



Layer 2 Spanning Tree Protocols

Redes de Comunicações II

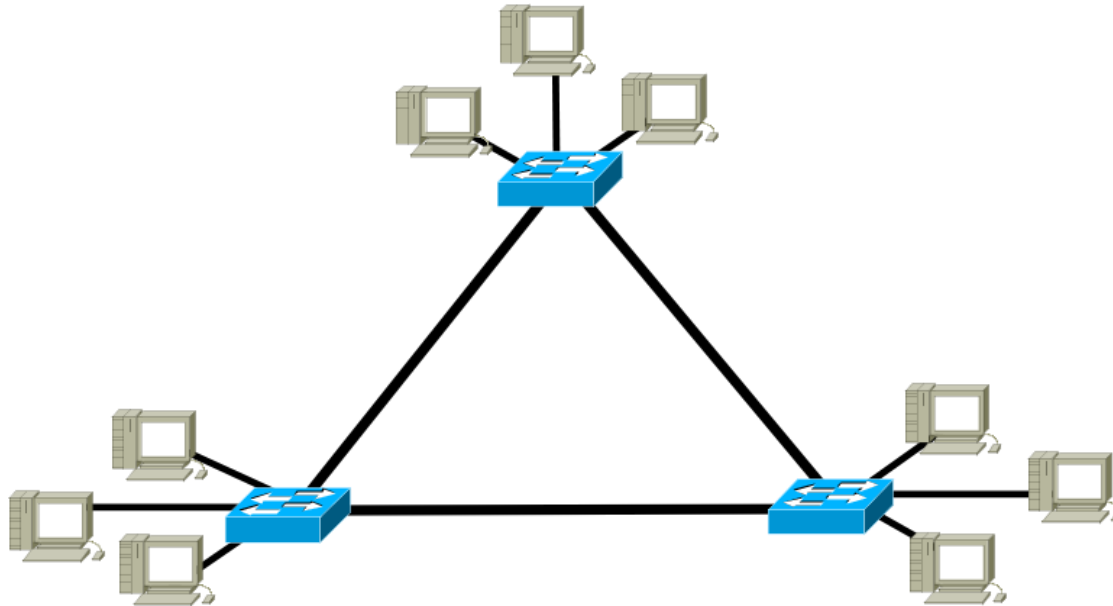
Licenciatura em Engenharia de
Computadores e Informática

Prof. Amaro de Sousa (asou@ua.pt)

DETI-UA, 2024/2025

Redundant Layer 2 network

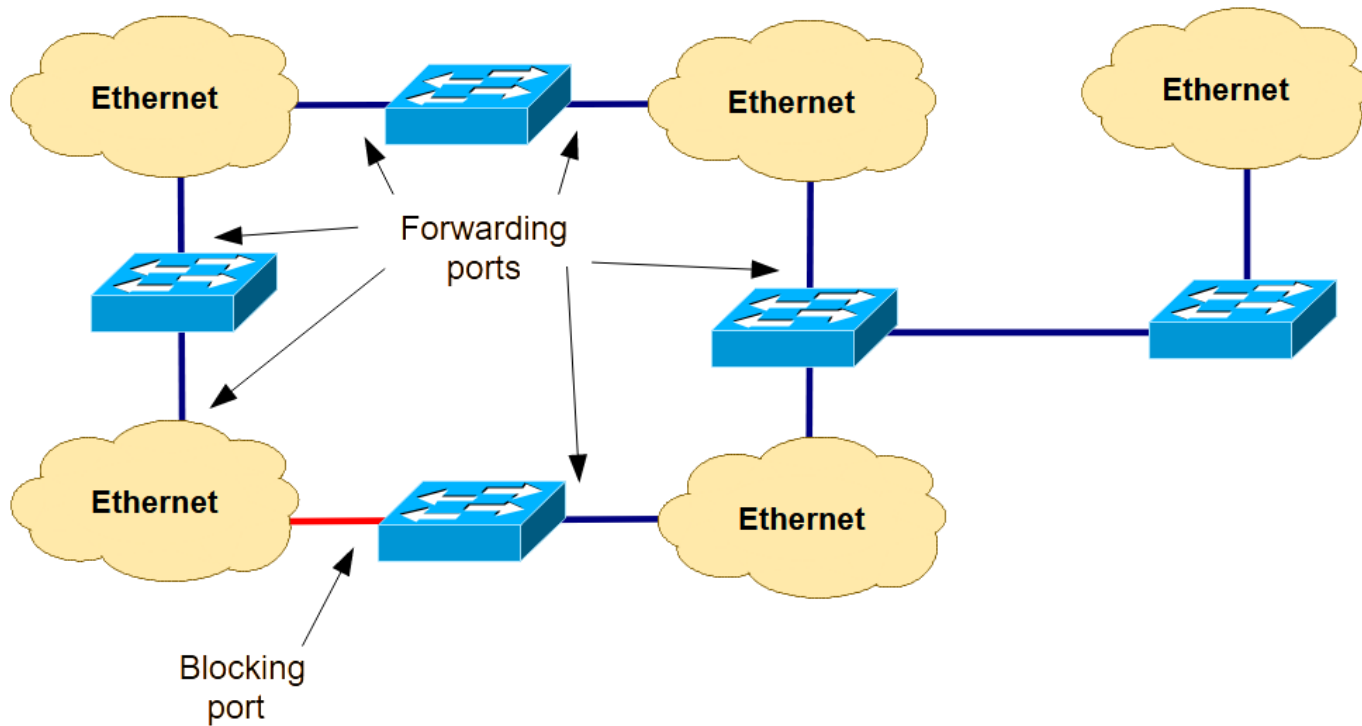
- Objective:** To allow the network for dynamically recovery from network failures.
- Problem:** Link redundancy creates Layer 2 loops. Causes the collapse of communications when MAC frames with broadcast address are sent by any host due to infinite frame flooding.



Spanning Tree Protocol (STP)

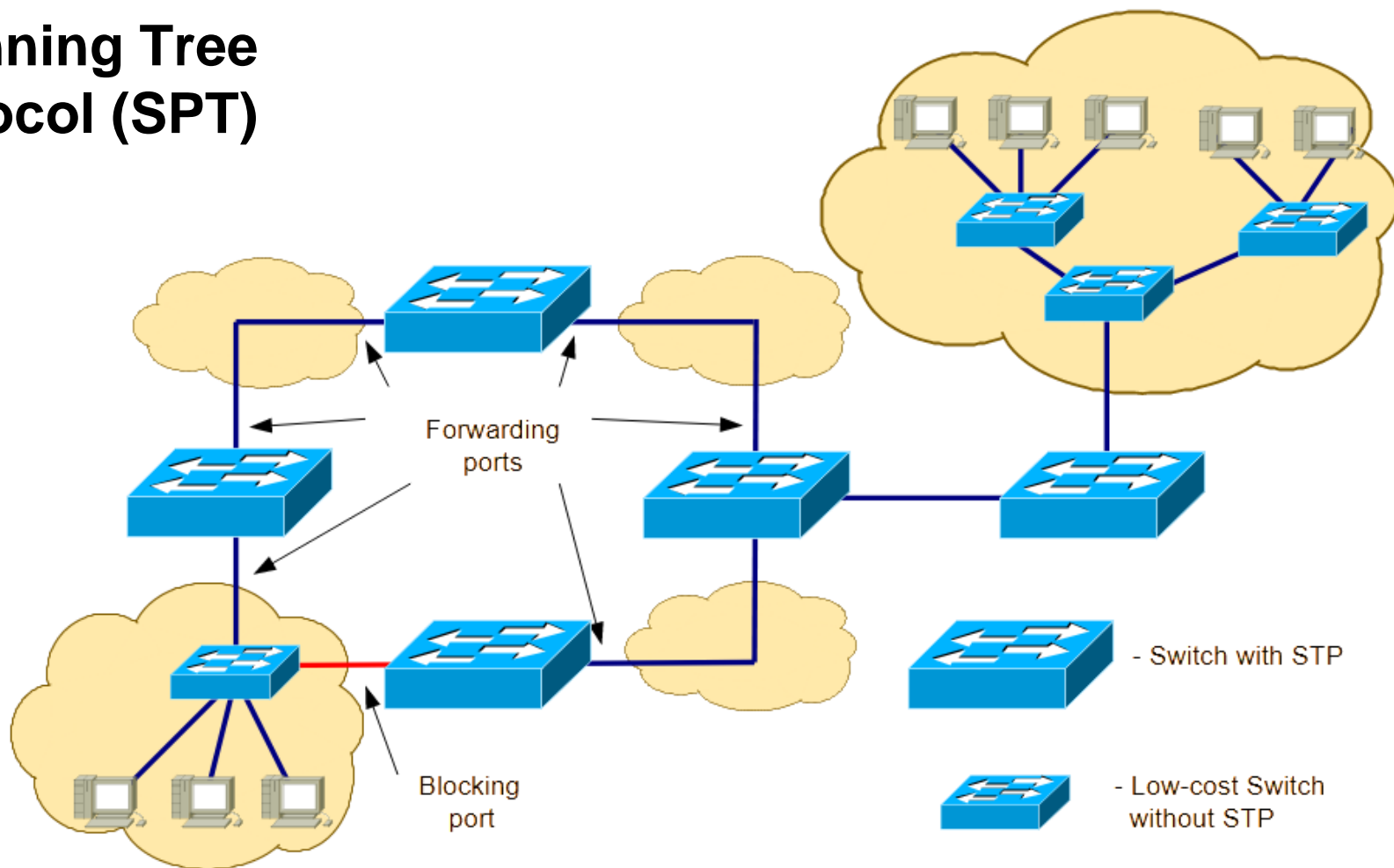
- STP enables the network to deterministically block ports and provide a fully connected loop-free topology in a network with redundant links.
- There are several STP Standards and Features:
 - STP is the original IEEE 802.1D version (802.1D-1998) that provides a loop-free topology in a network with redundant links.
 - RSTP, or IEEE 802.1W, is an evolution of STP that provides faster convergence of STP.
 - Multiple Spanning Tree (MST) is an IEEE standard. MST maps multiple VLANs into the same spanning-tree instance.
 - PVST+ (Per VLAN Spanning Tree Plus) is a Cisco enhancement of STP that provides a separate 802.1D spanning-tree instance for each VLAN configured in the network.
 - RPVST+ is a Cisco enhancement of RSTP that uses PVST+. It provides a separate instance of 802.1W per VLAN.

Spanning Tree Protocol (STP)



- The switches running STP exchange protocol messages between them to decide which ports are forwarding and which ports are blocking data frames.
- In the resulting spanning tree, all pairs of network segments are connected by a single forwarding path (i.e., there are no forwarding loops).

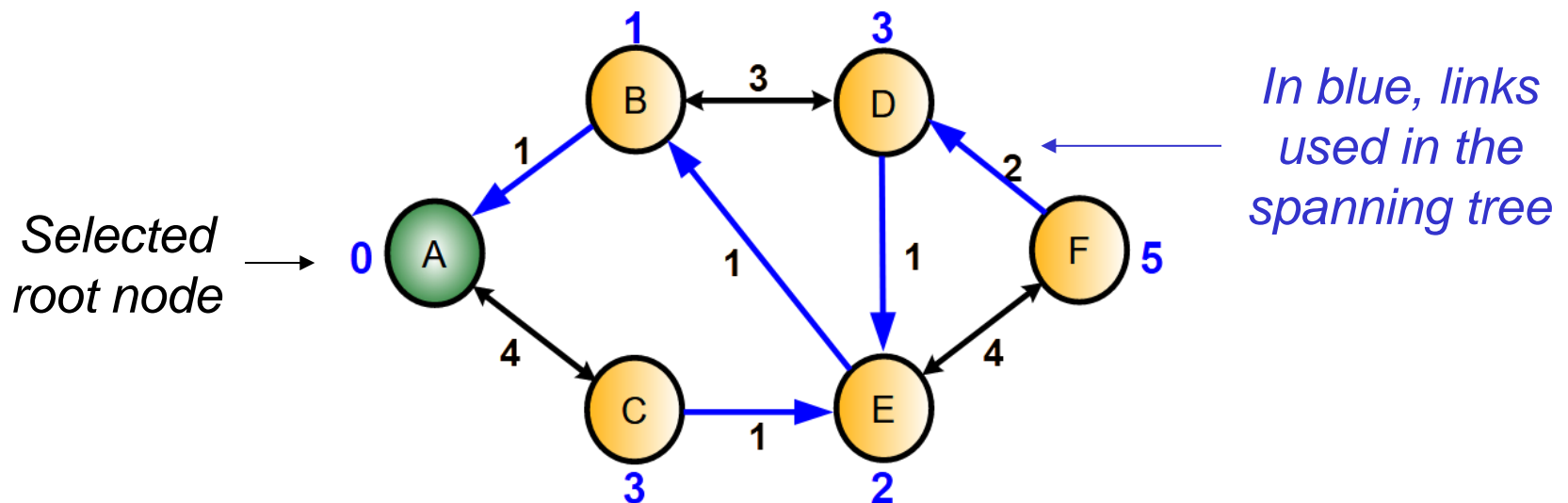
Spanning Tree Protocol (STP)



- The connections between Switches (supporting STP) might be:
 - Pont-to-point links
 - A network of switches (not supporting STP) with a loop-free topology

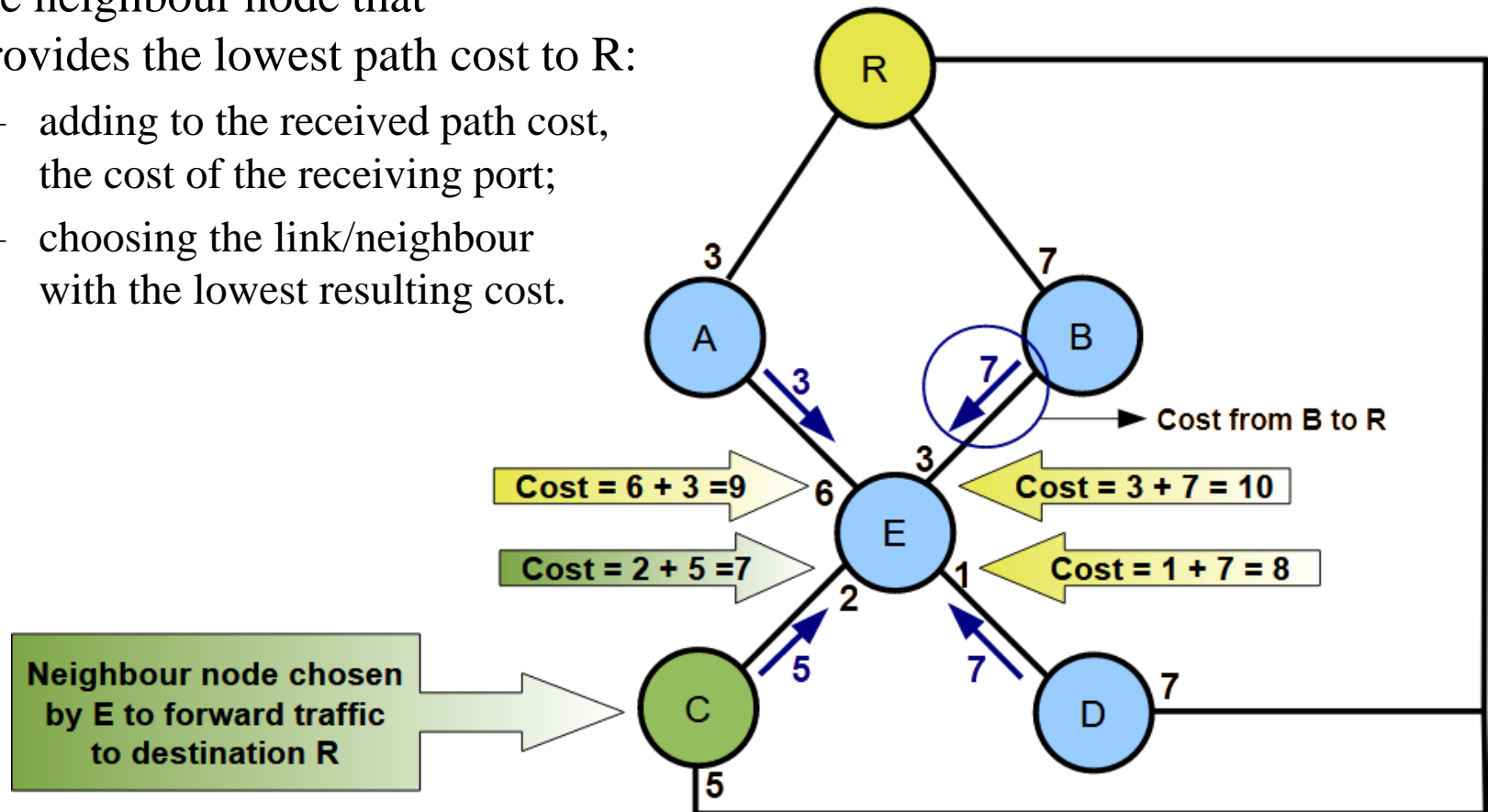
Spanning trees based on shortest paths

- The active links are the ones that belong to the shortest path from each node to a node selected as the root node
 - All nodes use the **Bellman-Ford Distributed and Asynchronous Algorithm** to calculate the neighbour node that provides the smallest path cost to the root node.
 - The set of links belonging to a shortest path is called the **Spanning Tree**.
 - A criterion to solve ties is required (as will be seen later).



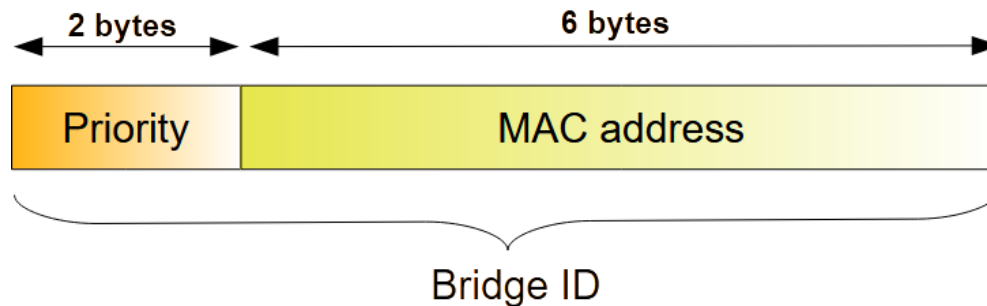
Bellman-Ford Distributed and Asynchronous Algorithm

- Each node transmits periodically (to its neighbours) the estimation of the path cost from it to the destination R.
- Upon reception of a message from a neighbour node, each node recalculates the neighbour node that provides the lowest path cost to R:
 - adding to the received path cost, the cost of the receiving port;
 - choosing the link/neighbour with the lowest resulting cost.



Spanning Tree Basic Concepts (1)

- **Bridge ID** – each switch is identified by an 8 bytes identifier:
 - **Priority** (2 bytes), configurable by the manager, default value: 32768 (8000h in hexadecimal)
 - **MAC address** (6 bytes), guaranteed by the manufacturer to be unique
 - **Priority** has precedence over the **MAC address**

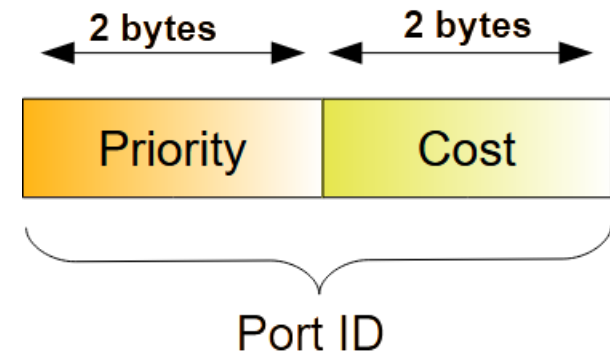


- **Root ID** – the switch with the lowest Bridge ID is selected as the **Root Bridge** and its ID is the Root ID
 - By default, the **Priority** is the same for all switches: in this case the **Root Bridge** is the switch with the lowest **MAC address**
 - The manager can change the selected Root Bridge by configuration of different **Priority** values on different switches

Spanning Tree Basic Concepts (2)

- **Port ID** – each interface is identified by a 4 bytes identifier:

- **Priority** (2 bytes), configurable by the manager, default value: 128 (80h in hexadecimal)
- **Cost** (2 bytes), configurable by the manager, default value depends on the interface speed
- **Priority** has precedence over the **Cost**

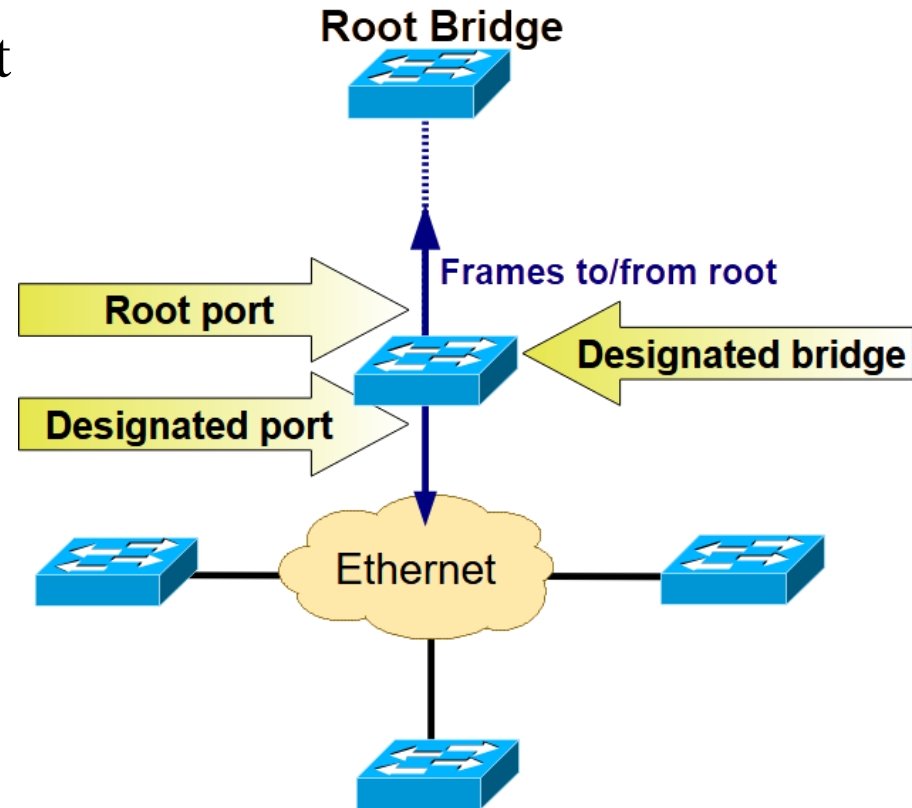


- The aim of the default **Cost** values is to include preferably the higher capacity links in the spanning tree
- The manager can change the **Cost** values in the different interfaces by configuration
- The need to change the **Priority** value of each interface is outside the scope of this course unit

Interface Speed	Default Cost Value
10 Mbps	100
100 Mbps	19
1 Gbps	4
10 Gbps	2
25 Gbps	1
40 Gbps	1
100 Gbps	1

Spanning Tree Basic Concepts (3)

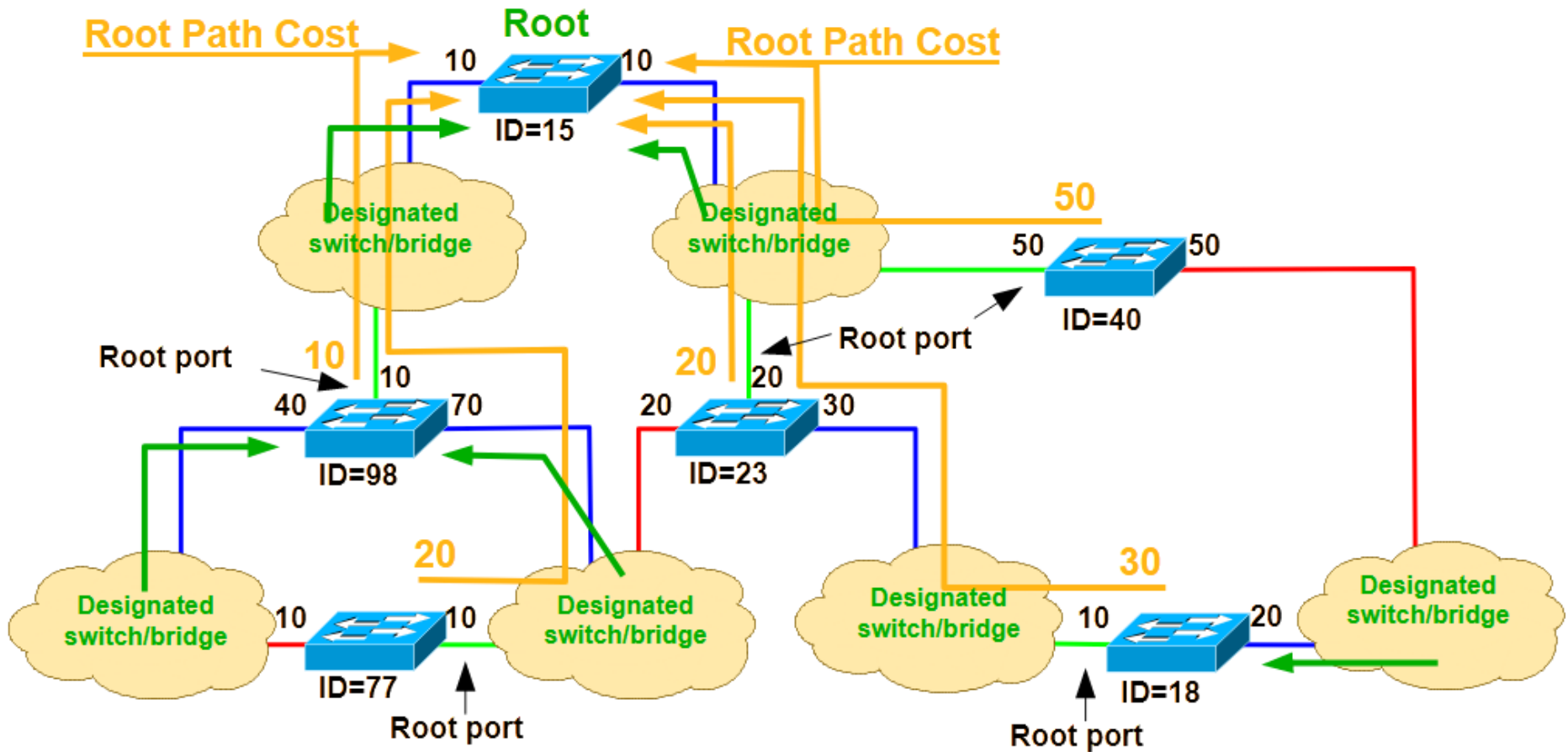
- **Designated Bridge** – Switch responsible to forward the packets from an Ethernet segment to and from the **Root Bridge**.
 - The root bridge is the designated bridge to all Ethernet segments connected to it.
- **Designated Port** – Port of the designated bridge of the Ethernet segment.
- **Root Port** – Port of the switch that provides the shortest path to the **Root Bridge**.
 - The Root Bridge has no Root Port.



Spanning Tree Basic Concepts (4)

- Each switch has an associated cost of the shortest path to the root (**Root Path Cost**), given by the sum of the costs of all root ports along the path to the root.
- The **Root Port** in each switch is the port with the lowest Root Path Cost.
 - If multiple ports have the same lowest Root Path Cost, the port connected to the neighbour switch with the lowest Bridge ID is chosen as Root Port
- The **Designated Bridge** of each Ethernet segment, is the switch with the lowest Root Path Cost among all switches connected to the segment.
 - If multiple switches connected to an Ethernet segment have the same lowest Root Path Cost, the switch with the lowest Bridge ID becomes the Designated Bridge of the segment.
- The **Designated Port** of each Ethernet segment is the port of its Designated Bridge.
- The Root Ports and the Designated Ports are put in the **Forwarding state**.
- All other ports are put in **Blocking state**.

Spanning Tree Basic Concepts Illustration



Spanning tree info in a Cisco switch

Running the command
show spanning-tree brief:

Root ID:

- 32768 - c201.1a70.0000

Root Path Cost: 38

Root Port:

- 56 (FastEthernet1/15)

Bridge ID:

- 32768 - c204.2dac.0000

Interface F1/14:

- Port priority: 128
- Port cost: 19
- State: Blocking

Interface F1/15:

- Port priority: 128
- Port cost: 19
- State: Forwarding

```
ESW1#show spanning-tree brief
```

```
VLAN1
```

```
Spanning tree enabled protocol ieee
```

```
Root ID    Priority    32768
           Address    c201.1a70.0000
           Cost      38
           Port      56 (FastEthernet1/15)
           Hello Time 2 sec  Max Age 20 sec  Forward Delay 15 sec
```

```
Bridge ID  Priority    32768
           Address    c204.2dac.0000
           Hello Time 2 sec  Max Age 20 sec  Forward Delay 15 sec
           Aging Time 300
```

Interface Name	Port ID	Prio	Cost	Sts	Designated Cost	Designated Bridge ID	Designated Port ID
FastEthernet1/14	128.55	128	19	BLK	19	32768 c203.24a4.0001	128.43
FastEthernet1/15	128.56	128	19	FWD	19	32768 c202.504c.0001	128.43

Designated Bridge on interface F1/14:

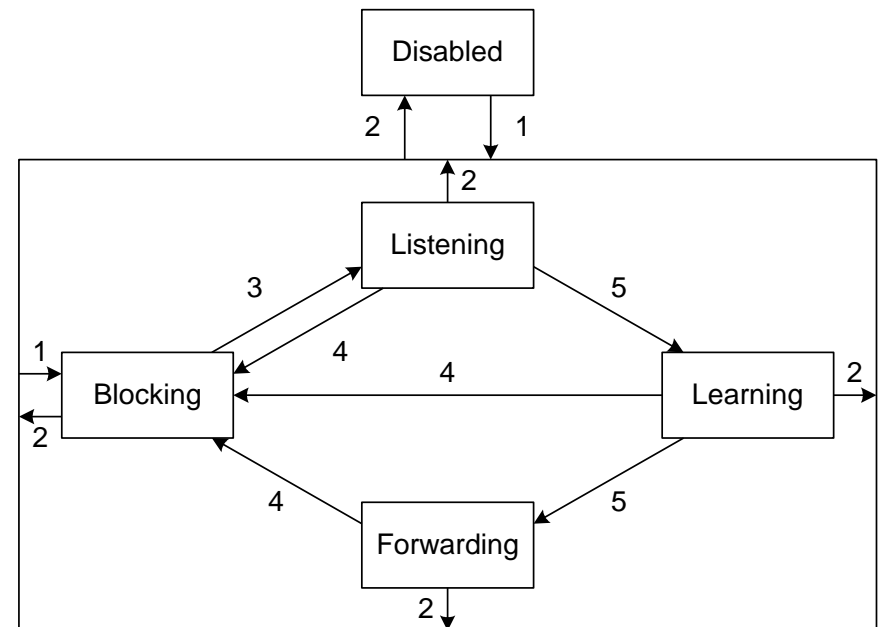
- 32768 - c203.24a4.0001 (with Port ID 128.43)

Designated Bridge on interface F1/15:

- 32768 - c202.504c.0001 (with Port ID 128.43)

Possible Port States (to avoid temporary cycles)

- Blocking state:
 - MAC address learning and packet forwarding are disabled
 - Receives and processes BPDU
- Listening state:
 - MAC address learning and packet forwarding are disabled
 - Receives and processes BPDU
 - When **Forward Delay** timer expires the port transits to Learning state
- Learning state:
 - MAC address is enabled but packet forwarding are disabled
 - Receives and processes BPDU
 - When **Forward Delay** timer expires the port transits to Forwarding state
- Forwarding state:
 - MAC address learning and packet forwarding are enabled
 - Receives and processes BPDU
- Disabled state:
 - MAC address learning and packet forwarding are disabled
 - Does not receive BPDU



IEEE 802.1D Protocol

BPDUs (Bridge Protocol Data Units)

- To build the spanning tree, switches exchange special messages between them called Bridge Protocol Data Units (BPDU).
 - There are two types: Configuration e TCN (Topology Change Notification).

IEEE 802.3 Ethernet

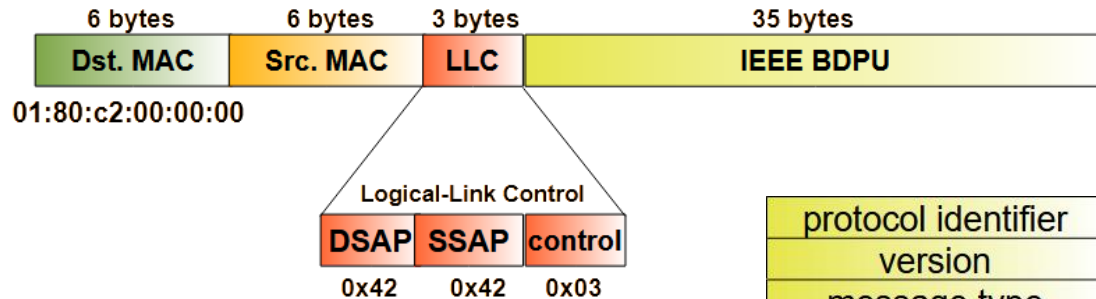
Destination: 01:80:c2:00:00:00 (01:80:c2:00:00:00)
 Source: 00:16:e0:9a:c3:92 (00:16:e0:9a:c3:92)
 Length: 39

Logical-Link Control

DSAP: Spanning Tree BPDU (0x42)
 SSAP: Spanning Tree BPDU (0x42)
 Control field: U, func=UI (0x03)

Spanning Tree Protocol

Protocol Identifier: Spanning Tree Protocol (0x0000)
 Protocol Version Identifier: Spanning Tree (0)
 BPDU Type: Configuration (0x00)
 Root ID: 32768 / 00:05:1a:4e:fd:58
 Root Path Cost: 24
 Bridge ID: 32768 / 00:16:e0:9a:c3:80
 Port ID: 0x8012
 Message Age: 1
 Max Age: 20
 Hello Time: 2
 Forward Delay: 15



protocol identifier		
version		
message type		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		

Set up of the Spanning Tree

- The setup of the Spanning Tree is done using the Configuration BPDU messages.

IEEE 802.3 Ethernet

Destination: 01:80:c2:00:00:00 (01:80:c2:00:00:00)

Source: 00:16:e0:9a:c3:92 (00:16:e0:9a:c3:92)

Length: 39

Logical-Link Control

DSAP: Spanning Tree BPDU (0x42)

SSAP: Spanning Tree BPDU (0x42)

Control field: U, func=UI (0x03)

Spanning Tree Protocol

Protocol Identifier: Spanning Tree Protocol (0x0000)

Protocol Version Identifier: Spanning Tree (0)

BPDU Type: Configuration (0x00)

Root ID: 32768 / 00:05:1a:4e:fd:58

Root Path Cost: 24

Bridge ID: 32768 / 00:16:e0:9a:c3:80

Port ID: 0x8012

Message Age: 1

Max Age: 20

Hello Time: 2

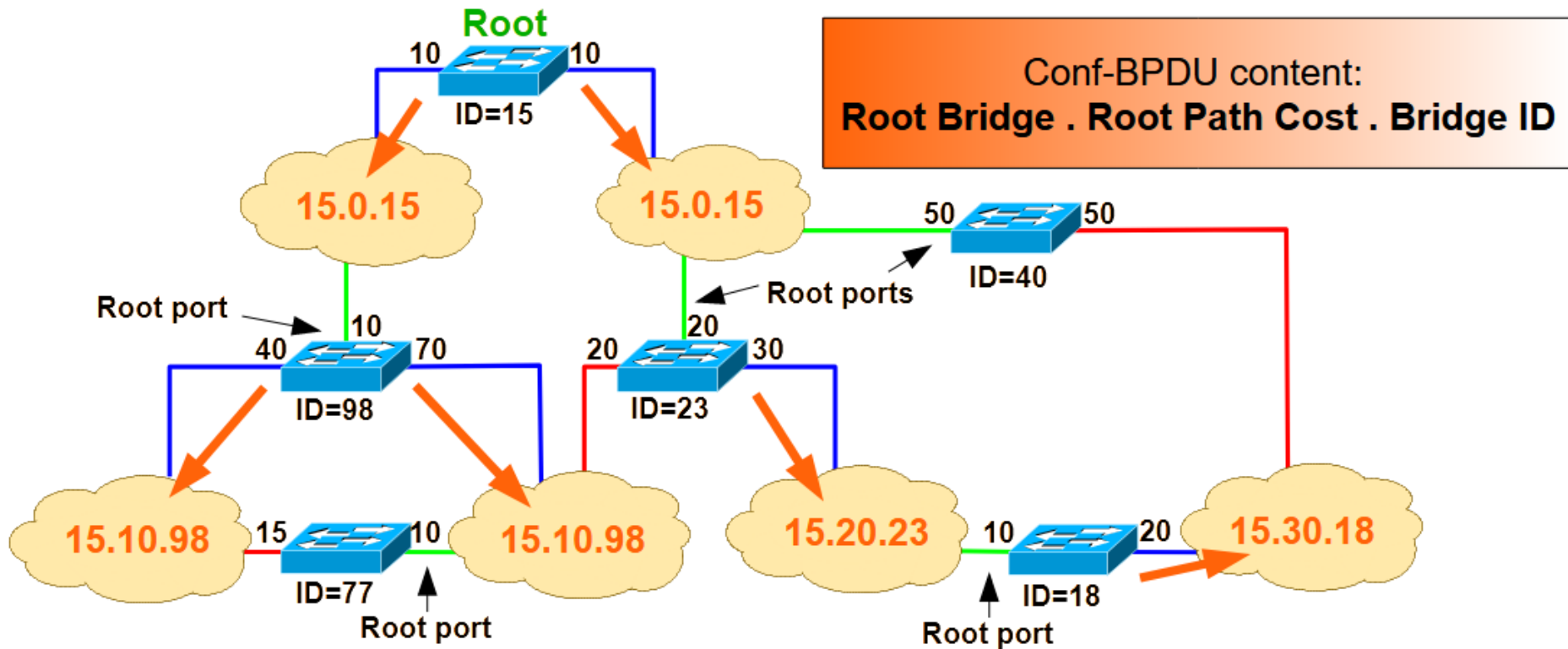
Forward Delay: 15

More relevant fields:

- **Root ID** of the current Root Bridge known by the sending switch.
- **Root Path Cost** of the sending switch.
- **Bridge ID** of the sending switch.

Spanning Tree maintenance

- Switches send periodically Conf-BPDUs by their Designated Ports in accordance with the Bellman-Ford Algorithm.
 - Periodicity of Conf-BPDU messages = Hello Time
 - Recommended Hello time: 2 seconds.
 - Defined at the Root Bridge.



Lifetime of MAC Address Table

Recall how MAC Address Tables are managed:

- A new MAC address is automatically inserted when a frame is received
- An existing MAC address is deleted when the **Lifetime** is reached without receiving any other frame from it

In general, the **Lifetime** value is a trade-off:

- Long Lifetime values – many frames will be lost when network is changing topology.
- Short Lifetime values – generate too much traffic due to frequent flooding.

There are two lifetime values:

- **Long Lifetime:** used by default (recommended value = 300 seconds)
- **Short Lifetime:** used when the Spanning Tree changes (recommended value = 15 seconds)

TCN (Topology Change Notification) BPDUs

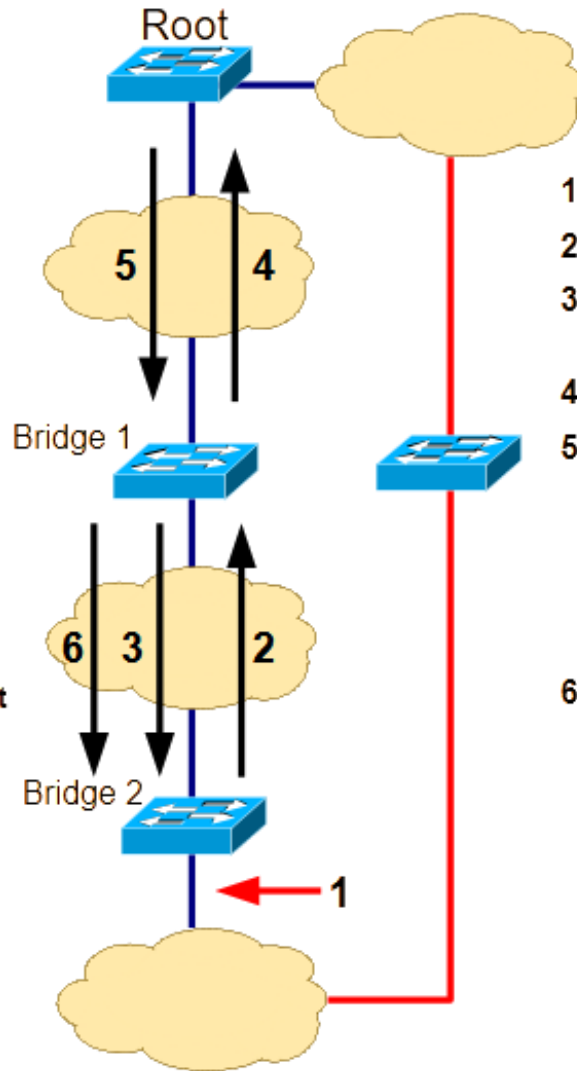
Conf (Configuration) BPDU

protocol identifier		
version		
message type = 0		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		

TCA - flag Topology Change Acknowledgment
TC - flag Topology Change

TCN (Topology Change Notification) BPDU

protocol identifier
version
message type = 1



1. Port changes state to disabled or blocking
2. Sends TCN-BPDU (periodicity = hello time)
3. Sends Conf-BPDU with TCA = 1 while receiving TCN-BPDU
4. Sends TCN-BPDU (periodicity = hello time)
5. Sends Conf-BPDU with TCA = 1 while receiving TCN-BPDU and with TC=1 for a period of time equal to *ForwardDelay + MaxAge*

Root bridge uses the forwarding table short lifetime during this period

6. Sends Conf-BPDU with TC=1

Bridge 1 uses the forwarding table short lifetime while receiving Conf-BPDU with TC=1

Bridge 2 uses the forwarding table short lifetime while receiving Conf-BPDU with TC=1

Other protocols: PVST+

- Cisco's proprietary versions of the STP.
- It create a different spanning tree for each VLAN:
 - By default, the default spanning tree parameters are used in all spanning trees (i.e., all spanning trees are equal)
 - Configuring different parameters on the different VLANs, different spanning trees be obtained
 - The spanning tree of each VLAN runs only on the links belonging to the VLAN

```
> Frame 4: 68 bytes on wire (544 bits), 68 bytes captured (544 bits) on interface -, id 0
> Ethernet II, Src: c2:02:4b:4c:f1:02 (c2:02:4b:4c:f1:02), Dst: PVST+ (01:00:0c:cc:cc:cd)
v 802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 2
    000. .... = Priority: Best Effort (default) (0)
    ...0 .... = DEI: Ineligible
    .... 0000 0000 0010 = ID: 2
    Length: 50
> Logical-Link Control
v Spanning Tree Protocol
    Protocol Identifier: Spanning Tree Protocol (0x0000)
    Protocol Version Identifier: Spanning Tree (0)
    BPDU Type: Configuration (0x00)
    > BPDU flags: 0x00
    > Root Identifier: 32768 / 0 / c2:01:1a:70:00:01
    Root Path Cost: 5
    > Bridge Identifier: 32768 / 0 / c2:02:50:4c:00:00
    Port identifier: 0x802b
    Message Age: 1
    Max Age: 20
    Hello Time: 2
    Forward Delay: 15
    v Originating VLAN (PVID): 2
        Type: Originating VLAN (0x0000)
        Length: 2
        Originating VLAN: 2
```

Example of a (PVST+) BPDU:

- Dest. MAC: 01:00:0c:cc:cc:cd
- 802.1Q tag: VLAN 2

Other protocols: RSTP

- IEEE 802.1w Rapid Spanning Tree Protocol

- Extension of IEEE 802.1D
- Speeds up the convergence time of the Spanning Tree in case of topology changes

- There are only three port states in RSTP that correspond to the three possible operational states.
- Adds two additional port roles to a port when in blocking state
 - Alternate port: possible alternative Root port.
 - Backup port: possible alternative Designated port.
- Adds a negotiated mechanism between switches.
 - Uses the reserved bits in the Conf-BPDU.

Conf (Configuration) BPDU

protocol identifier		
version		
message type = 0		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		