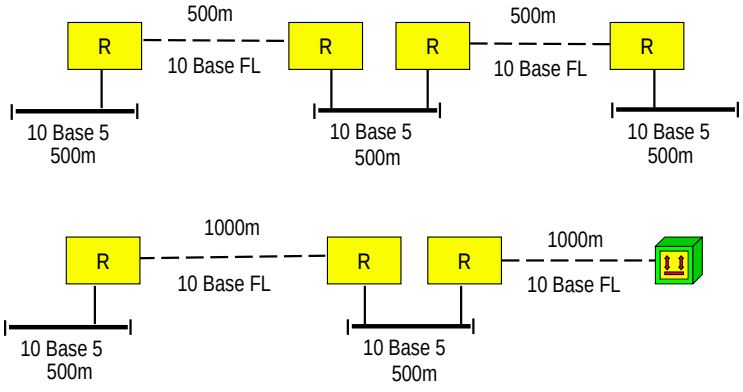
In 10BaseT, the Manchester code is sent symmetrically, with no DC offset. Collisions are detected in the repeater hub, which can observe when two or more devices are transmitting at the same time. Normally, the hub does not repeat a station's own signal back to the station on its receive cable pair. However, when a collision is noted, the hub does send a signal (the so-called "collision enforcement", or "jam") to the transmitting stations. The stations detect collisions by noting when they see a signal on their receive pair at the same time that they are transmitting on their transmit pair.

A repeater with more than two segments and different physics is called a multiport repeater. Multiport repeater in a star like topology is called a “Hub”. Be careful using this expression because it is also used for L2 Ethernet-Switch which is a packet switch but not an amplifier like a repeater.





## Repeater-Rules




 LAN and Ethernet Fundamentals, v5.2 50

The collision domain of a 10Mbit/s Ethernet LAN is limited by the slot-time of 51,2 microsecond.

Topology of repeaters must obey this budget. Therefore so called repeater rules existed for cabling an Ethernet LAN.

Maximal 5 segments over 4 repeater-sets are allowed, in this case 2 segments have to be link-segments (rest arbitrary), length of fibre optic link segments must not exceed 500m each

**-> results in a maximum diameter of 2500m**

On 4 segments with 3 repeater-sets, the length of a fibre optic link segment must not exceed 1000m, the segments may be mixed in any desired way

**-> results in a maximum diameter of 3000m**

# Agenda

- **Introduction**
- **IEEE 802**
- **Logical Link Control**
- **Ethernet**
  - Introduction
  - CSMA/CD
  - Elements and Basic Media-Types
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**







# IEEE 802.3 Frame Format


| | |
|---|---|
| Preamble ..... | for clock synchronization (64 bit)<br>(62 bits 10101.....01010 + 2bits 11 as SD,<br>bit synchronization within 18 bit times) |
| DA ..... | Destination MAC-address (48 bit) |
| SA ..... | Source MAC-address (48 bit) |
| Length ..... | of IEEE 802.3 frame (16 bit)<br>= bytes following without CRC (46-1500) |
| Data ..... | Payload |
| FCS ..... | Frame Check Sequence (32 Bit) |



Some Ethernet parameters:

Interframe gap between to Ethernet frames is 9.6 microsecond.

Jam size is 32 bit.

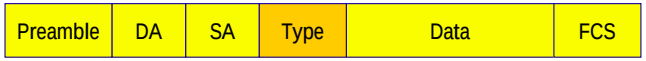
Slot time is 512 bits, minimal frame length 64 Byte

Maximal frame length 1518 byte (6+6+2+1500+4)

Maximum number of frames on a 10 Mbit/s Ethernet: 14880 frames of minimal length per second

(6+6+2+46+4, FCS counted, preamble not counted)

## Ethernet Version 2 (DEC, Intel, Xerox -> DIX)




Preamble ..... for clock synchronization
DA .....Destination Address (48 Bit)
SA .....Source Address (48 Bit)
Type ..... Protocol-type field (16 Bit)
(Ethernet Version II frame)
Data ..... payload
FCS ..... Frame Check Sequence (32 Bit)

- **Ethernet V2 and 802.3 can coexist on the same cable, but each associated sending and receiving station must use the same format.**
- **Fortunately all type-field values are larger than 1518 (max frame length), so any incoming frame can be recognized and handled properly.**

Remember IEEE 802.3 relies on LLC (802.2) and SAPs -> the protocol-type is indicated by SSAP and DSAP and the LLC control field can provide connectionless and connection-oriented services to the upper layers.

Ethernet Version 2 uses a protocol-type-field instead of the length field and lacks from any kind of HDLC like control field. Therefore only connectionless services can be provided by Ethernetv2 to the upper layers.





Datenkommunikation 384.081 - SS 2012

L06 - LAN and Ethernet Fundamentals (v5.2)

## DIX Type field

- **2-bytes Type field to identify payload (protocols carried)**
  - Most important: IP type 0x800
- **No length field**

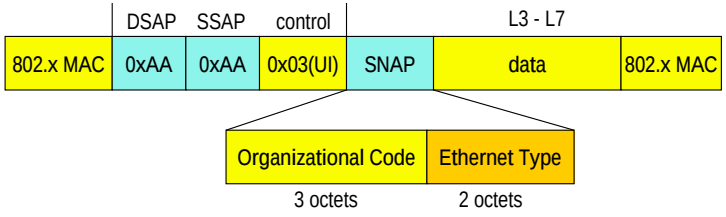
The Type field used by the DIX Eth2 frame format is 16 bit in length and allows therefore to address up 65 536 different layer 3 processes. The Type field only allows the addressing of the destination service access point. The indication of the source service access point is not supported by the DIX frame format. Typically only layer 3 processes of the same kind are able to communicate with each other.

Some Type field examples:

| | |
|---|---|
| Hex 0800..... | IP |
| Hex 0806..... | ARP |
| Hex 8035..... | RARP |
| Hex 814C..... | SNMP |
| Hex 6001/2..... | DEC MOP |
| Hex 6004..... | DEC LAT |
| Hex 6007..... | DEC LAVC |
| Hex 8038..... | DEC Spanning Tree |
| Hex 8138..... | Novell |

## SNAP

- **Demand for carrying type-field in 802.4, 802.5, 802.6, ... also !**
  - Convergence protocol was needed to transport Ethernet V2 type information over IEEE LANs -> SNAP
- **Subnetwork Access Protocol (SNAP) header introduced**
  - If DSAP=SSAP=0xAA and Ctrl=0x03
  - then a 5 byte SNAP header follows containing 3 bytes organizational code plus 2 byte DIX type field



The IEEE had problems to address all necessary layer 3 processes, due to the short (8 bit) DSAP and SSAP fields in the IEEE header. So they introduced a new frame format which was called Subnetwork Access Protocol (SNAP). The SNAP format was simply importing the DIX Type field by the backdoor. This new header format was then also used for technologies like Token Ring, Token Bus, DQDB, etc.

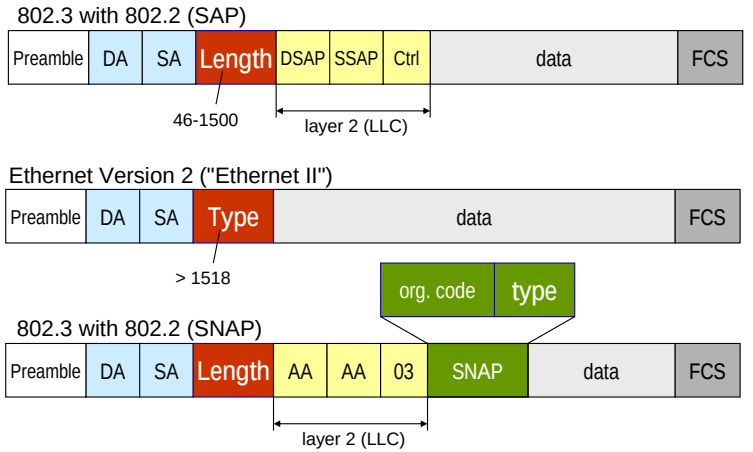
In the SNAP format the DSAP and the SSAP is set to the hex value of AA. This indicates an five byte extension to the standard 802.2 header, which is made up of a three byte long field called Organization Unique Identifier (OUI) and the two byte Type field.







## Frame Types Summary



So we end up with three different frame formats used in Ethernet systems. The 802.3 without SNAP, the DIX Eth2 format and the 802.3 with SNAP.

The DIX Eth 2 frame format is mainly used for the data transport of protocols that have the functionality of error recovery and flow control implemented in their protocol stack e.g. IP.

The 802.3 without SNAP frame format is used for protocols that need the functions of error recovery and flow control on layer 2 e.g. NetBeui, SNA.

The 802.3 with SNAP frame format is used by vendors to implement proprietary protocols, for example Cisco's CDP, VTP, CGMP, etc. protocols. For such purposes the OUI field is used to indicate the vendor and the type field value is chosen vendor specific.

# Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**







# What is Bridging?

- **Packet switching principle in connectionless mode applied to layer 2**
  - The already well known store and forward principle of WAN world (OSI layer 3 intermediate system) was taken by the LAN community
- **Adapted to work with MAC addresses**
  - Instead of unique and structured OSI layer 3 addresses
- **The bridging table or MAC address table**
  - Is used in the same way as the routing table of chapter network principles
  - Signposts to reach a given MAC address by pointing to the corresponding port
    - MAC address -> Port mapping



Remember: Packet switching in connectionless service mode relies on unique, structured addresses and routing table. The content of a routing table is used as “signposts” for reaching a given destination address.

The bridging table contains information about which port to be used to reach a certain MAC address.

## Why Packet Switching on LAN?

- **LAN was originally designed for shared media**
  - But too many clients cause performance problems
    - Higher probability for collisions in case of Ethernet
- **Bridge separates two (or more) shared-media LAN segments**
  - Only frames destined to the other LAN segment are forwarded, frames destined to own LAN segment are not forwarded
  - Number of collisions reduced (!)
- **Different bridging principles**
  - Ethernet: Transparent Bridging
  - Token Ring: Source Route Bridging

Bridges forward layer 2 packets (frames) according to their destination address. Hereby, those frames are filtered whose destination is not reachable on another port of the bridge. This filtering capability significantly enhances the total performance of a LAN as it is divided into multiple segments—multiple broadcast domains: The number of collisions is reduced!

IEEE defined bridges for all kind of LAN technologies. For example a Token Ring network relies on so-called source route bridging, while Ethernet uses "Transparent Bridging".

Transparent Bridging (Ethernet world) is easy for end systems but there are much more complex transferring-decisions (compared to source route bridging) done by a bridge -> dedicated hardware. Today's L2 switching technology (= Ethernet switch) is just a fast transparent bridge. This chapter only discusses Transparent Bridging.

Source Routing Bridging (Token Ring world, IBM) means more overhead for end systems (path finding) but source route bridges are less complex (task of such a bridge could be done e.g. by a PC with two network cards). We will not further cover this in this lecture.

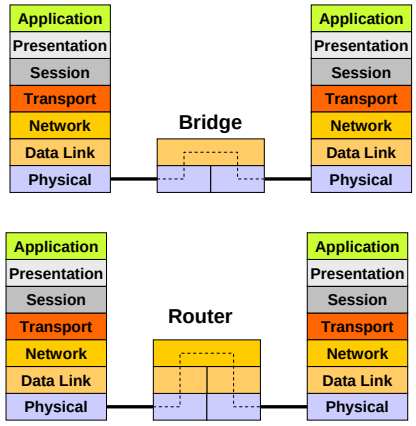
Partitioning of a local are network may additionally be used for error limitation, because of security reasons, increase the number of clients on A LAN or even to extend the geographical diameter of the LAN by placing a bridge every 2500m between two adjacent LAN segments (which have reached their maximum diameter by obeying the repeater rules).







## OSI Comparison




- **MAC addresses not routable**
  - NetBIOS over NetBEUI not routable (no L3)
- **Bridge supports different physical media on each port**
  - E.g. 10Mbit/s to 100Mbit/s
- **Router supports different layer-2 technologies**
  - E.g. Ethernet to Frame Relay

It is very important to understand the differences between bridges and routers. There are many implications related to the operating layer these devices support. As a rule of thumb any device is able to terminate all layers below the highest layer implemented.

Bridge is a packet switch implemented on OSI layer 2. Forwarding is based on unstructured MAC addresses, signposts are stored in MAC bridging table ( = routing table of L2 packet switch). Above OSI Layer 2 bridges are transparent for all higher layer protocol. For any two stations connected to a bridged Ethernet network the bridge is not even visible for them, hence transparent. No L2 address of a bridge has to be configured in an end-system in order to enjoy communication on a LAN.

A router is packet switch implemented on OSI layer 3. Forwarding is based on structured L3 addresses and any end system must speak the corresponding L3 „language" of the router, hence the router is a visible component the end system must know about (e.g. in IP you have to configure the IP address of the default router).

## Bridging versus Routing

- **Bridging works on OSI layer 2**
  - Forwarding of frames
  - Use MAC addresses only
  - Termination of physical layer (!)
  - Invisible for end-systems on a bridged LAN
- **Routing works on OSI layer 3**
  - Forwarding of packets
  - Use routable addresses only (e.g. IP)
  - Termination of both layer 1 and 2
  - Visible for end-systems (e.g. default router in IP)

There are many differences between bridging and routing! The only thing in common is the store and forwarding principle, based on some sort of destination address.

But a bridge forwards layer 2 frames while a router forwards layer 3 packets. Layer 2 frames use simple MAC addresses, having no logical structure, while layer 3 packets use structured addresses, revealing topology information. Only layer 3 addresses are routable. In order to understand the latter statement, it is important to understand the principles of routing and how a routing table works. We will discuss this soon.

Bridges terminate physical links. Thus, one port of the same bridge might support optical fiber transmission and another port might support twisted pair copper cabling.

On the other hand, routers terminate layer 2 links. That is, one interface might utilize Ethernet as link layer technology, another interface Frame Relay, and a third interface might run ATM. A router only forwards the packet—the layer 3 information—carried inside a frame.







## Bridge History

- **Bridges came after routers!**
- **First bridge designed by Radia Perlman**
  - Ethernet has size limitations
  - Routers were single protocol and expensive
- **STP (Spanning Tree Protocol)**
  - Is an important part of any bridged network
  - Ethernet has no hop count (or IP time-to-live field) to implement a kill mechanism in case of loops
- **IEEE 802.1D contains**
  - Transparent bridging and STP

Bridging is a fundamental part of the IEEE LAN standard. Actually bridges were invented relatively late—routers were invented a bit earlier. Radia Perlman, a pioneer in data communication designed the first bridge. The main reason was to extend the total network diameter of Ethernet and to provide a transport technique which supports multiple layer 3 technologies. She also invented the Spanning Tree Protocol (STP) because Ethernet had no hop count, thus any store and forwarding technology would suffer from broadcast storms, when broadcast destination addresses are used. But this issue is discussed in more detail in the following chapter. The IEEE standard 802.1D specifies bridging and spanning tree (and more).

## What is an Ethernet Switch?

- **A switch *is* basically a bridge, differences are only:**
  - Faster because implemented in HW
  - Multiple ports
  - Improved functionality (e.g. VLAN)


- **Don't confuse it with WAN Switching!**
  - Completely different !
  - Connection oriented (stateful) VCs

Now what is the difference between a bridge and a switch? Logically there is no difference. Ethernet switching is much more a marketing term to express something new to the customers but basically it is the same as a transparent bridge. An Ethernet switch is just faster and has much more ports than the good old bridge. Ethernet switches typically employ more than two ports, and the bridging functionality is implemented in hardware. Additionally other features like VLAN (Virtual LAN) support or NAC (Network Access Control) are added in modern Ethernet switches, depending on the vendor

Note: Don't confuse LAN switching with WAN switching. Unfortunately modern bridging is called switching but logically it is still bridging. The term "bridging" was originally defined to differentiate this technique strictly from WAN switching. The main characteristic of WAN switching is its connection oriented behavior—WAN switches are never transparent! In order to connect to a WAN switch the end system must comply to some specific User to Network Interface (UNI).

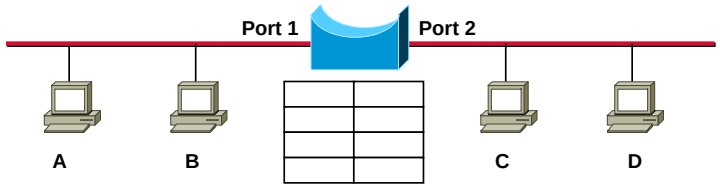






## How does it work?

- **Transparent bridging is like "plug & play"**
- **Upon startup a bridge knows nothing**
  - MAC address table is empty
  - Exception: Static entries are configured by the network admin -> hard work
- **Bridge is in learning mode**
  - Dynamic entries are built on the fly



The main advantage of transparent bridging over source route bridging (token ring) is the transparency or "plug & play" capability. No end station notices the presence of bridges. Bridge is invisible for end stations. LAN Left and LAN Right appear to the end systems like one single, logical, big LAN. But because of this transparency a bridge must receive and process every frame on a LAN. This means much more performance is needed for a bridge than for a router which is explicitly addressed. Also flow control between end systems and bridges was not defined in the original implementation.

Transparent bridge uses layer 2 MAC-addresses to decide if a given frame must be a forwarded or not -> destination-address of a frame is used for this decision.

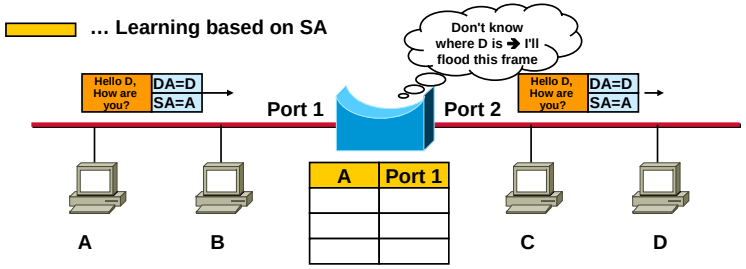
But in order to be invisible, bridges must also learn somehow where end stations are located.

MAC-addresses of all stations are registered in a bridging table either statically done by administrator e.g. for security reasons or dynamically done by a self-learning mechanism.

Following the self-learning mechanism upon startup, a bridge knows nothing and the bridging table is empty. At this time the bridge is in learning mode. For learning the source-address of a frame will be used.

## Learning

- **Once stations send frames the bridge notices the source MAC address**
  - Entered in bridging table
- **Frames for unknown destinations are flooded**
  - Forwarded on all ports







Assume we have a bridge with only two ports, each attached at one Ethernet segment. Assume the left station "A" sends one frame to "D" on the right side. Obviously the bridge learns the location of A but has no idea where D is. Thus the MAC address of A is entered in the bridging table and also the port number "1", on which A is reachable. Since the location of D is unknown, the bridge loods this frame over all ports, in our case only to port two (as there are no other ports).

This way, connectivity is granted even if there is no entry in the bridging table.

Destination address of a frame is used for MAC address table look up in order to decide what have to be done with the received frame. The following actions are possible:

**Filtering**: frame will be rejected if destination's home is on the LAN segment of the receiving port

**Forwarding:** a duplicate of the frame will be forwarded to the appropriate port if destination's home is registered in the table of another port

**Flooding:** during learning time; frame will be forwarded to all other ports (multiport-bridge) if there is no entry in the table (unknown destination).

**Frames with broadcast/multicast-address are always flooded**.

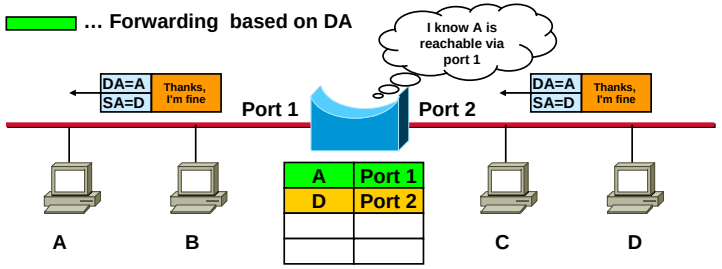






## Learning ➔ Table Filling (1)

- **If the destination address matches a bridging table entry, this frame can be actively**
  - Forwarded if reachable via other port
  - Filtered if reachable on same port

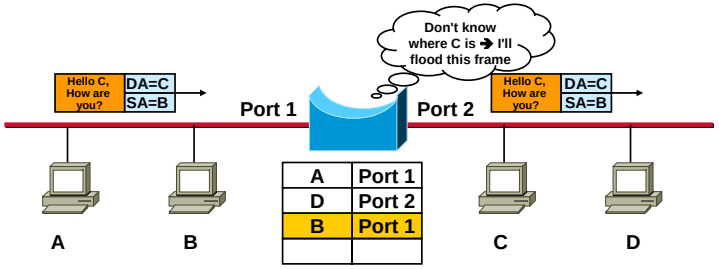






Now assume D replies to the message which has been received from A. The bridge knows already the port number over which A can be reached and forwards the frame accordingly. If A would be located on the same port as D then this frame would be filtered.

# Table Filling (2)



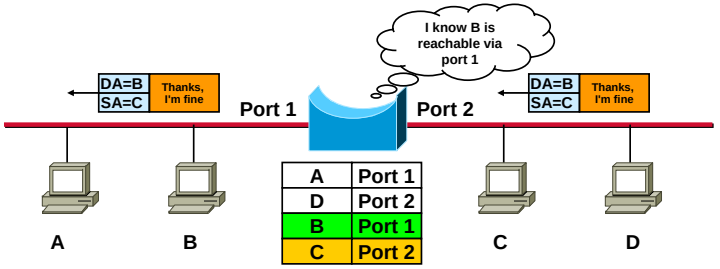







# Table Filling (3)

- After some time the location of every station is known – simply by listening!
- Now only forwarding and filtering of frames
  - Based on destination address

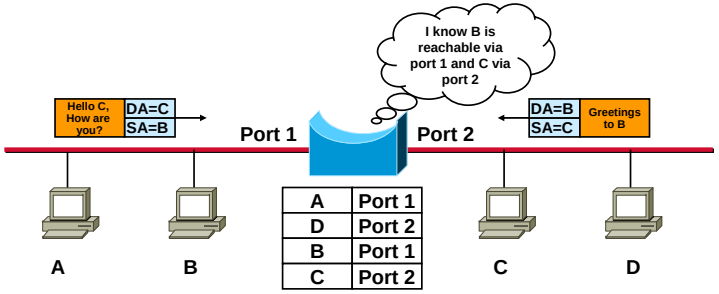





After some traffic observing time, the bridging table contains all host locations (addresses and port numbers). At this time the bridge enters the forwarding and filtering mode.

## Table Filling – Forwarding

- **Collision domains are separated**
  - Frames can travel in their LAN segments at the same time



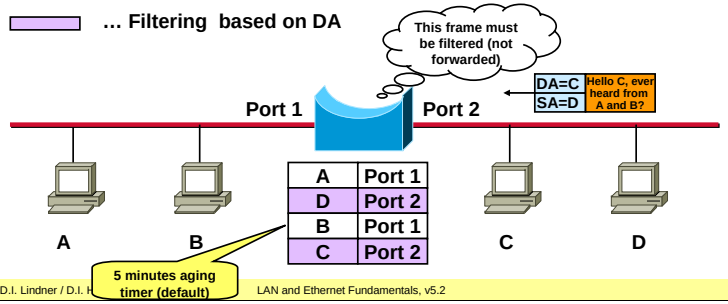
Since only frames are forwarded to other ports whose destination is really located there, the LAN is separated into as many collision domains as ports are available (and attached to a LAN segment).





## Table Filled – Filtering

- **Frames whose source and destination address are reachable over the same bridge port are filtered**
- **Entries times out**
  - If not refreshed within 5 minutes



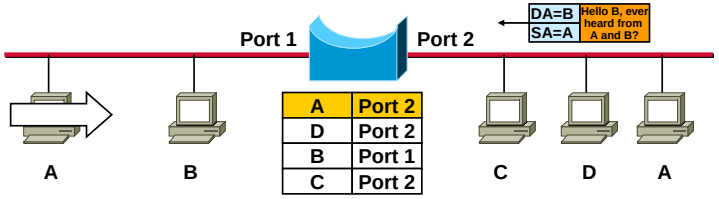
What if a host is removed from its location and attached at another place in the LAN? Obviously frames could be forwarded to the wrong port. Therefore each entry in the bridging table ages out after some time. The default aging time is 300 seconds or 5 minutes.

In case of a dynamic bridging table an aging mechanism allows for changes of MAC addresses in the network which may caused either by change of network card or by location change of end system. If an already registered MAC address is not seen within e.g. 5 minutes as source address of a frame the corresponding bridging table entry is deleted.

.

## Last Seen – Last Win

- **Now assume notebook with MAC A is moved**
- **Address is immediately relearned**
  - With the first frame containing source MAC A on the other port
- **Imagine the problem**
  - If there are duplicated MAC address in the LAN !!!



Duplicated MAC addresses would cause continuous table rewriting on a last seen – last win base.





## Most Important !

- **Bridge separates LAN into multiple collision domains**
- **A bridged network is still one broadcast domain**
  - Broadcast frames are always flooded
- **A switched network works the same way as a bridged network**
  - Ethernet switch instead good old bridge
- **A router separates the whole LAN into multiple broadcast domains**
  - LAN broadcast are therefore limited to a location and does not spread over the whole network

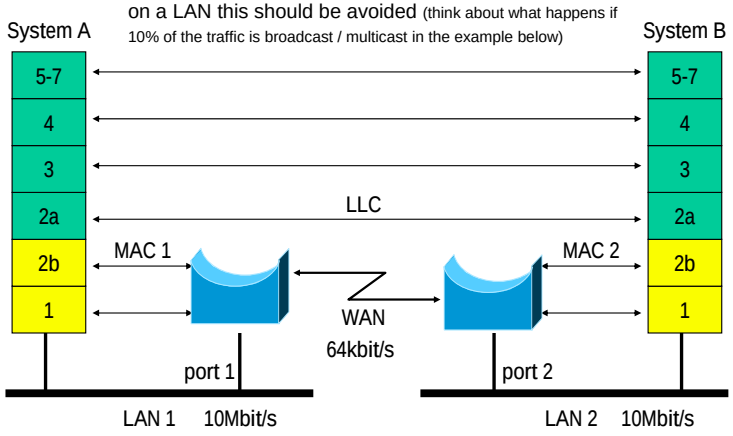
It is very important to understand the basic message which is given here:

The use of bridges results in a separation of multiple collision domains of the LAN. Still we have one single broadcast domain! That is, broadcast frames are always flooded throughout the network.

Only the use of routers results in a separation of multiple (layer 2) broadcast domains—or the use of VLANs, which will be discussed soon in this chapter.

## Remote Bridging ? (!!! TB means Broadcast Domain!!)

Because of inherent L2 broadcast / multicast percentage on a LAN this should be avoided (think about what happens if 10% of the traffic is broadcast / multicast in the example below)




 LAN and Ethernet Fundamentals, v5.2 73

Relay function of local bridge can be split in two half-bridges. Coupling of the half-bridges via slow WAN-connection is a problem in case of high amount of broadcast on a LAN. The mismatch of bit rate can cause a buffer overflow in the bridge and frames will get lost in this case.

Therefore transparent bridging over WAN or any other Ethernet tunneling technique should be avoided !!!





## Bridging Problems

- **Redundant paths lead to**
  - Endless cycling of frames
  - Continuous table rewriting
  - Blocking of buffer-resources
  - Stagnation of the LAN
  - Broadcast storms
- **To eliminate these unwanted effects**
  - STP (Spanning Tree Protocol) is necessary
- **No load sharing possible and no ability to select best path**
  - It is just bridging and not routing

 LAN and Ethernet Fundamentals, v5.2 74

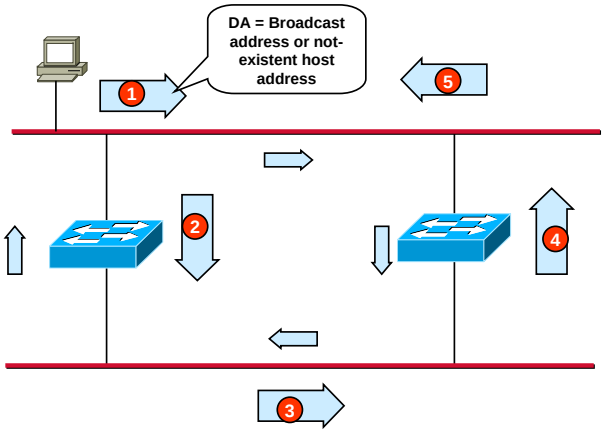
You might have noticed that bridges do not really learn the network topology. They only learn a simple destination to port association! Because of this there is no means to determine the best path, and furthermore frames might be caught in a loop.

Especially broadcast frames have no defined destination and would be forwarded over all parallel paths—endlessly! This results in endless circling of frames, or more dangerous, in a so-called "broadcast storm".

Also a continuous table rewriting might occur (this is not so widely known but also explained in the next pages).

Most people are not aware that frames might be stored up to 4 seconds inside the buffer of a switch—and it still complies to the IEEE standard. Although this would happen only in rare cases of congestion, transparent bridging is not suitable for hard realtime applications. Today the situation has changed, QoS features are included to assure bounded delays.

## Endless Circling



The picture above illustrates the endless circling phenomena. Assume a network with parallel paths between two LAN segments, realized by two bridges. Any frame with a broadcast destination address would be forwarded by both bridges to the other segment and back and forth and so on.

Obviously endless circling leads to congestion problems an is not desired. Remember that there is not hop count or time-to-live number within the Ethernet header.

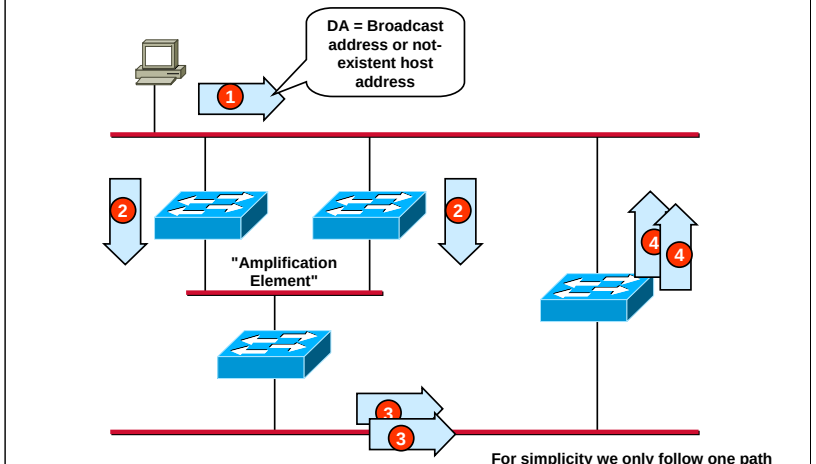
But endless circling is not the main problem... (see next slide)







## Broadcast Storm (1)

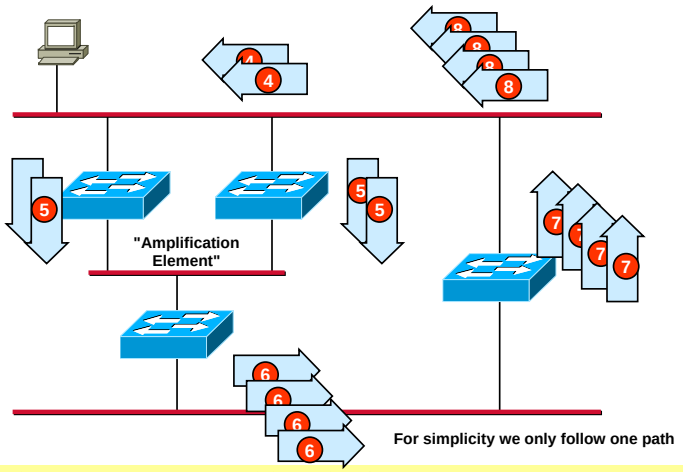


The most feared issue with bridging are broadcast storms. Broadcast storms can be considered as a dramatically "enhanced" endless circling problem. Broadcast storms appear when there is an "amplification" element within the network, such as those threefold parallel paths in the diagram above.

Within a very short time (e.g. 1 second) the whole LAN is overloaded with broadcast frames and nobody could transmit any useful frame anymore.

## Broadcast Storm (2)



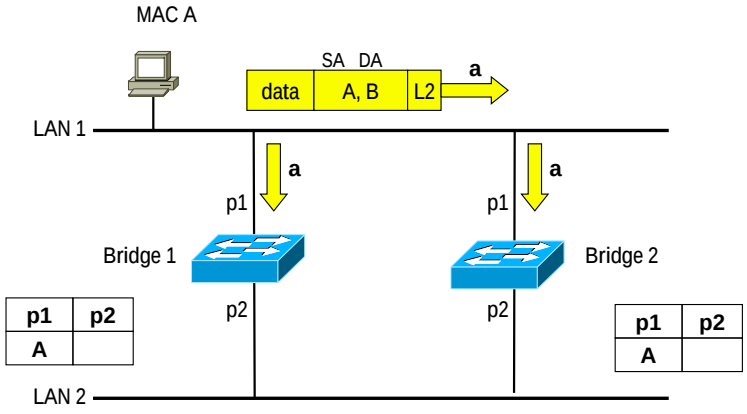

 LAN and Ethernet Fundamentals, v5.2 77

The picture above shows the amplification effect mentioned on the previous page.





## Table Rewriting (Unknown Destination) 1

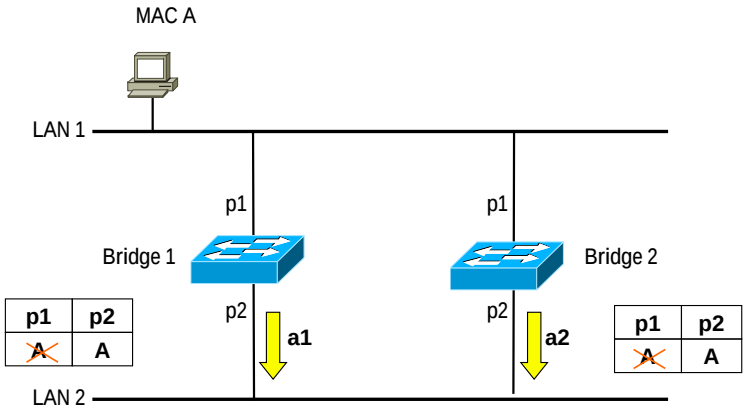



 LAN and Ethernet Fundamentals, v5.2 78

A relatively seldom known problem is the mutual table rewriting phenomena. This problem occurs with unicast frames and broadcast frames!

Assume that host A sends an unicast frame to an unknown destination B, both bridges learn the location of host A but B is unknown and hence must be flooded. If the other bridge receives this flooded frame it will change the table entry for MAC to be seen on the other port (last seen last win!). But again this frame has to be flooded and the game just go on ... ad infinitum!

# Table Rewriting (Unknown Destination) 2

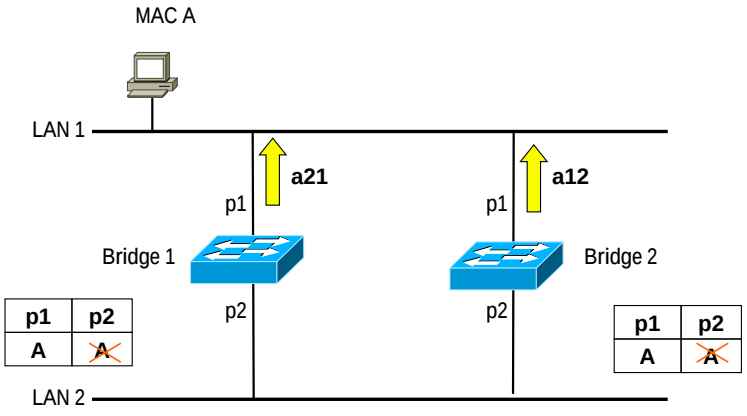









# Table Rewriting (Unknown Destination) 3



## Spanning Tree

- Invented by *Radia Perlman* as general "mesh-to-tree" algorithm
- A must in bridged networks with redundant paths
- Only one purpose:
  cut off redundant paths with highest costs



Now we have learned that active parallel paths lead to severe problems in a switched (i.e. bridged) network. Therefore we can only overcome this problem by deactivating any redundant path. This should be performed automatically in order to call Ethernet bridging still "Transparent" bridging.

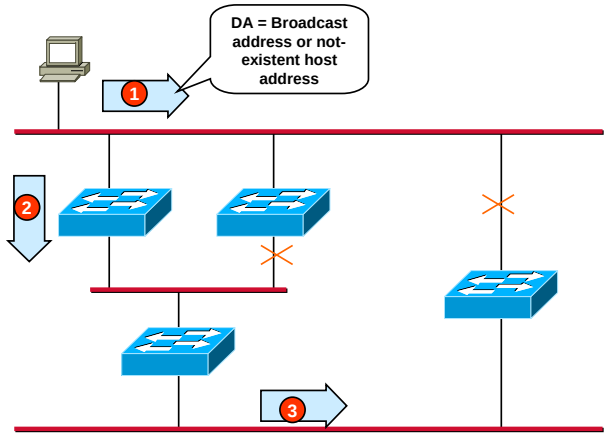
The inventor of bridging, Radia Perlman, also created an easy solution for the redundancy problem: The Spanning Tree Protocol (STP).

The STP is implemented in bridges only (not in hosts) and has only one purpose: To determine any redundant paths and cut them off! Hereby cost values are considered for each path in order to maintain the best paths.





## STP in Action (1)
## No Broadcast Storm

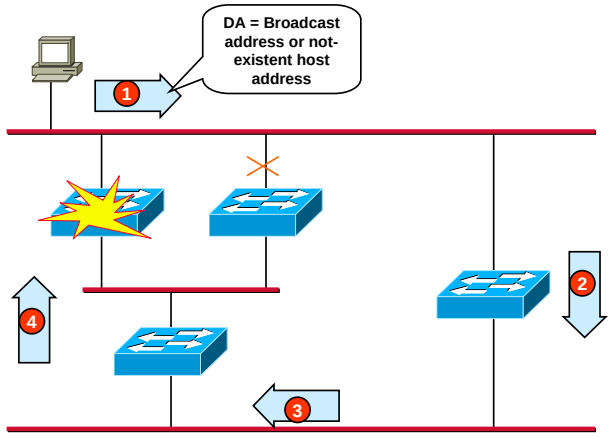





STP eliminates redundancy in a LAN bridged environment by cutting of certain paths which are determined by the STP parameters Bridge ID, Bridge Priority and interface Port Costs. An easy way to achieve this is built a tree topology. A tree has per default no redundancy or have you ever seen leafs of a tree which are connected via two or more branches to the same tree?

Spanning Tree Protocol (STP) takes care that there is always exact only one active path between any 2 stations implemented by a special communication protocol between the bridges using BPDU (Bridge Protocol Data Unit) frames with MAC-multicast address. The failure of an active path causes activation of a new redundant path resulting in new tree topology.

## STP in Action (2)
## Bridge Failure – New STP Topology




 LAN and Ethernet Fundamentals, v5.2 83

Additional task of STP is to recognizes any failures of bridges and to automatically build a new STP topology allowing any-to-any communication again.

Here you can also see one main disadvantage of STP: Redundant lines or redundant network components cannot be used for load balancing. Redundant lines and components come only into action if something goes wrong with the current active tree.



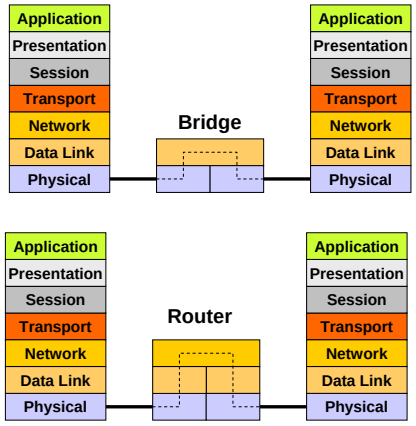


## Agenda

- **Introduction to LAN**
- **IEEE 802**
- **Logical Link Control**
- **Legacy Ethernet**
  - Introduction
  - CSMA/CD
  - Repeater, Link Segments
  - Framing
- **Transparent Bridging**
- **Bridging versus Routing**

 LAN and Ethernet Fundamentals, v5.2 84

# Bridging versus Routing




- **Bridging is**
  - Connectionless packet switching on OSI layer 2 using unique but unstructured MAC addresses without any topology information
  - Signpost in the MAC address table
- **Routing is**
  - Connectionless packet switching on OSI layer 3 using unique and structured addresses which contain topology information
  - Signpost in the routing table







# Bridging versus Routing

| | Bridging | | Routing |
|---|---|---|---|
| ⊕ | Depends on MAC addresses only | ⊖ | Requires structured addresses (must be configured) |
| ⊕ | Invisible for end-systems; transparent for higher layers | ⊖ | End system must know its default-router |
| ⊖ | Must process every frame | ⊕ | Processes only packets addressed to it |
| ⊖ | Number of table-entries = number of all devices in the whole network | ⊕ | Number of table-entries = number of subnets only |
| ⊖ | Spanning Tree eliminates redundant lines; no load balance | ⊕ | Redundant lines and load balance possible |
| ⊖ | No flow control (may be changed by usage of MAC Pause command) | ⊕ | Flow control is possible in theory (router is seen by end systems) |



The list shown above summaries all pro and cons of bridging (switching) and routing.

## Bridging versus Routing

| Bridging | Routing |
|---|---|
| − No LAN/WAN coupling because of high traffic (broadcast domain!) | + Does not stress WAN with subnet's broad- or multicasts; commonly used as "gateway" |
| − Paths selected by STP may not match communication behavior/needs of end systems | + Router knows best way for every destination a packet is sent for |
| + Faster, because implemented in HW; no address resolution | − Slower, because usually implemented in SW; address resolution (ARP) necessary |
| + Location change of an end-system does not require updating any addresses | − Location change of an end-system requires adjustment of layer 3 address |
| − Spanning tree necessary against endless circling of frames and broadcast storms | − Routing-protocols necessary to determine network topology |

The list shown above summaries all pro and cons of bridging (switching) and routing (continued from previous slide).

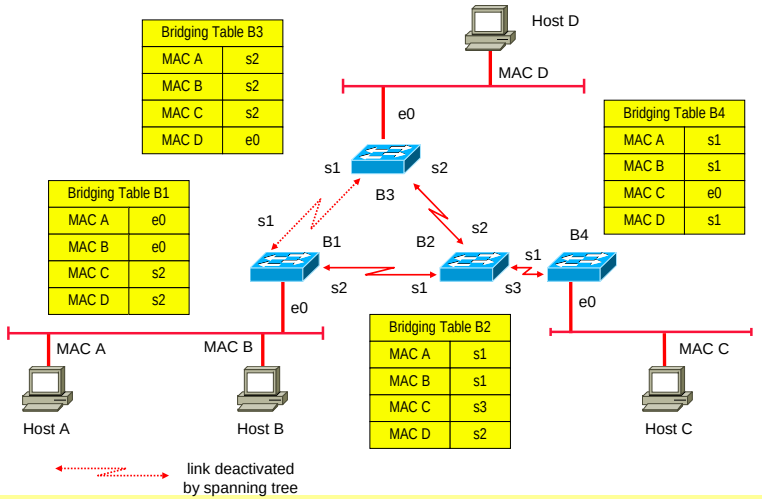




Datenkommunikation 384.081 - SS 2012

L06 - LAN and Ethernet Fundamentals (v5.2)

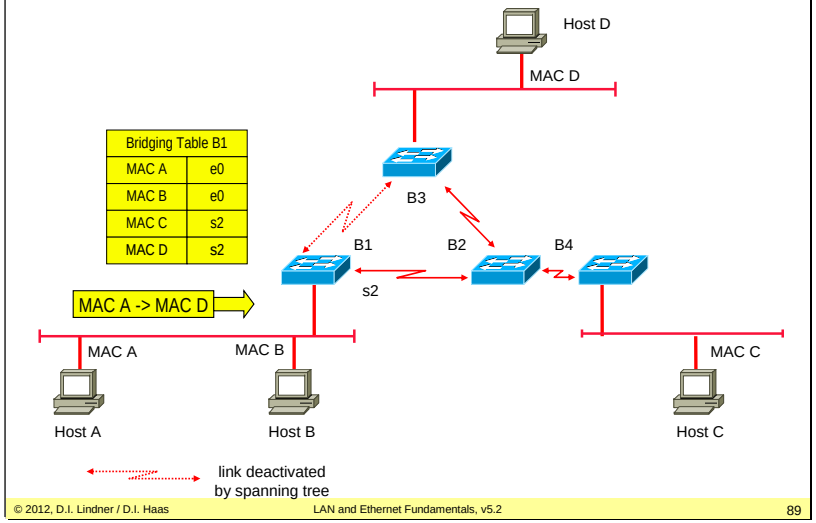
## Example Topology: Bridging



© 2012, D.I. Lindner / D.I. Haas

# Frame MAC A to MAC D (1)

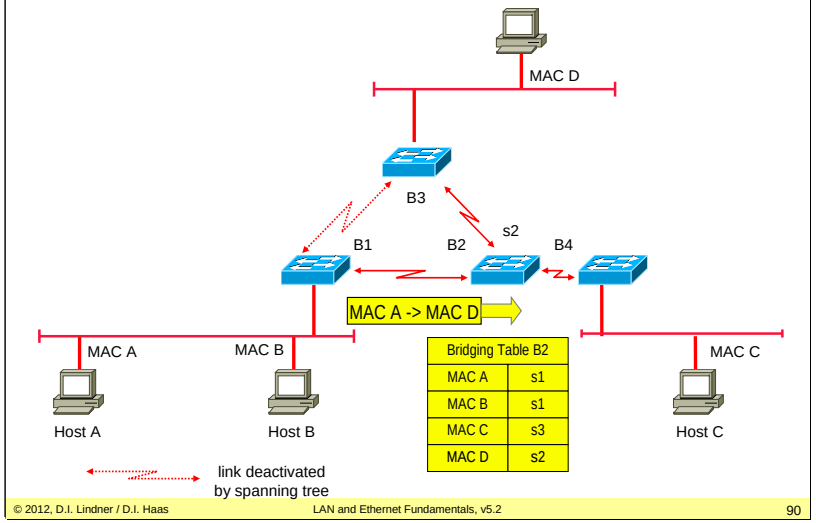




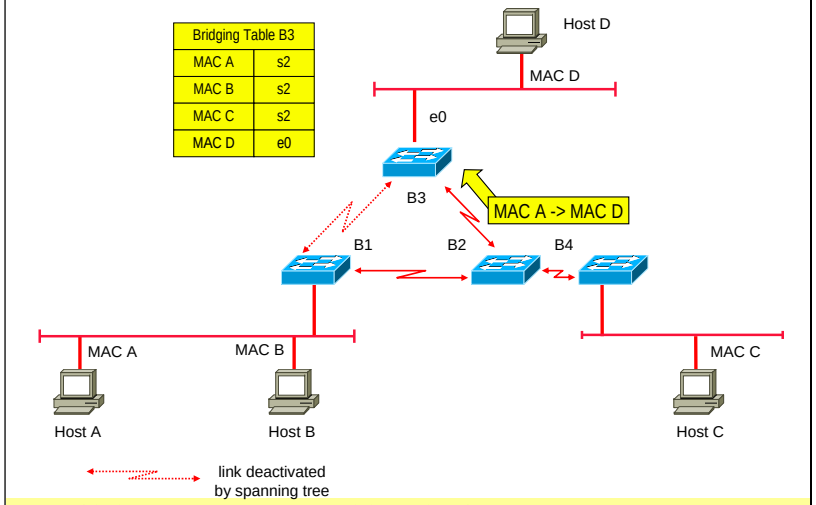





# Frame MAC A to MAC D (2)



# Frame MAC A to MAC D (3)

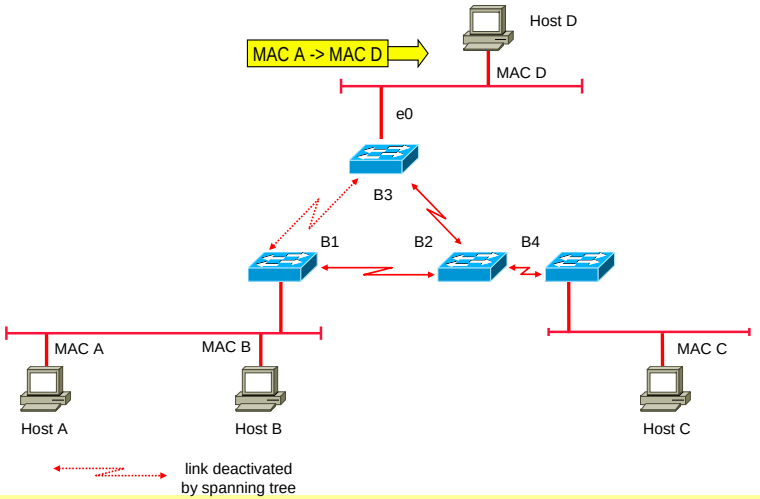



 LAN and Ethernet Fundamentals, v5.2 91



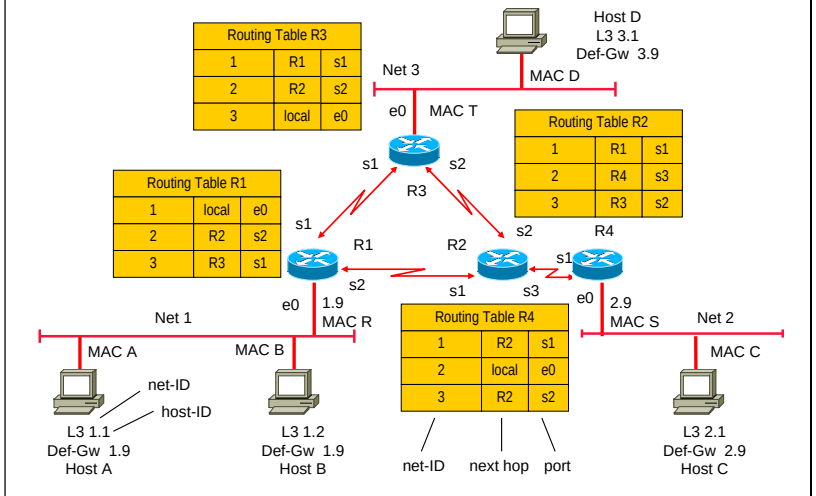


# Frame MAC A to MAC D (4)




 LAN and Ethernet Fundamentals, v5.2 92

# Example Topology: Generic Routing

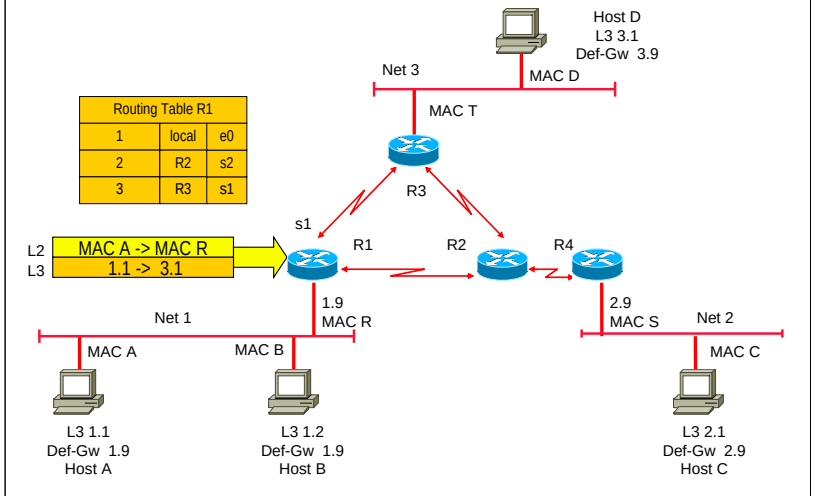




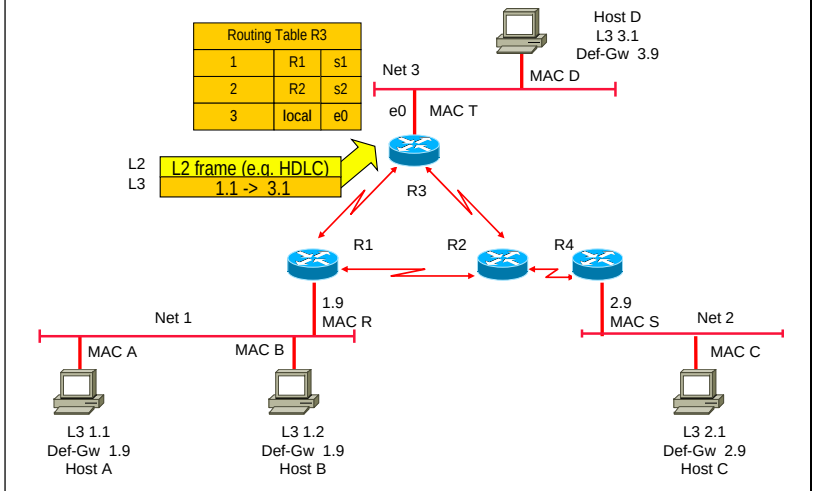





# Frame 1.1 to 3.1 (1)



## Frame 1.1 to 3.1 (2)

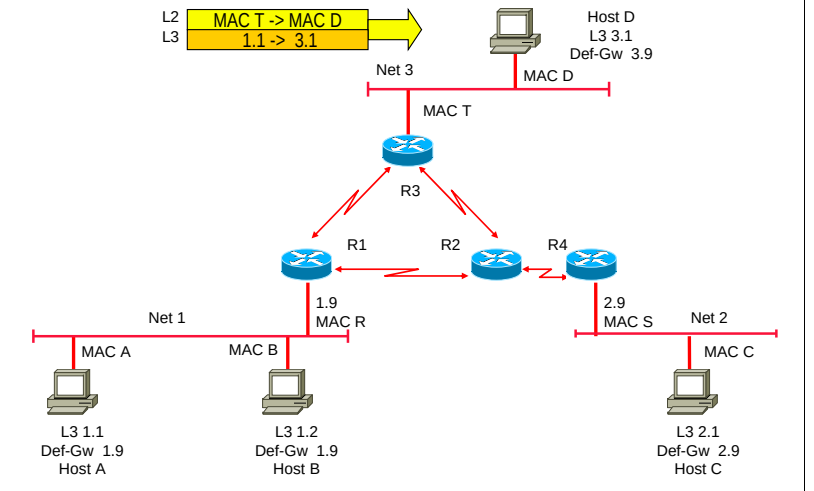









## Frame 1.1 to 3.1 (3)



## Requirements for Routing

- **Consistent layer-3 functionality**
  - For entire transport system
  - From one end-system over all routers in between to the other end-system
  - Hence routing is not protocol-transparent
    - all elements must speak the same „language“
- **End-system**
  - Must know about default router
  - On location change, end-system must adjust its layer 3 address
- **To keep the routing tables consistent**
  - Routers must exchange information about the network topology by using routing-protocols







## Routing Facts 1

- **In contrast to bridges**
  - Router maintains only the subnet-part of the layer 3 addresses in its routing table
  - The routing table size is direct proportional to the number of subnets and not to the number of end-systems
- **Transport on a given subnet**
  - Still relies on layer 2 addresses
- **End systems forward data packets for remote destinations**
  - To a selected router using the router's MAC-address as destination
  - Only these packets must be processed by the router

## Routing Facts 2

- **Flow control between router and end system is principally possible**
  - End systems knows about the local router
- **Broadcast/multicast-packets in the particular subnet**
  - Are blocked by the router so broad/multicast traffic on the subnets doesn't stress WAN connections
- **Independent of layer 1, 2**
  - so coupling of heterogeneous networks is possible
- **Routers can use redundant paths**
  - meshed topologies are usual
- **Routers can use parallel paths for load balancing**