Computer - networks
Inter-network cooperation

2025/2026, 1st semester

Dr. Szilveszter Kovacs

E-mail: szilveszter.kovacs@uni-miskolc.hu

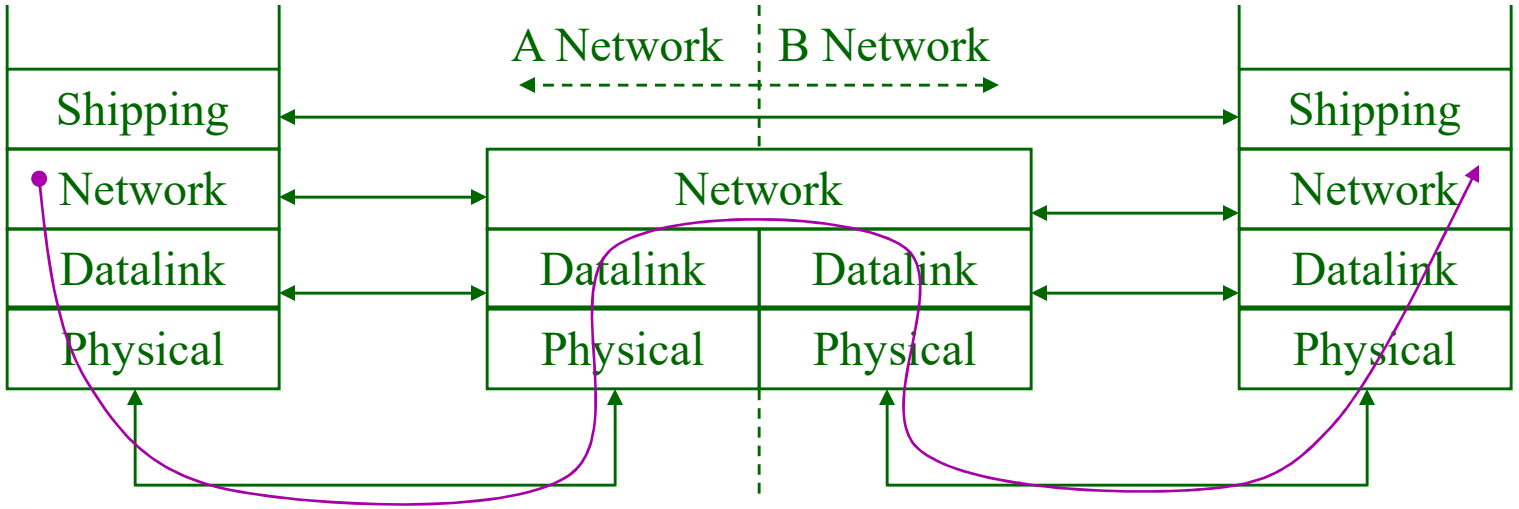
www.iit.uni-miskolc.hu/~szkovacs

Institute of Information Technology 107/a.

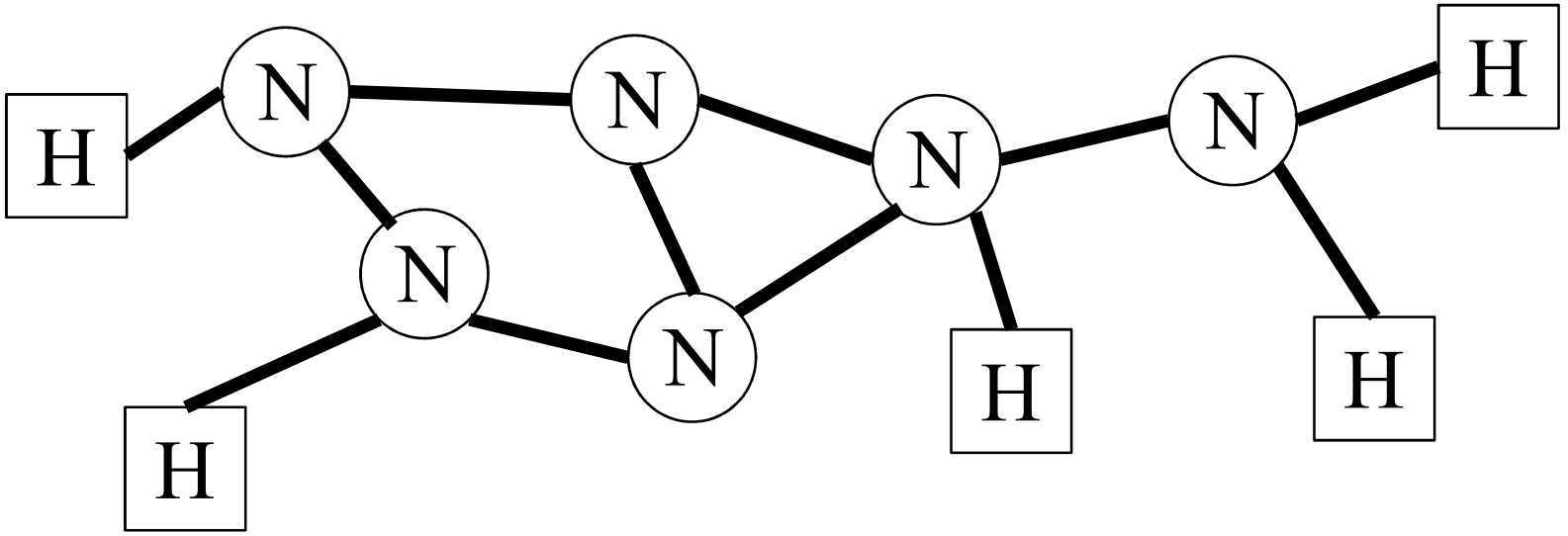
Phone: +36 46 565-111 / 21-07

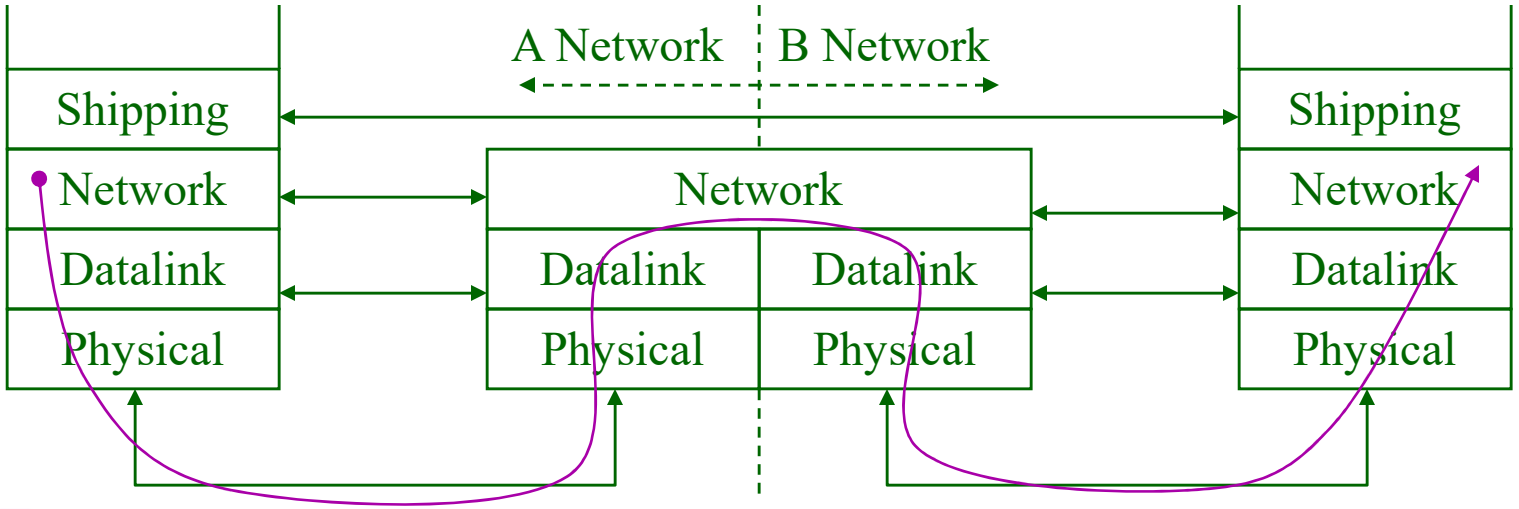
Motivation

- **Connecting different networks**
  ⇒ Creating a heterogeneous network
  ⇒ Expanding the (size) **of the network**
- **According to the OSI model, this can only happen at layer 3 (network)**
  (traffic control, congestion control)
- **The OSI model of inter-network cooperation:**

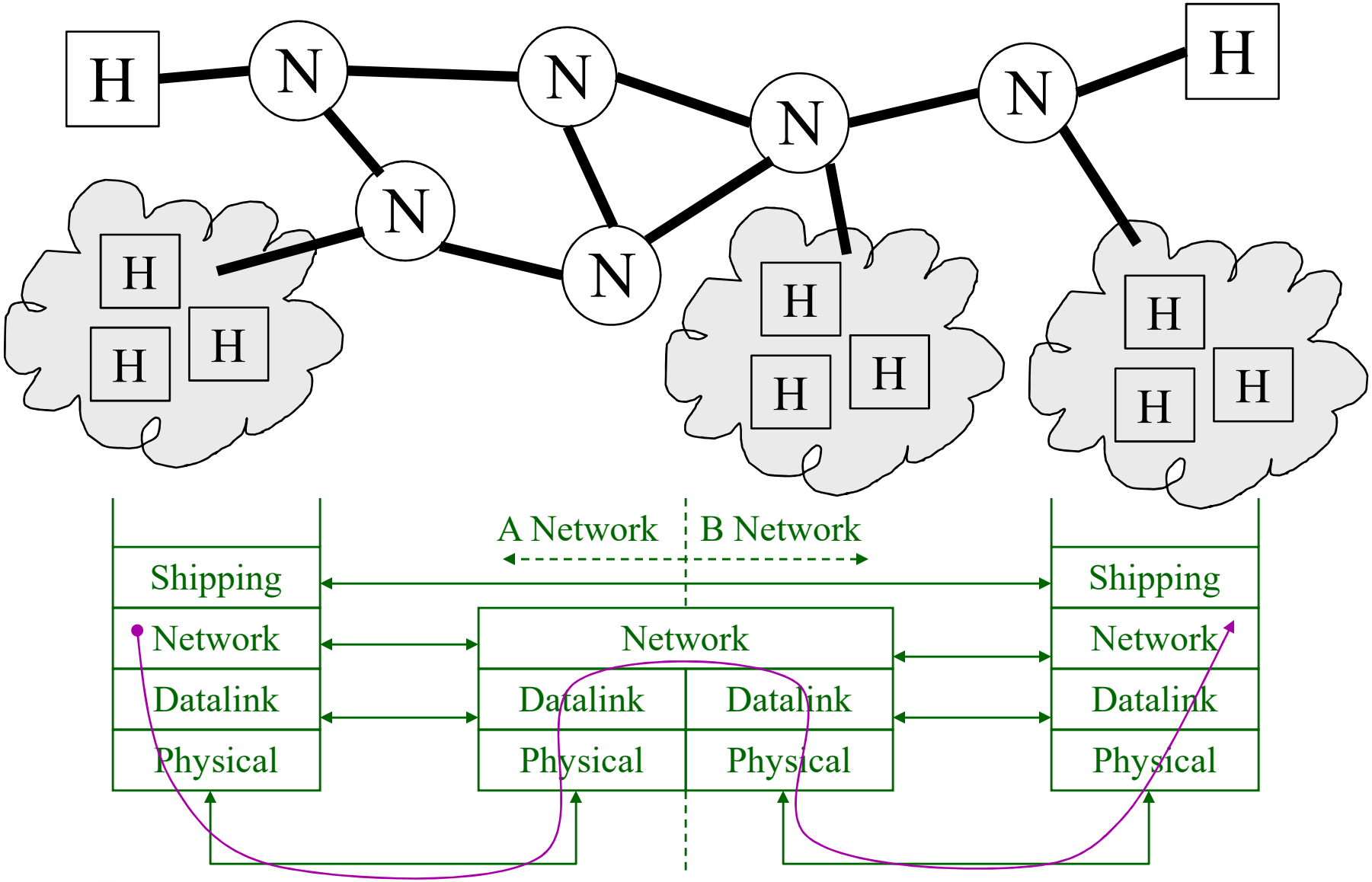


Inter-network cooperation

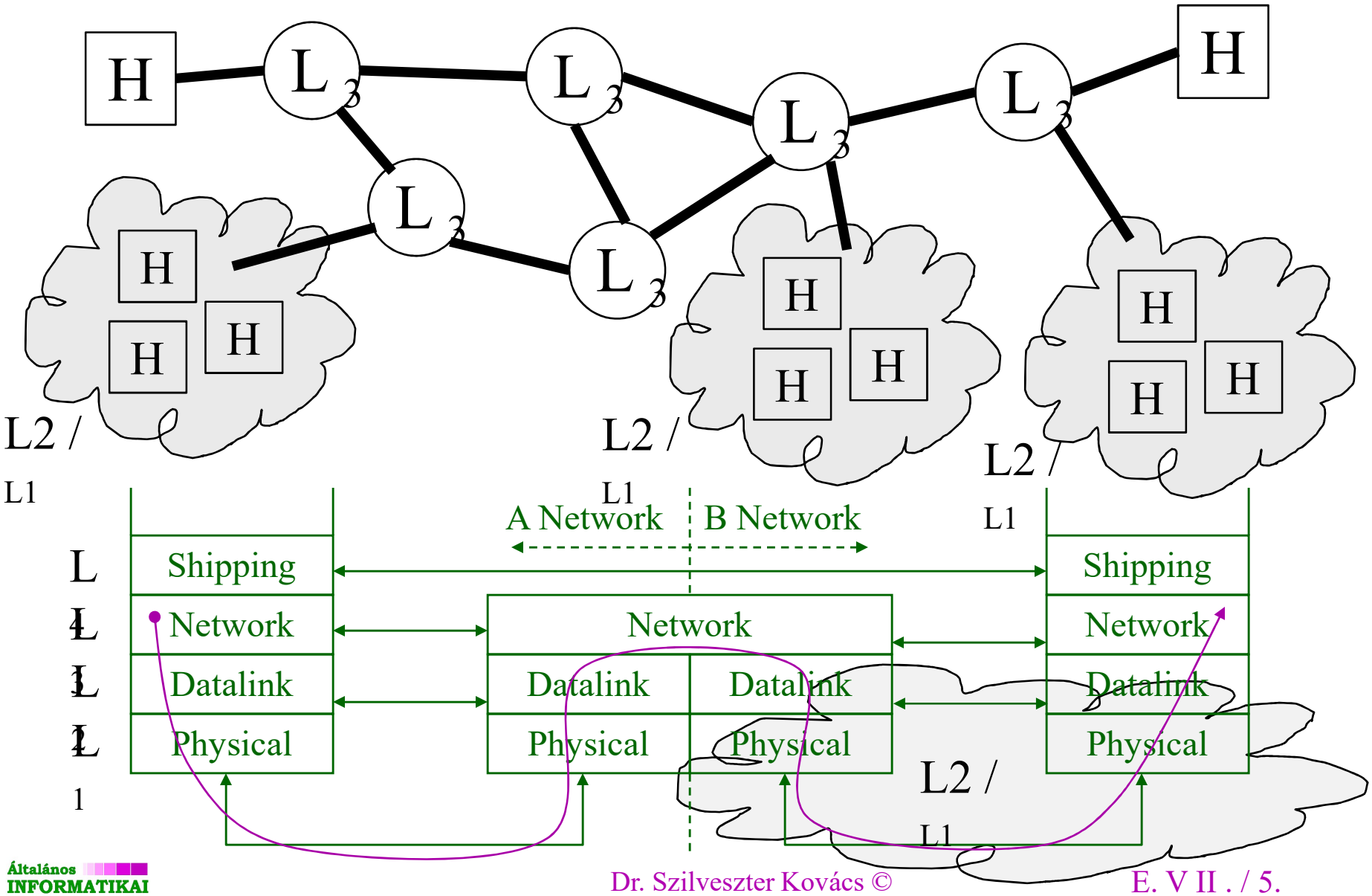




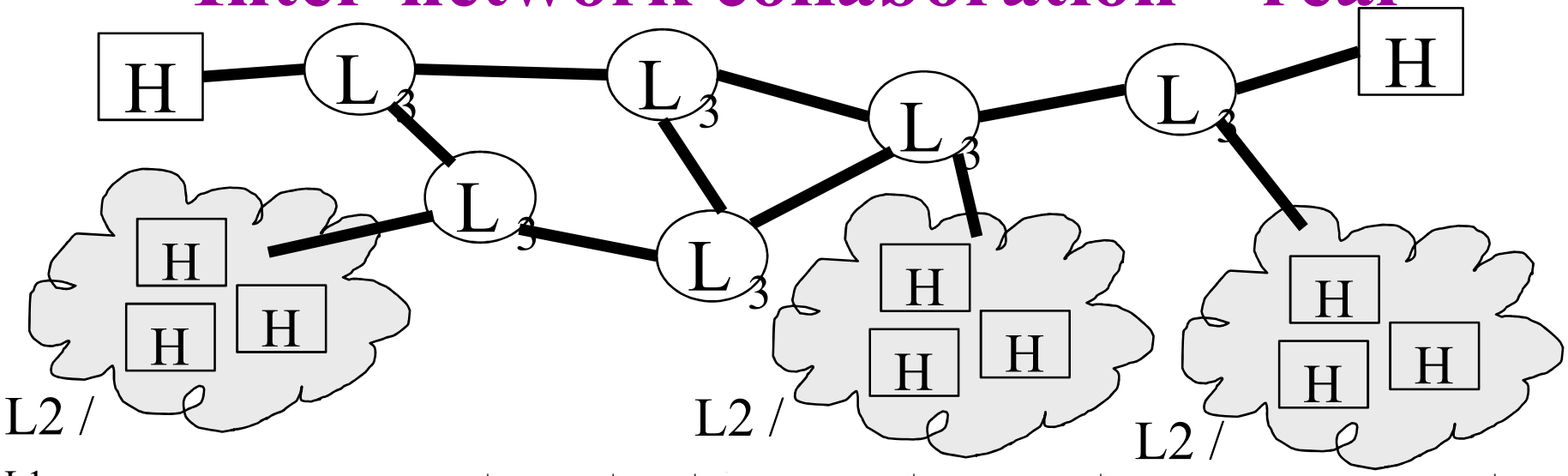
Inter-network collaboration – real

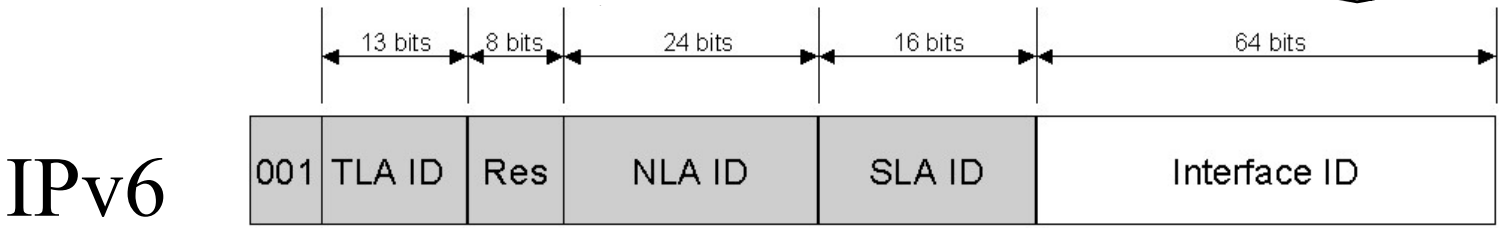


Inter-network collaboration – real



Inter-network collaboration – real



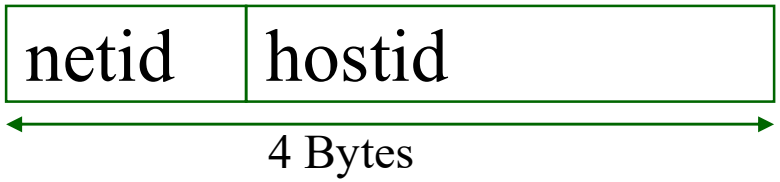


IPv6

:

IPv4

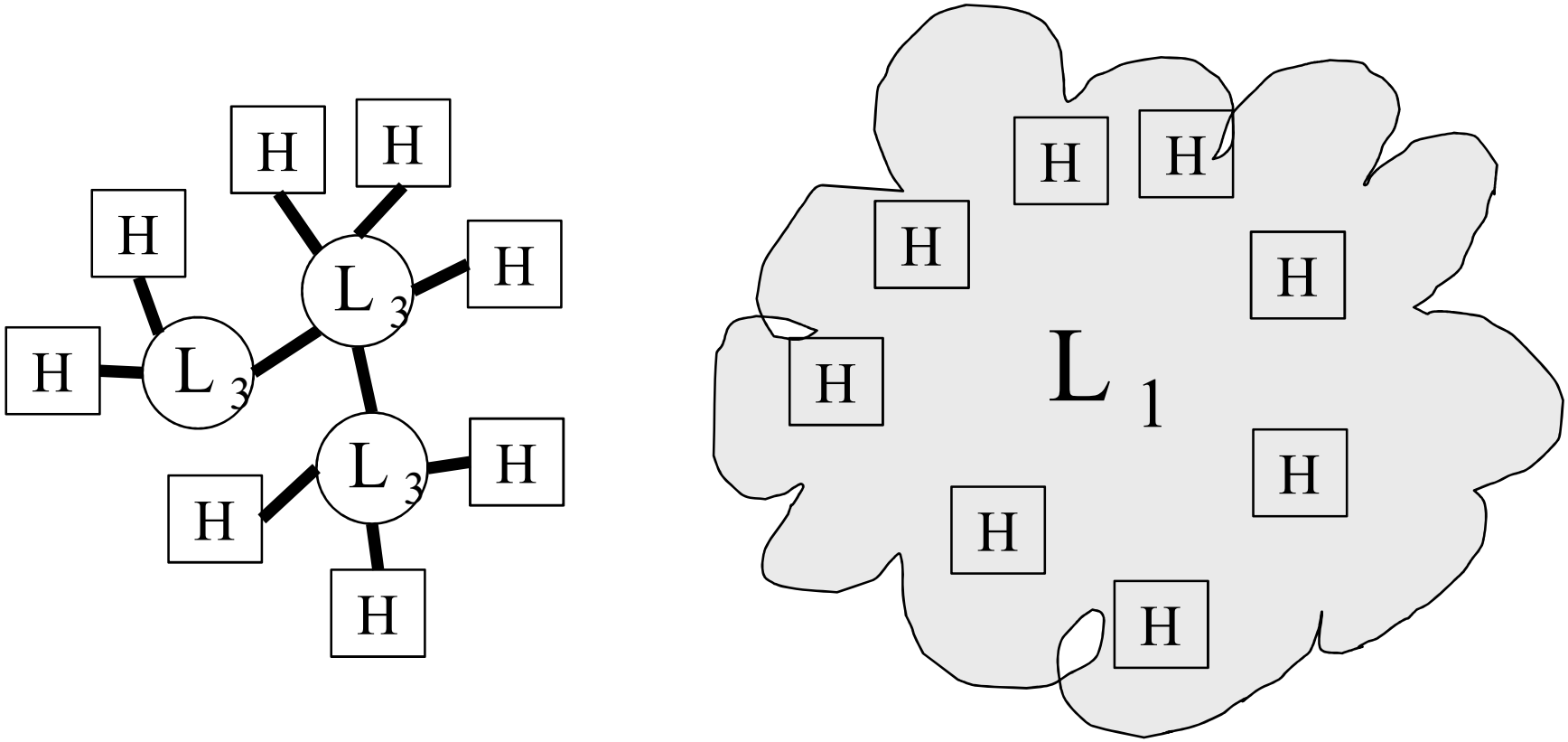
:



IPX: &00000025 %080002001216 !5

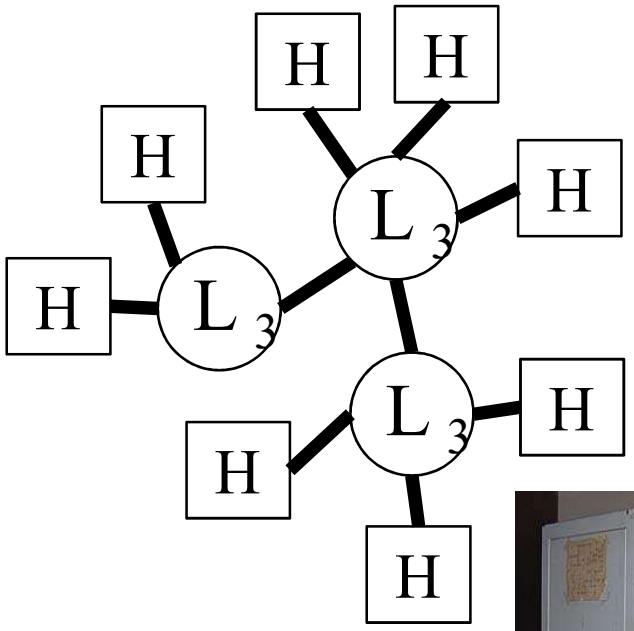
Network Number 4B Host Number 6B Socket Number 2Byte

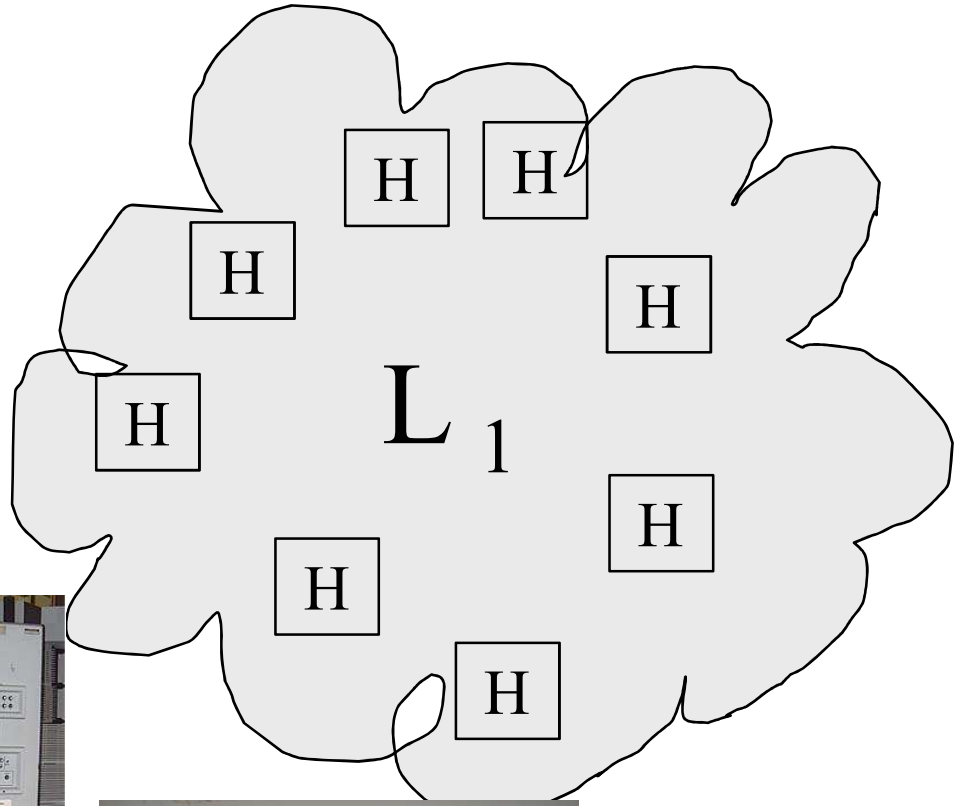
Inter-network collaboration – Why?



In terms of overall throughput, it's much worse!

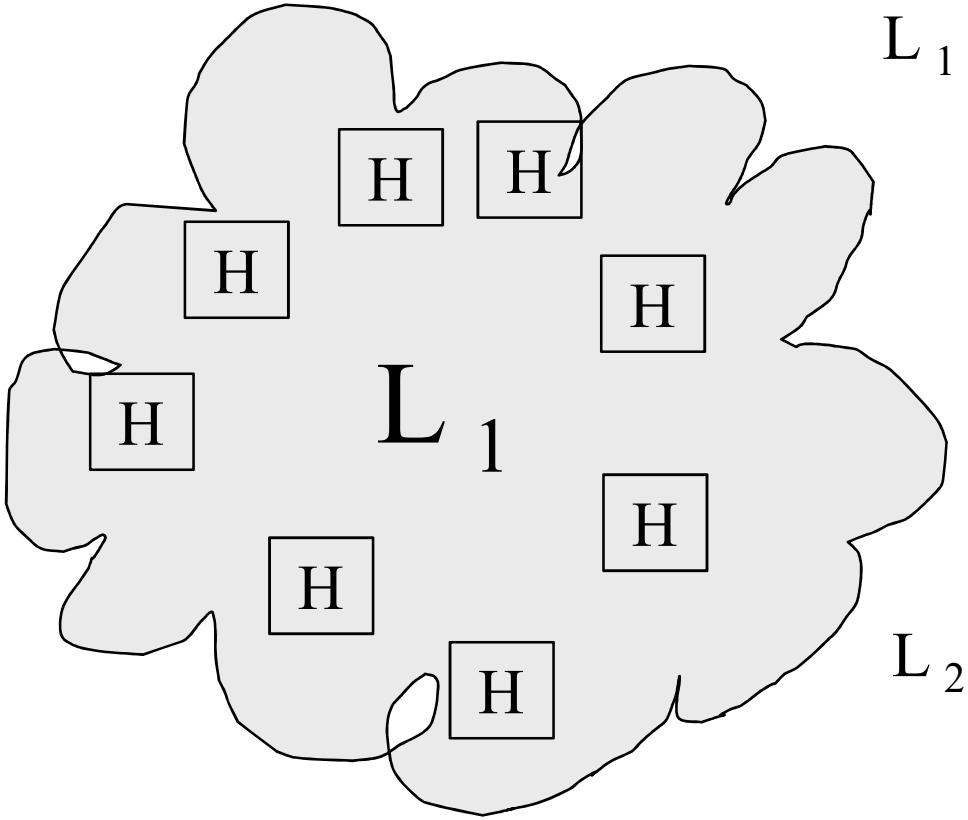
Inter-network collaboration – Why?





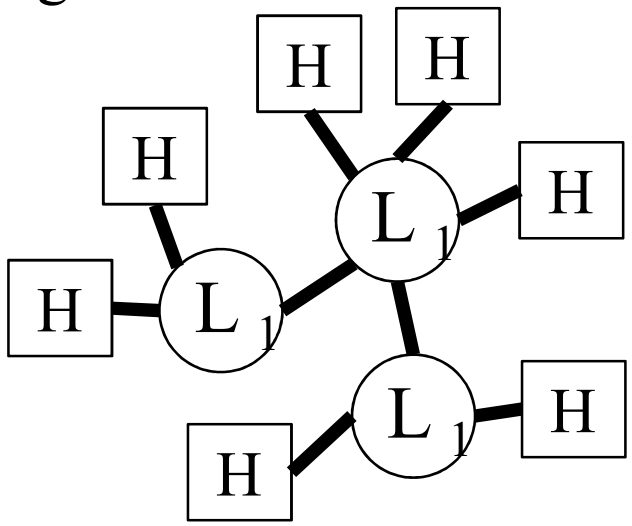
Reason:
IMP (December 1969)

Inter-network collaboration – How?

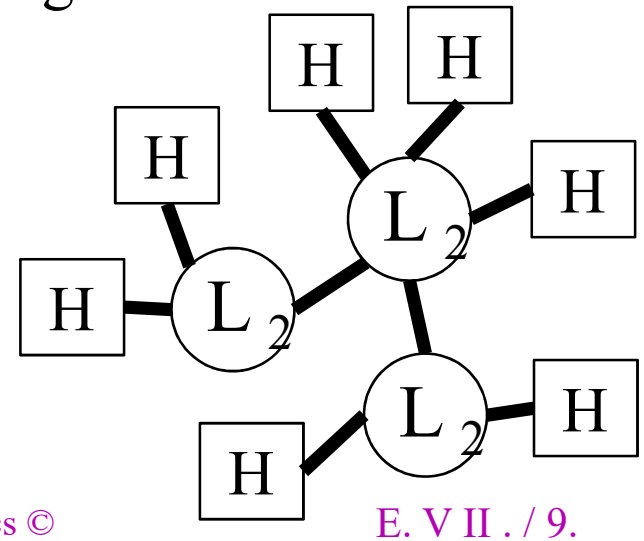


"Real" broadcasting medium

L 1 : flooding with bits

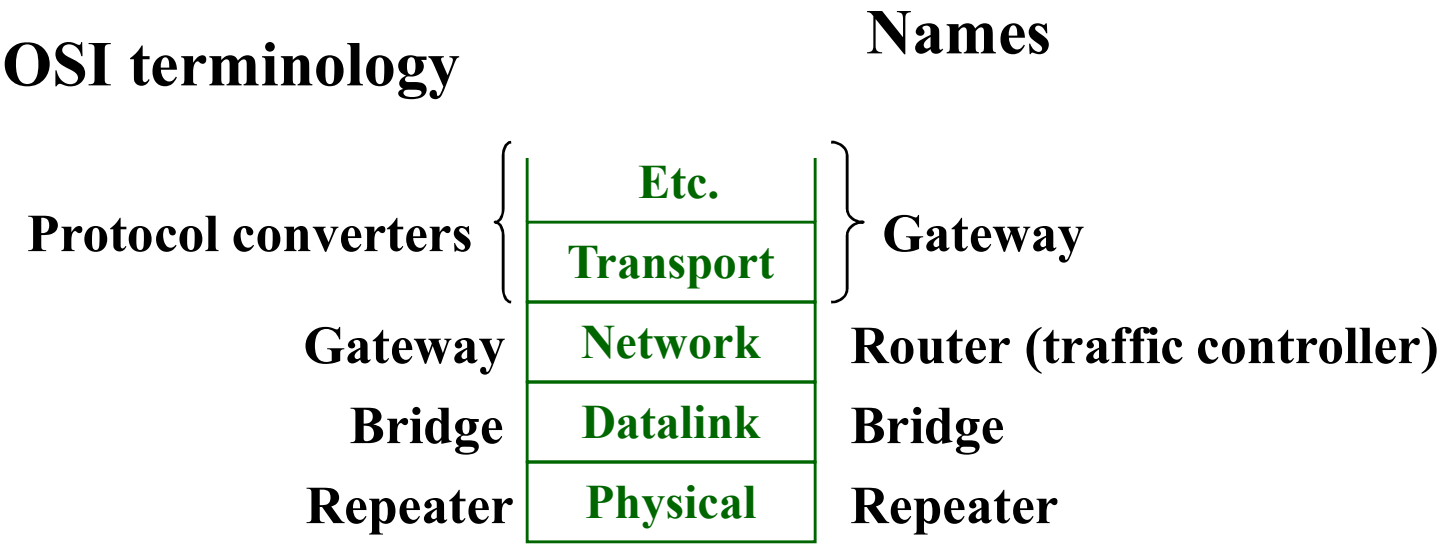


L 2 : flooding with frames



Inter-network cooperation

- **In general, it is possible to connect networks in multiple layers** (connected networks can also be of the same type).
- **General purpose of connecting networks**
  ⇒ network expansion
- **Can be grouped according to layers :**



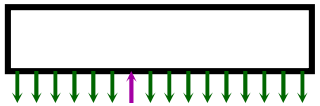
Today's main topics

- **Repeaters**
  Other names:
  - **active (passive** (see Novell resistance network)**) hub,**
  - **media converter**
- **Bridges**
  Other names:
  - **Switch, or**
  - **Layer 2 Switch**
- **Routers**
  Other names:
  - **Layer 3 Switch**
- **Gateways (protocol converter) – (not a topic today)**
  Other names (not exactly the same purpose, but similar device):
  - **Proxy**
  - **Firewall**

Inter-network cooperation – tasks of the devices

- **Expanding technological boundaries**
  (e.g.: increase the max. number of connected stations)
- **Connecting a greater distance**
  (expand network size)
- **Traffic separation**
  (load shedding (broadcasting))
- **Overcoming heterogeneity**
  (connecting different types of networks)
- **Security considerations**
  (traffic isolation, traffic filtering, firewall (Proxy))

Repeater – physical layer



Functions:

- **Overcoming limitations due to transmission medium attenuation.**
- **Converting multiple point-to-point connections into a single broadcast channel** (e.g. Ethernet on UTP).

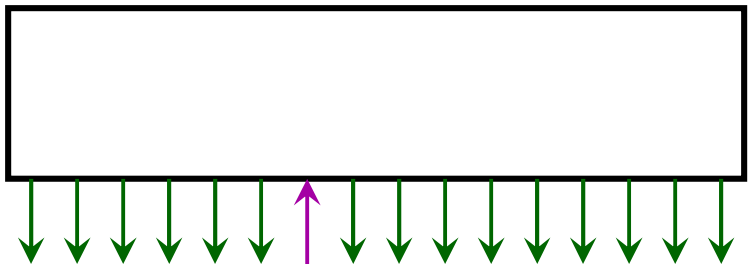
How it works:

- **Received frames** (bitstream) **after signal refresh** (in the same bit time as the reception – **without storage**) **is transmitted on all outputs** (except where it came from).
- **It does not perform any protocol or medium access functions** (except for collision forwarding in the case of CSMA/CD).
- **Connects two or more networks.**
- **Suitable for connecting different physical media (see as “ media converter”).**

Repeater – physical layer

Features:

- **It can connect different physical media, but**
- **can only connect networks with the same MAC** (Media Access Control) **procedure**
  (Does not perform media access functions)
- **Transparent to upper protocols**

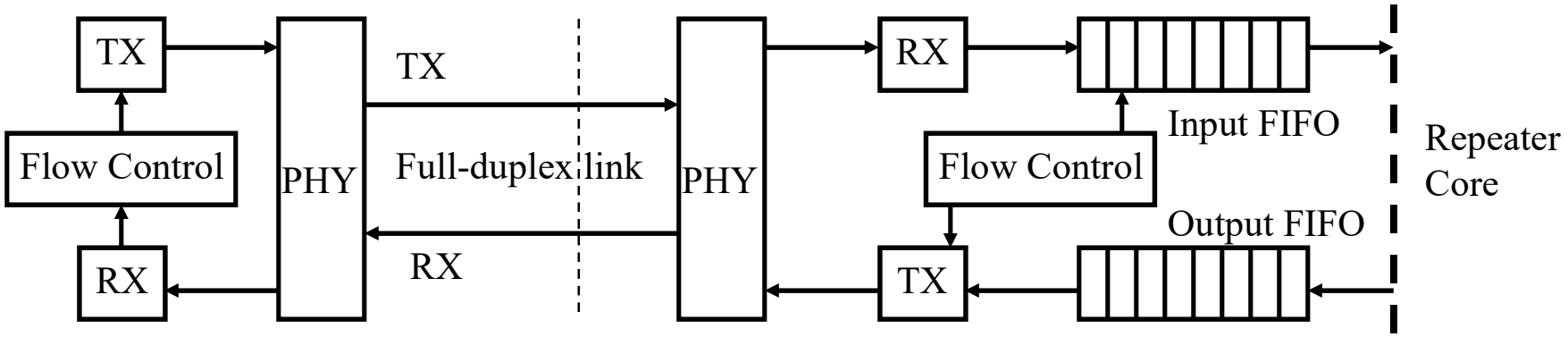


Repeater – physical layer

Special features :

- **Converting point-to-point connections into a broadcast medium** (e.g. UTP)
- **Physical media type change** (media converter)
- **Special security features on some repeaters e.g.:**
  - Usually 1 station is connected to its ports
  - **A given port will only accept frames from a** switch **whose MAC address** (6 bytes) **has been set using administrative tools. (Filtering out unauthorized station connections.)**
  - **The frame with the known destination address is only forwarded in its original form to the station with the destination MAC address**, and only a random signal of the same length as the frame is forwarded to the other connected stations. **(Preventing unauthorized interception of a broadcast channel – "Need to Know" security)**

Repeater – Full Duplex link

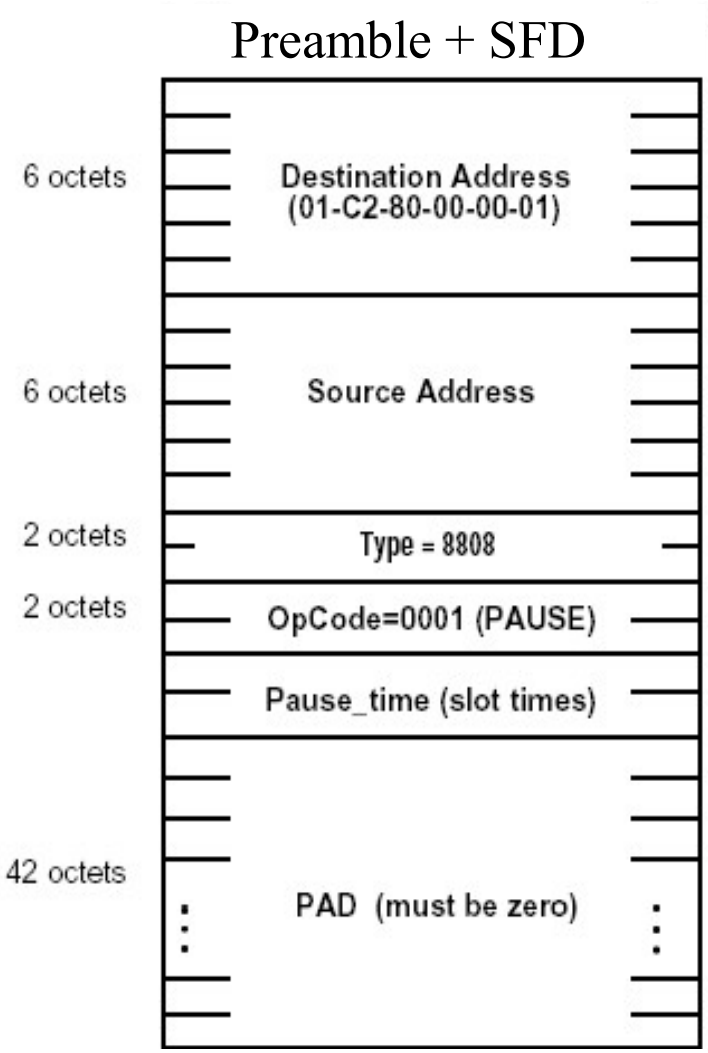


- **Buffered Distributor** (both 10/100/1000 Ethernet)
- Each port has an Input and Output FIFO queue
- A frame arriving on the Input queue is forwarded to all Output queues (except the one on which it arrived)
- CSMA/CD arbitration (scheduling) takes place within the Buffered Distributor, resulting in frames being placed in the Output queues.
- Since there are no collisions on the links, the maximum length limit of the links depends only on the physical medium (there is no round trip time limit).
- Since the sender can easily flood the FIFO, frame-level flow control (802.3x – pause frame) is used between the port and the sending station.
- A relatively inexpensive device (compared to a switch) that can handle full duplex traffic on links.

used to be like this ~2000

(802.3x – Pause Frame)

- **Devices that want to "stop" the data stream send a Pause Frame.**
- **The Pause Frame contains the time, calculated in "slot time", until the transmitter should suspend its transmission.**
- **This duration can be modified (deleted, extended) by sending additional Pause Frames (Additional Pause frames will overwrite the current pause process.)**
- **DA: 01:80:C2:00:00:01 IEEE MAC-specific Contro Protocols group address**
  IEEE 802.1D bridges do not forward
- **SA:** MAC address of the station sending the frame
- **Length/Type:** 8808.
  "MAC Control of CSMA/CD LANs"
- **Opcode:** 0001 - Pause
- **Parameters:** Pause_time. 0-65535 unsigned int.
  Calculated in 512 bit times



e.g. 1000 for 512,000 bit time, which for Gigabit Ethernet is 512µsec
(max. 65535 * 512 = 33,553,920 bit time, which is 33.554ms for Gigabit Ethernet).

Bridge – data link layer

Functions:

- **It overcomes the latency limitation of media access procedures.**
- **It connects networks that are independent from a media access control (MAC) perspective .**

How it works:

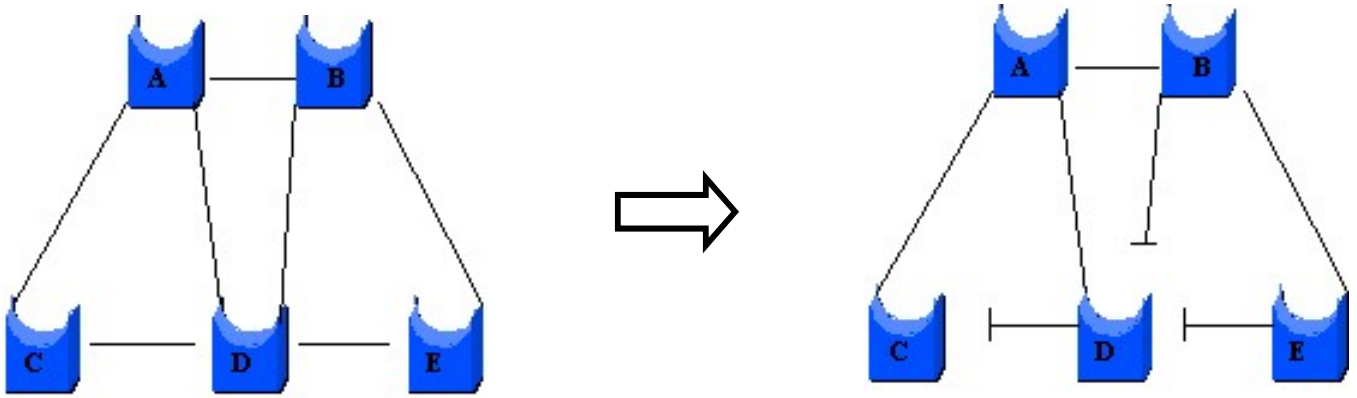
- **You can connect two or more networks .**
- **It is present as a separate station on each network** (separate media access).
- **A frame received on a network connected to it to all the other** (or one) **networks.**

Types:

- **bridges**, or **spanning** tree bridges – this is what is used today.
- **Source routing bridge** e.g. IBM Token Ring – hardly used nowadays.

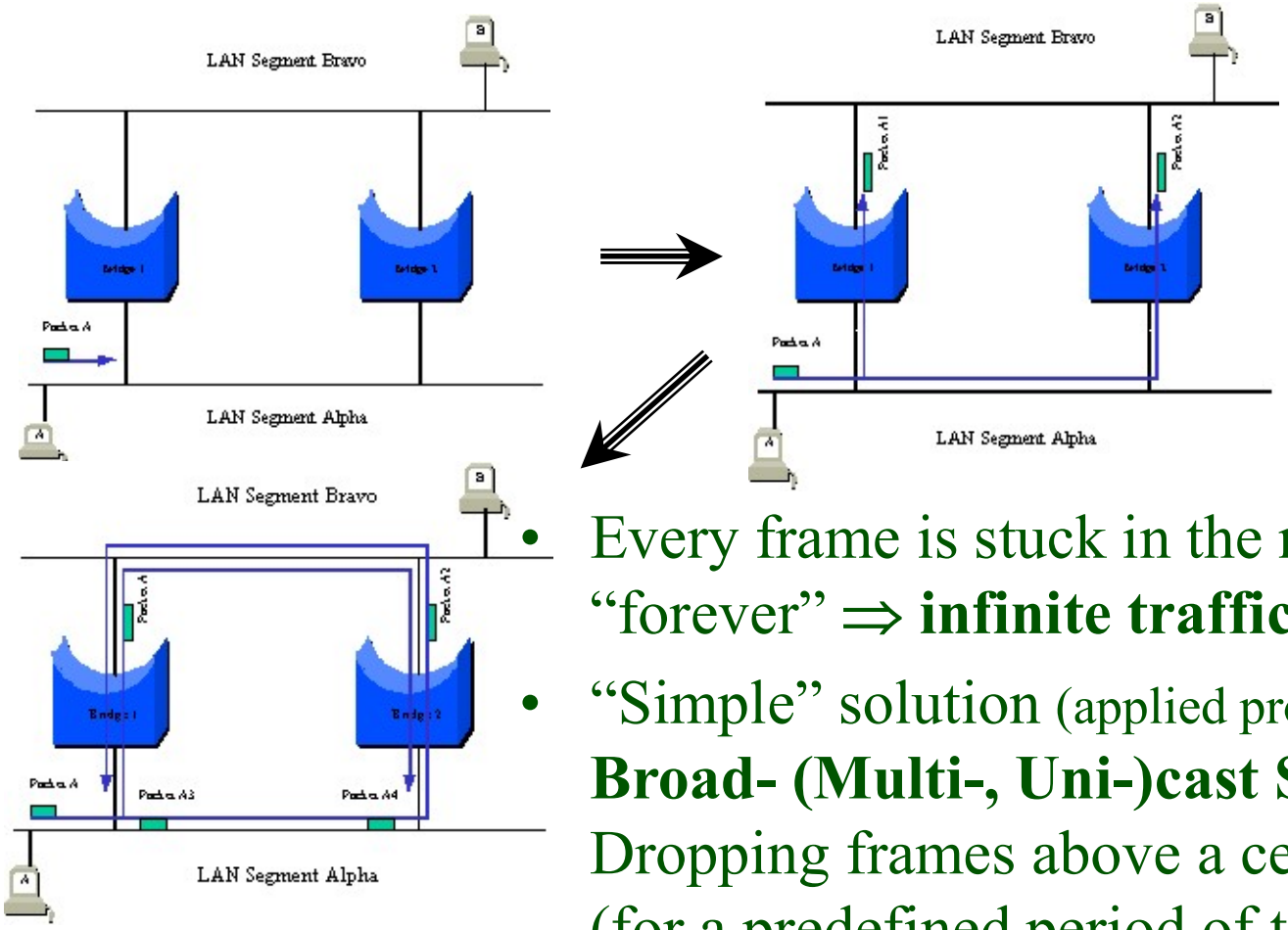
Bridge – Transparent bridges

- **(Spanning tree bridges)**
- **Goal: transform a multiply connected network into a singly connected one**
- **Solution: a "spanning tree" is fitted to the network, which contains all nodes, but has a tree** (singly connected) **topology**



Bridge – Transparent bridges

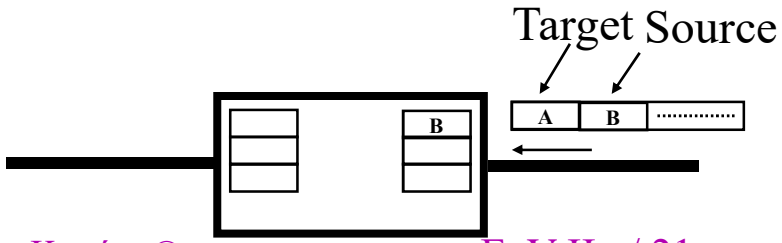
- **Redundant topology ⇒ infinite repetition**
  (no hop counter like in IP)



- Every frame is stuck in the network "forever" ⇒ **infinite traffic**
- "Simple" solution (applied protection):
  Broad- (Multi-, Uni-)cast Storm Control:
  Dropping frames above a certain load level (for a predefined period of time).

Transparent bridges – How they work

- **Backward learning**
- **It forwards in the direction corresponding to the destination using tables (hash tables).**
- **It assigns a table to each port containing the MAC addresses accessible through that port.**
- **The tables are filled based on the source addresses of packets received on that port (they are initially empty).**
- **Entries become old if x time (a few minutes) has passed since the last received sender, deleting the entry from the table** (e.g. station is moved elsewhere).
- **If the recipient is unknown, it forwards in all directions.**



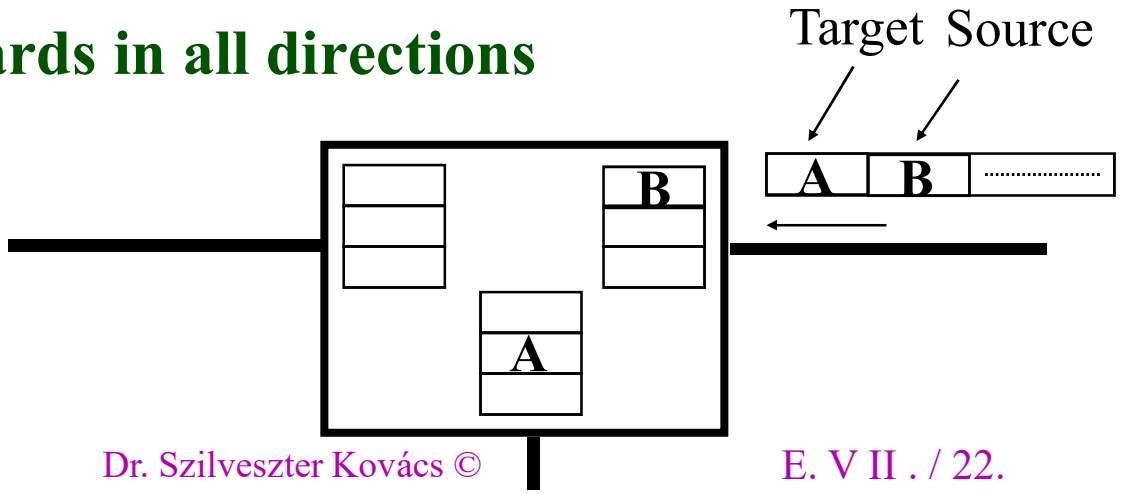
Transparent bridges – How they work

In case of two ports: only checks whether the destination address is in the hash table of the port on which the frame was received

- included ⇒ no need to forward
- not included ⇒ must be forwarded

In case of multiple ports: it also checks which port's hash table it is in if it needs to be forwarded

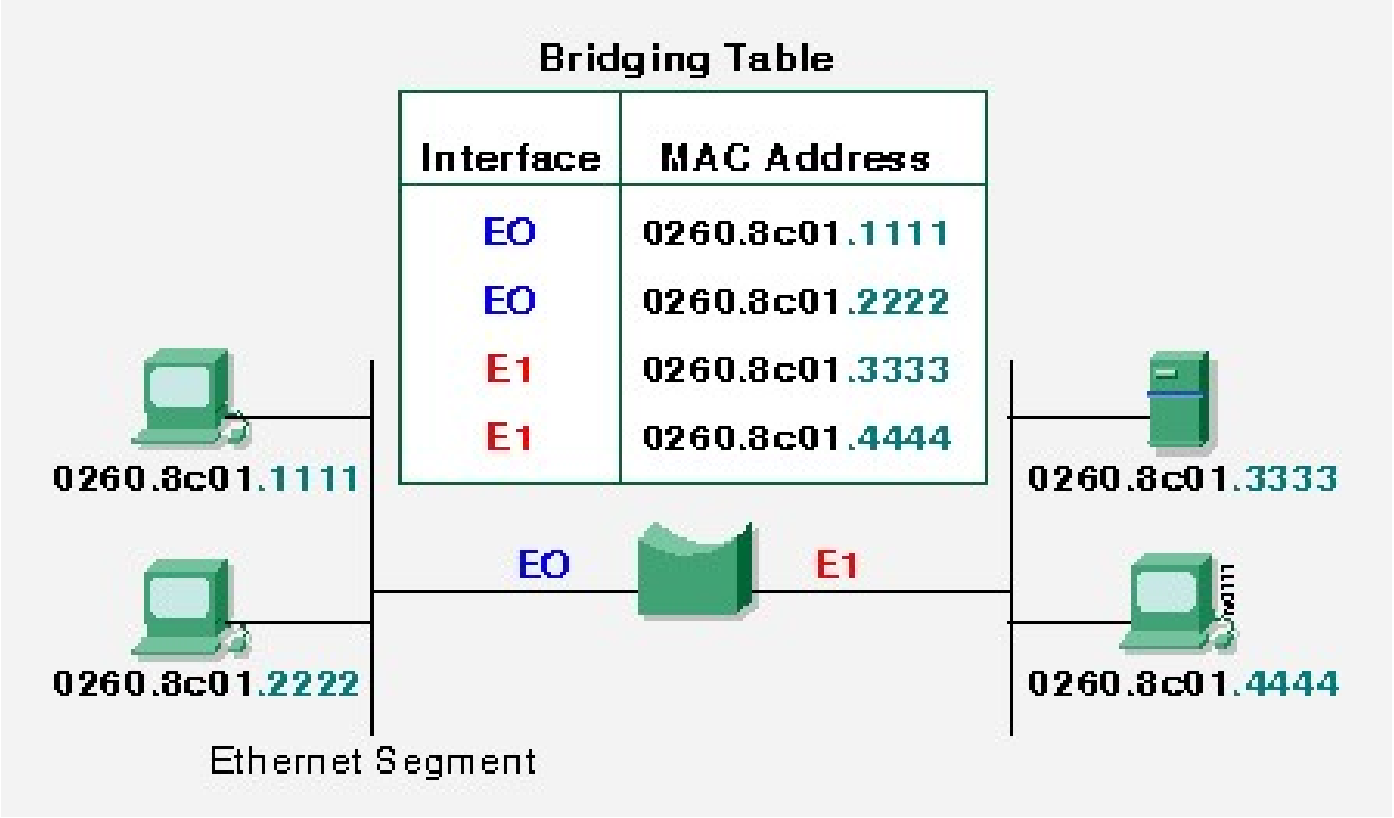
- finds ⇒ forwards it only to
- can't find it ⇒ forwards in all directions



Transparent bridges – How they work

Hash tables (**CAM** Content Addressable Memory) **:**

in practice it is just a single MAC address – port table
(an address can only be assigned to one port anyway).



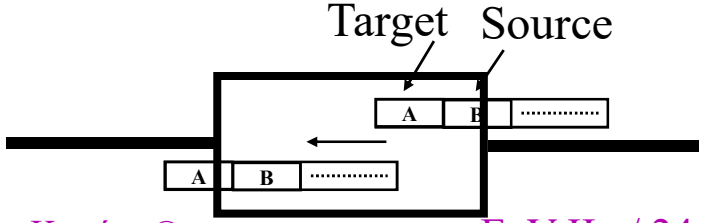
Transparent bridges – Operating modes

Store-and-forward :

- It receives the entire frame, checks it for errors (CRC) and forwards it if it is error-free.
  ⇒ Slower (longer delay), but generates less unnecessary traffic.

Cut-through :

- Fast Forward: starts sending immediately (interleaved) after receiving the target address and processing time. ⇒ lower latency, but unnecessary traffic in case of a faulty packet.
- Fragment Free: only transmits after reading 64 bytes

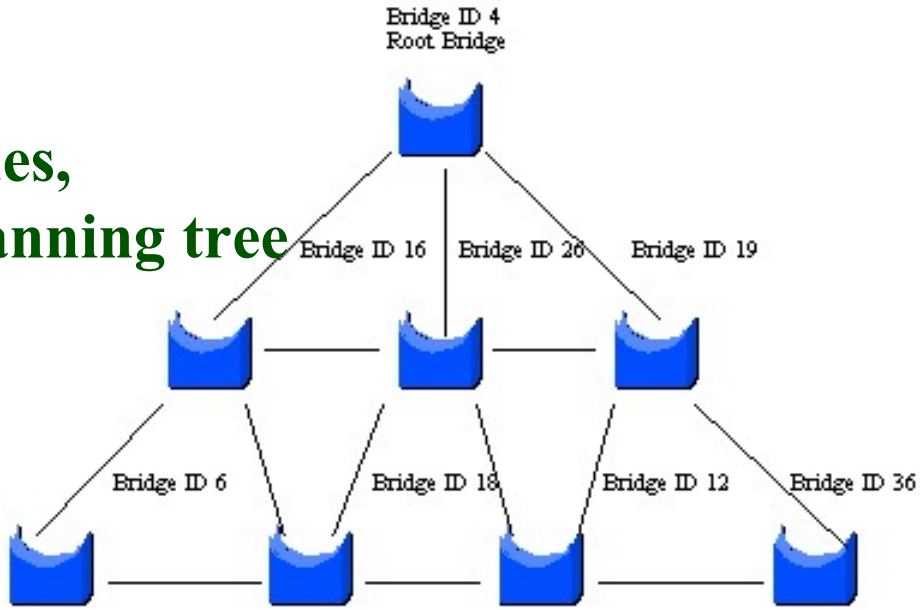


Spanning Tree Protocol (STP 802.1D)

- **Transparent bridges always start with building the spanning tree** (possible loops would be so dangerous)

How it works:

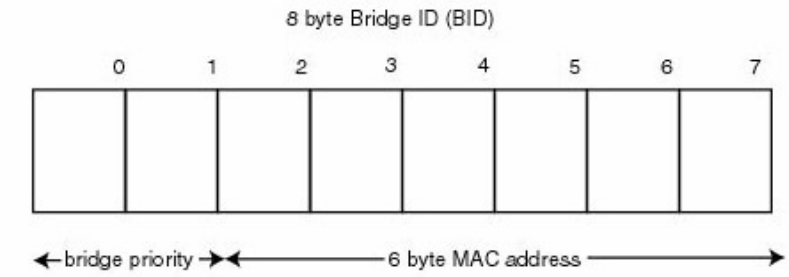
- **Starting up, each bridge sends a Bridge Protocol Data Unit (BPDU DA: 01:80:C2:00:00:00 – IEEE Bridge Group address, Nearest Customer Bridge group address) to all its ports with its own identifier.**
  (Nobody forwards the BPDU)
- **A distributed algorithm decides, who will be the root of the spanning tree (E.g. smallest number)**
- **They exchange BPDUs in pairs to see who thinks who is the root.**

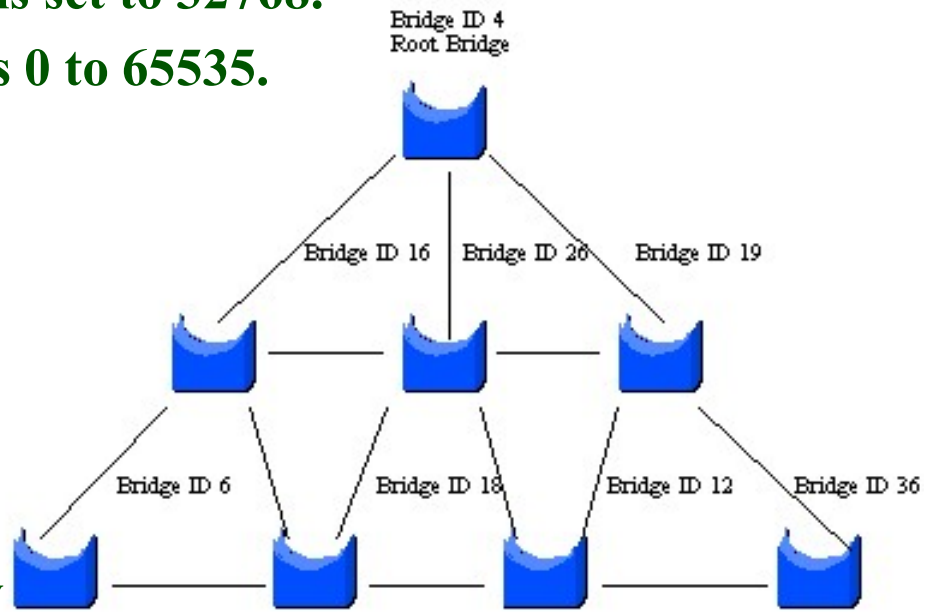


Spanning Tree Protocol (STP 802.1D)

Spanning-Tree Bridge Identifier

- In spanning tree, the 2-octet field is prepended to the 6-octet MAC address to form an 8-octet bridge identifier.
- The device with the lowest bridge identifier is considered the highest priority bridge and becomes the root bridge.
- By default, the bridge priority is set to 32768.
- The range for bridge priority is 0 to 65535.

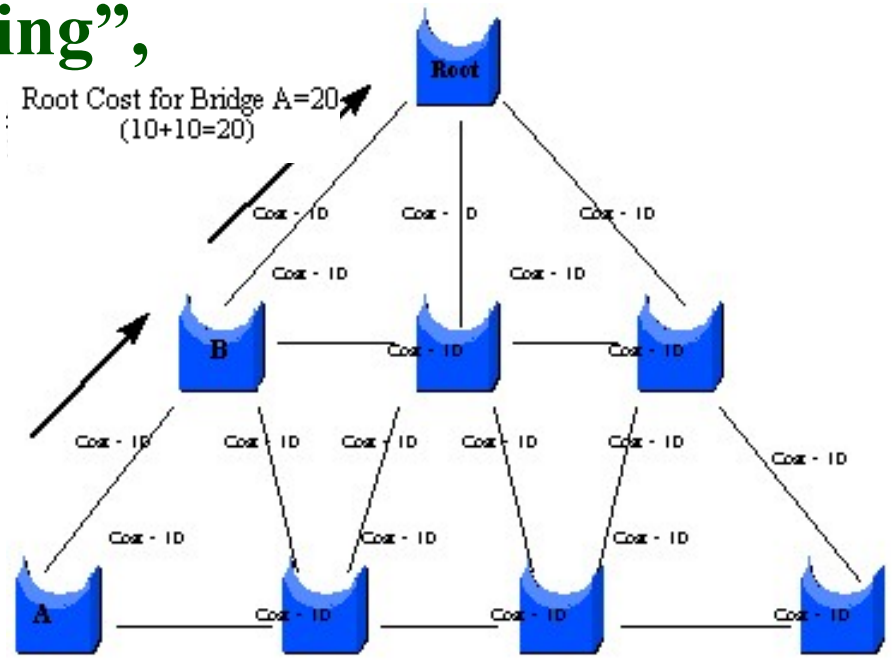




- Extended System ID:
  the lower 12 bits of the priority are the VLAN ID

Spanning Tree Protocol (STP 802.1D)

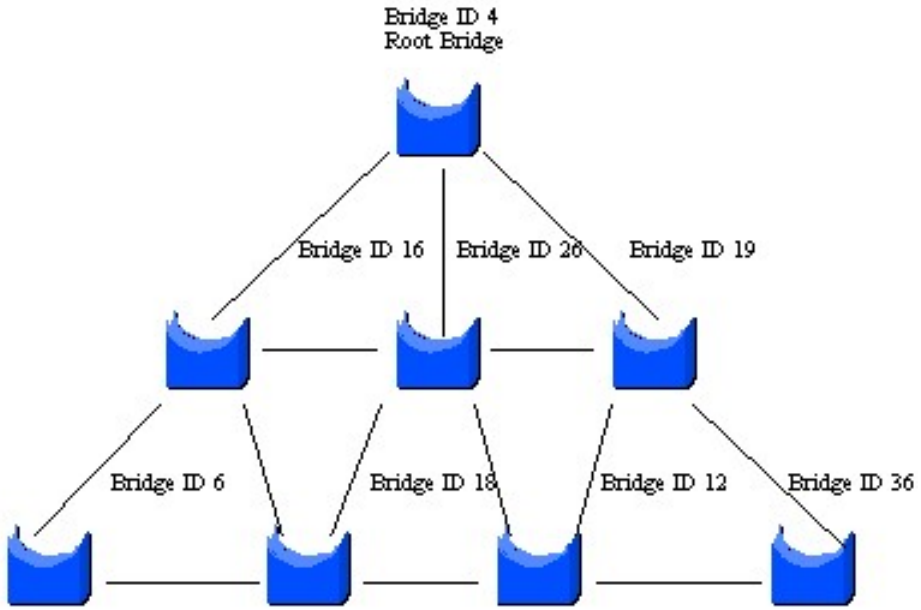
- **Once the root of the tree is found, the other bridges choose the shortest path to the root – always comparing themselves with their neighbors (in case of a match, e.g. the one with the smaller ID wins) .**
- **All links that do not fit into the shortest paths will be inactive and will be "blocking", while the links on the shortest path will be in "forwarding" state.**
  ⇒ The spanning tree is built.



Spanning Tree Protocol (STP 802.1D)

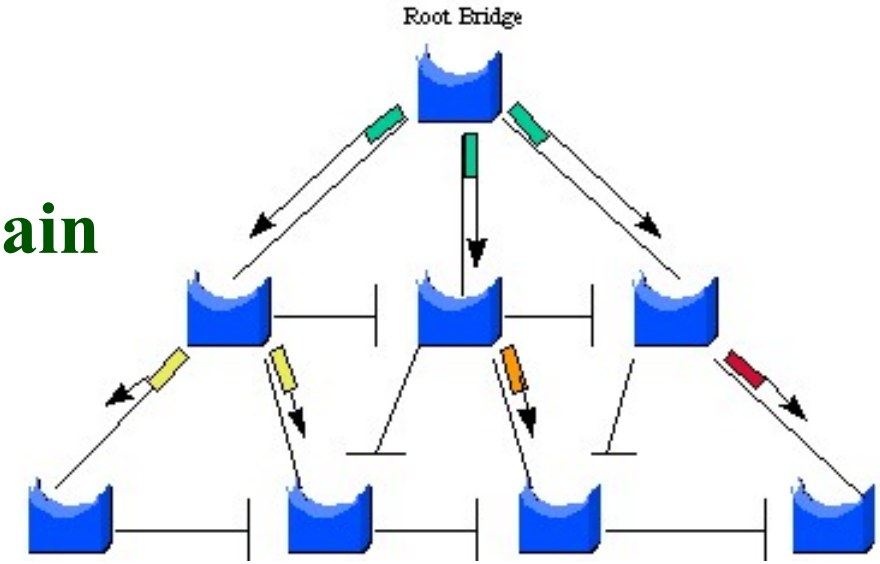
Default STP path cost of an interface for a given data rate:

- **Data rate STP Cost**
- **4 Mbps 250**
- **10 Mbps 100**
- **16 Mbps 62**
- **45 Mbps 39**
- **100 Mbps 19**
- **155 Mbps 14**
- **200 Mbps 12**
- **622 Mbps 6**
- **1 Gbps 4**
- **2 Gbps 3**
- **10 Gbps 2**



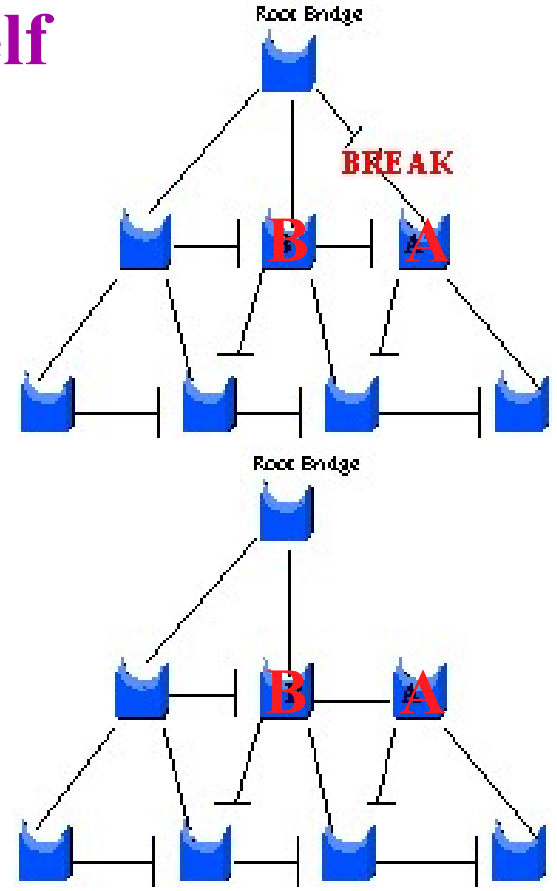
Spanning Tree Protocol (STP 802.1D)

- **Only after the spanning tree has been securely built do the bridges start forwarding the actual traffic.**
- **The root bridge at regular intervals (hello time) sends "hello" BPDUs with its own configuration.**
- **As a result, the other bridges will also be downgraded** (with their own config). **Sends "hello" BPDUs.**
- **If doesn't receive a "hello" within "max age", they start the algorithm again ("fallen from the tree")**



Spanning Tree Protocol (STP 802.1D)

- In case of topology change – if a link is broken
- The one whose "hello" timer expires (A) deletes its previous configuration and sends a BPDU to its neighbors in which it marks itself as root .
- B compares this with its own configuration, which it sees as a much better root – and sends this to A.
- A based on this BPDU, recalculates its configuration and finds that the shortest path is directed to B and subsequently advertises this.



Spanning Tree Protocol (STP 802.1D)

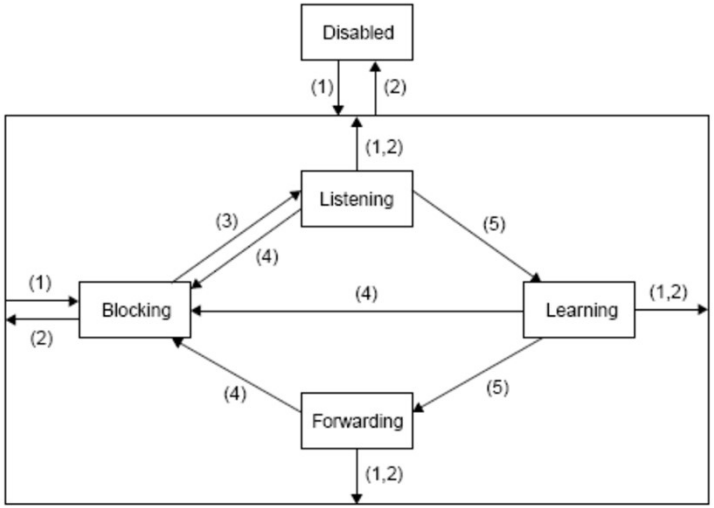
Port roles:

- **Root port: the port that receives the best BPDU on a bridge is the root port. This is the port that is the closest to the root bridge in terms of path cost. The root bridge is the only bridge in the network that does not have a root port. All other bridges receive BPDUs on at least one port.**
- **Designated port: A port is designated if it can send the best BPDU on the segment to which it is connected. On a given segment, there can only be one path towards the root bridge .**

Spanning Tree Protocol (STP 802.1D)

- **Port states:**
  The blocking and listening states are very similar, except that listening is a designated, or root , port that will soon be in forwarding state.
- **Learning is already collecting MAC addresses, but not yet forwarding.**

STP(802.1D) Port State	RSTP(802.1w) Port State	Is Port Included in Active Topology	Is Port Learning Mac Addresses?
Disabled	Discarding	No	No
Blocking	Discarding	No	No
Listening	Discarding	No	No
Learning	Learning	No	Yes
Forwarding	Forwarding	Yes	Yes

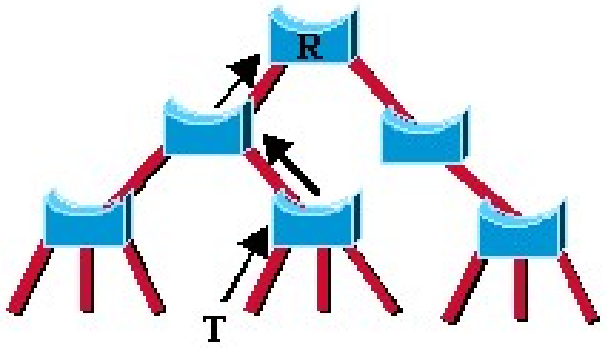


Spanning Tree Protocol (STP 802.1D)

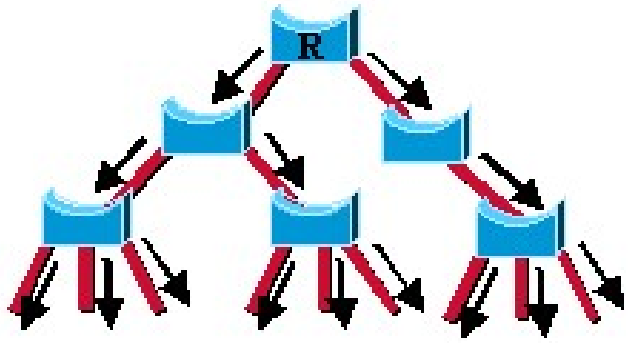
- **STP timers:**
- **Hello: The hello time is the time between each bridge protocol data unit (BPDU) that is sent on a port. This time is equal to 2 seconds (sec) by default, but you can tune the time to be between 1 and 10 seconds.**
- **Forward delay: The forward delay is the time that is spent in the listening and learning state. This time is equal to 15 sec by default, but you can tune the time to be between 4 and 30 sec.**
- **Max age: The max age timer controls the maximum length of time that passes before a bridge port saves its configuration BPDU information. This time is 20 sec by default, but you can tune the time to be between 6 and 40 sec.**

Spanning Tree Protocol (STP 802.1D)

- **When a bridge detects a change on any of its ports (the monitored ports are configurable), it sends a Topology Change Notification (TCN) BPDU to the root bridge directly (unicast) and does so until it receives an acknowledgement of receipt.**
- **The monitored ports can be configured!!!! The root bridge then uses a bit in the configuration BPDU to indicate to the network (Topology Change TC bit) that a change is taking place and everyone should reduce the validity time of the entries in the CAM (Content Addressable Memory) table (to the Forwarding Delay time).**



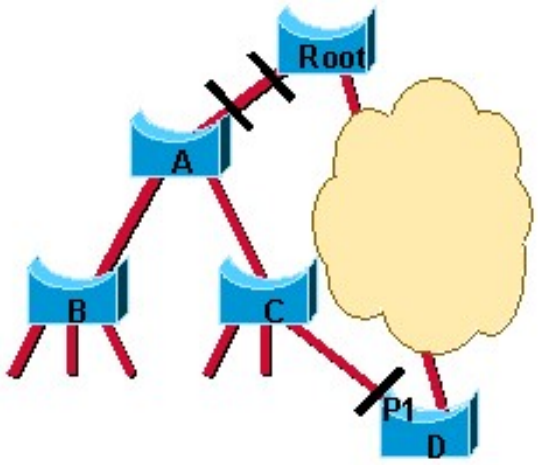
A topology change is generated on point T.
1st step: A TCN is going up to the root.



2nd step: the root advertises the TC for max-age+ forward delay.

Spanning Tree Protocol (STP 802.1D)

- When a new link appears in the network (A – Root link added)
- The STA (Spanning Tree Algorithm) blocks a port and disables the bridging loop.
- First, as they come up, both ports on the link between the root and Bridge A are put in the listening state. Bridge A is now able to hear the root directly.
- It immediately propagates its BPDUs on the designated ports, towards the leaves of the tree.
- As soon as Bridges B and C receive this new superior information from Bridge A, they immediately relay the information towards the leaves.
- In a few seconds, Bridge D receives a BPDU from the root and instantly blocks port P1.
- After waiting 2*forward delay (30 sec), the new link will enter to forwarding state (listening/learning until the change is safe)
  (no feedback that the new tree is ready) .



Very quickly, the BPDUs from the root reach D that immediately blocks its port P1. The topology has now converged, though, the network is disrupted for twice forward_delay.

Rapid Spanning Tree Protocol (RSTP 802.1W)

- **Problems with STP:**
  - **Slow convergence (MaxAge+2*ForwardDelay=20s+2x15s)**
  - **All ports are the same**
- **RSTP:**
  - **Compatible with STP**
  - **There is a chance for a faster transition to the forwarding state**
    - **The edge port type (host connection) can go directly from the blocked state to the forwarding state**
    - **Point-to-Point connection acceleration using BPDU handshake**

Rapid Spanning Tree Protocol (RSTP 802.1W)

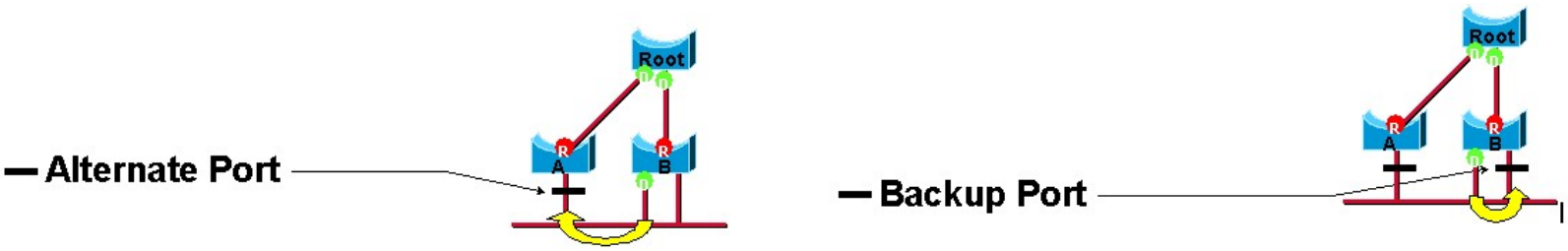
- **Fewer port status**

STP(802.1D) Port State	RSTP(802.1w) Port State	Is Port Included in Active Topology	Is Port Learning Mac Addresses?
Disabled	Discarding	No	No
Blocking	Discarding	No	No
Listening	Discarding	No	No
Learning	Learning	No	Yes
Forwarding	Forwarding	Yes	Yes

Rapid Spanning Tree Protocol (RSTP 802.1W)

Newer port roles :

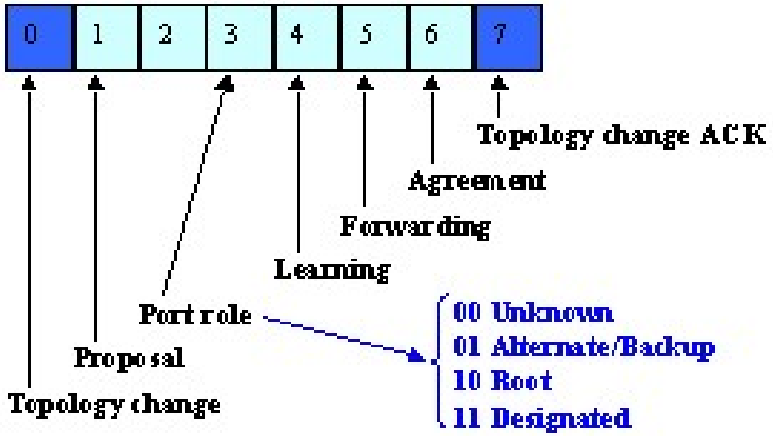
- **Root port, Designated port**
- **Alternate port :**
  An alternate port receives more useful BPDUs from another bridge and is a port blocked.
- **Backup port :**
  A backup port receives more useful BPDUs from the same bridge it is on and is a port blocked.
- **Acceleration: If you break away from the root, you can immediately choose these**



Rapid Spanning Tree Protocol (RSTP 802.1W)

New BPDU flag bits:

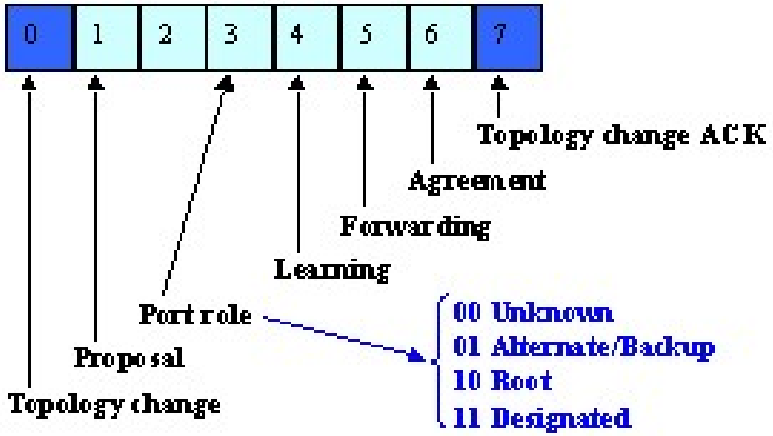
- In STP only: Topology Change (TC) and TC Acknowledgment (TCA) bits
- Encode the role and state of the port that originates the BPDU
- Handle the proposal/agreement mechanism



Rapid Spanning Tree Protocol (RSTP 802.1W)

New BPDU handling :

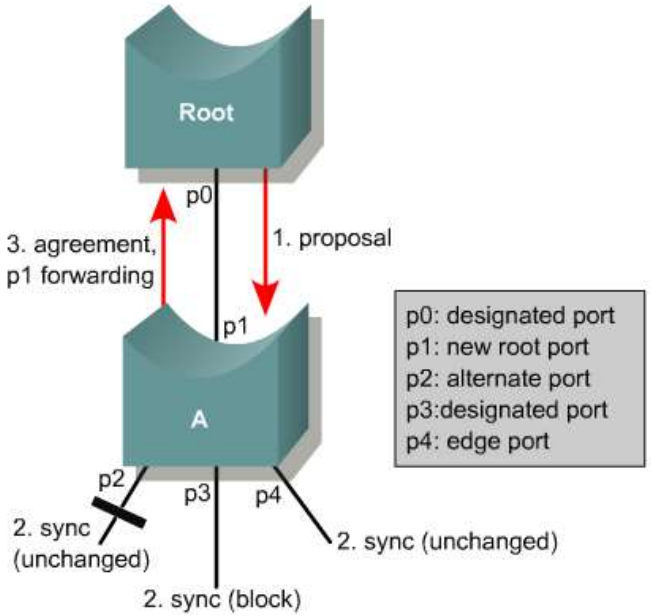
- In STP, HELLO BPDUs are sent only by the Root, the others repeat them
- In RSTP, BPDU is sent by everyone in Hello time
- And then if someone does not receive BPDUs for 3 Hellos, the connection is considered to be lost.

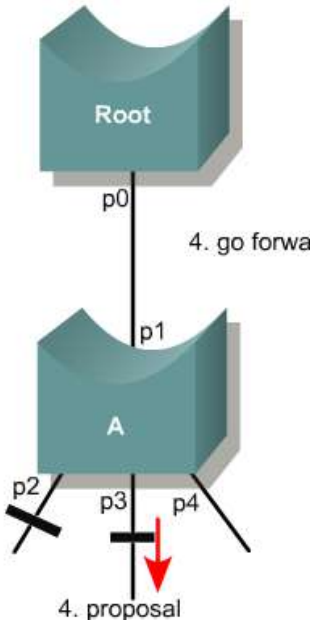


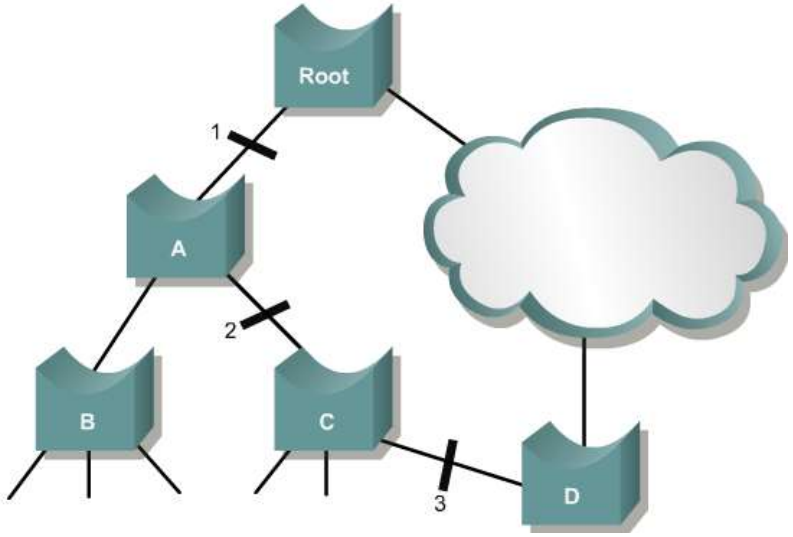
Rapid Spanning Tree Protocol (RSTP 802.1W)

Faster convergence – new link appears
proposal/agreement sequence:

- **The designated port will be discarding and will send a "proposal", if it receives an "agreement" (repeat of the proposal, only with agreement flag), it will go into forward state.**
- **If it receives and accepts a proposal, it blocks its previous forward port (sync) and then sends agreement**



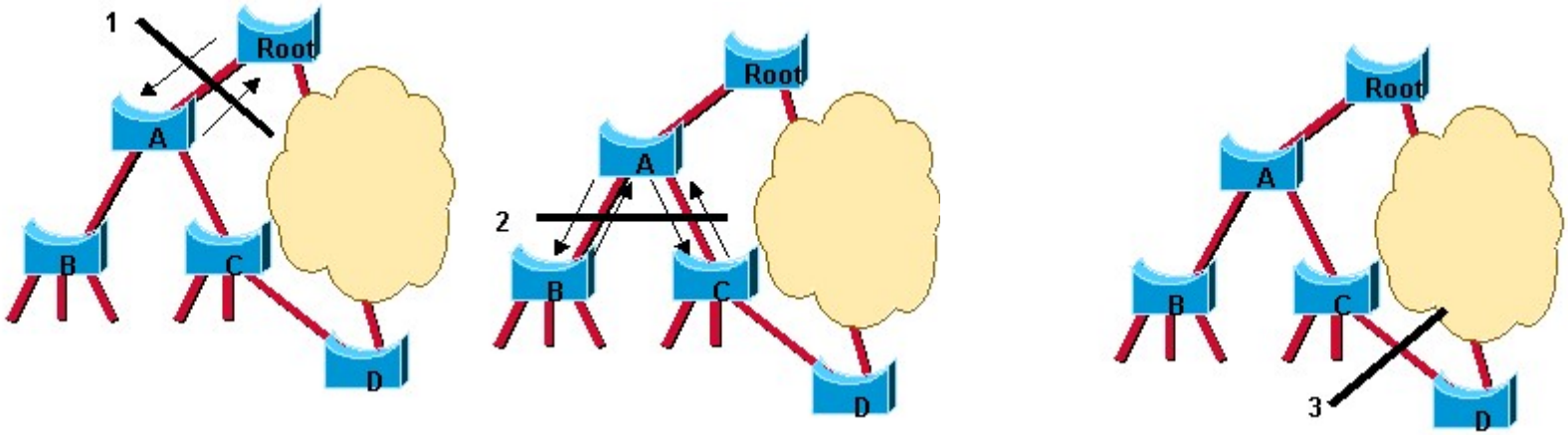




Rapid Spanning Tree Protocol (RSTP 802.1W)

Faster convergence – proposal/agreement sequence:

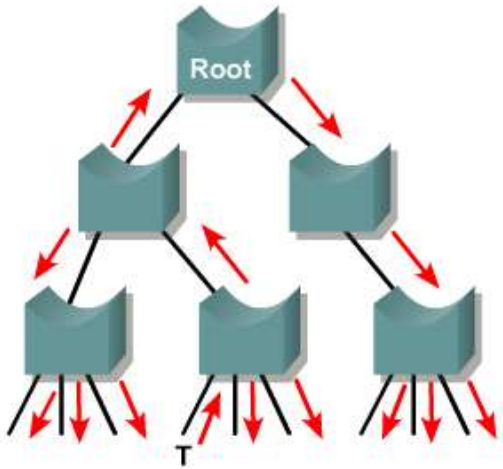
- **Starts a "handshake wave"**
- **Converges much faster than STP**



Rapid Spanning Tree Protocol (RSTP 802.1W)

RSTP topology change :

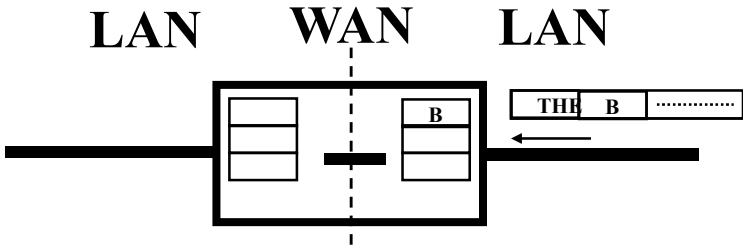
- Only the change in non-edge ports is interesting!
- Not only the root sends messages periodically, but all bridges
- Due to topology change (change):
  - Starts a timer (twice the hello time)
  - It waits for BPDUs with the TC bit set on all designated ports and the root port until the timer expires. It clears the CAM on these ports.
  - Whoever received this does the same. (TC spreads everywhere) But it does not empty the CAM belonging to the incoming port)
  - (No need to go to the root and then back to STP)



Transparent bridges – other

"Dual Speed (10/100) Hub":

- **Two repeaters connected by a bridge**



Remote bridge:

- **A two-port bridge with a hash table on only one port**
  $\Rightarrow$ it does not filter on the other port, it always forwards what comes from there (it is already filtered)
  E.g.: connecting two remote networks, LAN-WAN connection

Security features

- **You can specify the maximum number of different MAC addresses that can be used on a port – if there are more than this, then either**
- **deletes the oldest and learns the newer one, or**
- **it discards the newer ones, in this case it can be configured to block the affected port when the new address appears**
  - Disconnect Unauthorized Device (DUD)

Transparent bridges – other

Spanning Tree can be disabled (if the topology doesn't require it, don't waste your time with it)

Channel:

- **Channeling different ports – distributes load between the ports of the channel.**
  (The ports of the channel are considered a single link from the perspective of Spanning Tree.)

Managing priorities:

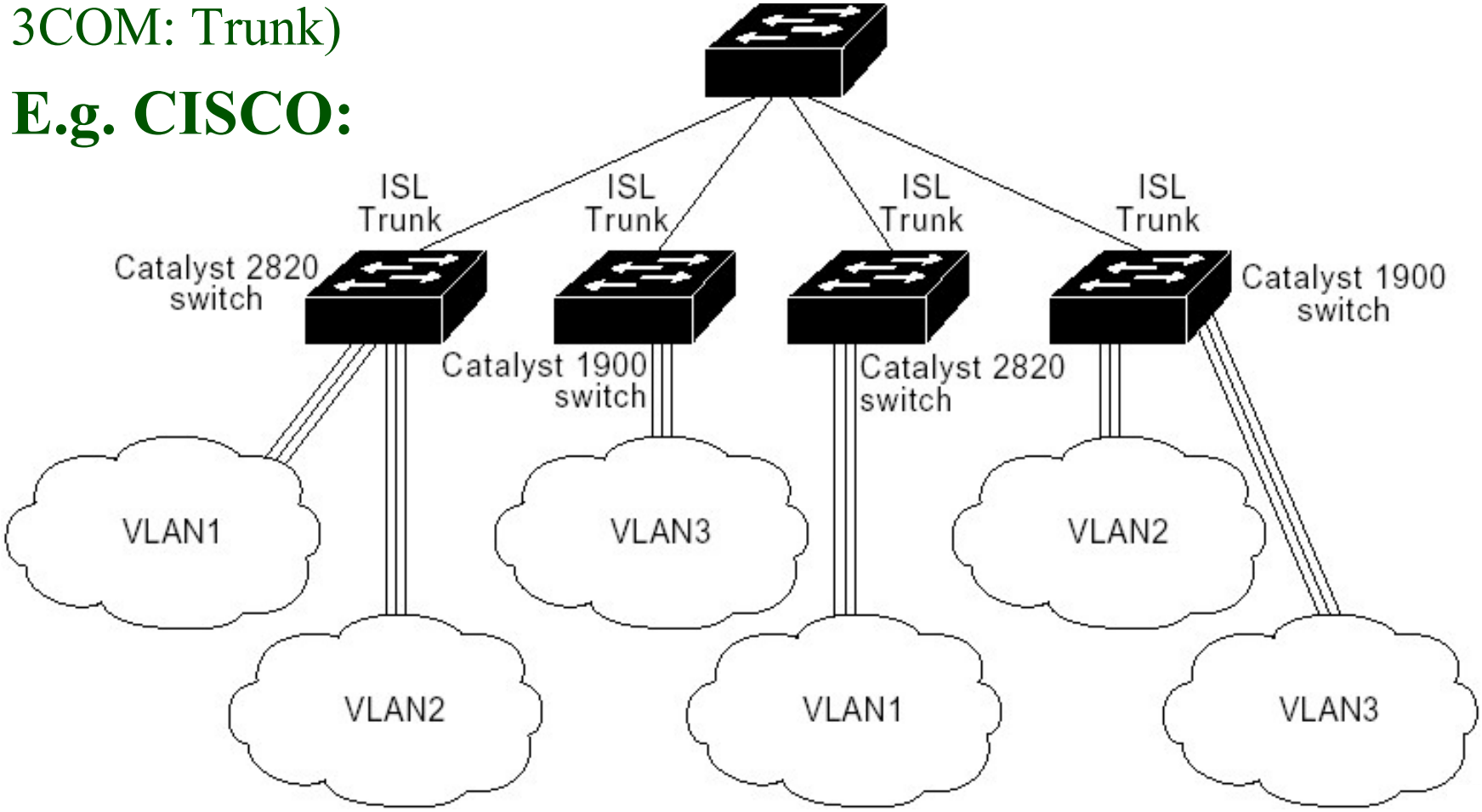
- **Each priority class can have a separate queue on each port** (typically two) **and allows higher priority ones to pass first.**
  (This is the layer, where priority is first handled – queues are first appearing (except for duplex repeaters).)

Transparent bridges – VLAN

- **Virtual LANs (VLANs) are managed at the bridge level.**
- **Each port can be assigned to a VLAN.**
- **If a port belongs to multiple VLANs, the frames are marked with VLAN Tags, to indicate the VLANs they belong to. (Cisco: ISL, IEEE: 802.1q)**
- **If a port belongs to only one VLAN, the switch can remove the appropriate tag from frames going out on the port and add it to frames going in.**
  $\Rightarrow$ VLANs are transparent to such hosts.
- **Tables are created and switched separately for each VLAN (virtually independent LANs on the same infrastructure)**

Transparent bridges – VLAN

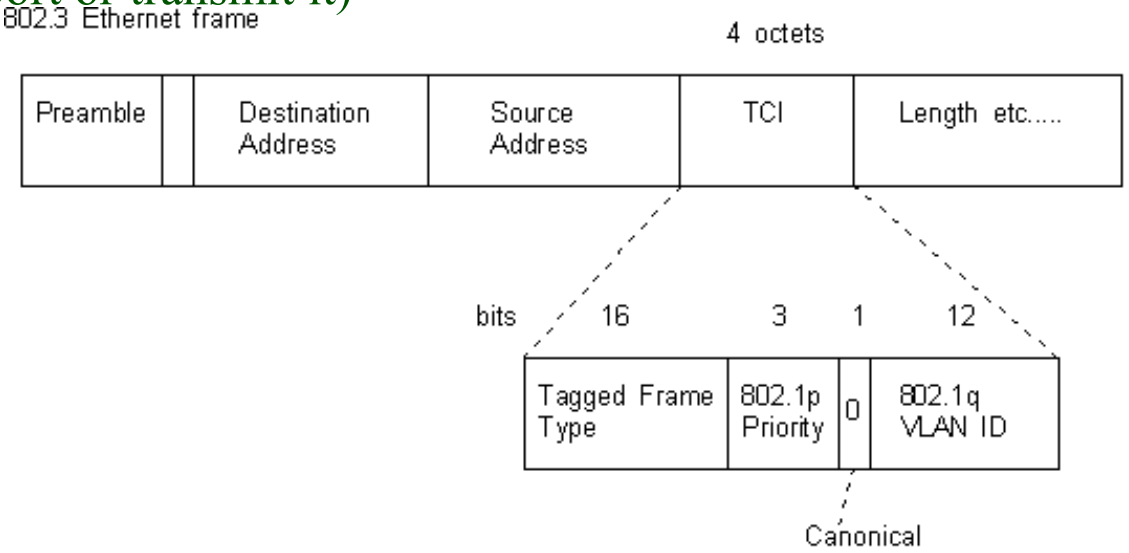
- **The name of the link marked with VLAN Tags:**
  CISCO: Trunk (the names of the merged links: CISCO: Channel, 3COM: Trunk)
- **E.g. CISCO:**



Transparent bridges – VLAN (802.1p, 802.1q)

- **Class of Service and VLANs (802.1p, 802.1q)**
- **Tag Control Info (TCI)**
  - **Extra 4 bytes only** ⇒ maximum frame size 1518 → 1522 (not all devices support or transmit it)

Cisco: Inter Switch Link (ISL) frame encapsulation (completely different)



- **Tagged Frame Type** – Tag type, currently always for Ethernet frames **0x8100**.
- **802.1p Priority code point (PCP)** – from low priority binary 000 (0) to high priority binary 111 (7) (“**class of service**”)
- **Drop eligible indicator (DEI) (formerly “ Canonical” 0)** – Together with PCP, “class of service”, indicates that the frame can be discarded in case of congestion
- **802.1q VLAN ID** - the VLAN ID on VLAN trunks.
- The **CRC must be recalculated at the end of the frame.**

Transparent bridges – VLAN (ISL)

- **Inter-Switch Link (ISL) - Cisco proprietary protocol**
- **ISL is an Ethernet frame encapsulation (not an inner, but an outer envelope)**
- **ISL encapsulation: 26 byte header + 4-byte CRC**
- **10-bit VLAN ID**

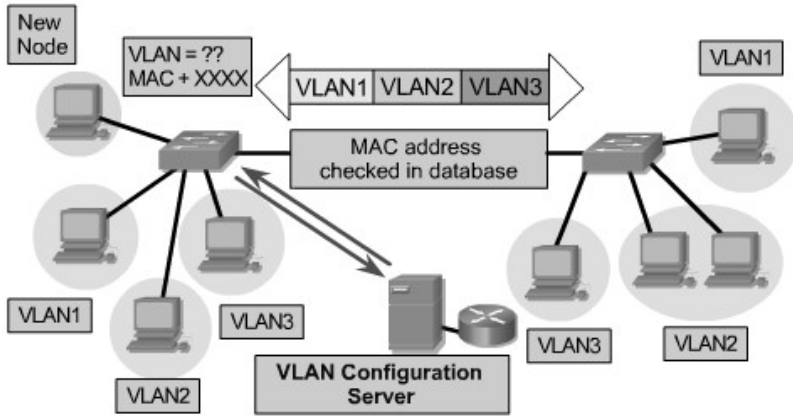
40 bits	4 bits	4 bits	48 bits	16 bits	24 bits	24 bits	15 bits	1 bits	16 bits	16 bits	Variable length	32 bits
DA	TYPE	USER	SA	LEN	SNAP/ LLC	HSA	VLAN ID	BPDU/ CDP	INDX	Reserved	Encapsulated Frame	FCS (CRC)

Transparent bridges – VLAN (ISL)

Octet	Description
DA	A 40-bit multicast address with a value of 0x01-00-0C-00-00 that indicates to the receiving Catalyst that the frame is an ISL encapsulated frame.
Type	A 4-bit value indicating the source frame type. Values include 0 0 0 0 (Ethernet), 0 0 0 1 (Token Ring), 0 0 1 0 (FDDI), and 0 0 1 1 (ATM).
User	A 4-bit value usually set to zero, but can be used for special situations when transporting Token Ring.
SA	The 802.3 MAC address of the transmitting Catalyst. This is a 48-bit value.
Length	The LEN field is a 16-bit value indicating the length of the user data and ISL header, but excludes the DA , Type, User, SA, Length, and ISL CRC bytes.
SNAP	A three-byte field with a fixed value of 0xAA-AA-03.
HSA	This three-byte value duplicates the high order bytes of the ISL SA field.
VLAN	A 15-bit value to reflect the numerical value of the source VLAN that the user frame belongs to. Note that only 10 bits are used.
BPDU	A single-bit value that, when set to 1, indicates that the receiving Catalyst should immediately examine the frame at an end station because the data contains either a Spanning Tree, ISL, VTP, or CDP message.
Index	The value indicates what port the frame exited from the source Catalyst.
Reserved	Token Ring and FDDI frames have special values that need to be transported over the ISL link. These values, such as AC and FC, are carried in this field. The value of this field is zero for Ethernet frames.
User Frame	The original user data frame is inserted here incuding the frame's FCS.
CRC	ISL calculates a 32-bit CRC for the header and user frame. This double-checks the integrity of the message as it crosses an ISL trunk. It does not replace the User Frame CRC.

Transparent bridges – VLAN

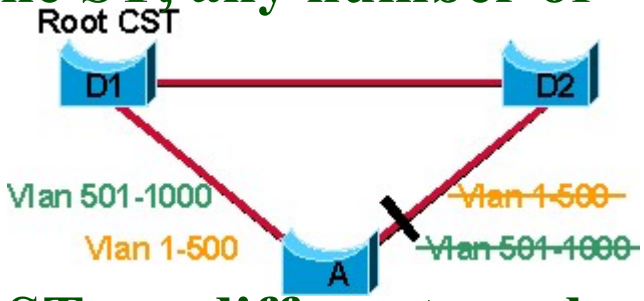
- **Static VLAN :**
  - **Port-based, if there is no VLAN tag on the link, then the port clearly belongs to a pre-configured VLAN.**
- **Dynamic VLAN :**
  - **MAC address based, the MAC address of the connected station determines which VLAN it is connected to.**
  - **MAC-VLAN mapping must be pre-configured in a VLAN Management Policy Server (VMPS).**

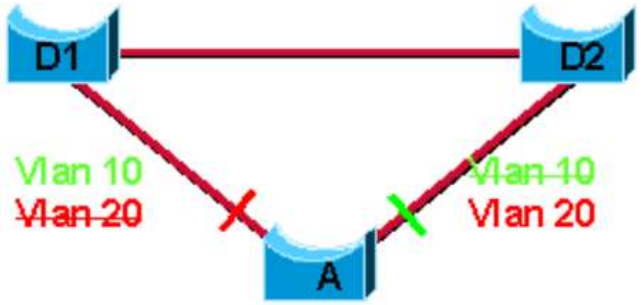


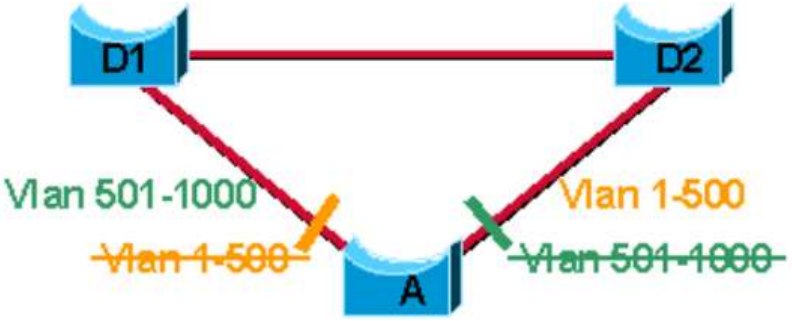
Transparent bridges – VLAN and STP

Problem: if there are multiple VLANs, it would be good to take advantage of the topology redundancy.

- **CST: Common Spanning Tree only one ST, any number of VLANs (802.1q default)**
- **PVST: Per-VLAN Spanning Tree as many STs as there are VLANs**
- **MST (802.1s) : Multiple ST, as many STs as different can be in the topology and the VLANs are distributed to them**



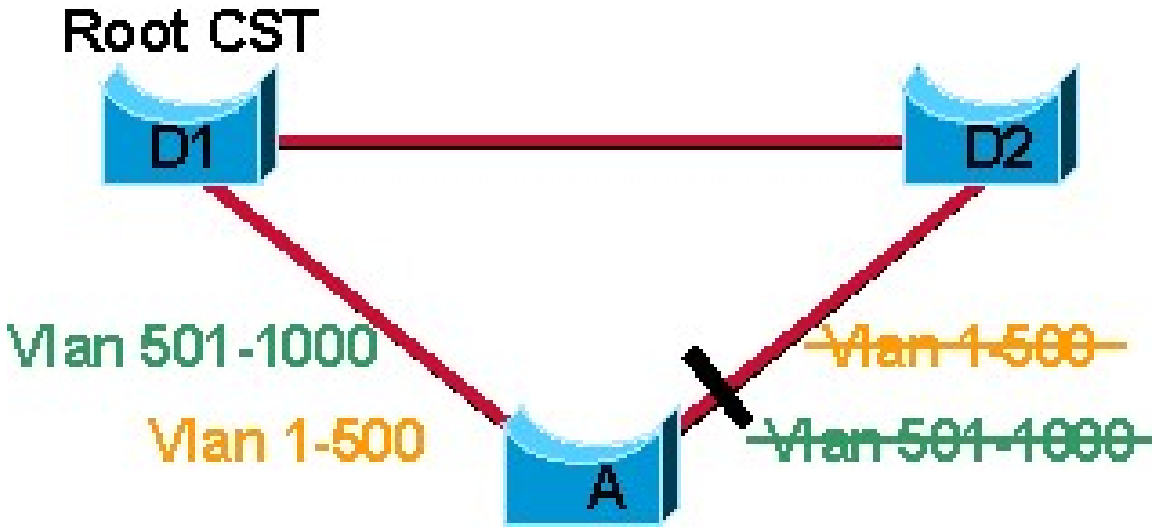




Transparent bridges – PVST

CST: Common Spanning Tree

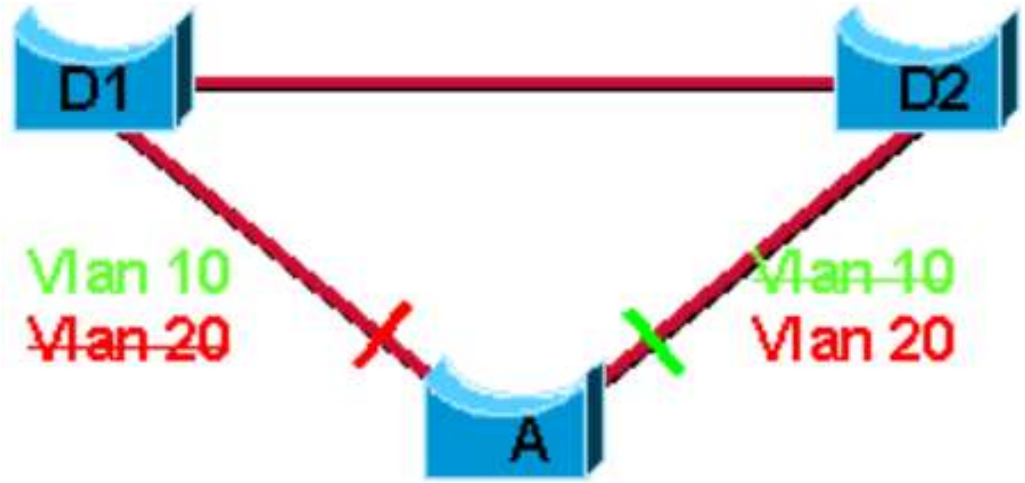
- **Only one ST, any number of VLANs (this is the 802.1q default)**
- **No network load balancing available**



Transparent bridges – PVST

PVST: Per-VLAN Spanning Tree

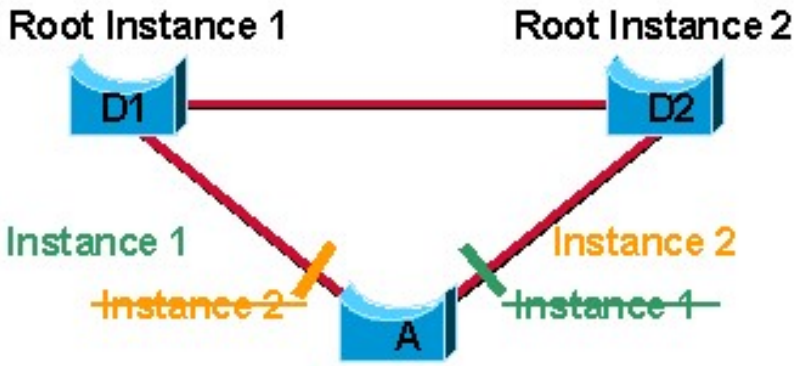
- **As many STPs as VLANs**
- **Each STP can have a different root bridge**
- **High CPU load, but there is an option to distribute the network load**

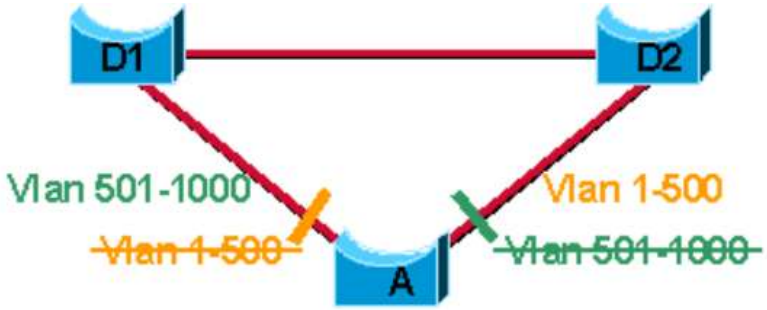


Transparent bridges – MST (802.1s)

MST (802.1s) : Multiple ST

- **Multiple ST instances (MST instances), as many as different are possible in the topology and the VLANs are distributed to them**
- **Manages VLAN groups**
- **A VLAN can only belong to one MST instance.**
- **One switch can have multiple MST instances**
- **Low CPU load, network load balancing possible**





Transparent bridges – MST (802.1s)

Each MST bridge stores the following:

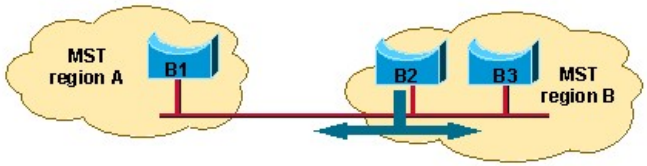
- An alphanumeric configuration name (32 bytes)
- A configuration revision number (two bytes)
- A 4096-element table that associates each of the potential 4096 VLANs supported on the chassis to a given MST instance ((4096 different VLANs possible) table for VLAN - MST instance (RSTP) about binding

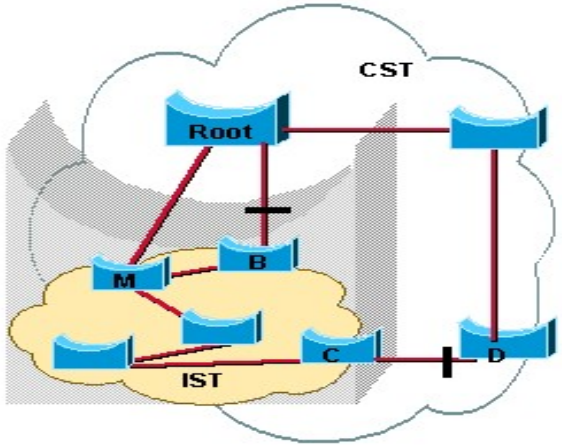
MST region :

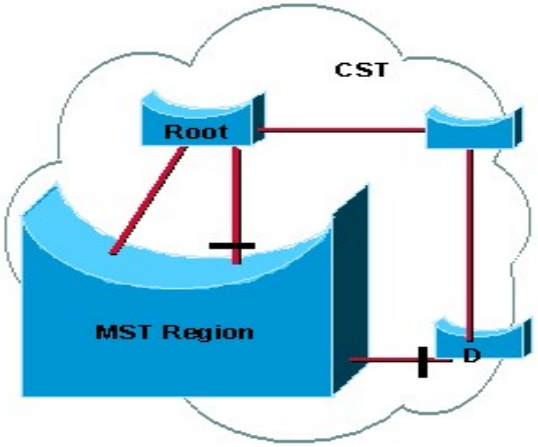
- Bridges that contain the same configuration belong to the same region.
- There is no recommendation for the distribution of this configuration…

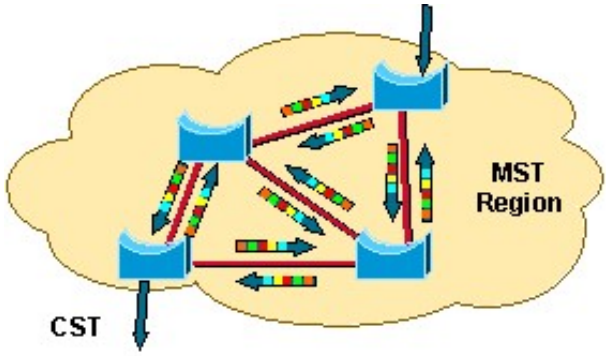
To operate, you need to know the exact limits:

- The configuration digest is also included in the BPDU
- If this is different from your own on a port, then that port is a border port.



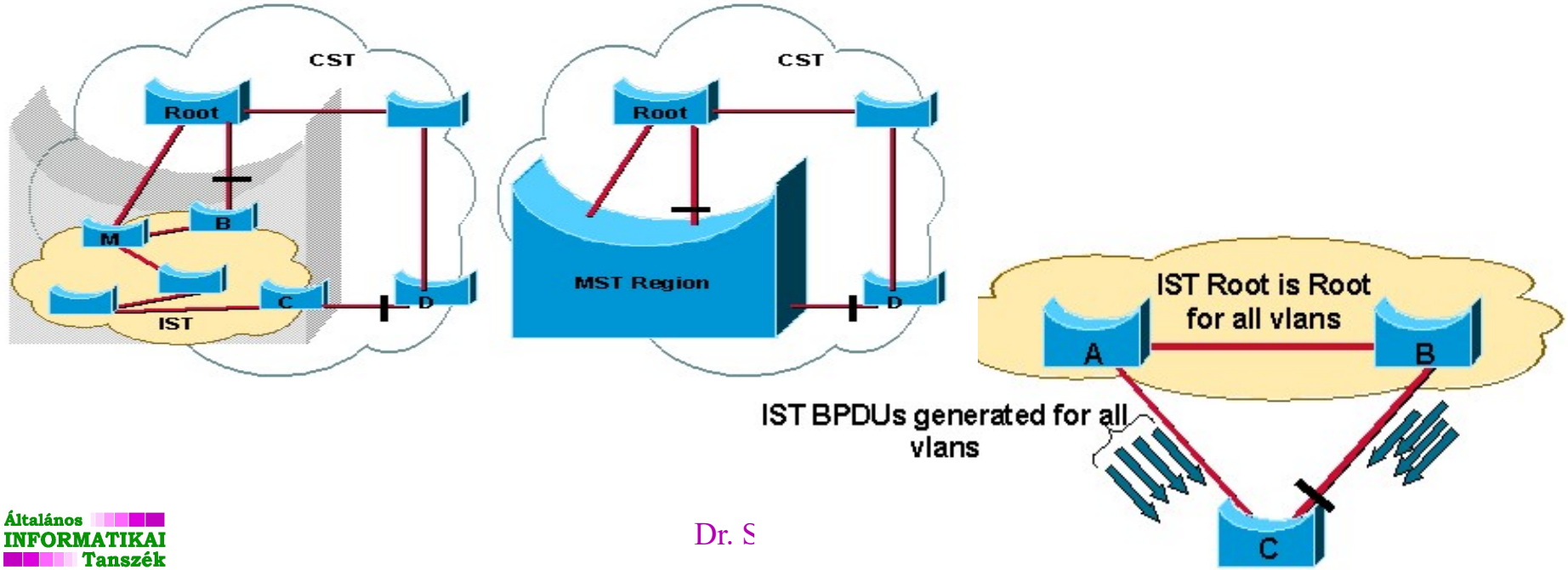






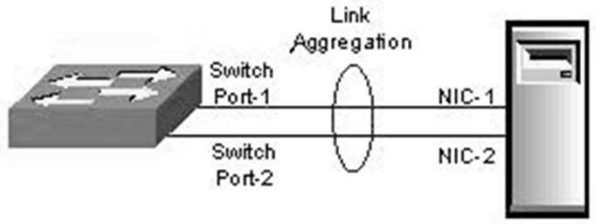
Transparent bridges – MST (802.1s)

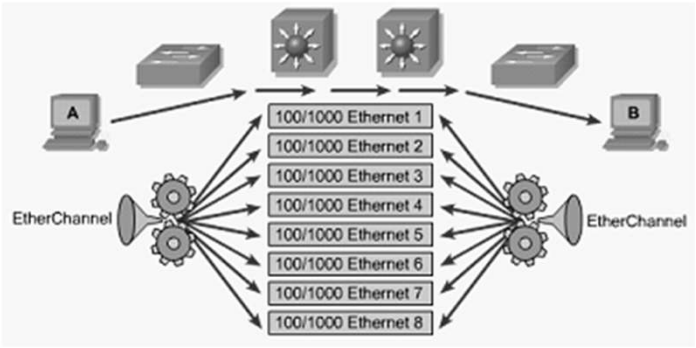
- Any number of MSTIs (Multiple Spanning Tree Instances)
- Each MSTI has its own ROOT within the MST region
- An IST (Internal Spanning Tree) is also available (from the outside, the entire region appears as a bridge, which communicates according to IST (as the MST region now replicates the IST BPDUs on every VLAN at the boundary), so it is compatible with both plain CST (802.p) and PVST.
- The ROOT may be outside the IST, or it may not even know MST (it gets lost at MST borders)
- Recommendation: IST ROOT should have the highest priority (so it will be ROOT for all PVSTs)



Transparent Bridges – Link Aggregation

- "Line merging"
- STP, RSTP, MSTP
  - Puts the redundant paths in a blocked state
  - No dynamic load sharing
- If the links belong to a high-speed category (10BaseT, Fast Ethernet, Giga, 10G)
  - We can combine several lines into one line (e.g. max 8)
  - Can only be used for point-to-point connections between two devices
  - Can only be used in full-duplex mode
- Advantages:
  - Dynamic load distribution
  - High availability
  - Fast, automatic reconfiguration (~1s)
  - Transparent for the upper layers
  - Deterministic
  - Configurable





Transparent Bridges – Link Aggregation

IEEE: 802.3ad, Link Aggregation Control Protocol (LACP)

- **10,100,1000 Mbps**
- **Usually, for example, in the case of 3COM this is called "Trunk", Cisco: "Channel", Linux: "Bonding"**

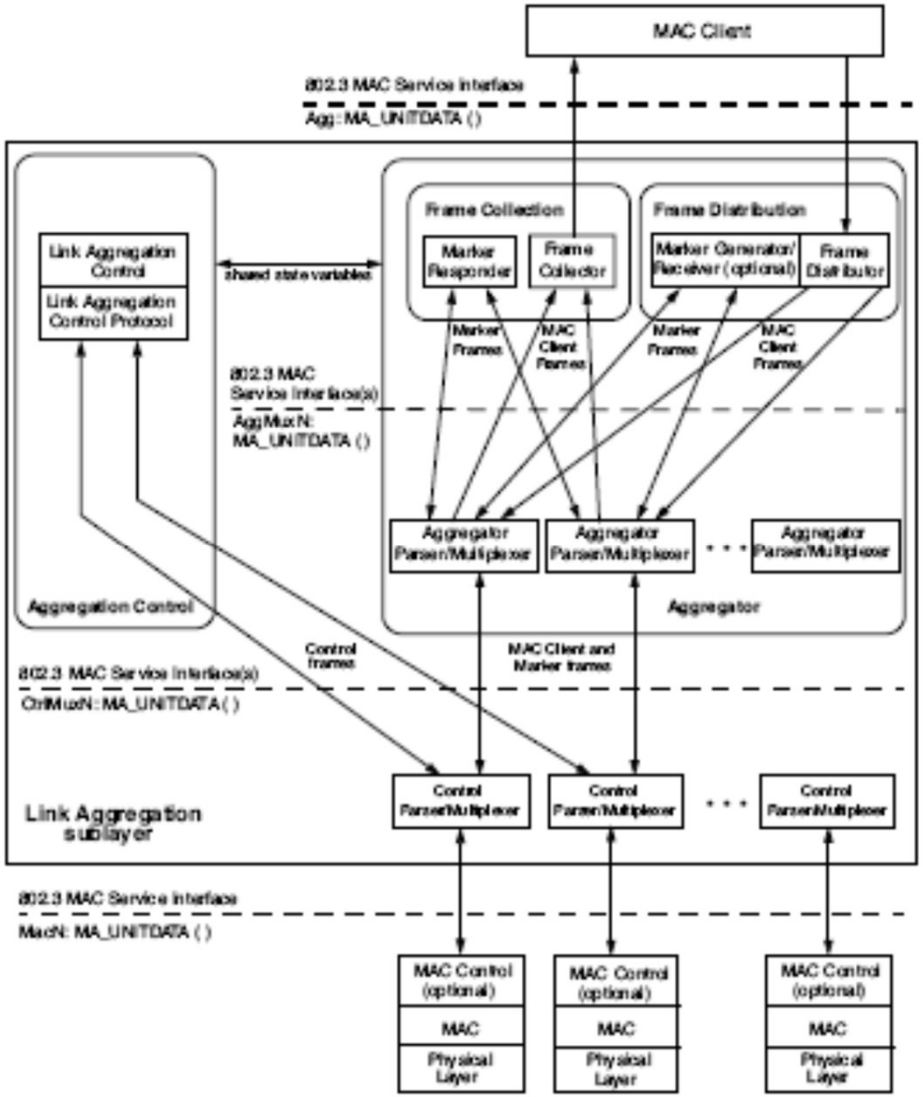
Other proprietary trunking protocols (preceded 802.3ad):

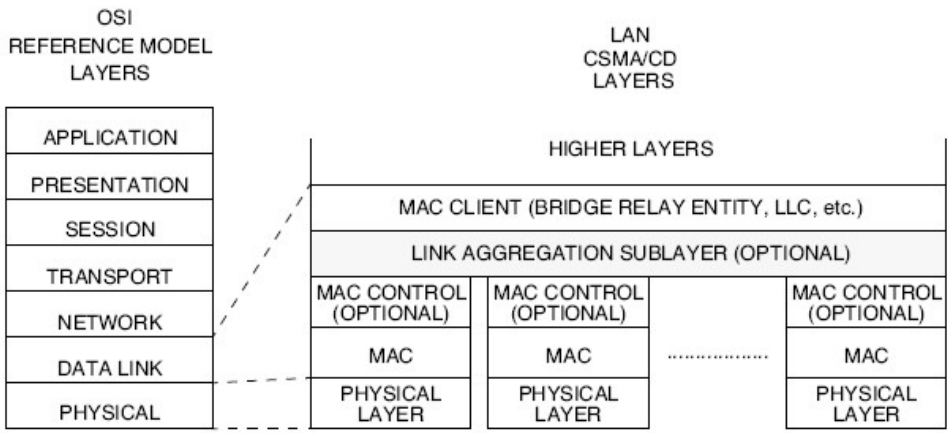
- **Cisco: EtherChannel , Port Aggregation Protocol (PAgP)**
  - **EtherChannel**
  - **FastEtherChannel**
  - **GigaEtherChannel**
  - **10GigaEtherChannel (0.16TBit/s!!)**
- **Adaptec : Duralink trunking**
- **Nortel Multi-link Trunking (MLT) ...**
- **They are usually only compatible with devices from the same manufacturer, and sometimes only with certain devices from the same manufacturer**

Transparent Bridges – Link Aggregation

The architecture of 802.3ad

- Frame Collector/Distributor
- Aggregator
- Aggregation Control





Transparent Bridges – Link Aggregation

Frame distribution between merged links

- **IEEE 802.3ad does not specify**
- **Cisco:**
  - **L2: Source/Destination by MAC address**
  - **L3: Source/Destination by IP address**
  - **L4: By port**
- **They can cause problems: e.g. frame order swapping**

Transparent Bridges – Link Aggregation

Negotiation :

- **IEEE: Link Aggregation Control Protocol (LACP) - 802.1ad**
- **Active : sends LACP PDUs, its counterpart on the other end of the line:**
- **Passive : if it receives, it responds to it, but he does not initiate it**

Mode	Description
`On`	Forces the port to channel without LACP. With the on mode, a usable EtherChannel exists only when a port group in on mode is connected to another port group in on mode.
`Off`	Prevents the port from channeling.
`Passive`	Similar to the automode for PAgP, places a port into a passive negotiating state, in which the port responds to LACP packets it receives but does not initiate LACP packet negotiation. This is the default.
`Active`	Similar to the desirable mode for PAgP, places a port into an active negotiating state, in whch the port initiates negotiation with other ports by sending LACP packets.

Transparent Bridges – Link Aggregation

Negotiation :

- **Cisco: Port Aggregation Protocol (PAgP)**
  - **The switch automatically discovers the capabilities of the other side and builds the optimal number/type of link groups**

Mode or Keyword	Description
`On`	The mode that forces the port to channel without PAgP. With the on mode, a usable EtherChannel exists only when a port group in on mode is connected to another port group in on mode.
`Off`	The mode that prevents the port from channeling.
`Auto`	The mode that places a port into a passive negotiating state, in which the port responds to PAgP packets it receives but does not initiate PAgP packet negotiation. This is the default setting.
`Desirable`	The mode that places a port into an active negotiating state, in which the port initiates negotiations with other ports by sending PAgP packets.
`silent`	The keyword that is used with auto or desirable mode when no traffic is expected from the other device. This options prevents the link from being reported to the Spanning Tree Protocol as down, this is the default, secondary PAgP setting.
`non-silent`	The keyword that is used with auto or desirable mode when traffic is expected from the other device.

Transparent bridges – characteristics

- **It is also suitable for connecting networks with different MAC protocols (**in practice, it is not used for this unless necessary, because different frame formats cause problems - e.g. different max. frame lengths - but e.g. the **802.11 WLAN Access Point** does exactly this)
  Solution :
  - **encapsulation bridging ⇒ one frame format is "wrapped" into the other, and then it is "unwrapped" upon exit or arrival at the destination. E.g.: FDDI-Ethernet bridge** (wraps Ethernet into FDDI)
- **Transparent to most protocols**
- **Limited ability to separate traffic** (floods during learning)
- **Suitable for covering long distances**
  ⇒ no theoretical limit, only practical ones e.g. max. delay
- **There is no theoretical limit to the size of the network or the number of stations** (any size network can be built from it), **but there are practical limits:**
  Broadcast Storm (Broadcast Domain), maximum Hash table sizes.

Source Routing Bridge

- **Source Routing Bridge** (e.g. IBM Token Ring – hardly used nowadays)
- **It assumes that each station knows the complete path to the destination address and writes it into the frame to be transmitted (*Directed Frame*).**
  - **(** E.g. IBM Token Ring frame: **Routing Information Field (RIF) - in this case the first bit of the Source MAC address is 1 (like the multicast MAC), max. 15 route entries)**
- **Bridges then forward frames according to the (RIF) list .**
- **Exploring the routes – source by flooding:**
  - **Explorer frame *Frame*) sends to the destination.**
    - **ARE: all-routes explorer (all directions)** IBM TR, max. 15 hops
    - **SRE: single-route explorer (along spanning tree)** TR max.15 hops
  - **forwarding, the bridges insert their own identifier** (Route Descriptor (Bridge ID + Ring ID) into the Explorer Frame RIF) into the discovery frames.
  - **When the first reconnaissance group arrives at the finish line ⇒ contains the optimal route.**
  - **This is sent back to the source (*Directed Frame*) and both the destination and source record the route associated with the source/ destination address in a table .**

Source Routing Bridge – characteristics

This advantage is :

- **Optimal !**
- **It also works on multiple connected topologies !**

Disadvantages:

- **not transparent, stations must know the topology exactly** (manage tables).
- **It is difficult to manage the many tables at the stations.**
- **Slow to adapt to topology changes** (discovery frames are needed again) .
- **Flooding the discovery frames can generate too much traffic.** (e.g. starting at the same time in the morning)**.**

Router (network layer)

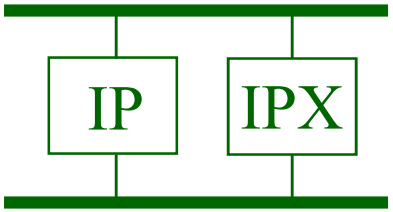
- **Router - traffic controller**

Function:

- **Connecting separated networks**

How it works

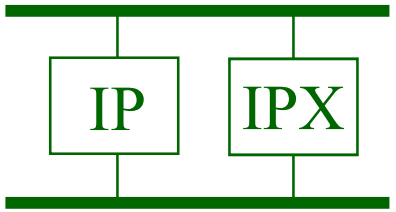
- **Traffic routing (packet routing) based on router tables.**
- **It works at the network layer, therefore: network-protocol dependent .**



Routers - Types

They can be:

- **Routers that handle a single protocol**
- **Multiprotocol router**
  - **Understands packet formats of multiple protocols**
  - **Connects in parallel according to different protocols**
- **Brouter (bridge router)**
  - **If the protocol is recognized ⇒ router**
  - **If the protocol is not recognized ⇒ acts as a bridge**



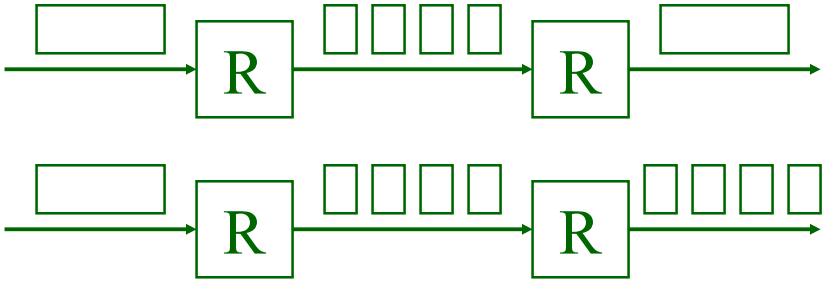
Router - other

Problem :

handling packets with a length that exceeds the maximum packet size of an intermediate subnet

- **Treatment**
  - **refusal to forward (and feedback), or**
  - **packet fragmentation**

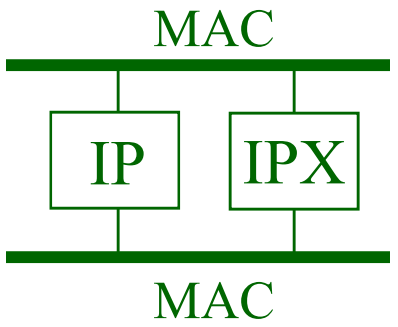
  The fragmentation can be
  - **Transparent:**
    routers also put it together.
  - **Not transparent:**
    only the destination station assembles it.
    (After the fragmentation, everyone forwards it in pieces.)



Routers - Their Features

Features:

- **You can connect different MAC networks.**
- **Protocol-dependent device.**
- **It is capable of complete traffic separation.**
  - **(Only the protocol required for traffic management represents additional traffic.)**
- **Suitable for covering long distances.**
- **There is no theoretical limit to the size or number of stations of the network.**



Additional features may include:

- **Data/network protection – packet filtering firewall.**
- **User management – allow/disable access (e.g. ISP dial-up point).**
- **Connection and route management – connection enable/disable, route selection, redundant (backup) path management (increasing reliability).**

Routers - Layer 3 Switch

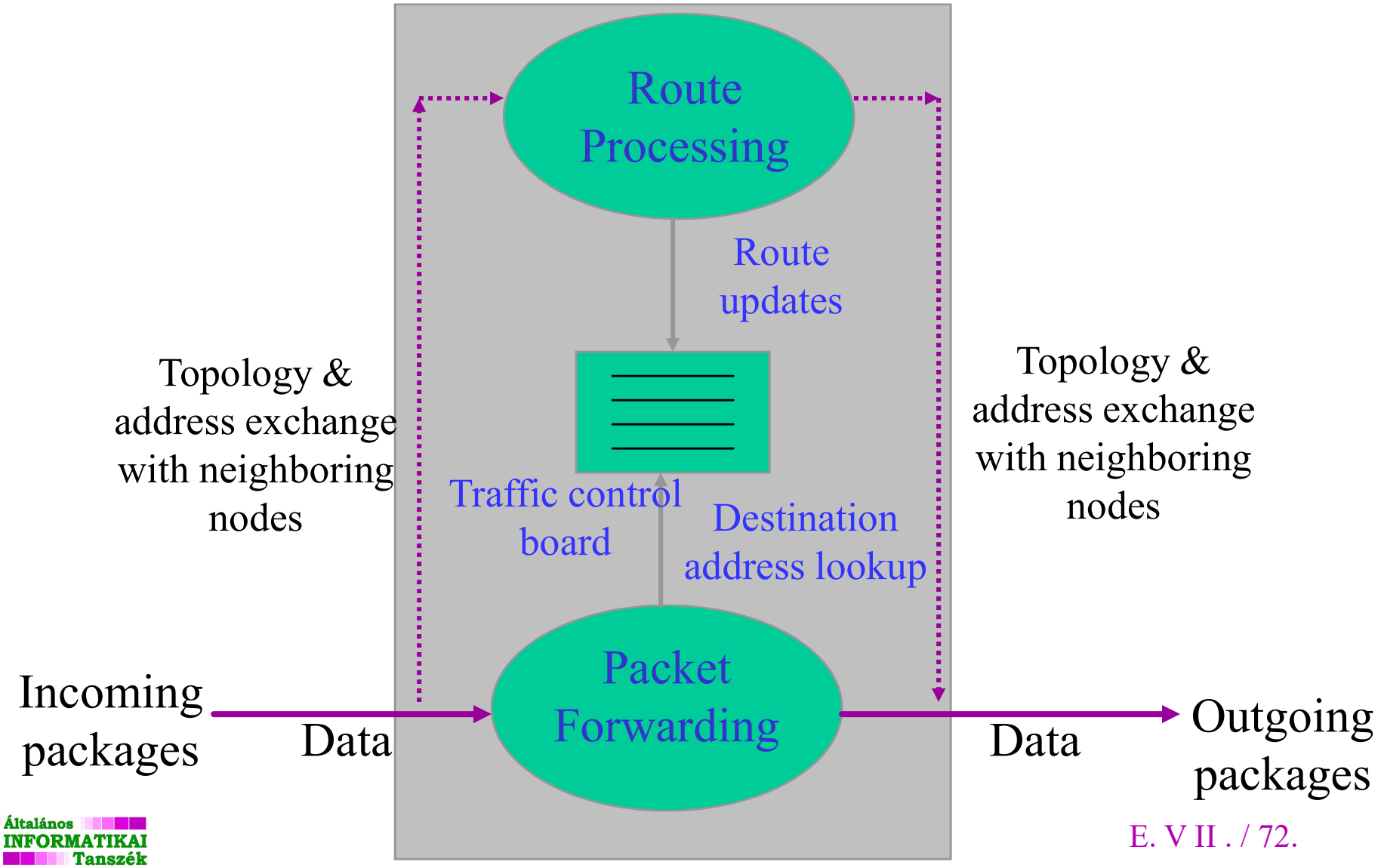
- **High-speed router**
- **Usually special purpose architecture**

Basic idea:

- **A source-destination connection usually consists of multiple packets and they should be treated similarly.**
- **After performing traffic routing on the first packet, the received direction is loaded into a table along with the characteristics of the expected packet sequence (layer 3 content, functions).**
- **Further packets of this packet sequence are routed (much faster) and changed (if necessary, e.g. TTL, CRC recalculation) based on the characteristics** ("fingerprint") stored in the table. (If there is no match to the table ⇒ complete processing again)

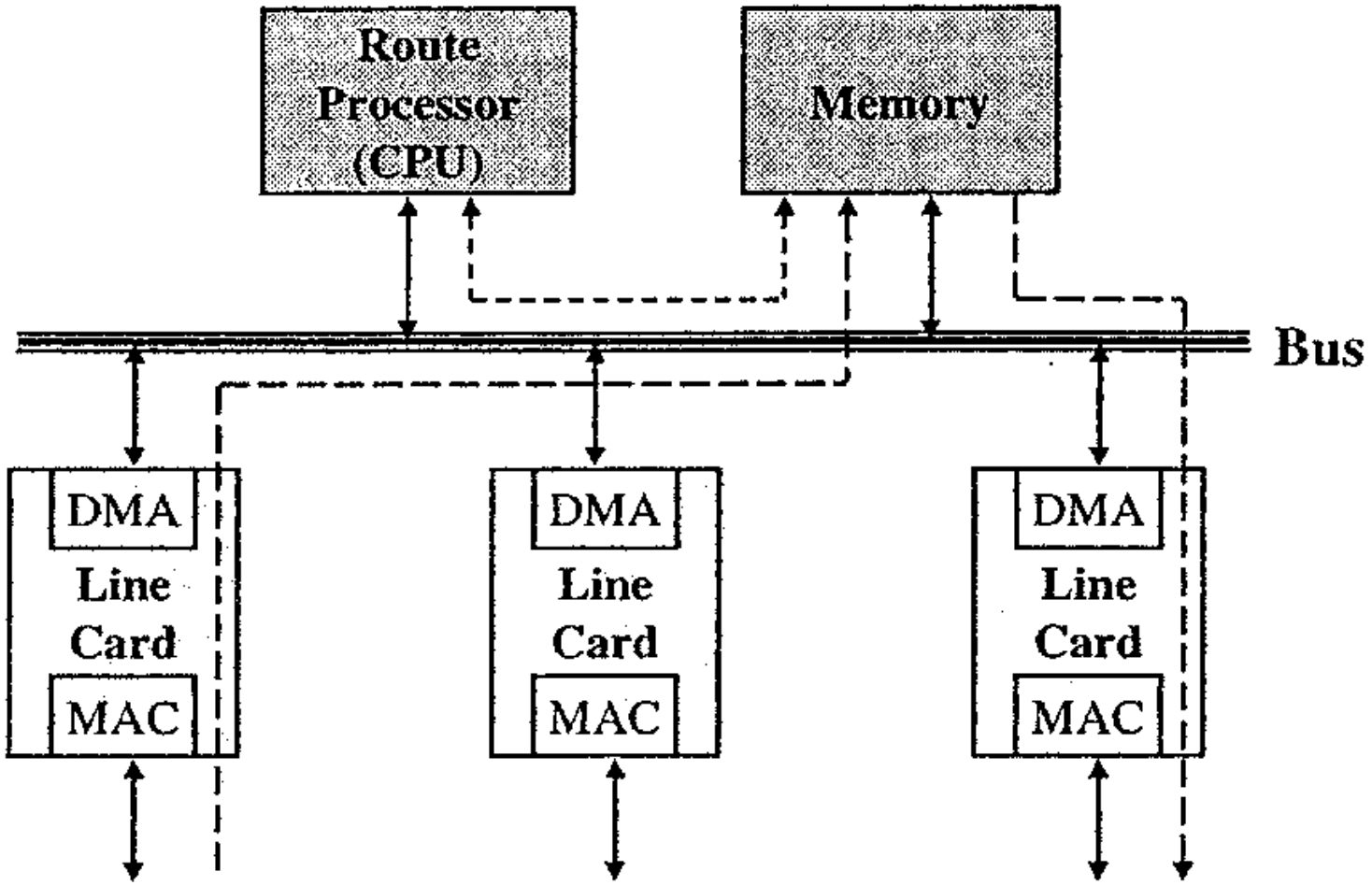
Routers - Layer 3 Switch

- **Traffic controller, main components**



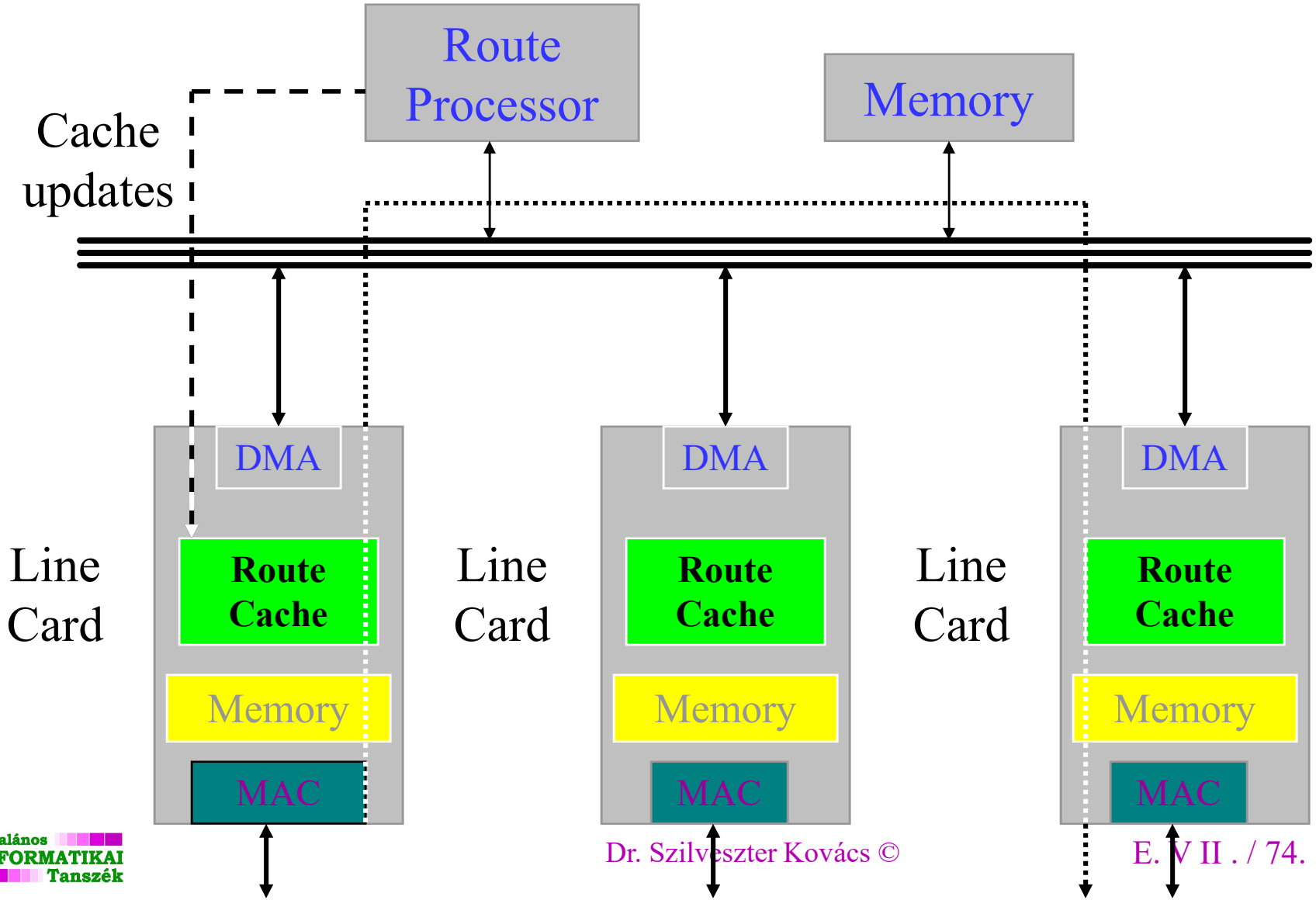
Routers - Layer 3 Switch

- **Classical architecture**



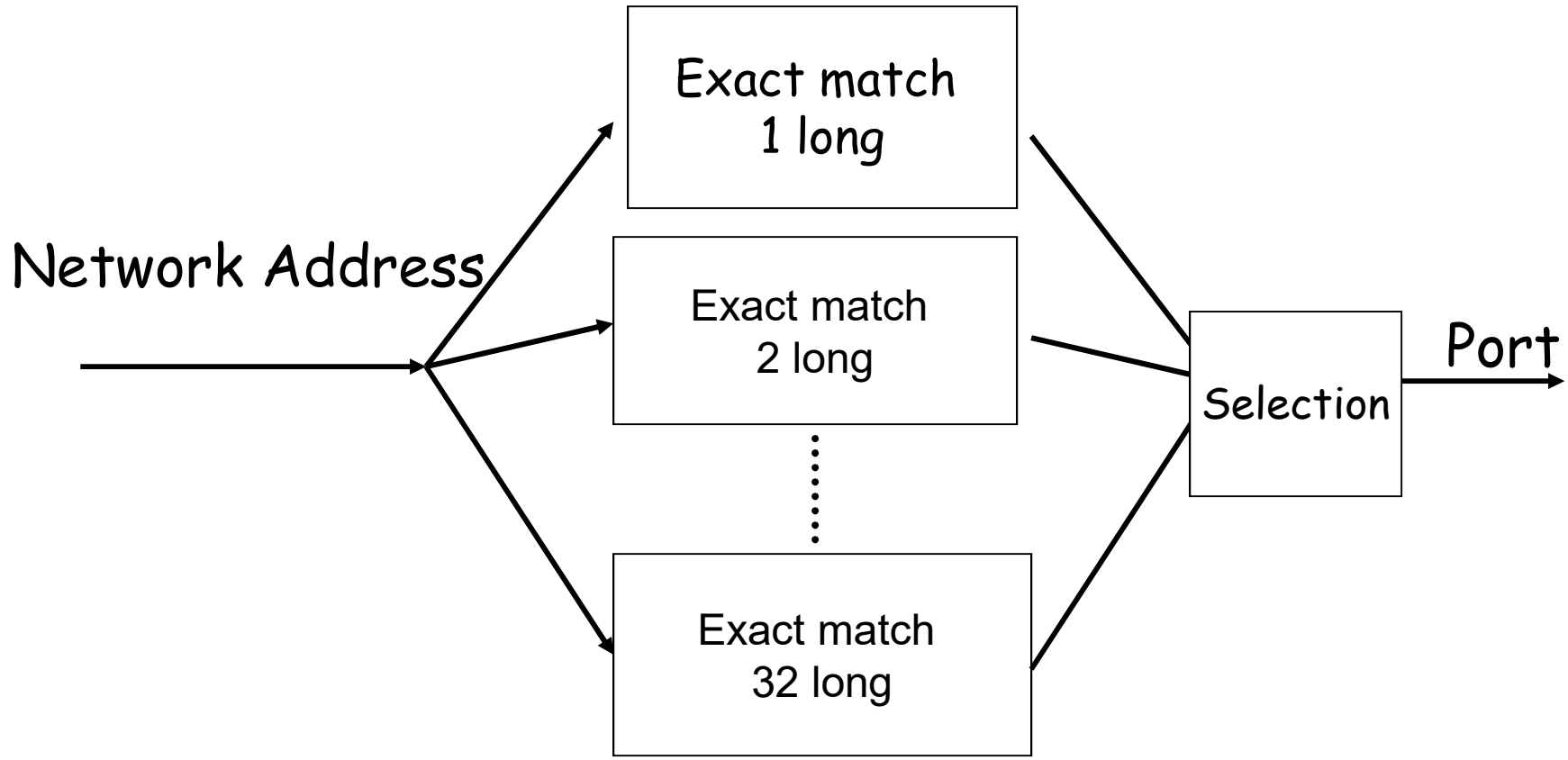
Routers - Layer 3 Switch

- **Route cache**



Routers - Layer 3 Switch

- **Parallel processing of the “longest match” (VLSM)**



Non-routable protocols

- **Protocols that do not have a "real" network layer – and hierarchical addressing (e.g., they use the global addresses of the MAC layer as "network" layer addresses), they cannot be routed !**
  - **For example, DEC Lat, NetBIOS, NetBEUI, etc.**
- **Non-routable protocols can only be forwarded using a bridge or repeater.**

Bridge -Router

- **In some applications, both device may be equally suitable for connecting networks.**
  (For example, for some other reason, a router is not necessarily required.)
- Virtually all routers available today **are Brouters (bridge routers).**
- **However, Bridges (Layer 2 Switches) with the same performance are significantly cheaper.**
- **When is it enough to install Bridge?**

Bridge – Advantages

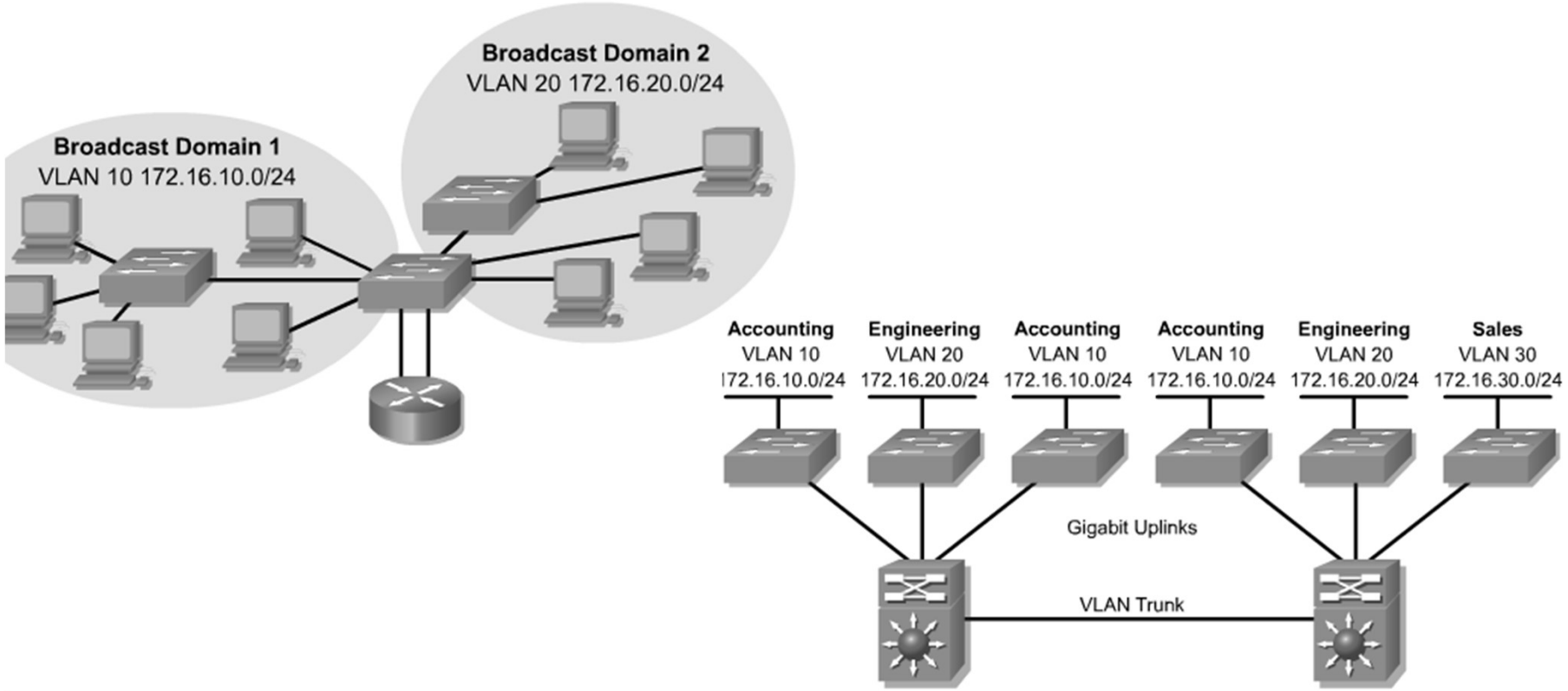
- **Easy to install** (plug and pay/pray/play) **.**
  (possibly port, VLAN, management configuration)
- **Transparent, can be placed anywhere where there was nothing before, or where there was a repeater.**
- **Network protocol independent :**
  - **If the network protocol is not routable, networks can only be connected using bridges/repeaters.**
  - **A new, previously unknown protocol can be introduced without changing the communication infrastructure.**
- **A network connected solely by bridges/repeaters appears to be one logical network (a "broadcast domain") :**
  - **A station can be moved without changing its network address.**
- **Good price/performance ratio .**

Bridge – Disadvantages

- **Not capable of load sharing between redundant paths – spanning tree** (bridges are in some cases capable of load sharing between parallel connections between bridge pairs (**Ethernet channel**)).
- **In certain situations, it can cause a lot of traffic congestion: unknown MAC address: broadcast ⇒ in case of large network: broadcast storm** $p_{empty}=(1-p_{broadcast})^n$, for n stations $n \to \infty$ $p_{empty} \to 0$ In a large network built from bridges, "fool-proof" applications (e.g. non-routable NetBEUI) can flood the system ("broadcast storm control" - throws it above one level) .
- **The traffic of each network part is partially mixed, making it difficult to control the traffic and find errors (attacks).**
- **Any part of the network traffic can be intercepted** (flooding tables (many fake source addresses) → broadcast, although this is partly an implementation error and can be remedied (limiting the memory area that can be reserved from a port)).

Bridge - Router

- **A Router (or Gateway) can provide transit between VLANs.**
- **Broadcast Domain Segmentation**



Router – Advantages

- **Complete traffic separation.**
  ⇒ A truly long-distance network can only be built using routers.
- **Load sharing between alternative routes.**
- **Flexible configuration options, traffic management rules**
  ⇒ "packet filter" firewall.

Router – Disadvantages

- **It needs to be configured.**
- **Protocol dependent.**
- **Slightly slower than Bridge**
  (even Layer 3 Switch than Layer 2 Switch).
- **In the case of non-routable protocols, it also only acts as a bridge.**

Bridge -Router

When is it enough to install Bridge ?

- **Almost always in “small” networks.**
- **If a router is needed for the WAN connection, a “smaller” router that can handle the load on the WAN link is usually sufficient.**
  (It can also be a firewall or NAT.)

When do you need a Router ?

- **If you have to.**
- **Where the protocol used is routable and it makes sense to talk about a “backbone network”.**
- **In WAN networks.**

References

- **STP:**
  http://www.cisco.com/warp/public/473/5.html
- **RSTP:**
  http://www.cisco.com/warp/public/473/146.html
- **MSTP:**
  http://www.cisco.com/warp/public/473/147.html
- **STP Timers:**
  http://www.cisco.com/warp/public/473/122.html