



Bridge & Switch

Pietro Nicoletti
Piero[at]studioreti.it



Copyright note

- These slides are protected by copyright and international treaties. The title and the copyrights concerning the slides (inclusive, but not only, every image, photograph, animation, video, audio, music and text) are the author's (see Page 1) property.
- The slides can be copied and used by research institutes, schools and universities affiliated to the Ministry of Public Instruction and the Ministry of University and Scientific Research and Technology, for institutional purpose, not for profit. In this case there is not requested any authorization.
- Any other complete or partial use or reproduction (inclusive, but not only, reproduction on discs, networks and printers) is forbidden without written authorization of the author in advance.
- The information contained in these slides are believed correct at the moment of publication. They are supplied only for didactic purpose and not to be used for installation-projects, products, networks etc. However, there might be changes without notice. The authors are not responsible for the content of the slides.
- In any case there can not be declared conformity with the information contained in these slides.
- In any case this note of copyright may never be removed and must be written also in case of partial use.

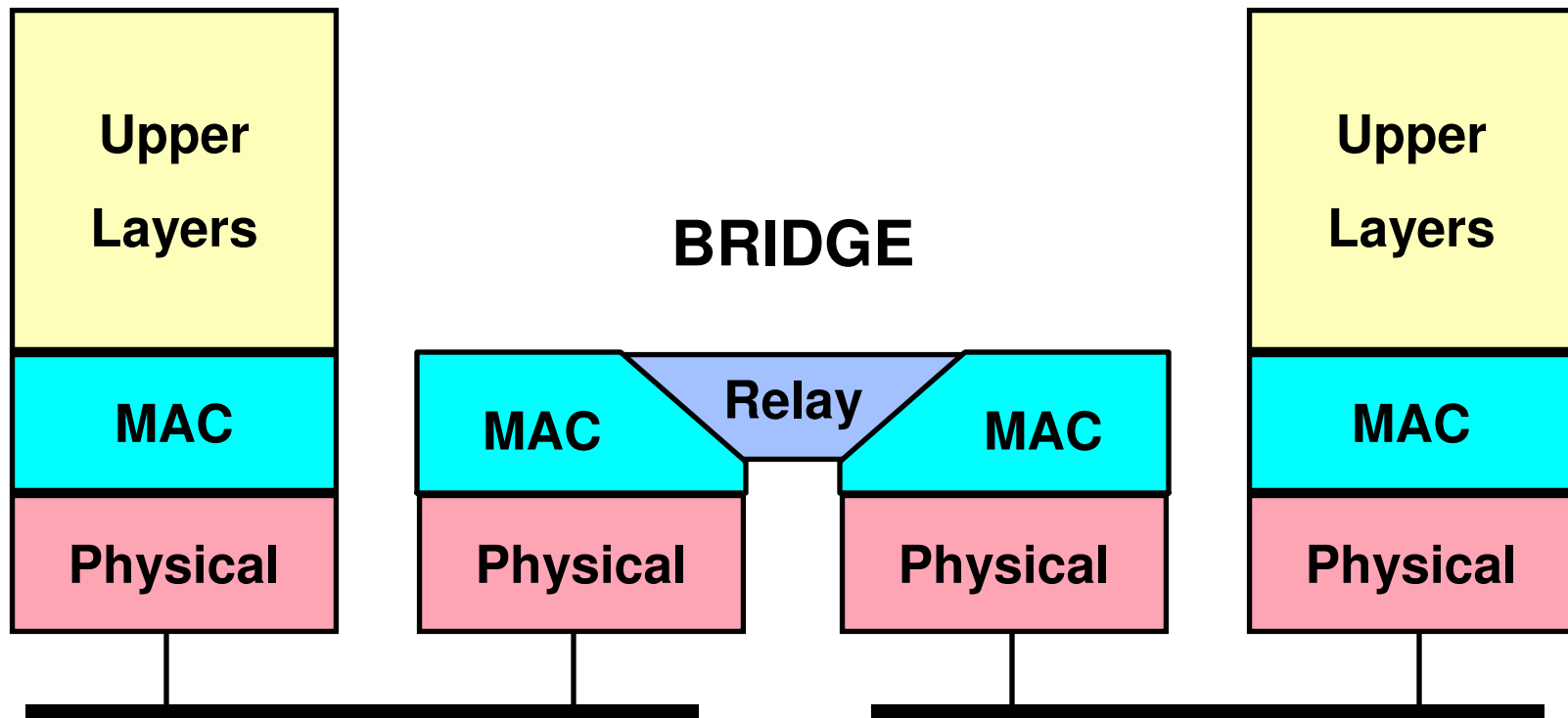


Bridge

- Called also MAC Bridge
 - operate at layer 2
 - use simple forwarding algorithm
 - forwarding technique based on MAC destination address
 - are normally used for local connections

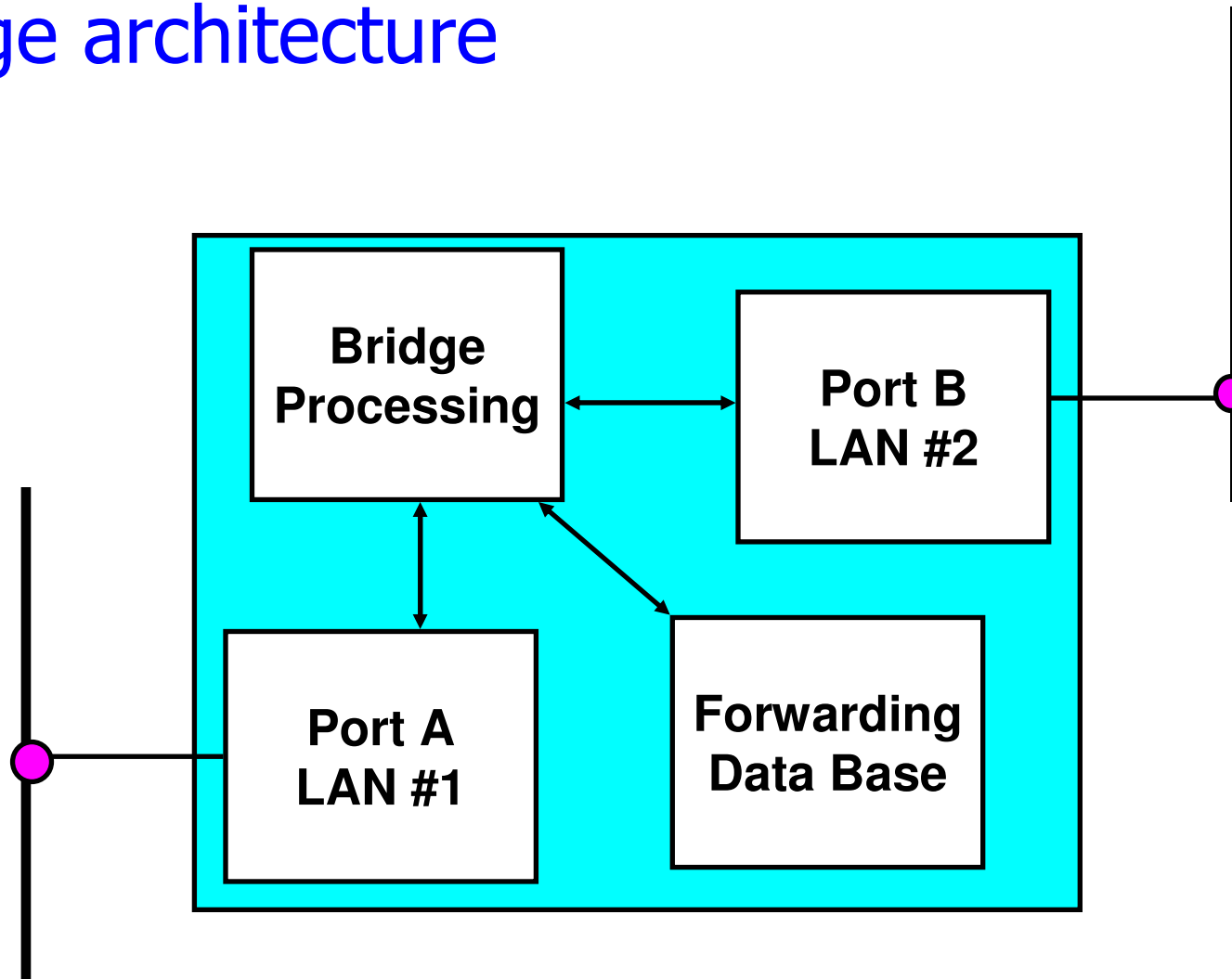


Bridge model





Bridge architecture





Forwarding

- Forwarding table is calculated by:
 - backward learning
 - looking the Source MAC address
- Backward learning
 - work only over LAN with tree topology
 - other topologies are transformed to tree topology by spanning tree algorithm
- Spanning tree algorithm operate periodically (any second)
 - decide which port set to forwarding state and which port set to blocking state

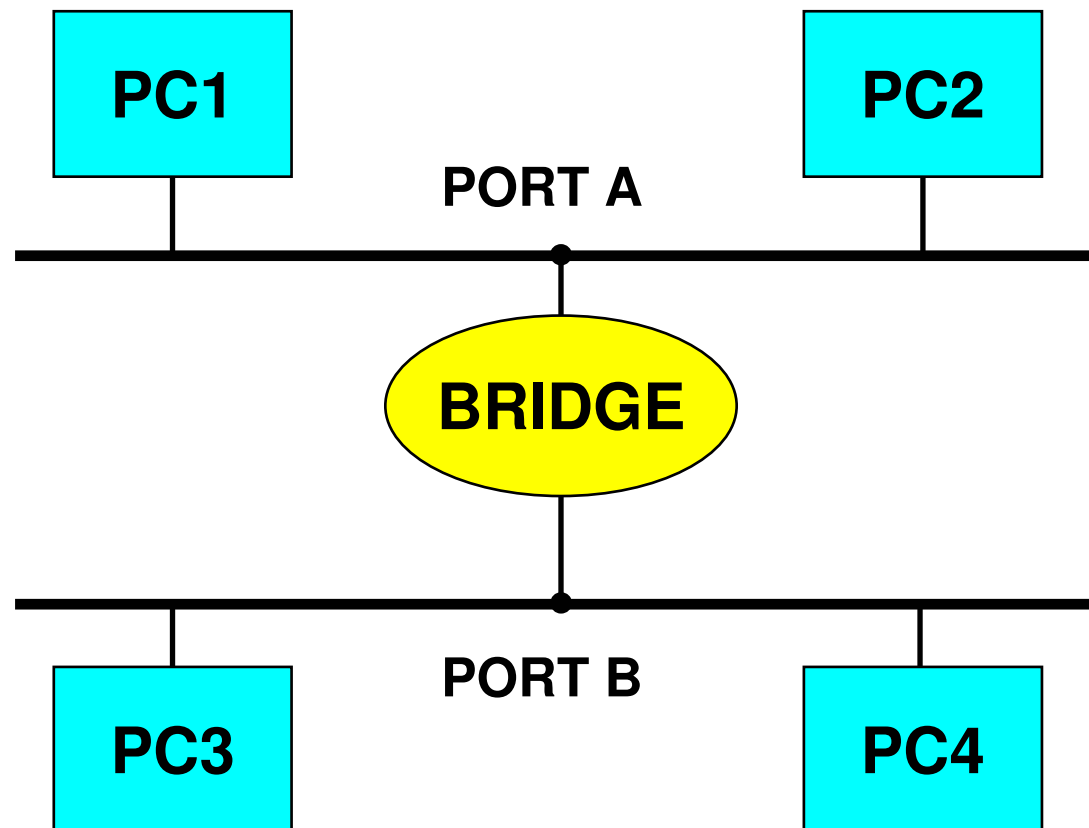


Transparent Bridge

- IEEE standardized the bridge functions by 802.1D standard
- The bridge 802.1D work in transparent mode, for this reason are also called transparent bridge:
 - are derived from Ethernet
 - have a local forwarding table
- The main functions of the bridge are:
 - frame forwarding
 - **Singlecast** frames forwarding based on *Destination MAC Address*
 - **Broadcast** frames are forwarded in every ports except the port where as been received
 - stations auto-learning
 - loop detection by spanning tree algorithm

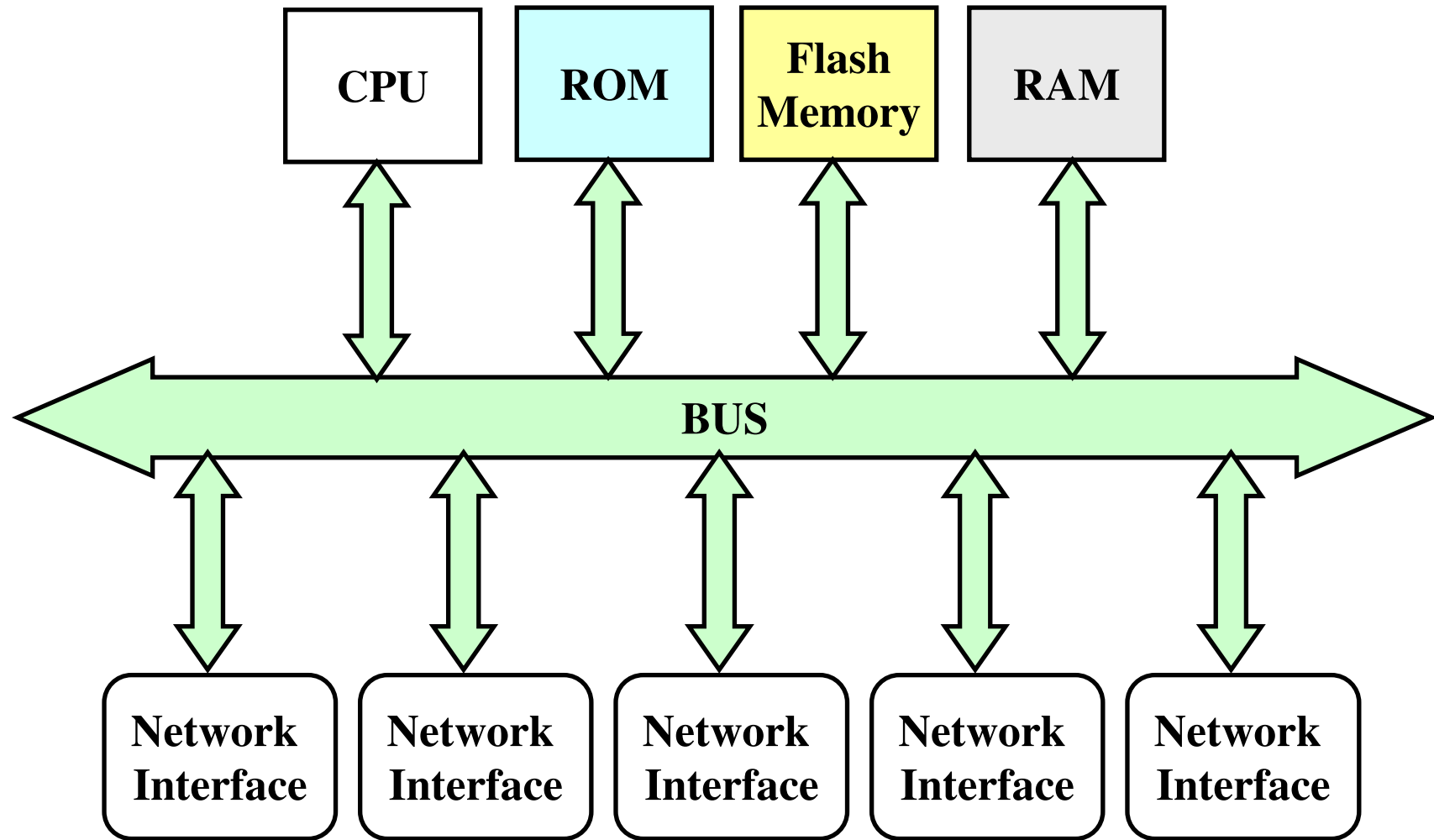


Transparent Bridge



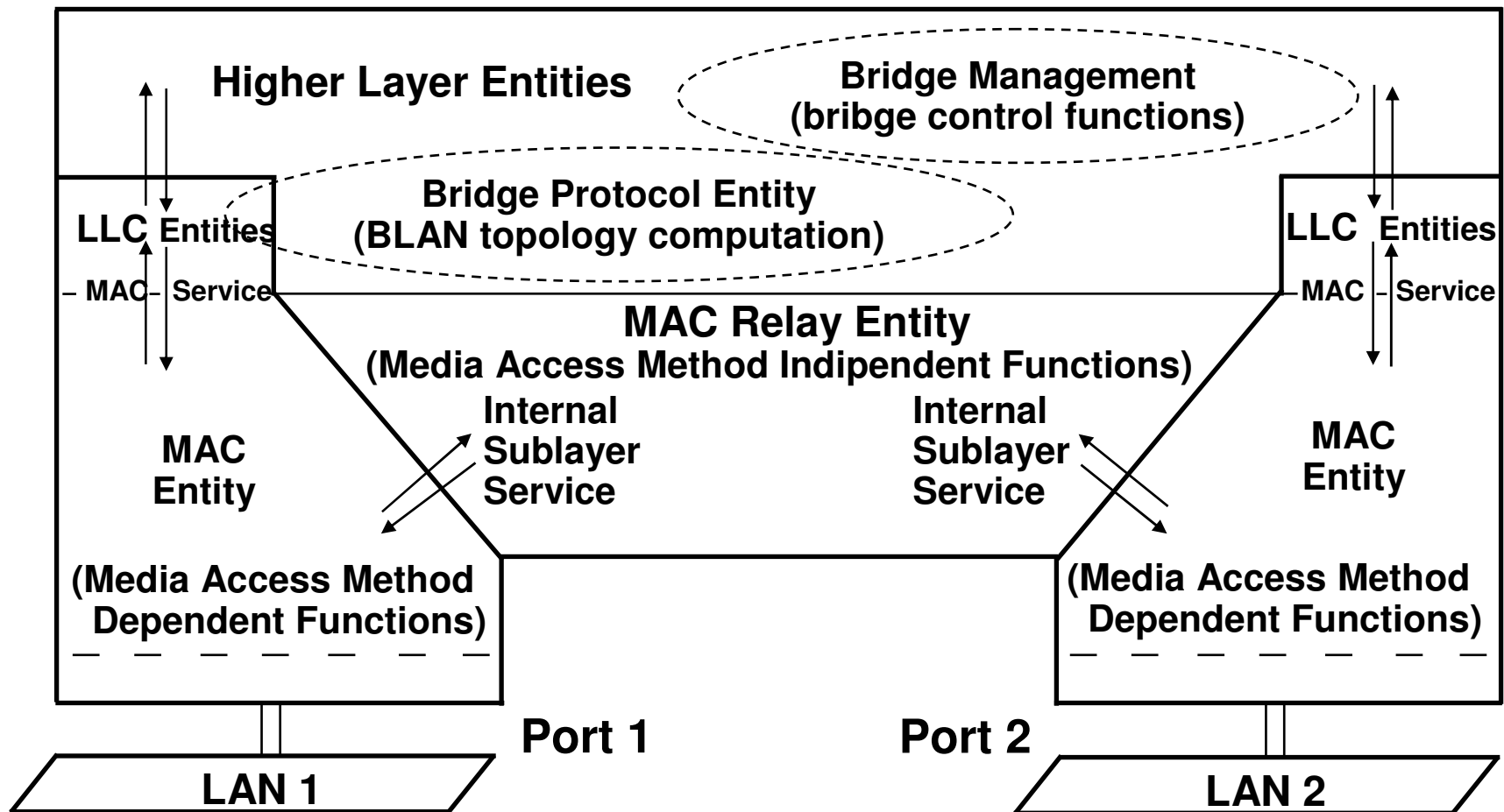


Bridge physical architecture



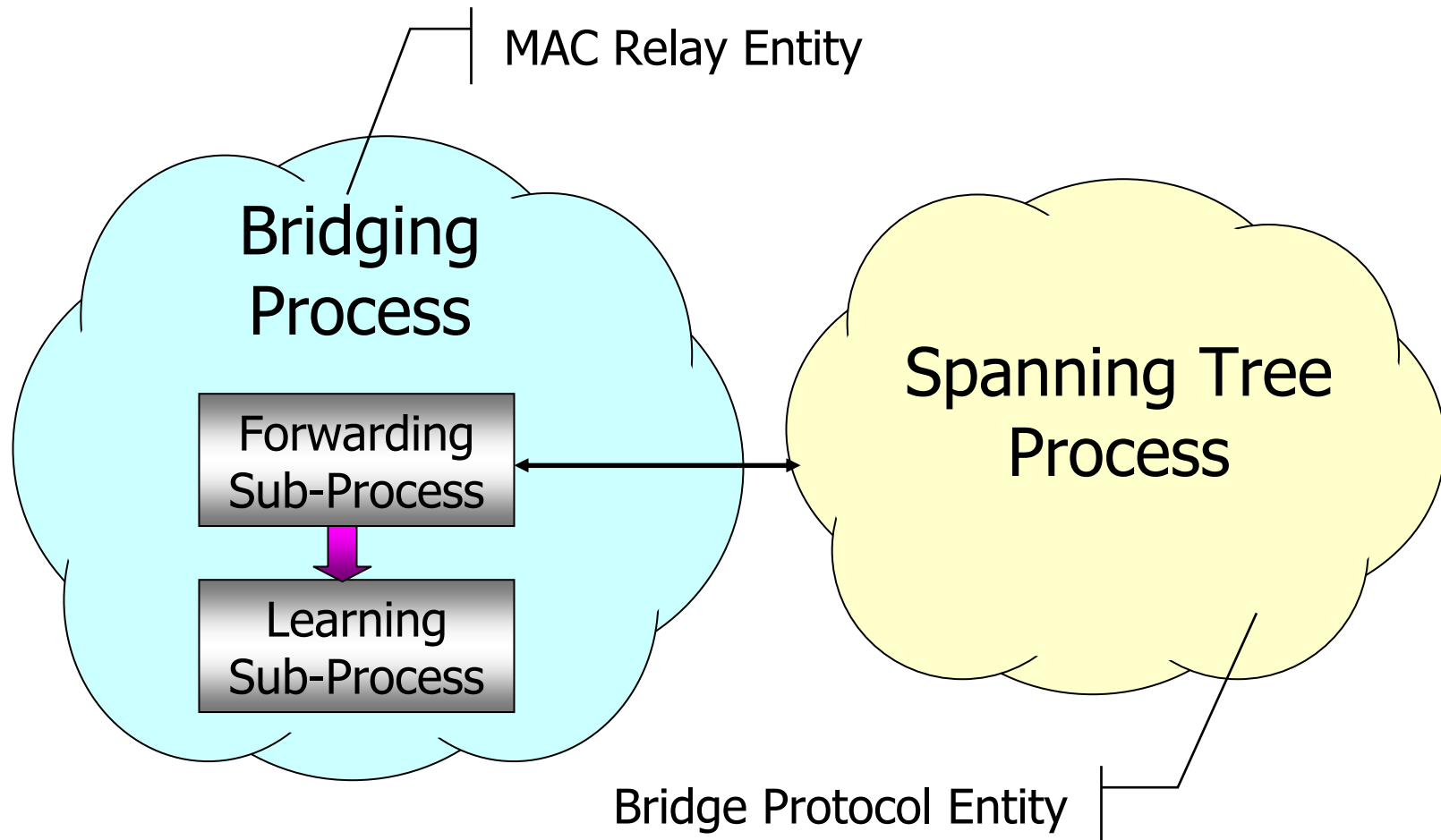


Bridge logical architecture





Bridge functions





Filtering data base

- ***Entry*** fields
 - MAC address
 - Destination port
 - Ageing time (default expire after 300 s)
 - Port status (depending by spanning tree protocol)
 - Entry type
 - dynamic
 - Updated by learning process
 - Max entries: 1024 ÷ 65 K
 - static
 - Not updated by learning process
 - 256 entries typically
- Lookup by CAM (Content Addressable Memory)



Filtering database example

filtering database show command

Switch-1> show cam dynamic

* = Static Entry. + = Permanent Entry.

= System Entry X = Port Security Entry

Dest MAC Address	Ports	Age
00-00-86-1a-a6-44	1/1	1
00-00-c9-10-b3-0f	1/1	0
00-00-f8-31-1c-3b	1/2	4
00-00-f8-31-f7-a0	1/1	2
00-01-e7-00-e3-80	2/2	0
00-02-a5-84-a7-a6	2/1	1
00-02-b3-1e-b4-aa	2/1	5
00-02-b3-1e-da-da	2/5	1
00-02-b3-1e-dc-fd	2/4	2

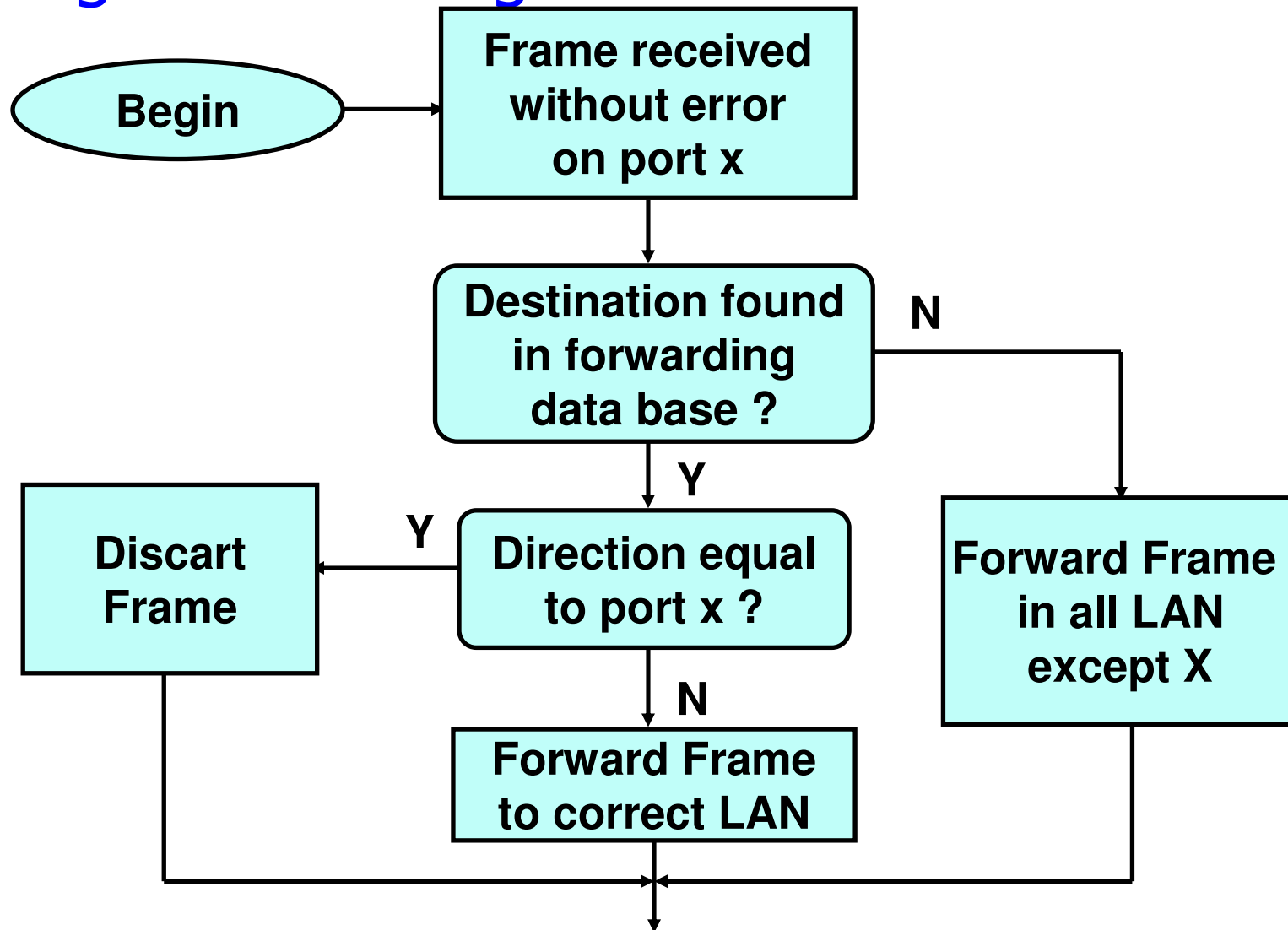
Modular switch: slot/port

Ageing time

Stations MAC address

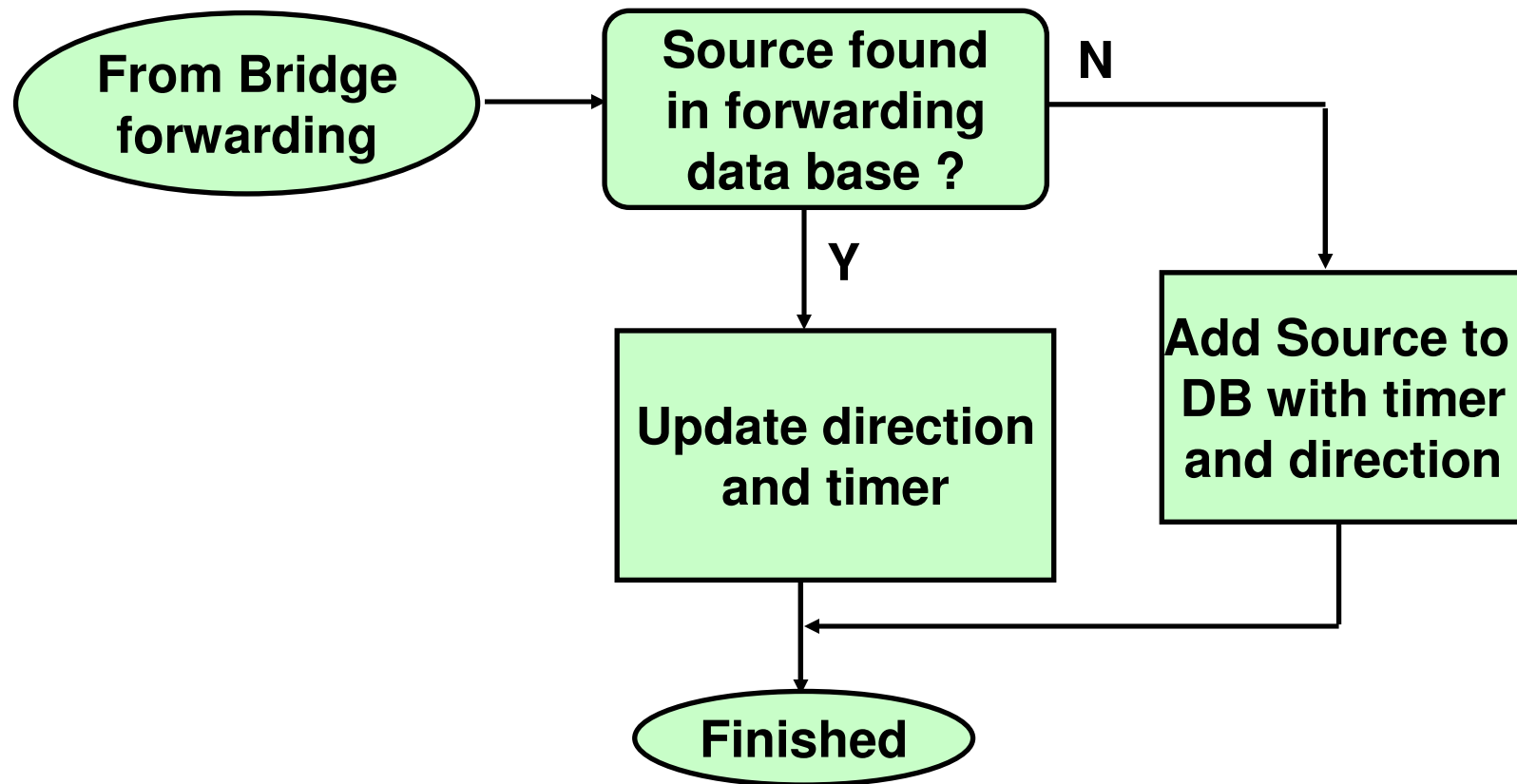


Bridge Forwarding



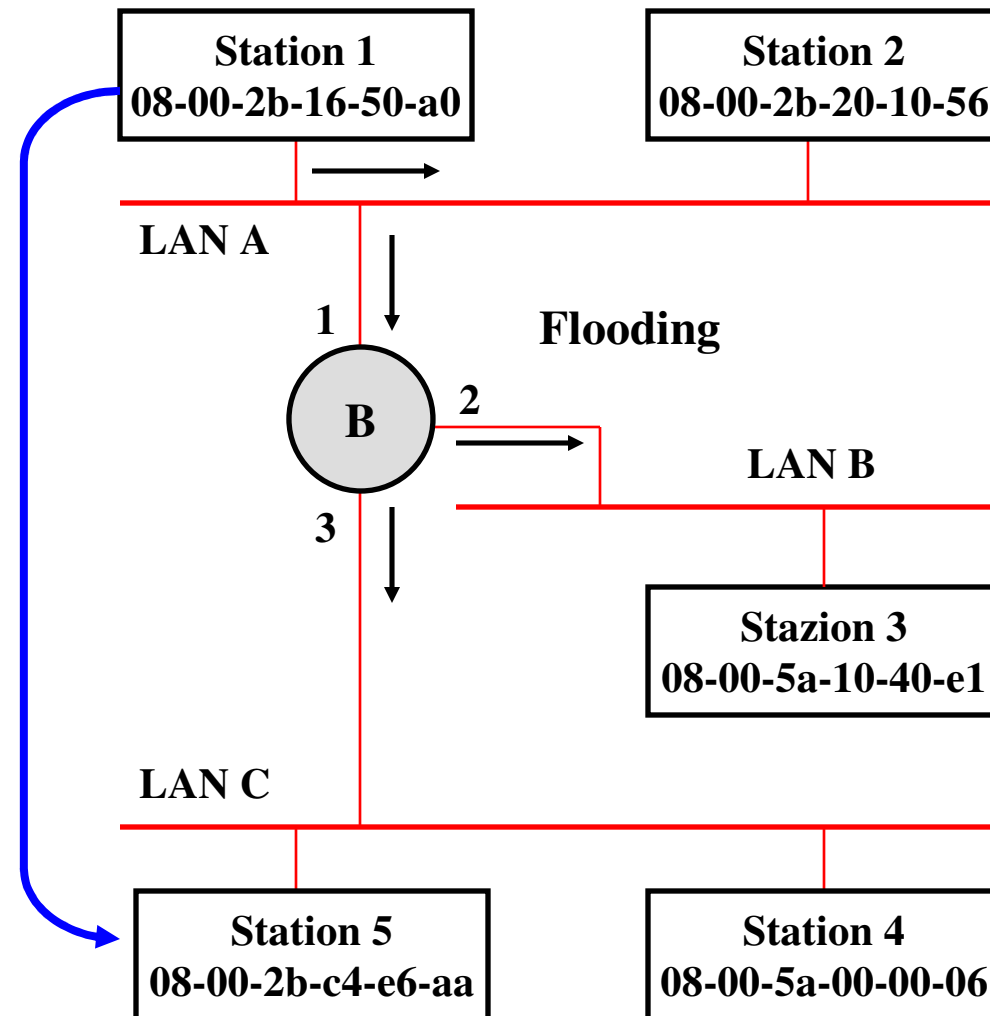


Bridge Learning





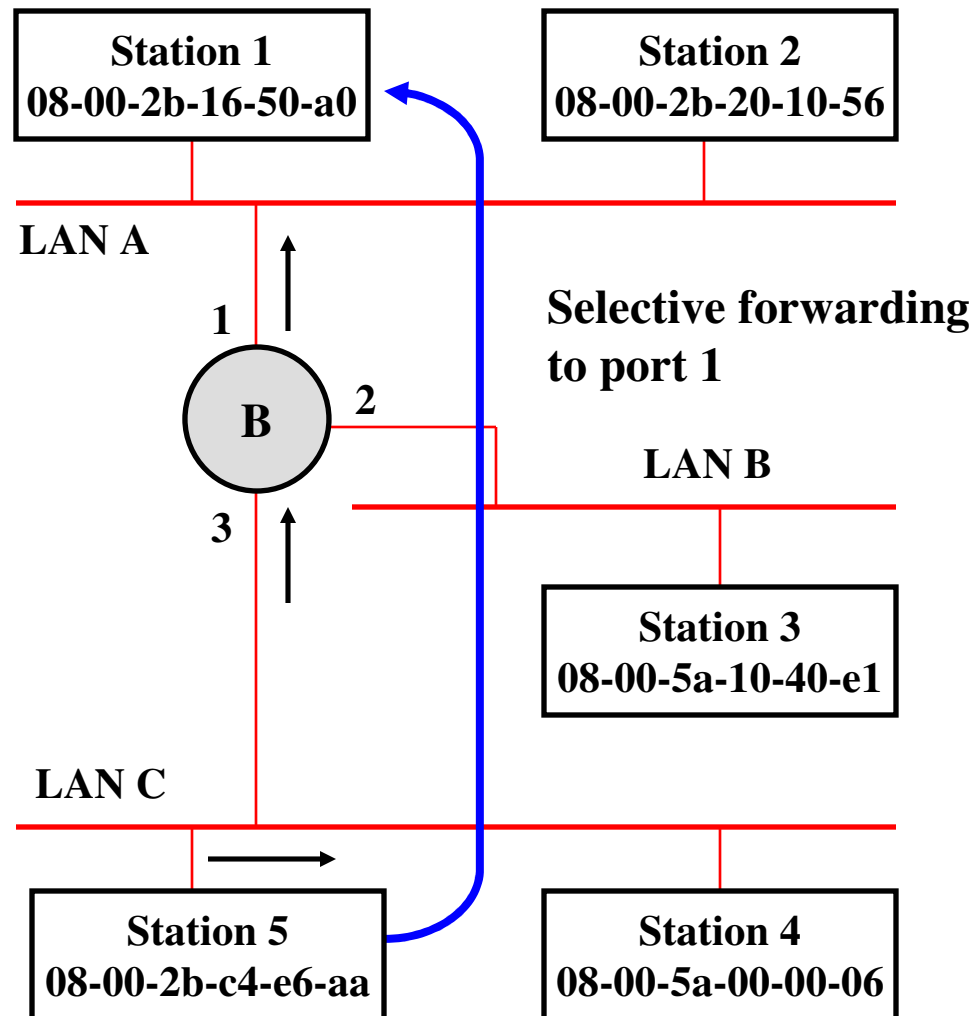
Forwarding frame to unknown destination



Port	MAC adress	Ageing time
1	08-00-2b-16-50-a0	0



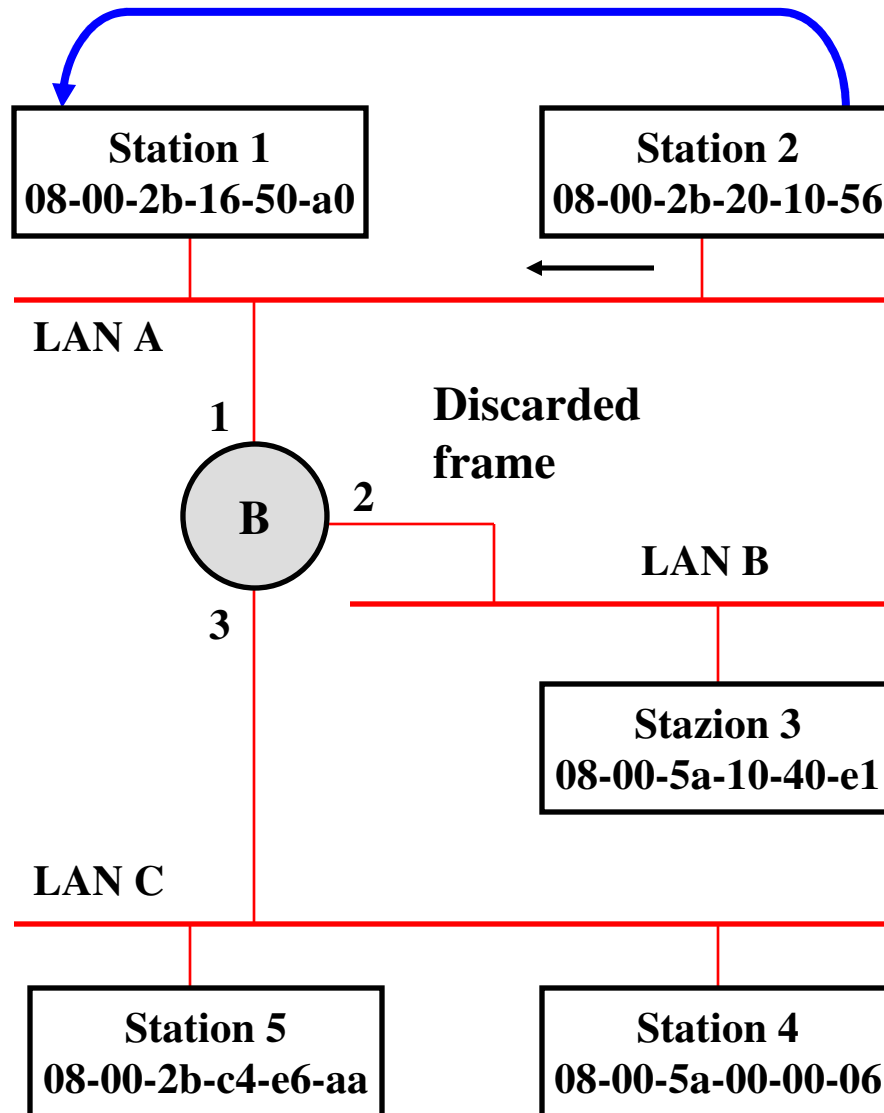
Forwarding frame to known address destination



Port	MAC adress	Ageing time
1	08-00-2b-16-50-a0	5
3	08-00-2b-c4-e6-aa	0



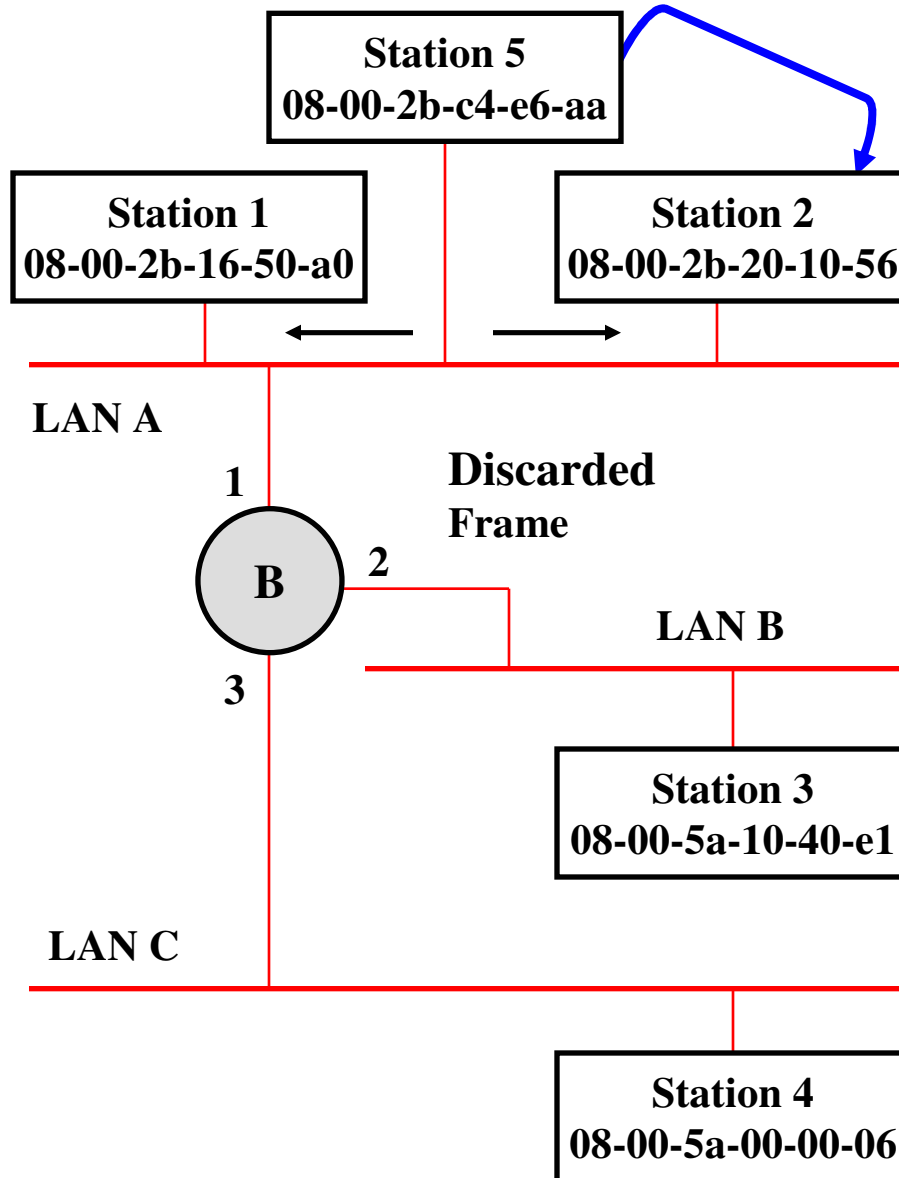
Traffic limitation



Port	MAC adress	Ageing time
1	08-00-2b-16-50-a0	12
3	08-00-2b-c4-e6-aa	13
1	08-00-2b-20-10-56	0



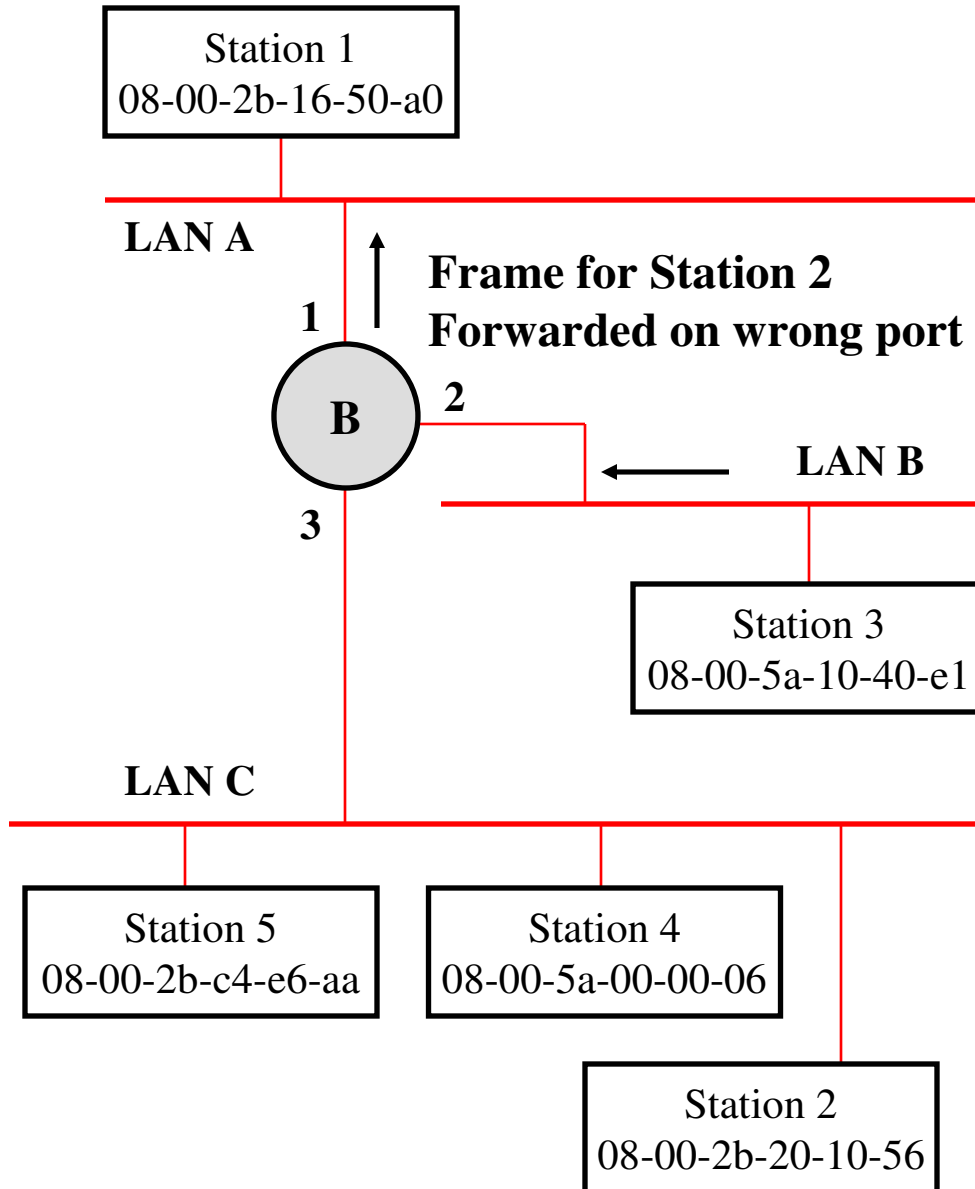
Station 5 move and Filtering Data Base update



Port	MAC adress	Ageing time
1	08-00-2b-16-50-a0	40
1	08-00-2b-c4-e6-aa	0
1	08-00-2b-20-10-56	20



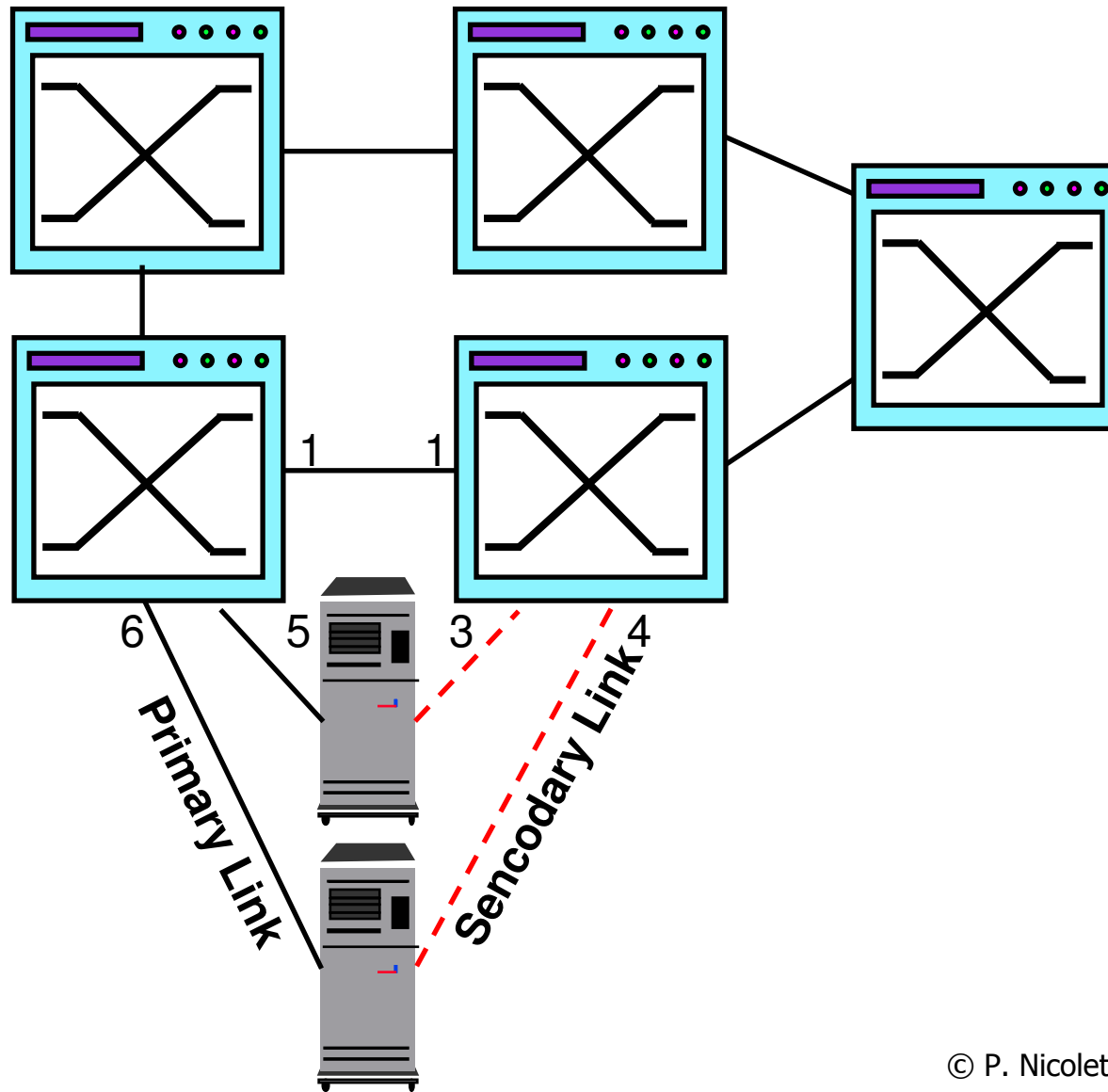
Station 2 move and forwarding error



Port	MAC address	Ageing time
1	08-00-2b-16-50-a0	50
3	08-00-2b-c4-e6-aa	51
1	08-00-2b-20-10-56	40
2	08-00-5a-10-40-e1	0



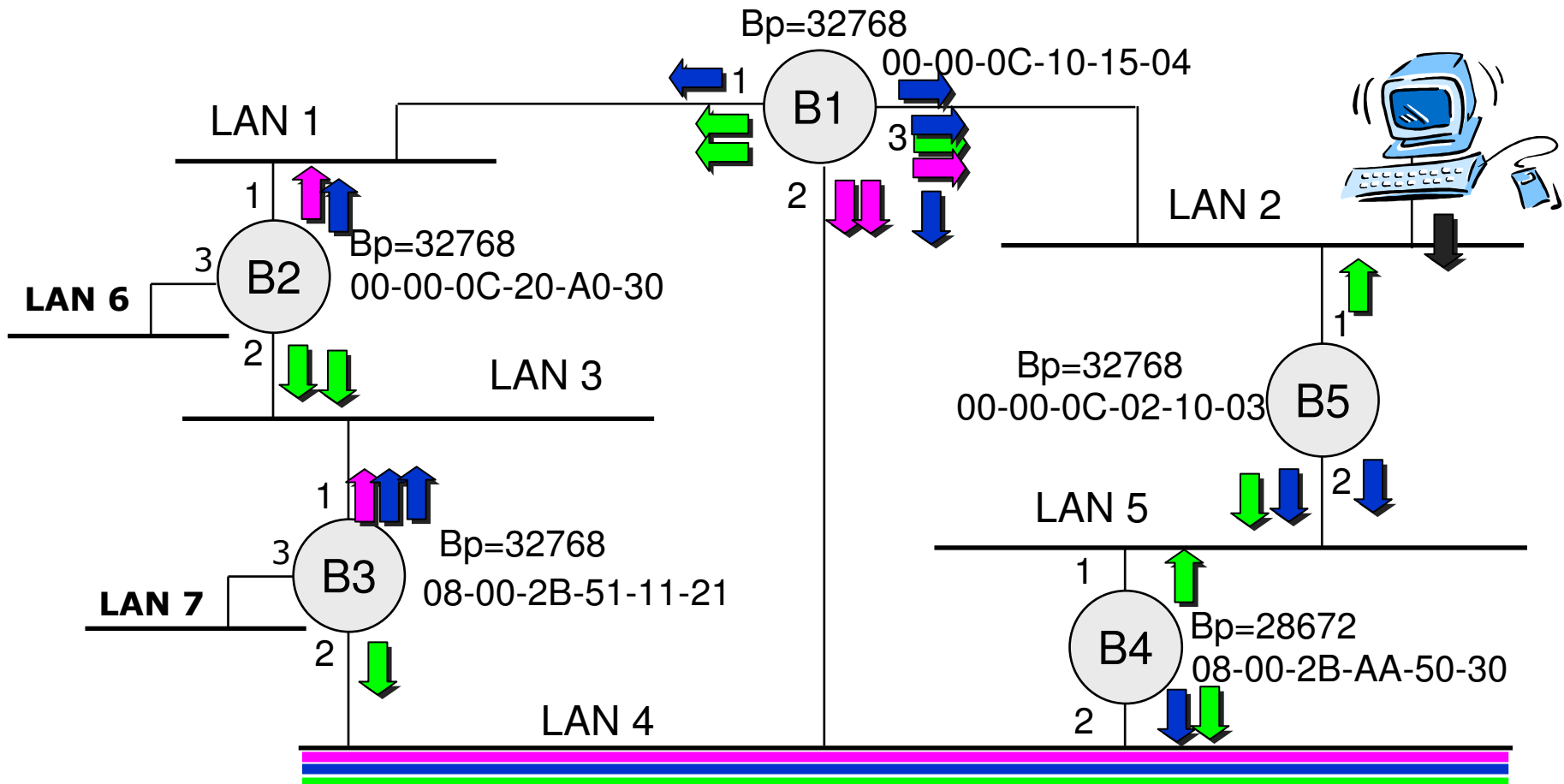
Fault Tolerant network and Filtering Data Base Update





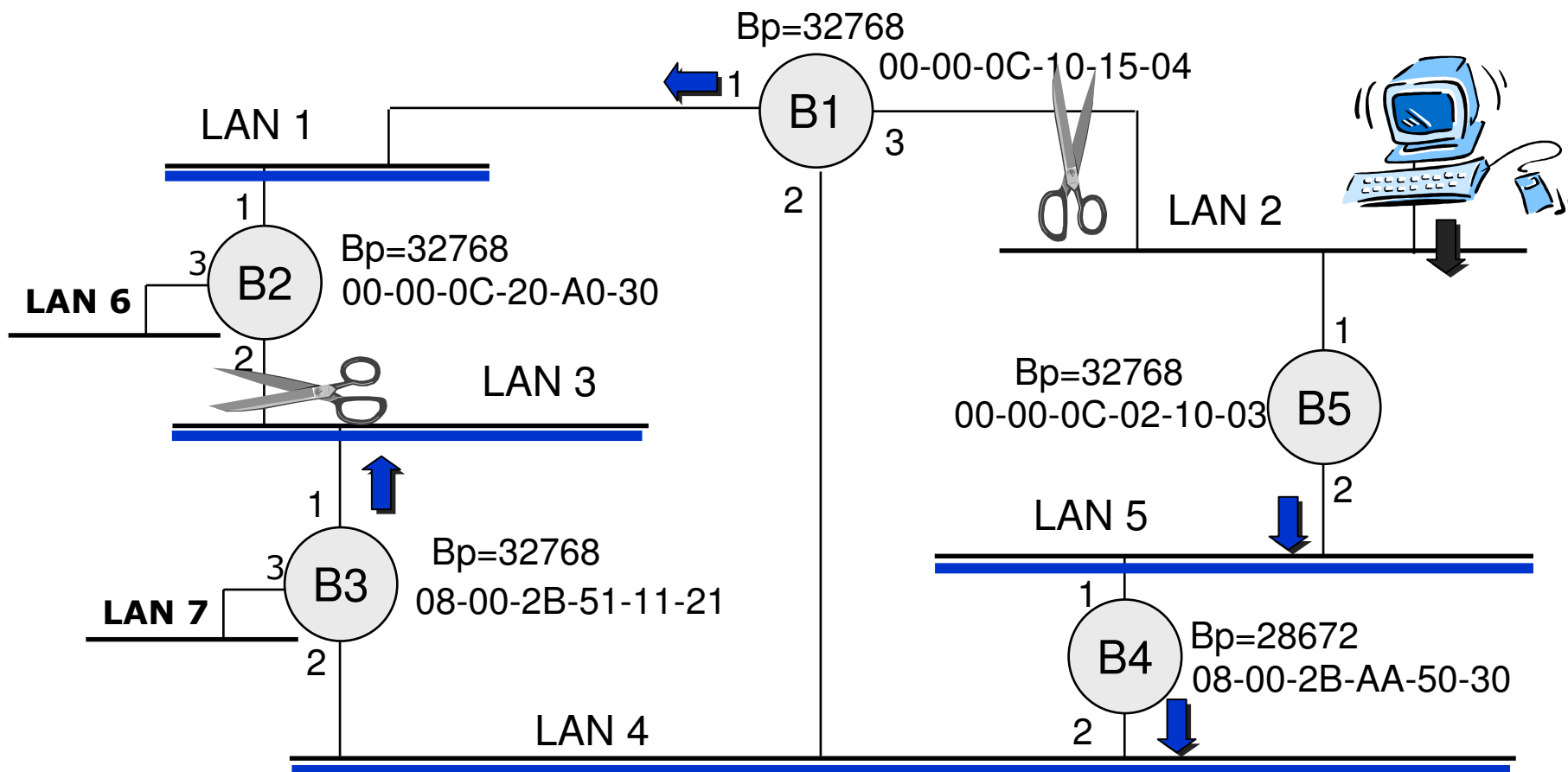
Loop on the network

- Broadcast are not filtered: **broadcast storm**





Solution: cut the loops



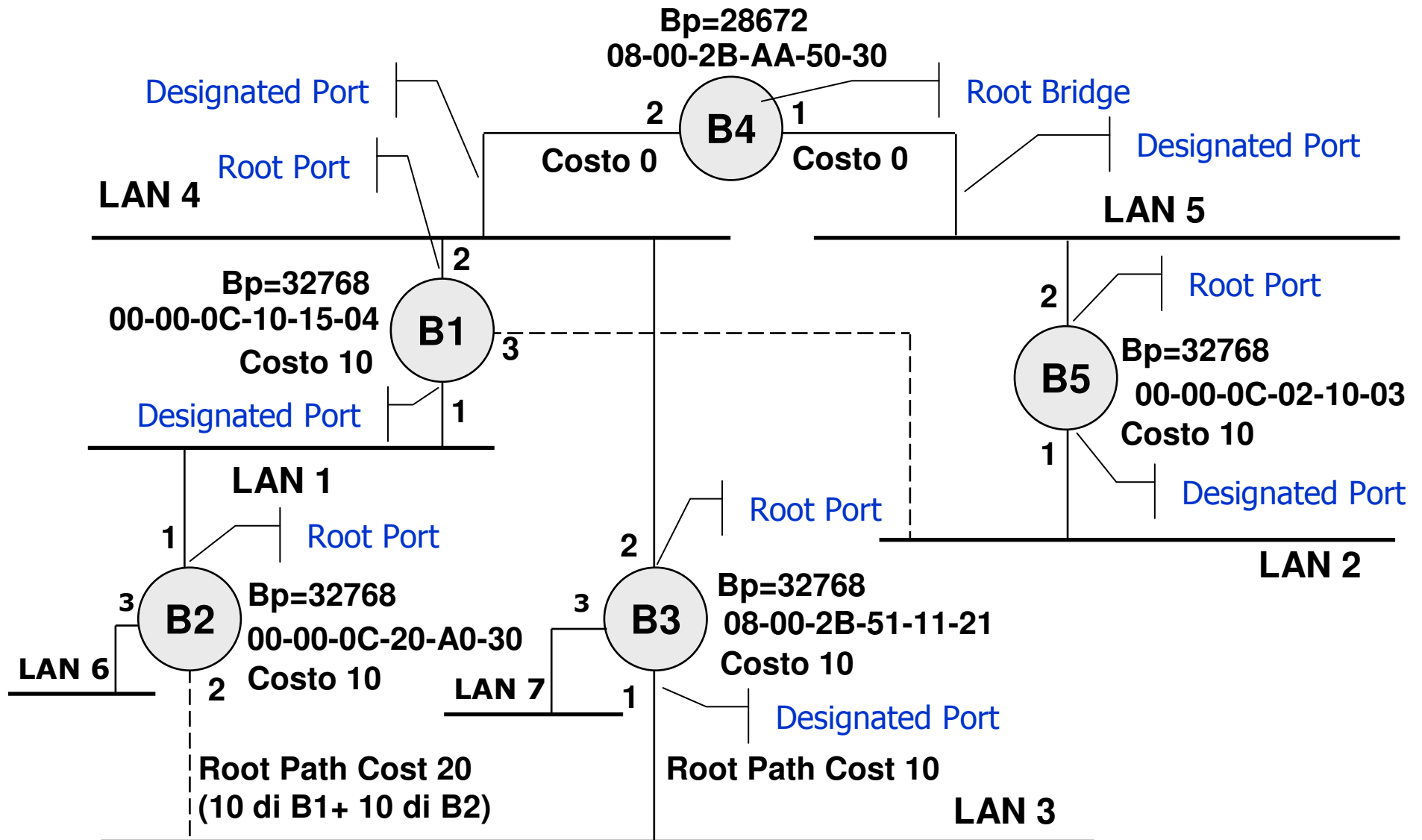


Spanning Tree

- All bridge in the network must be IEEE 802.1D spanning tree compliant
- The spanning tree algorithm operate with the following steps:
 - Root Bridge election
 - Root Port selection
 - Designated/Blocking Port selection
- The spanning tree protocol use the BPDU (Bridge Protocol Data Unit)



Spanning tree protocol result





Bridge Protocol Data Unit (BPDU)

- **Configuration** BPDU
- **Topology Change Notification** BPDU

Dest. Addr.	Source Addr.	Length	DSAP	SSAP	Control	BPDU	
Multicast 01-80-C2 00-00-00	Singlecast MAC Bridge Address	XY	042H	042H	XID	Configuration BPDU or Topology Change Notification BPDU	FCS

BPDU: Bridge Protocol Data Unit
DSAP: Destination Service Access Point
SSAP: Source Service Access Point



Configuration BPDU

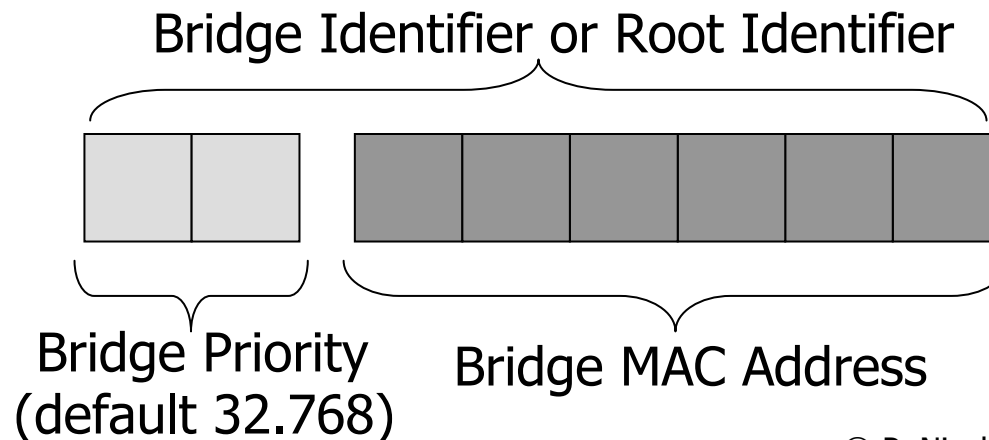
- **Root Identifier**
 - bridge identifier of root bridge
- **Root Path Cost**
 - Cost to reach the root bridge
 - Zero on BPDU transmitted by root bridge
- **Bridge Identifier**
- **Port Identifier**

Byte 1÷2	Protocol Identifier:	00-00
3	Protocol Version Identifier:	00
4	BPDU Type:	00
5	TC	TCA
Flags		
6÷13	Root Identifier	
	First 2 byte = Bridge Priority	
	Last 6 byte = Root Bridge MAC Address	
14÷17	Root Path Cost	
18÷25	Bridge Identifier	
	First 2 byte = Bridge Priority	
	Last 6 byte = Bridge MAC Address with transmit BPDU	
26÷27	Port Identifier	
	First byte = Port Priority	
	second byte = port number	
28÷29	Message Age	
30÷31	Max Age	
32÷33	Hello Time	
34÷35	Forward Delay	



Bridge Identifier

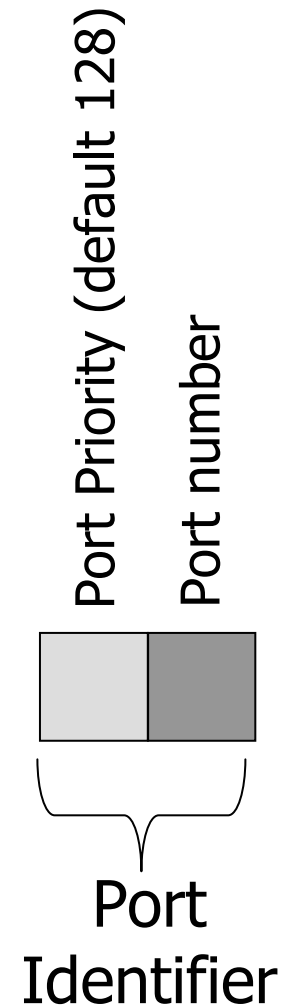
- MAC Address assigned to the Bridge
- Priority assigned to the Bridge: ***bridge priority***
 - Configuration
 - Default value to guarantee “plug&play” operation
- Became root bridge the bridge with less bridge identifier
 - Less bridge priority determine the root bridge election
 - If Bridge priority is the same in every bridge the lowest MAC address determine the root bridge election





Port Identifier

- Identify the port witch transmit the BPDU on connected LAN





The algorithm principle

- Root Bridge election
 - Is selected as root bridge the bridge with lowest Bridge Identifier
- Root port
 - The port (only one) connection to the root bridge
- Designated Port
 - The port that forwards frames from that LAN towards the Root
 - A bridge may one or more designated ports



The algorithm principle: Root & Designated Port selection

- Rule for Root and Designated Port selection
 - **Lowest Root Path Cost**
 - In case of equal Root Path Cost the algorithm look for **Lowest Bridge Identifier**
 - In case of equal Root Path Cost and equal Bridge Identifier the algorithm look for **Lowest Port Identifier**
 - If there is only one port to connect a LAN that port is immediately elected as Designated Port
 - All the active port of the root Bridge are selected as Designated Port
- The ports that are not been selected as root or designated go to *Blocking State*



Port State

■ Listening

- Receive frames
- Do not forward frames
- Do not update the forwarding data base
- Compute received BPDU
- Transmit BPDU

■ Forwarding

- Receive frames
- Forward frames
- Update the forwarding data base
- Compute received BPDU
- Transmit BPDU

■ Learning

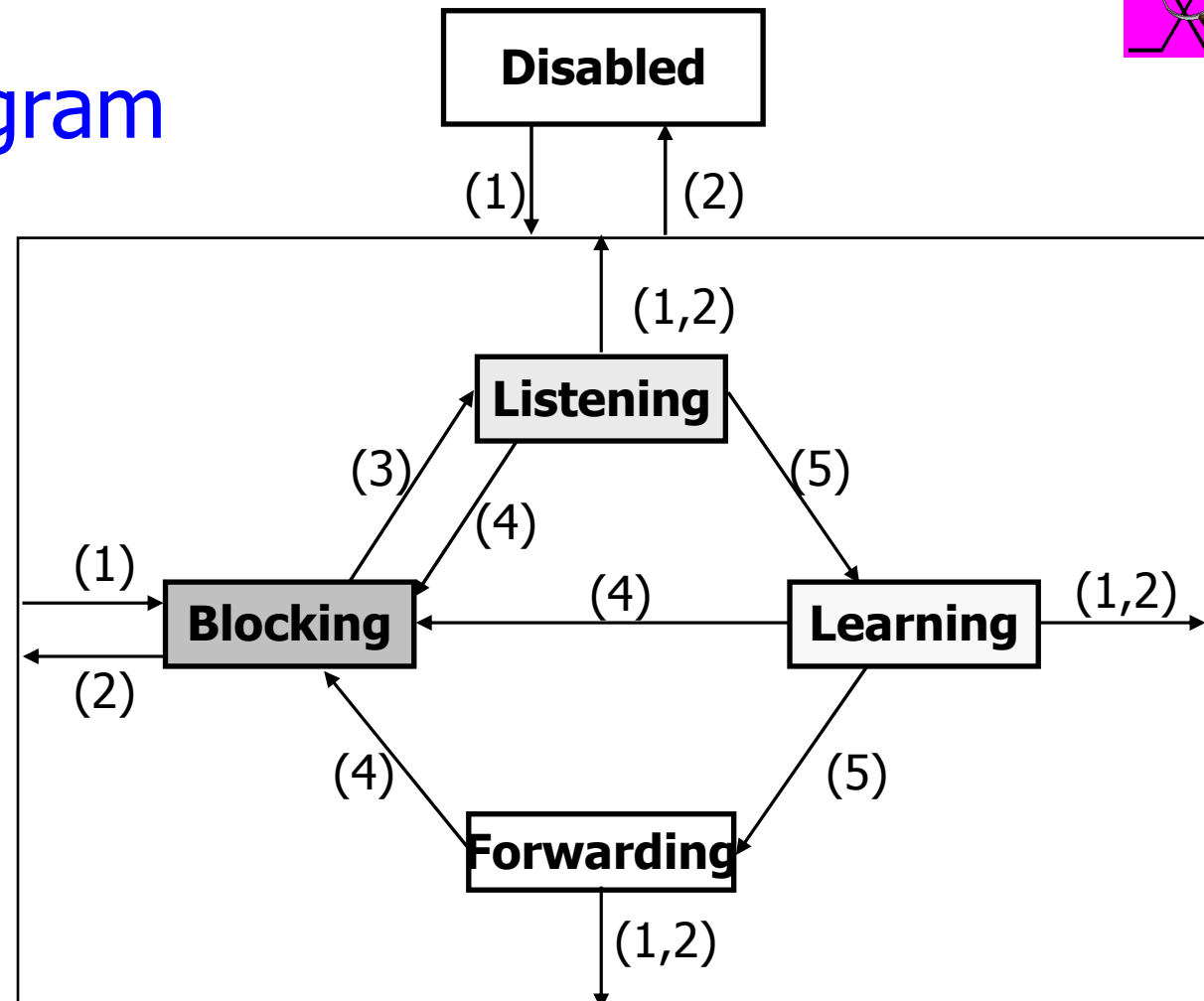
- Receive frames
- Do not forward frames
- Update the forwarding data base
- Compute received BPDU
- Transmit BPDU

■ Blocking

- Receive frames
- Do not forward frames
- Do not update the forwarding data base
- Compute received BPDU
- DO not transmit BPDU



Port state diagram



- 1) Port enabled, by management or initialization
- 2) Port disabled, by management or failure
- 3) Algorithm selects as Designated or Root Port
- 4) Algorithm selects as Alternate Port
- 5) Protocol timer expiry (Forwarding Timer)



Topology Change

- If a bridge detect a topology change send a Topology Change Notification BPDU trough root port
 - A bridge that receive a TCN BPDU forward that frame trough root port
- Root bridge answer with a Configuration BPDU with **topology change** bit set
 - The bridge receiving that BPDU from root port forward a BPDU with **topology change acknowledgment** bit set
- A bridge detecting or informed about a topology change set a short time to remove the entries equal to forward delay timer

Byte

1÷2	Protocol Identifier: 00-00
3	Protocol Version Identifier: 00
4	BPDU Type: 80

TCN



Timers values

- Hello time
 - Configuration BPDU generation interval
 - Range: 1 - 10 seconds - Recommended: 2 seconds
- Forward delay timer
 - Delay port state change: from listening to learning, from learning to forwarding state
 - short time to remove the entries from filtering database
 - Range: 4 - 30 seconds - Recommended: 15 seconds
- Max age
 - The maximum age of received protocol information before it is discarded (unblocking timer port).
 - Range: 6 - 40 seconds - Recommended: 20 seconds
- The timers are announced by the root bridge on the configuration BPDU



Timers limit

- The recommended timers assure a STP convergence of 50 seconds
- The recommended timers assure a correct STP operation with a maximum bridge diameter up 7 bridge
 - The **maximum bridge diameter** of the Bridge LAN is the maximum number of Bridges between any two points of attachment of end stations.
- Changing the timers may:
 - Increase or decrease the STP convergence
 - Increase or decrease the maximum bridge diameter



STP Configuration parameters

- The bridges are plug&play
 - They may work with default configuration parameters
- Bridge priority
 - Range: 0 - 61440
 - Default/recommended: 32768
 - Suggested increment (IEEE 802.1t): 4096
- Port priority
 - Range: 0 - 240
 - Default/recommended: 128
 - Suggested increment (IEEE 802.1t): 16
- Port Path cost
 - Range: 0 - 65535
 - Recommended (IEEE 802.1D): $1000 / (\text{Speed in Mb/s})$



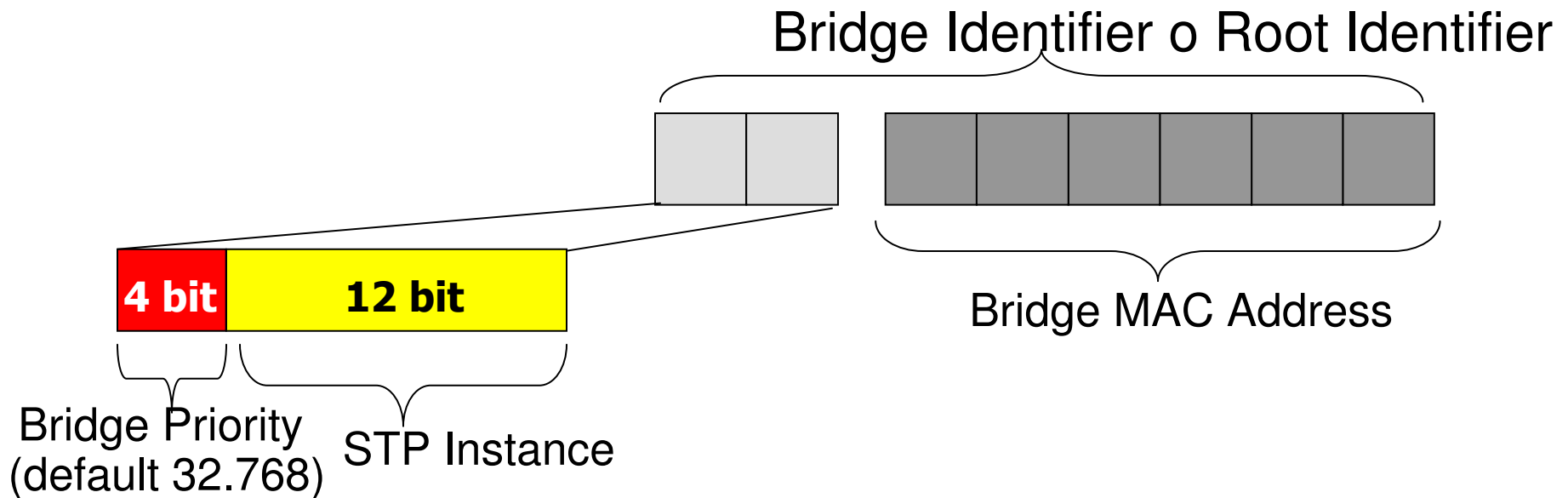
Recommended Port Path Cost by IEEE 802.D rev1998

Port speed	Recommended values	Recommended range values	Accepted range values
4Mb/s	250	100 - 1000	1 - 65535
10 Mb/s	100	50 - 600	1 - 65535
16 Mb/s	62	40-400	1 - 65535
100 Mb/s	19	10 - 60	1 - 65535
1 Gb/s	4	3 - 10	1 - 65535
10 Gb/s	2	1 - 5	1 - 65535



IEEE 802.1t: new STP parameters adopted by 802.1w e 802.1s standards

- Extended Port Path Cost in a range from 1 to 200.000.000
- Extended System Identifier



Path Cost IEEE 802.1t & 802.1w

June 2006



Port speed	Recommended value	Recommended range	Accepted range values
<= 100 Kb/s	200.000.000	20.000.000 - 200.000.000	1 - 200.000.000
1 Mb/s	20.000.000	2.000.000 - 200.000.000	1 - 200.000.000
10 Mb/s	2.000.000	200.000 - 200.000.000	1 - 200.000.000
100 Mb/s	200.000	20.000 - 2.000.000	1 - 200.000.000
1 Gb/s	20.000	2.000 - 200.000	1 - 200.000.000
10 Gb/s	2000	200 - 20.000	1 - 200.000.000
100 Gb/s	200	20 - 2000	1 - 200.000.000
1 Tb/s	20	2 - 200	1 - 200.000.000
10 Tb/s	2	1 - 20	1 - 200.000.000



Rapid Spanning Tree Protocol (RSTP) 802.1w

- Fast convergence Spanning tree:
 - Less than 1 second
- Interoperable with STP IEEE 802.1D
 - Without fast convergence
- Modern solution for mission critical BLAN
- Operate only on point-to-point links
 - Direct connections (no hub)



Port rules

- Blocking state substitute by Alternate & Backup port rules
- Alternate
 - Port that offers an alternate path in the direction of the Root Bridge to that provided by the Bridge's own Root Port.
- Backup
 - Acts as a backup for the path provided by a Designated Port in the direction of the leaves of the Spanning Tree.
 - Backup Ports exist only where there are two or more connections from a given Bridge to a given LAN.
- Root
- Designated
- Disabled



Edge Port

- Edge port is a terminating that connect a end-station
- Edge port do not work as other port:
 - As soon detect a Link Integrity Test go immediately to Forwarding state without Listening e Learning state
 - A port changing state do not cause a TCN BPDU transmission trough root port



Diagrammatic conventions

Legenda	
Graphic Symbol	
<p style="text-align: center;">Bridge Identifier</p> <p style="text-align: center;">Port Identifier & Port Cost</p> <div style="text-align: center;"> </div> <p style="text-align: center;">Root Identifier</p> <p style="text-align: center;">Root Path Cost</p>	BRIDGE
	LAN



Diagrammatic conventions

Port rule	Port state	Legenda
Designated	Discarding	●+———
	Learning	●++———
	Forwarding	●———
Edge Port	Forwarding	●◊———
Root Port	Discarding	○+———
	Learning	○++———
	Forwarding	○———
Alternate	Discarding	——+———
Backup	Discarding	——++———
Disabled	-----	—/———
BPDUs trasmesse		
Designated		————→
Designated Proposal		————→
Root		————▷
Root Agreement		————▷

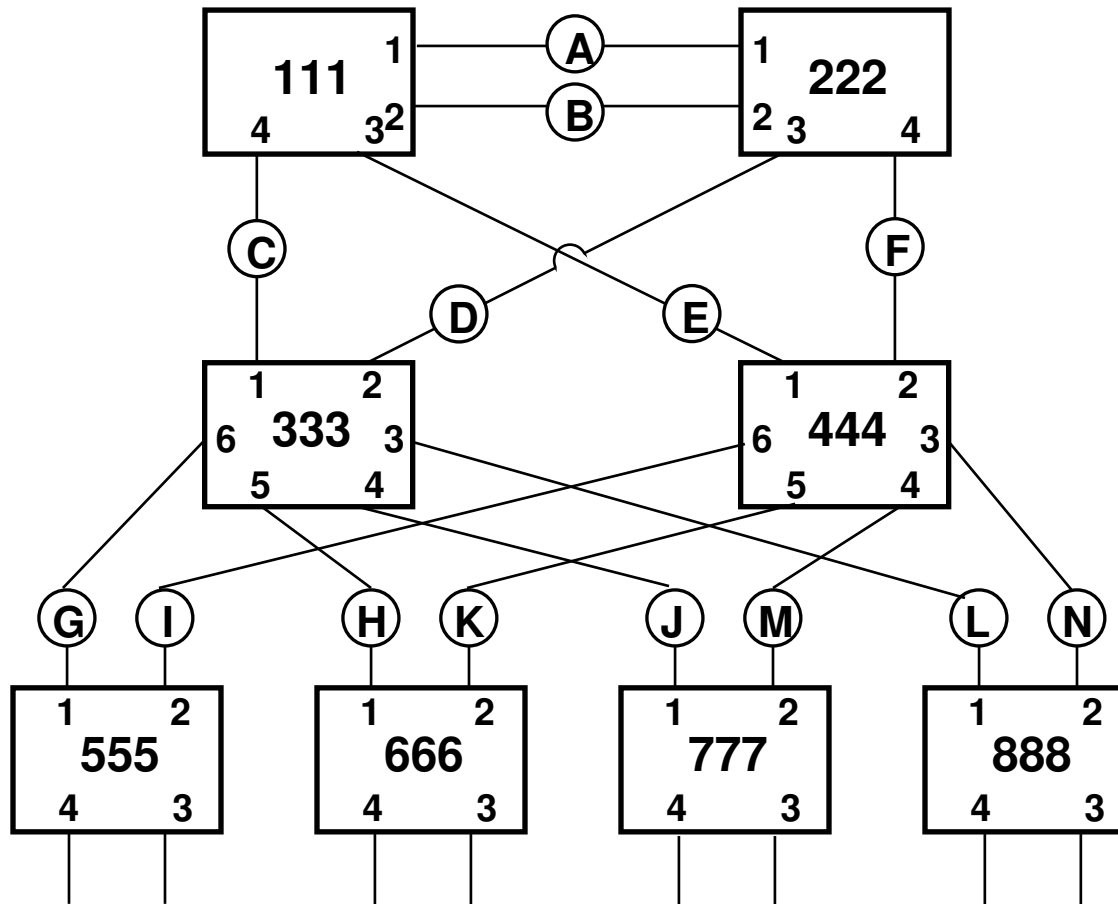


The algorithm principle

- Same as STP
- Root Bridge election
- Root port
- Designated Port
- The port not selected as root or designated are selected as:
 - Alternate if connected to a port on different bridge
 - Backup if connected to a different port of the same bridge

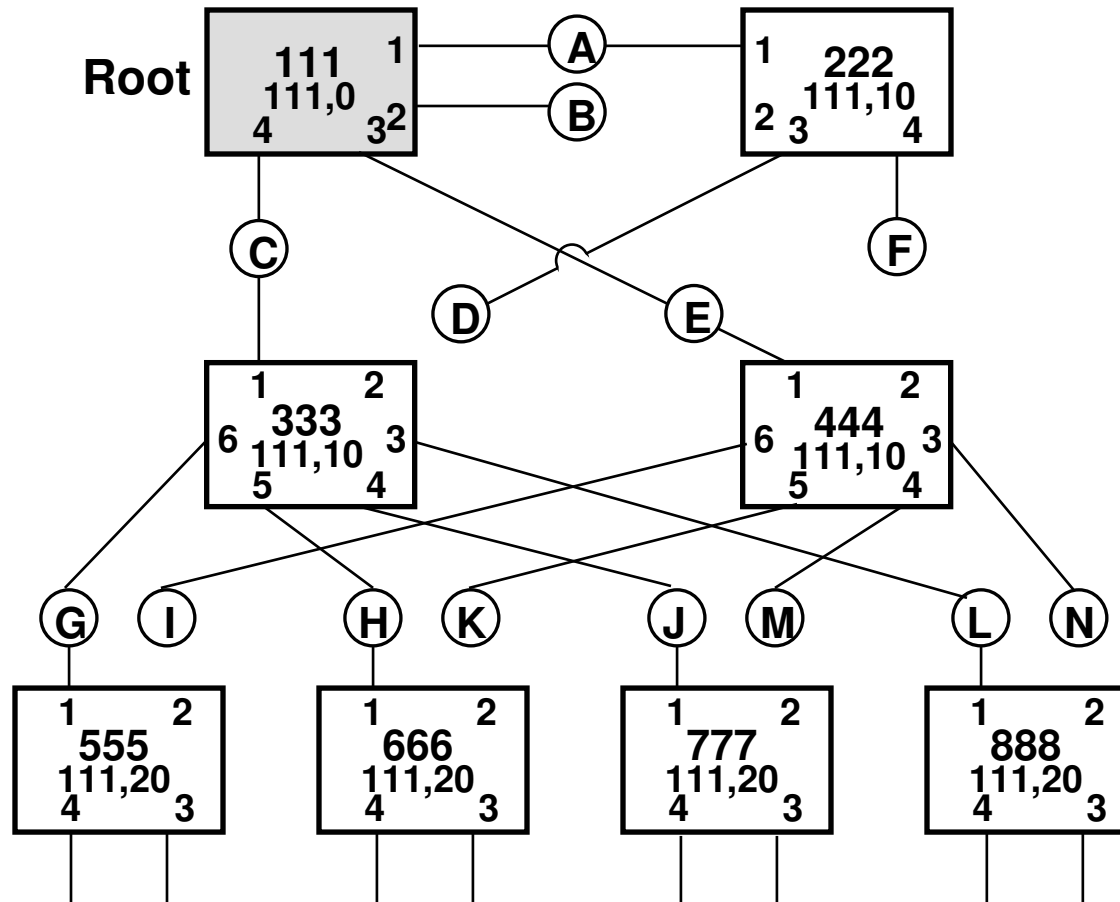


Physical redundant star topology



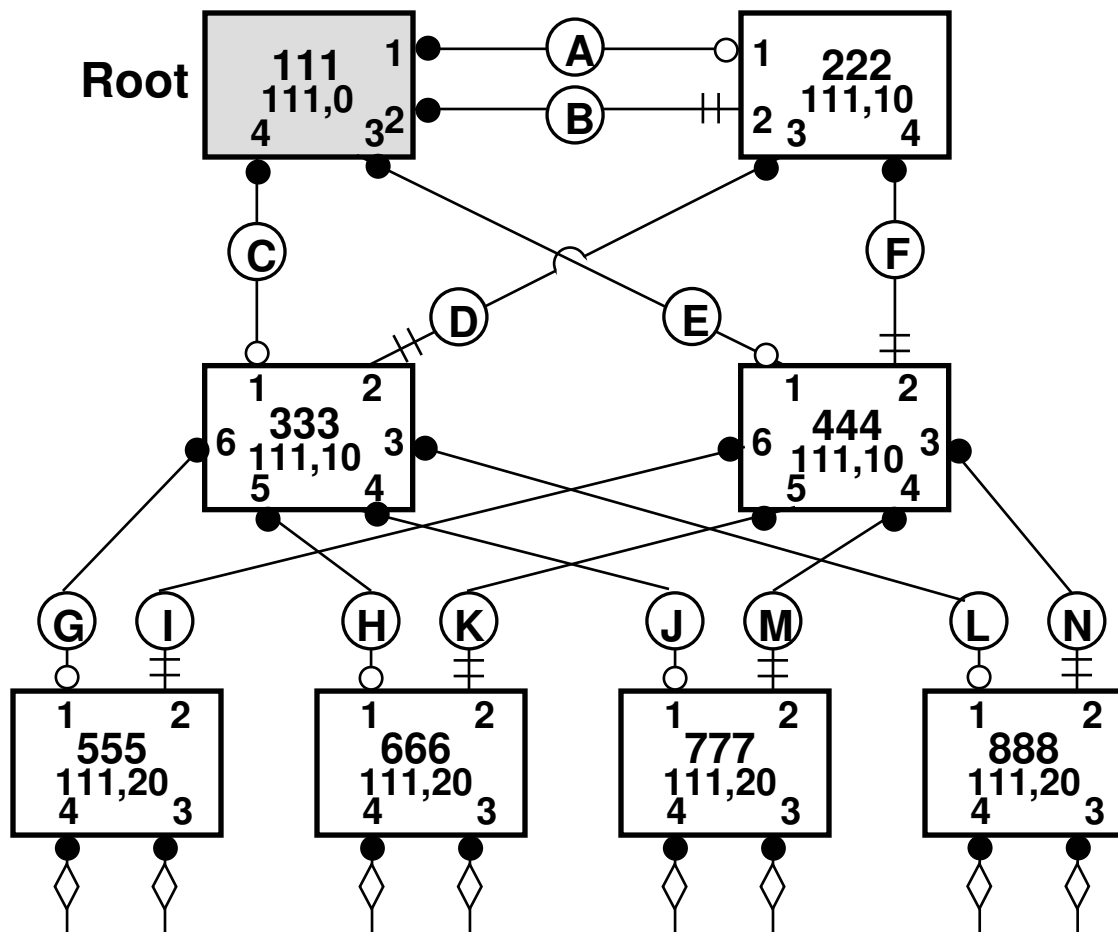


RSTP active topology



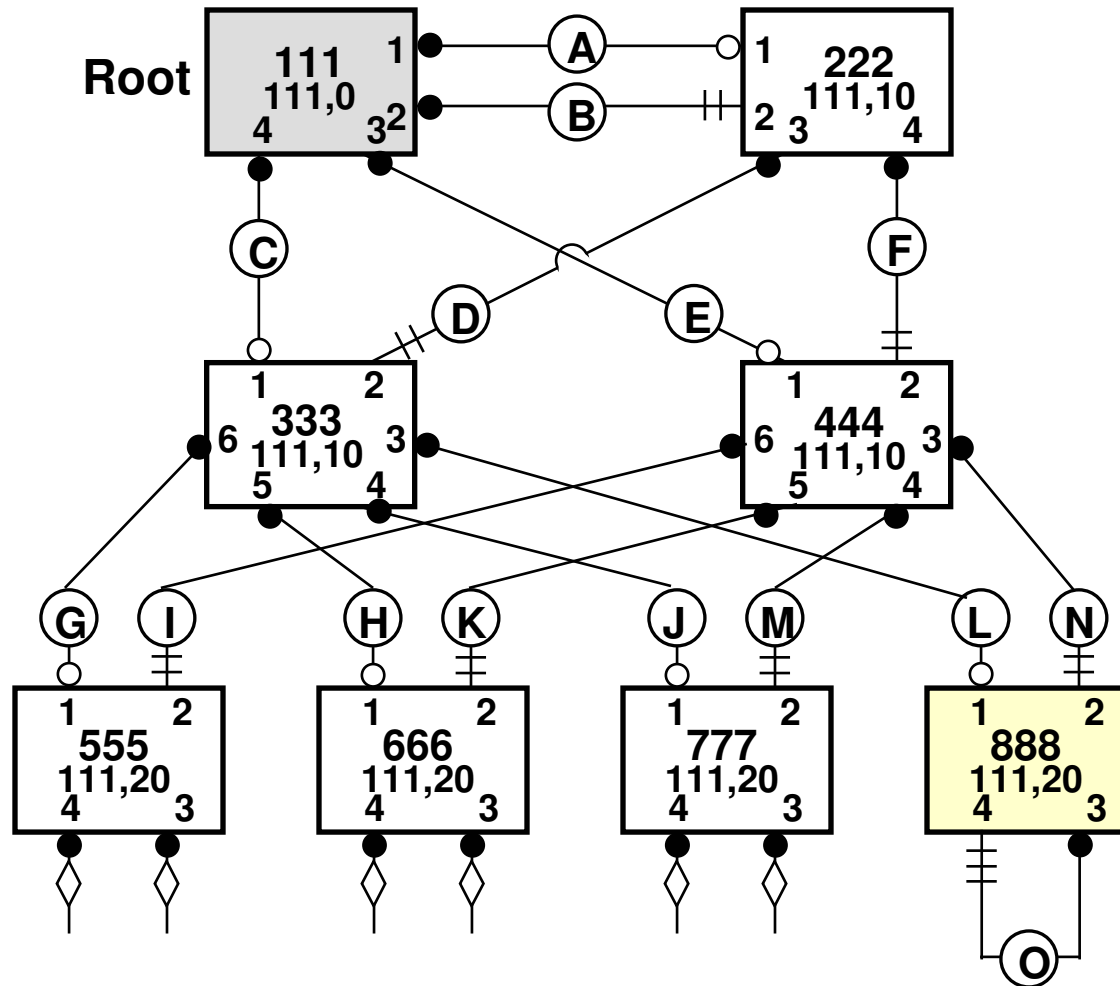


RSTP tree & port state



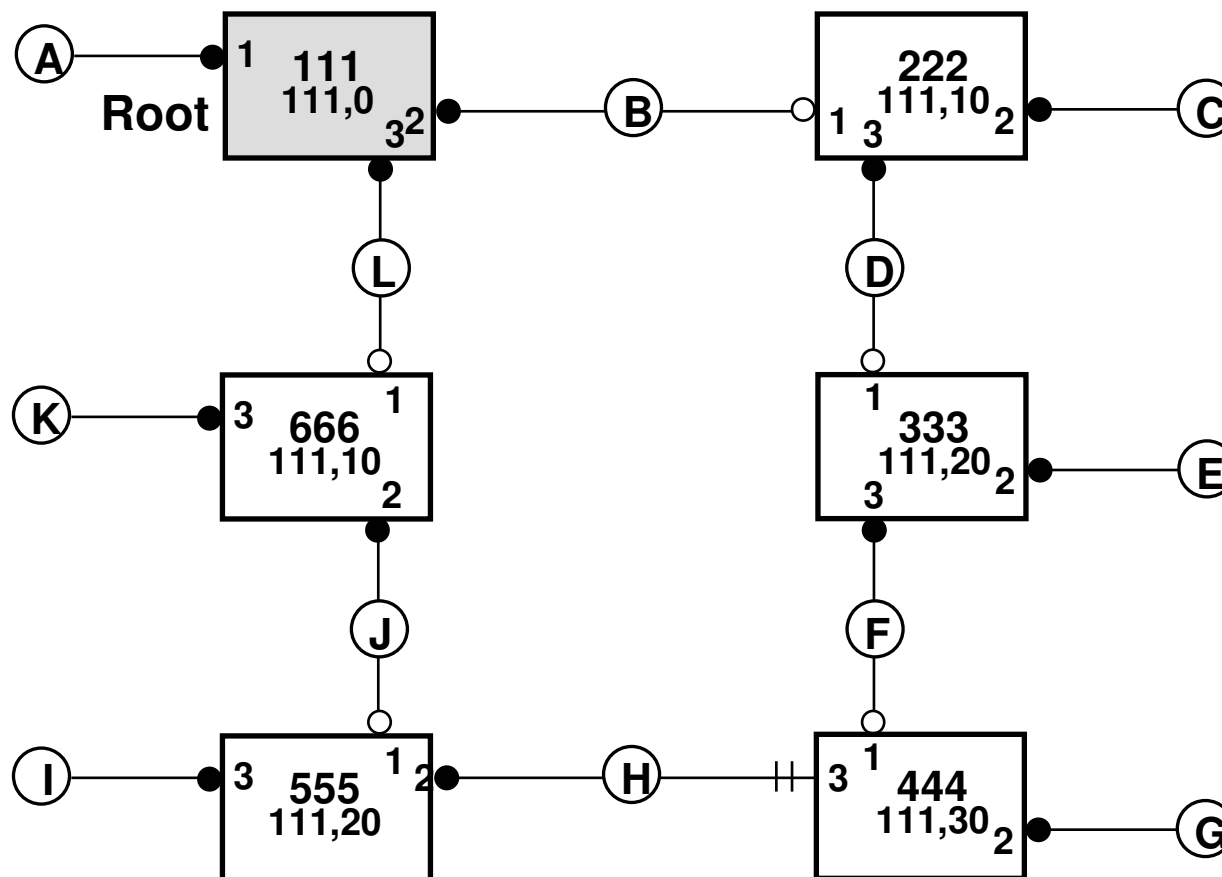


Backup port example





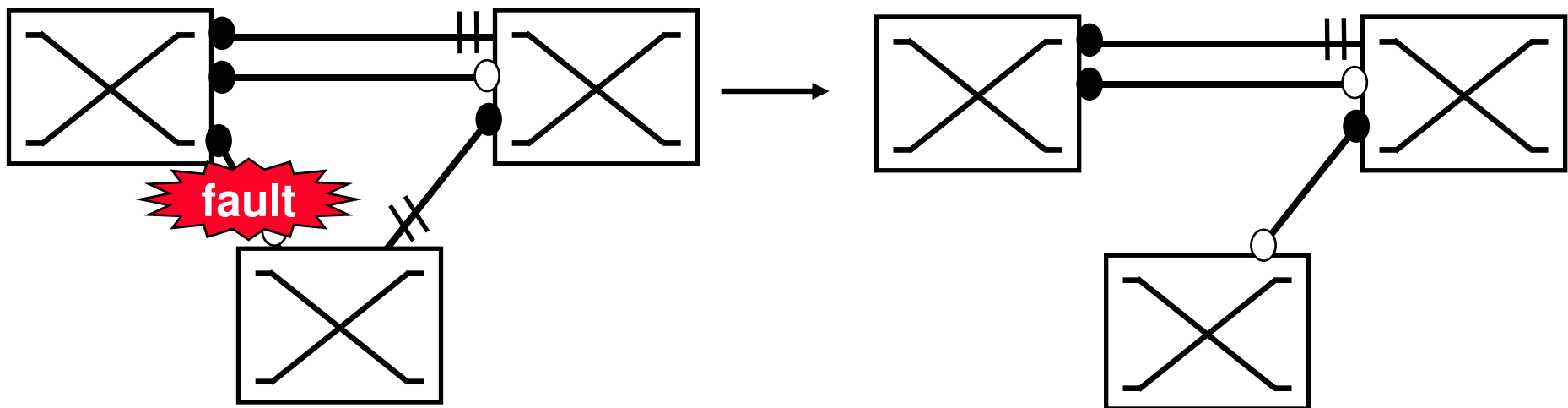
Ring topology





Fault recovery based on physical level fault detection

- If the switch loose a root port activate immediately the alternate port that became a new root port
 - Fault detection and recovery based on physical layer



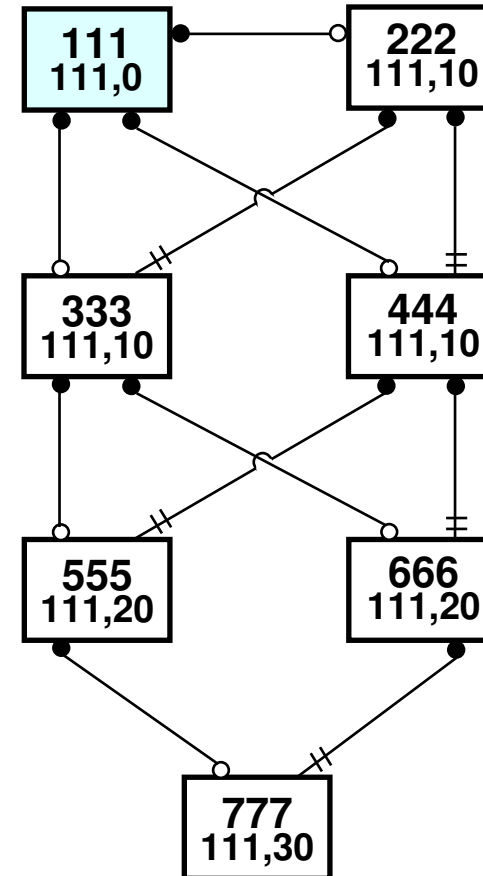
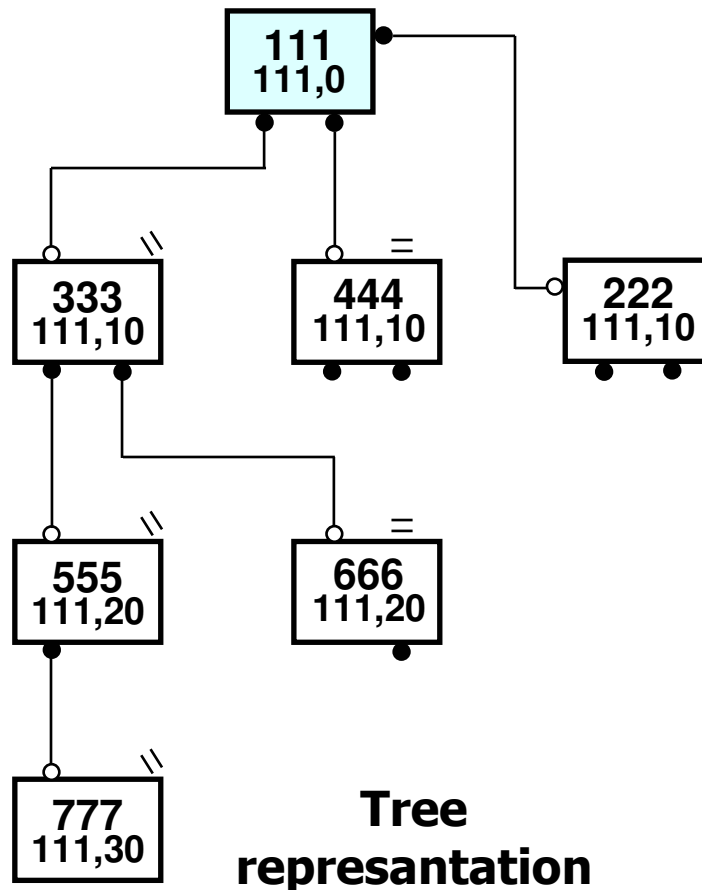


Fault detection and flush entry

- Local flush entries
 - The bridge flush immediately the entries that have a faulty port as destination port
 - Local flush entries on bridge ports at end of faulty link
- A bridge that change an Alternate Port to a Root Port send a ***Topology Change Notification*** (TCN) BPDU trough root port
- A bridge that receive a TCN BPDU:
 - Flush the entries on every ports except the port were the TCN as been received
 - Forward TCN on every ports in forwarding state except the port were the TCN as been received



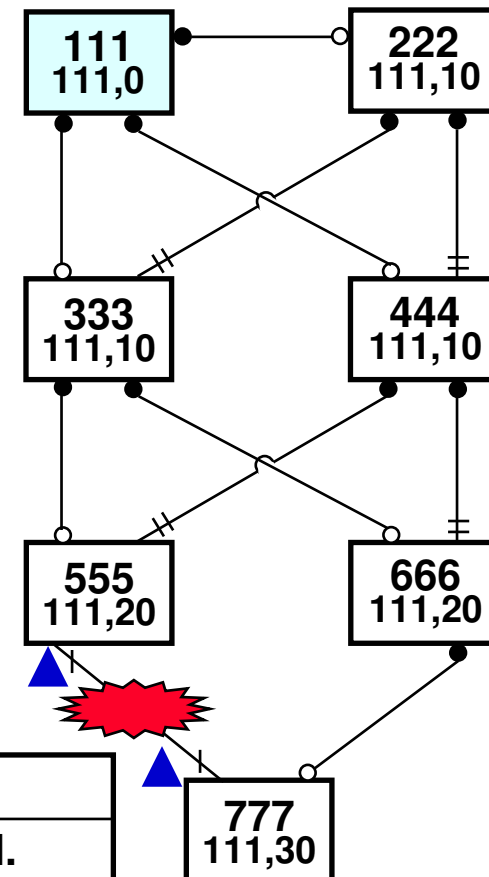
Flush entry example: initial configuration








Flush entry: step 1

- The link 555-777 fails, and both 555 and 777 notice. It can be seen that all the Ports indicated will need to be flushed.



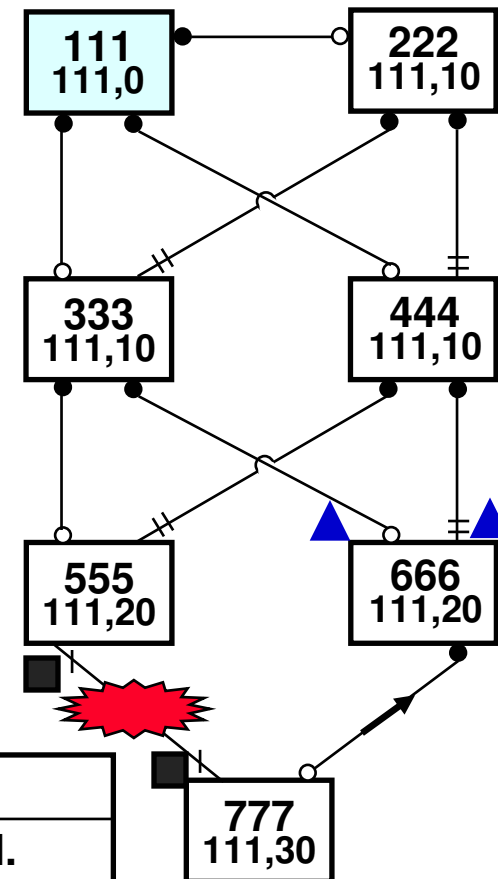
Legenda

-  Addresses learnt on these Ports need to be flushed.
-  Addresses learnt on these Ports have been flushed.
-  TCN transmitted in the direction of the arrow



Flush entry: step 2

- 555 and 777 flush addresses for their failed Ports.
- 777 change the port state from alternate to root
- 777 sends a TCN message (an RSTP BPDUs with TC set).



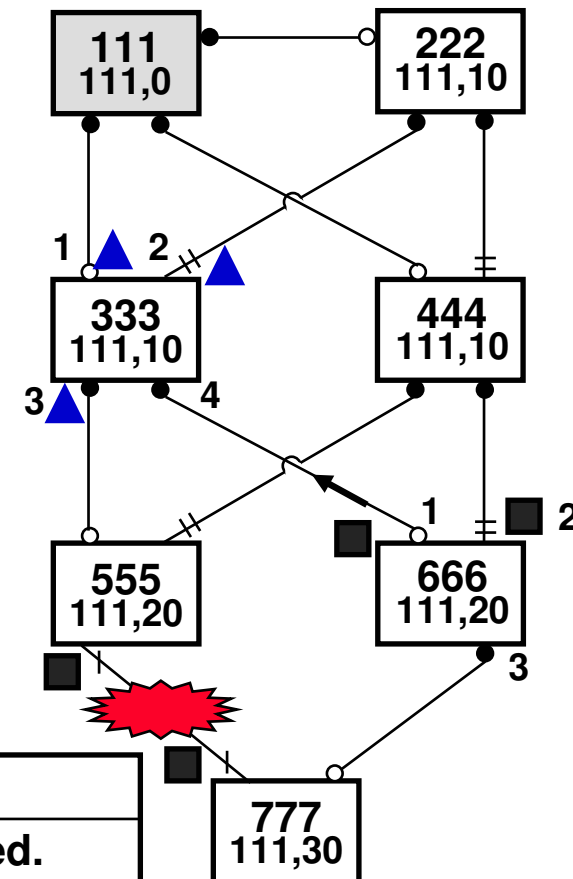
Legenda

- ▲ Addresses learnt on these Ports need to be flushed.
- Addresses learnt on these Ports have been flushed.
- TCN transmitted in the direction of the arrow



Flush entry: step 3

- 666 receives the TCN message, flushes addresses on all other Ports, and forwards a TCN.



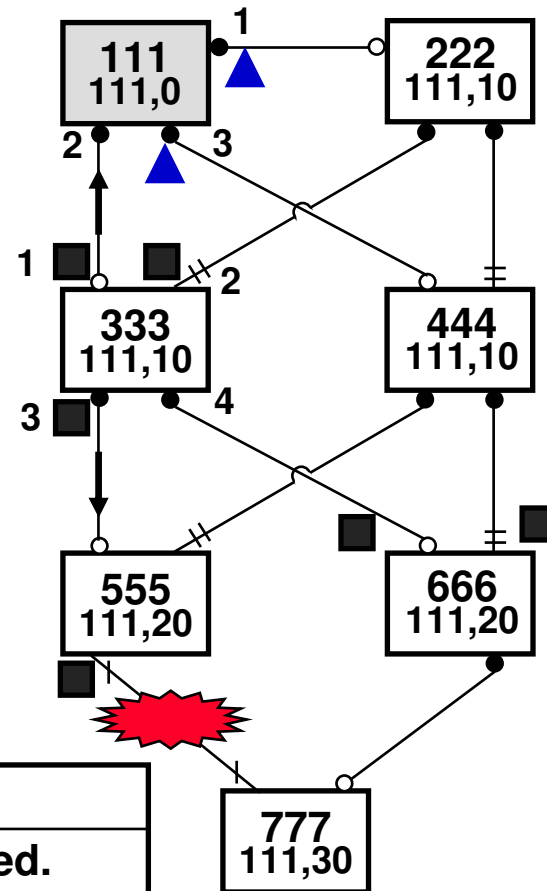
Legenda

- Addresses learnt on these Ports need to be flushed.
- Addresses learnt on these Ports have been flushed.
- TCN transmitted in the direction of the arrow



Flush entry: step 4

- 333 receives the TCN message, flushes, and forwards the TCN to 111 and 555.

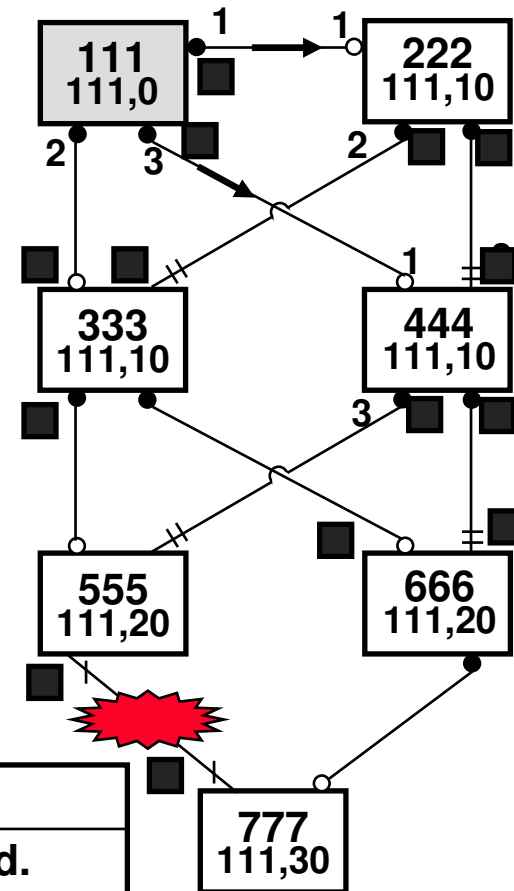


Legenda	
	Addresses learnt on these Ports need to be flushed.
	Addresses learnt on these Ports have been flushed.
	TCN transmitted in the direction of the arrow

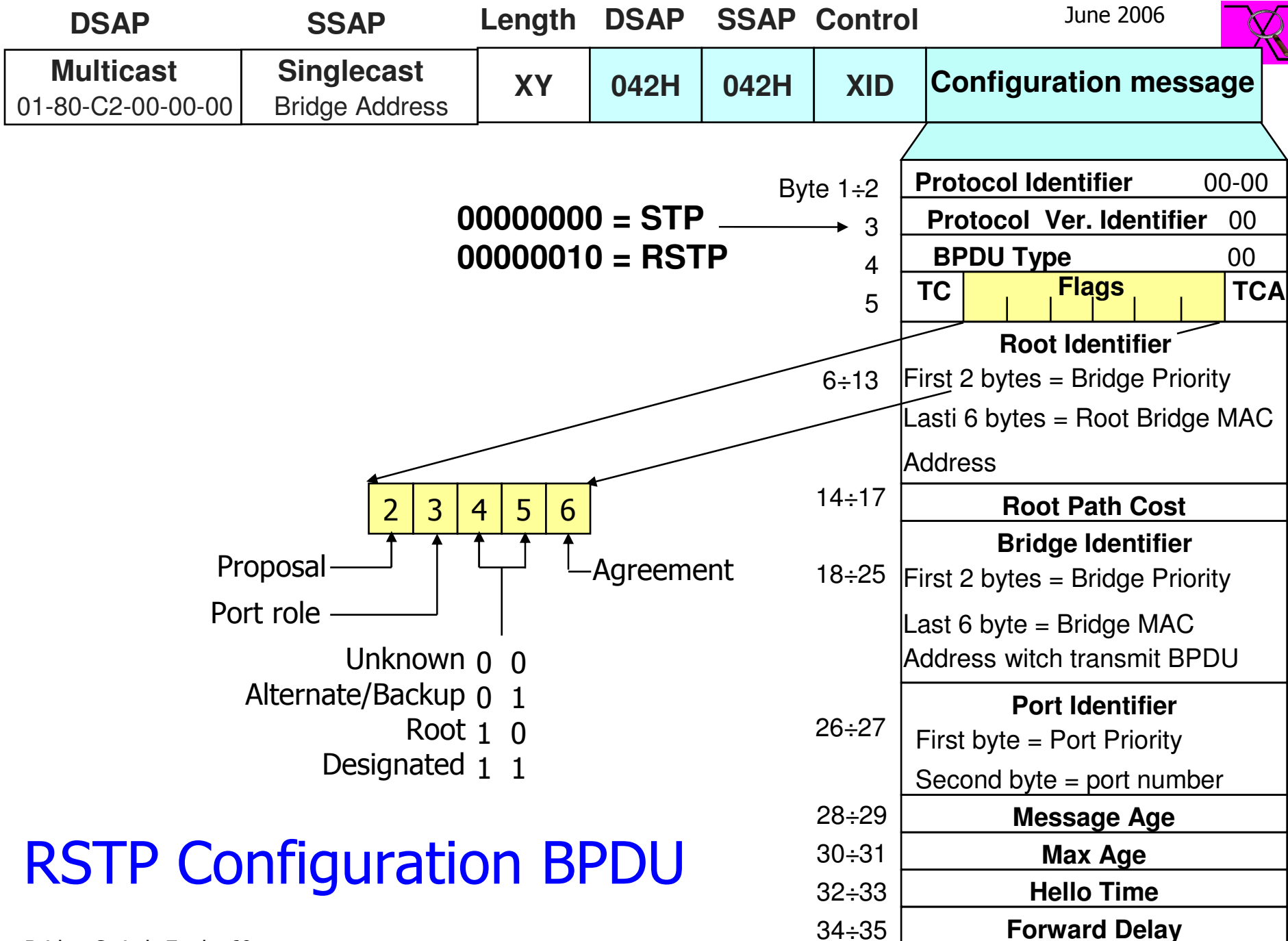


Flush entry: step 5

- 111 receives the TCN, flushes and sends TCN messages on all other Ports.
- Receipt of the TCN will cause 222 and 444 to flush their Ports.



Legenda	
▲	Addresses learnt on these Ports need to be flushed.
■	Addresses learnt on these Ports have been flushed.
→	TCN transmitted in the direction of the arrow



RSTP Configuration BPDUs



Ring topology and fast recovery

■ Fault on the ring

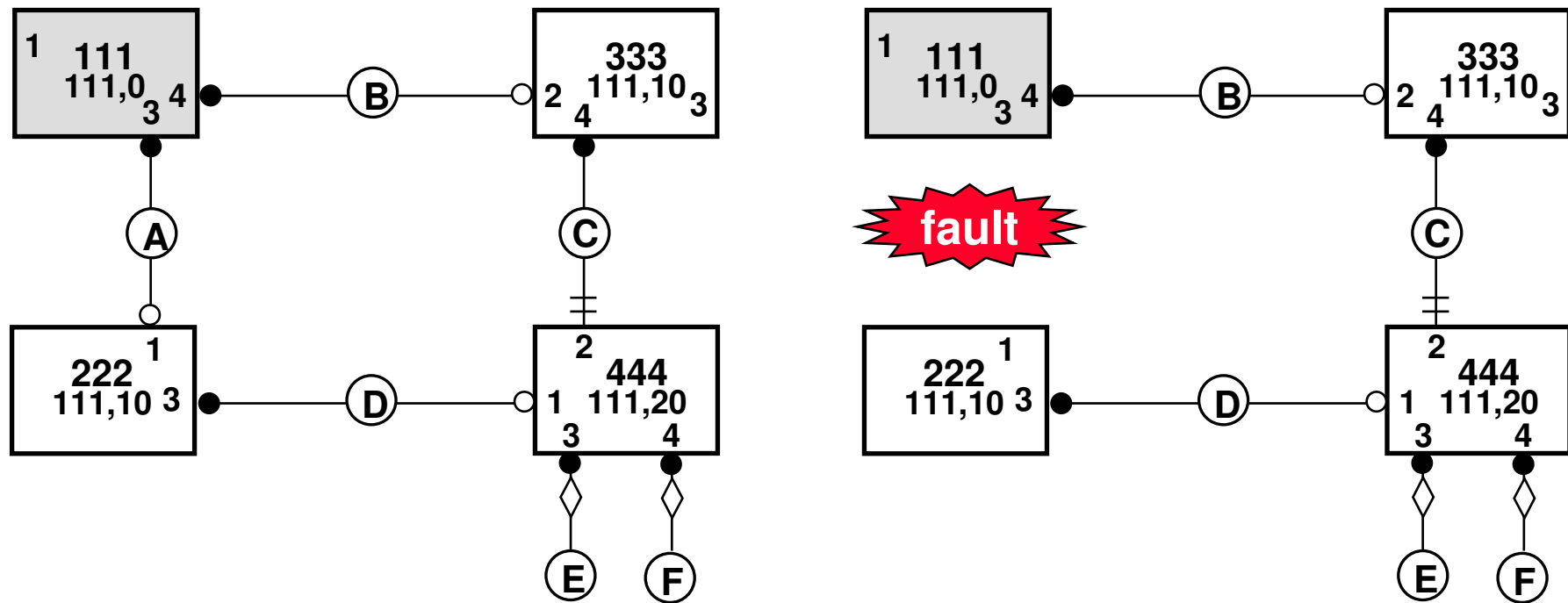
- If the bridge have an alternate port change immediately the state from alternate to root state
 - Recovery based physical layer
- If the bridge don't have an alternate port propose it self as Root Bridge
 - Recovery based on new protocol handshaking
 - Fast BPDU Proposal and Agreement type handshaking

■ Sync & Agree operations

- Port receiving BPDU with *Designated Proposal flag* execute a *Sync operation*
 - Discard all packet received on non-edge ports
- Port discarding packet (Sync) go immediately on Forwarding State when receive a BPDU with *Designated Agreement flag*

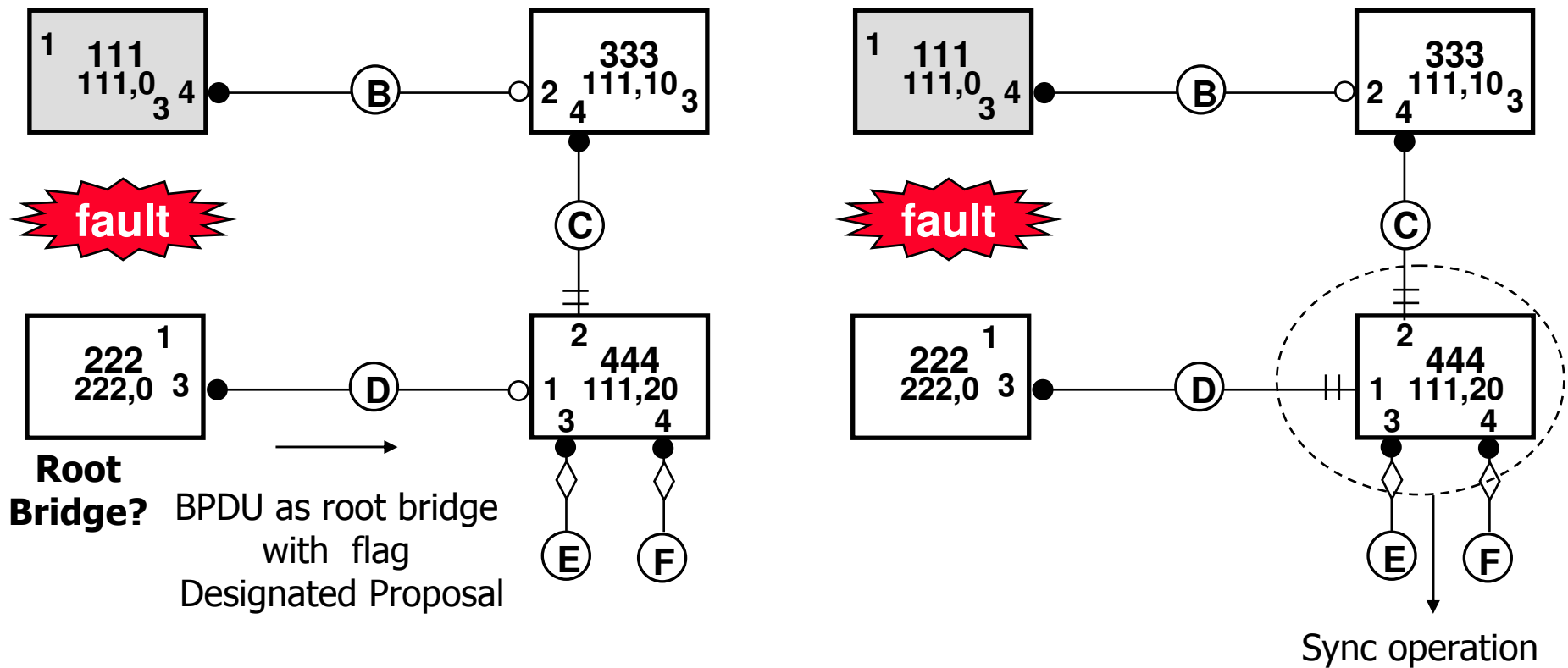


Ring topology fault: step 1





Ring topology fault and new Root proposal





Ring topology fault: Sync Operation

