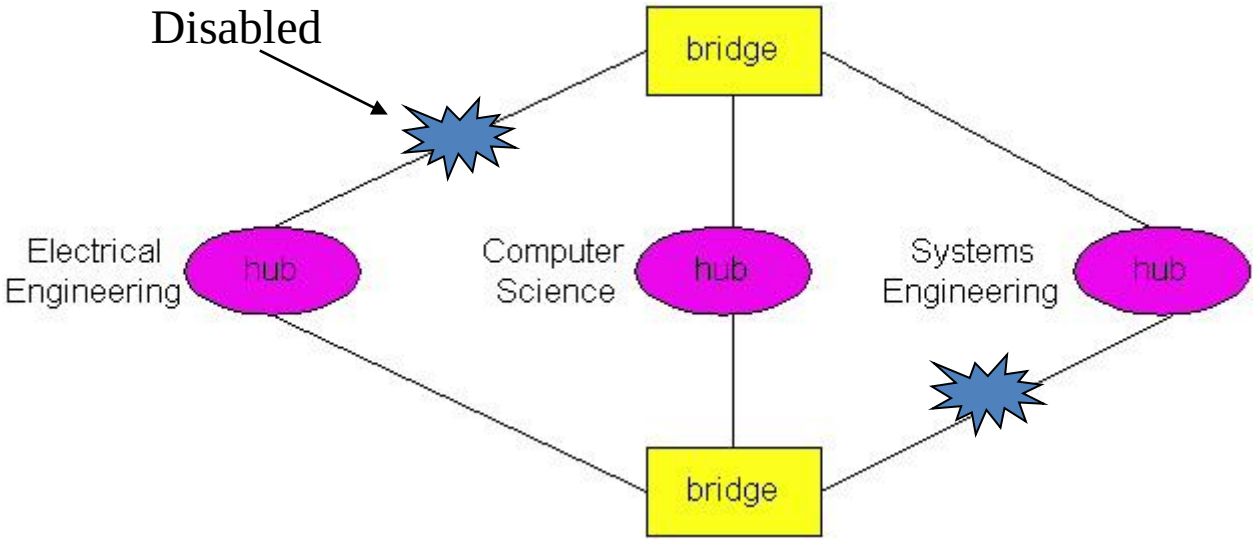
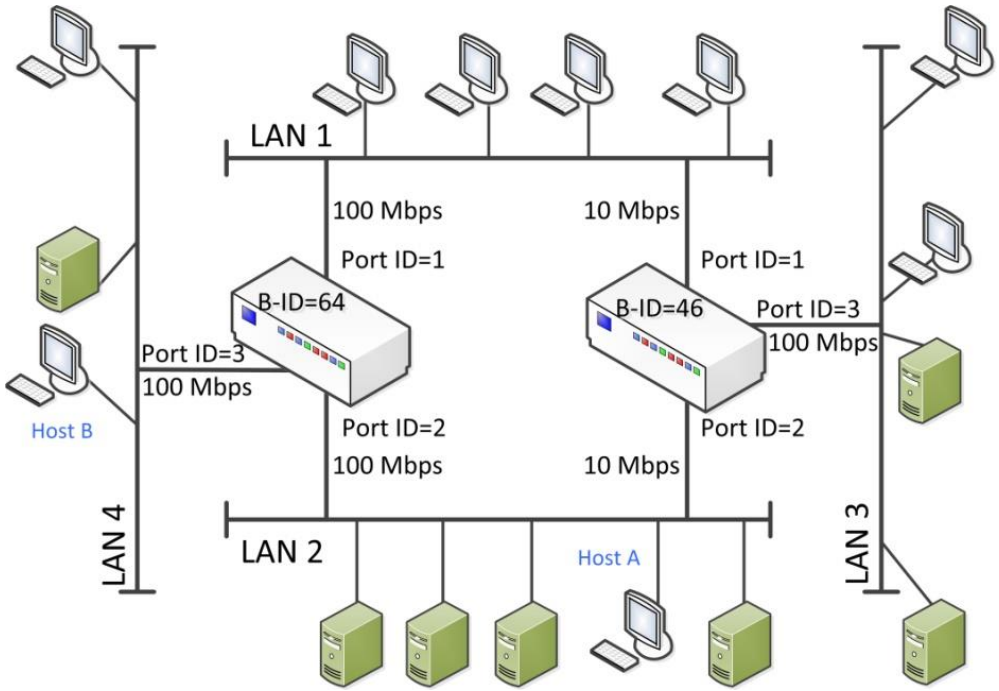
Bridges Spanning Tree

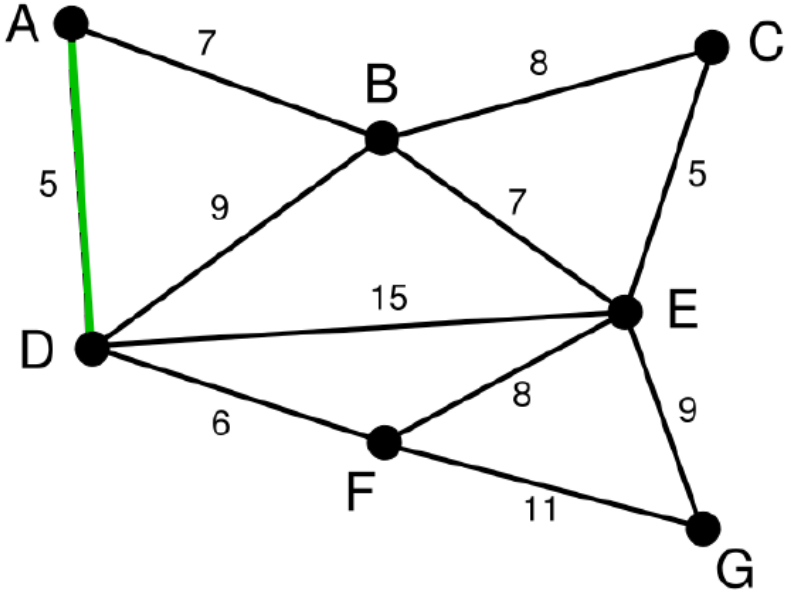
- for increased reliability, desirable to have redundant, alternate paths from source to dest
- with multiple simultaneous paths, cycles result - bridges may multiply and forward frame forever
- solution: organize bridges in a spanning tree by disabling subset of interfaces

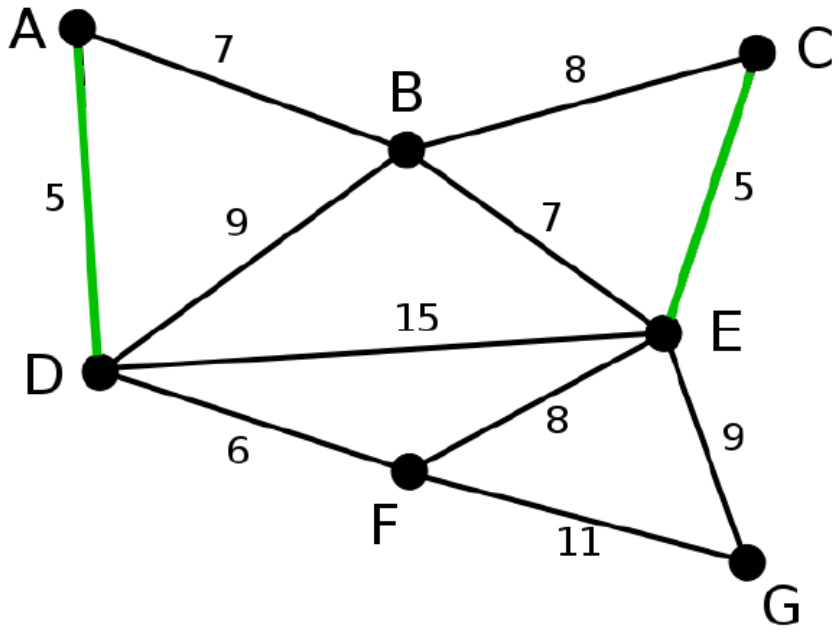


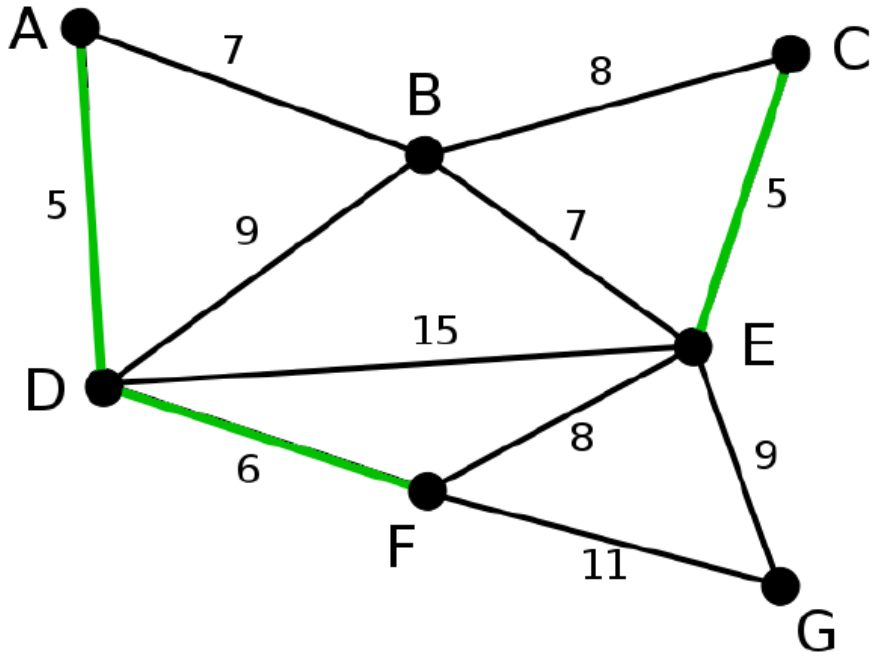
Ποιο είναι το πρόβλημα με τις Γέφυρες

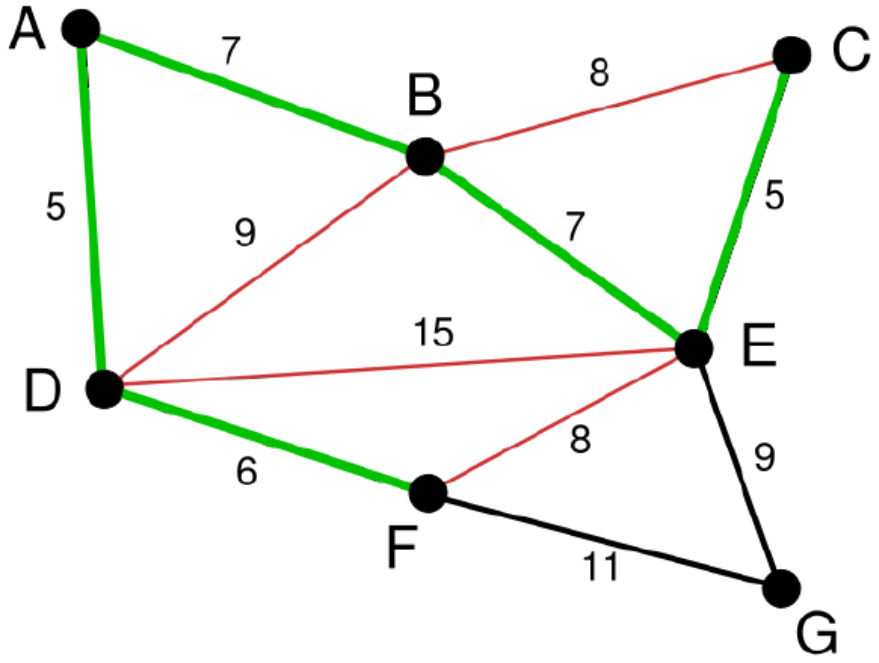


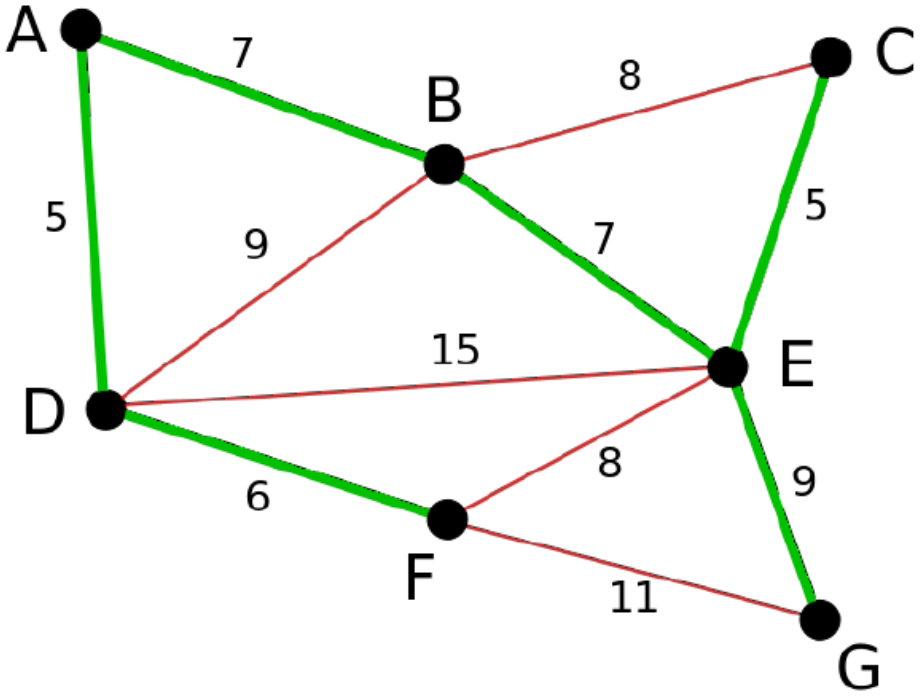
Spanning Tree Algorithm
(Kruskal)




A
B
C
D
E
F
G
7
8
5
9
7
5
15
8
9
6
11

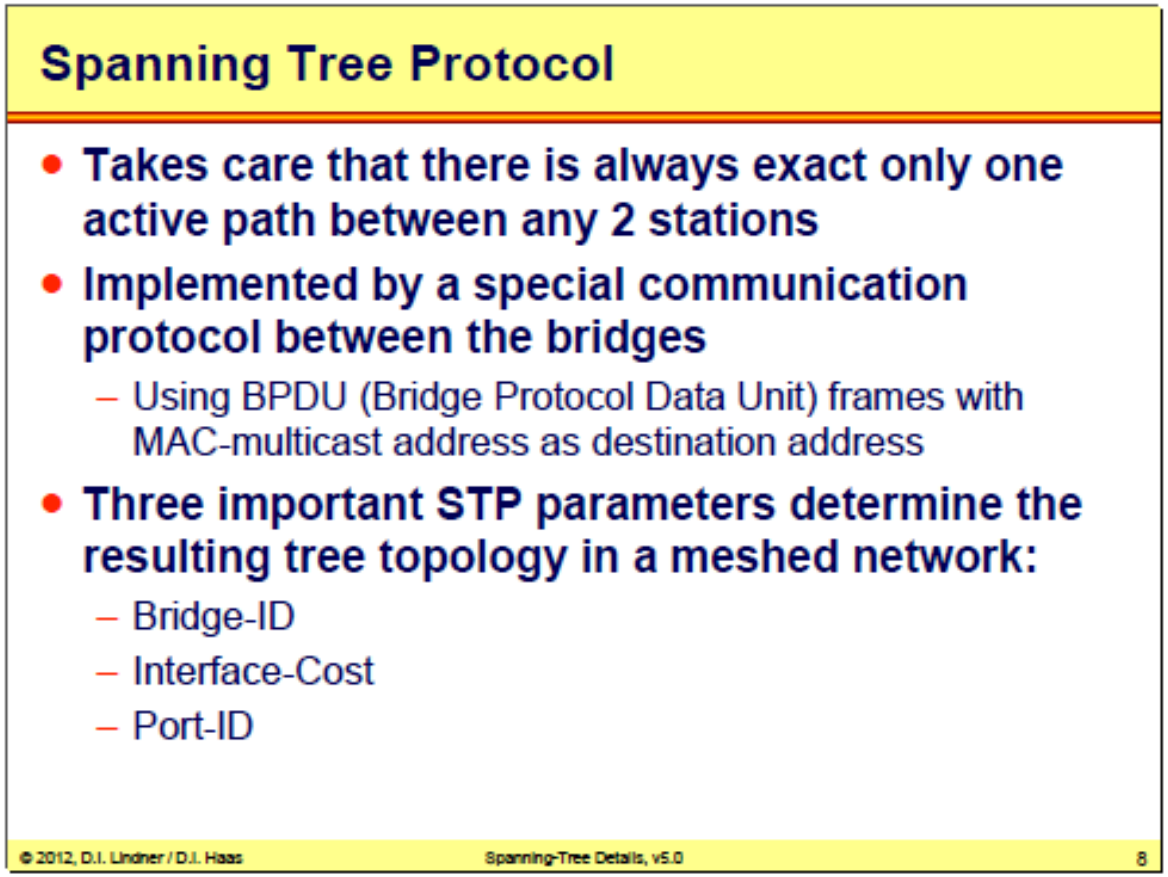

A
B
C
D
E
F
G
7
8
5
9
7
5
15
8
9
6
11


A
B
C
D
E
F
G
7
8
5
9
7
5
15
8
9
6
11



This is the Spanning Tree!!!

IEEE 802.1d Spanning Tree Protocol (STP)



What do we need for STP to work? First of all this protocol needs a special messaging means, realized in so-called **Bridge Protocol Data Units (BPDUs)**. BPDUs are simple messages contained in Ethernet frames containing several parameters described in the next pages.

By D. I. Linder and D. I. Haas
Austria

Parameters for STP 1

- **Bridge Identifier (Bridge ID)**
  - Consists of a priority number and the MAC-address of a bridge
    - Bridge-ID = Priority# (2 Byte) + MAC# (6 Byte)
  - Priority number may be configured by the network administrator
    - Default value is 32768
  - Lowest Bridge ID has highest priority
  - If you keep default values
    - The bridge with the lowest MAC address will have the highest priority

Each bridge is assigned one unique **Bridge-ID** which is a combination of a 16 bit priority number and the lowest MAC address found on any port on this bridge. The Bridge-ID is determined automatically using the default priority 32768. Note: Although bridge will not be seen by end systems, for bridge communication and management purposes a bridge will listen to one or more dedicated (BIA) MAC addresses. Typically, the lowest MAC-address is used for that. The Bridge-ID is used by STP algorithm to determine root bridge and as tie-breaker to when determine the designated port.

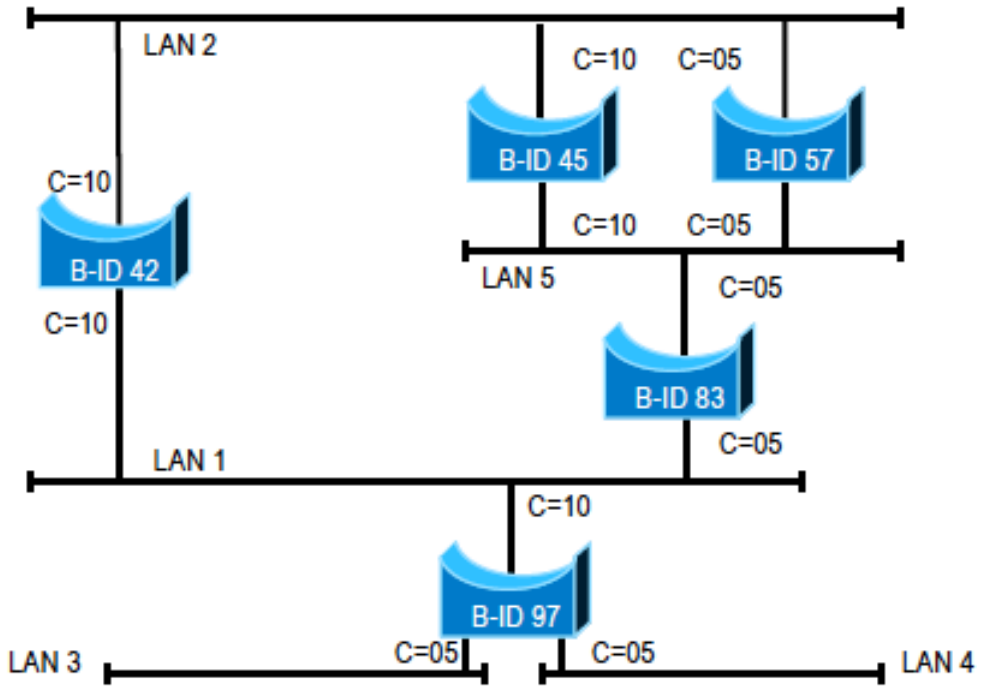
Parameters for STP 2

- **Port Cost (C)**
  - Costs in order to access local interface
  - Inverse proportional to the transmission rate
  - Default cost = 1000 / transmission rate in Mbit/s
    - With occurrence of 1Gbit/s Ethernet the rule was slightly adapted
    - May be configured to a different value by the network administrator
- **Port Identifier (Port ID)**
  - Consists of a priority number and the port number
    - Port-ID = port priority#.port#
    - Default value for port priority is 128
    - Port priority may be configured to a different value by the network administrator

Each port is assigned a **Port Cost**. Again this value is determined automatically using the simple formula Port Cost = 1000 / BW, where BW is the bandwidth in Mbit/s. Of course the Port Cost can be configured manually. Port Cost are used by STP algorithm to calculate Root Path Cost in order to determine the root port and the designated port

Each port is assigned a **Port Identifier.** Only used by STP algorithm as tie-breaker if the same Bridge-ID and the same Path Cost is received on multiple ports.

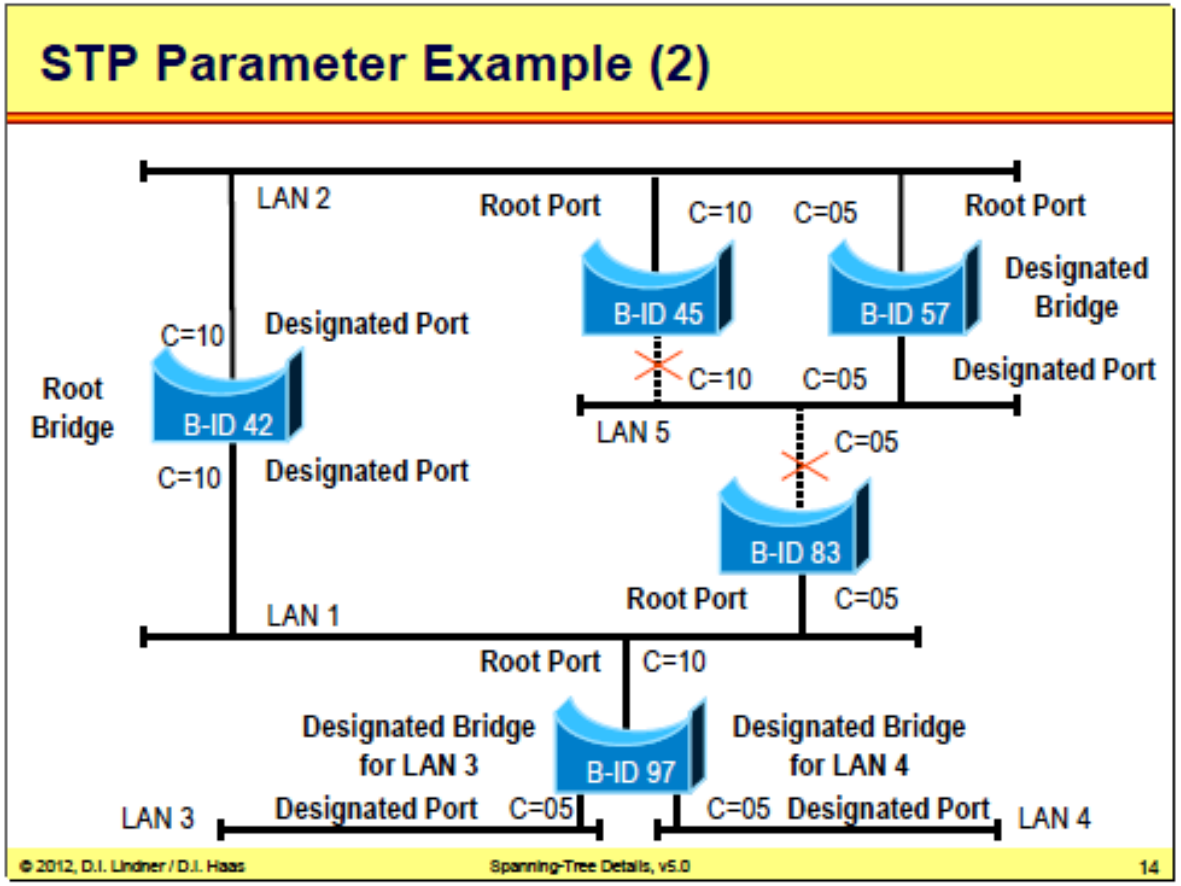
STP Parameter Example (1)



Spanning Tree Algorithm Summary

- **Select the root bridge**
  - Bridge with the lowest Bridge Identifier
- **Select the root ports**
  - By computation of the shortest path from any non-root bridge to the root bridge
  - Root port points to the shortest path towards the root
- **Select one designated bridge for every LAN segment which can be reached by more than one bridge**
  - Bridge with lowest root path costs on the root port side
  - Corresponding port on other side is called designated port
- **Set the designated and root ports in forwarding state**
- **Set all other ports in blocking state**

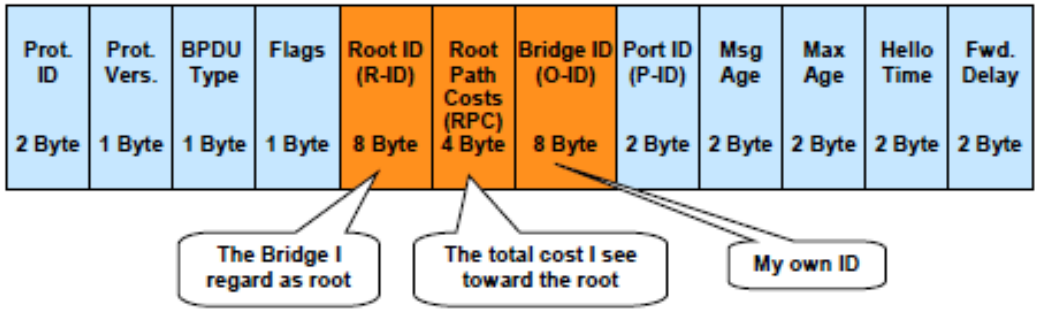
These creates single paths from the root to all leaves (LAN segments) of the network.



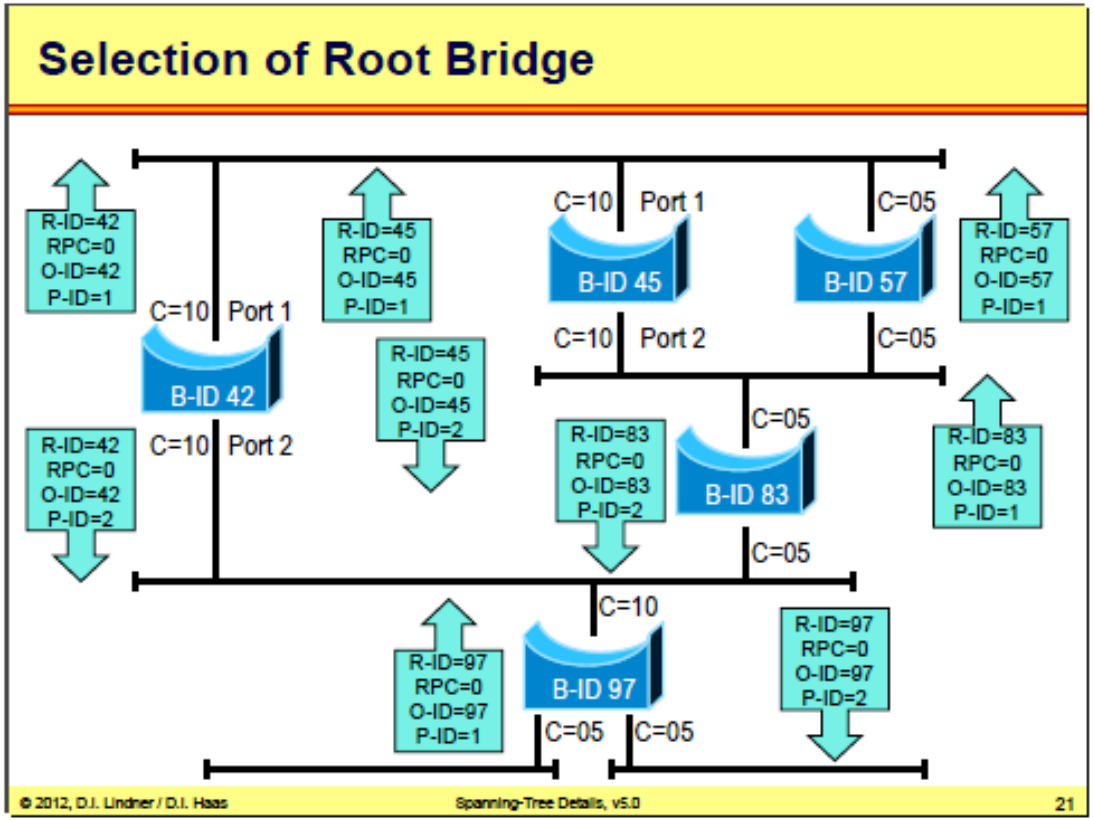
Data rate	STP Cost (802.1D-1998)	RSTP Cost (802.1D-2004 / 802.1w)
4 Mbit/s	250	5,000,000
10 Mbit/s	100	2,000,000
16 Mbit/s	62	1,250,000
100 Mbit/s	19	200,000
1 Gbit/s	4	20,000
2 Gbit/s	3	10,000
10 Gbit/s	2	2,000

BPDU Format

- **Each bridge sends periodically BPDUs carried in Ethernet multicast frames**
  - Hello time default: 2 seconds
- **Contains all information necessary for building Spanning Tree**



Just for your interest, the above picture shows the structure of BPDUs. You see, there is no magic in here, and the protocol is very simple. There are no complicated protocol procedures. BPDUs are sent periodically and contain all involved parameters. Each bridge enters its own "opinion" there or adds its root path costs to the appropriate field. Note that some parameters are transient and others are not.



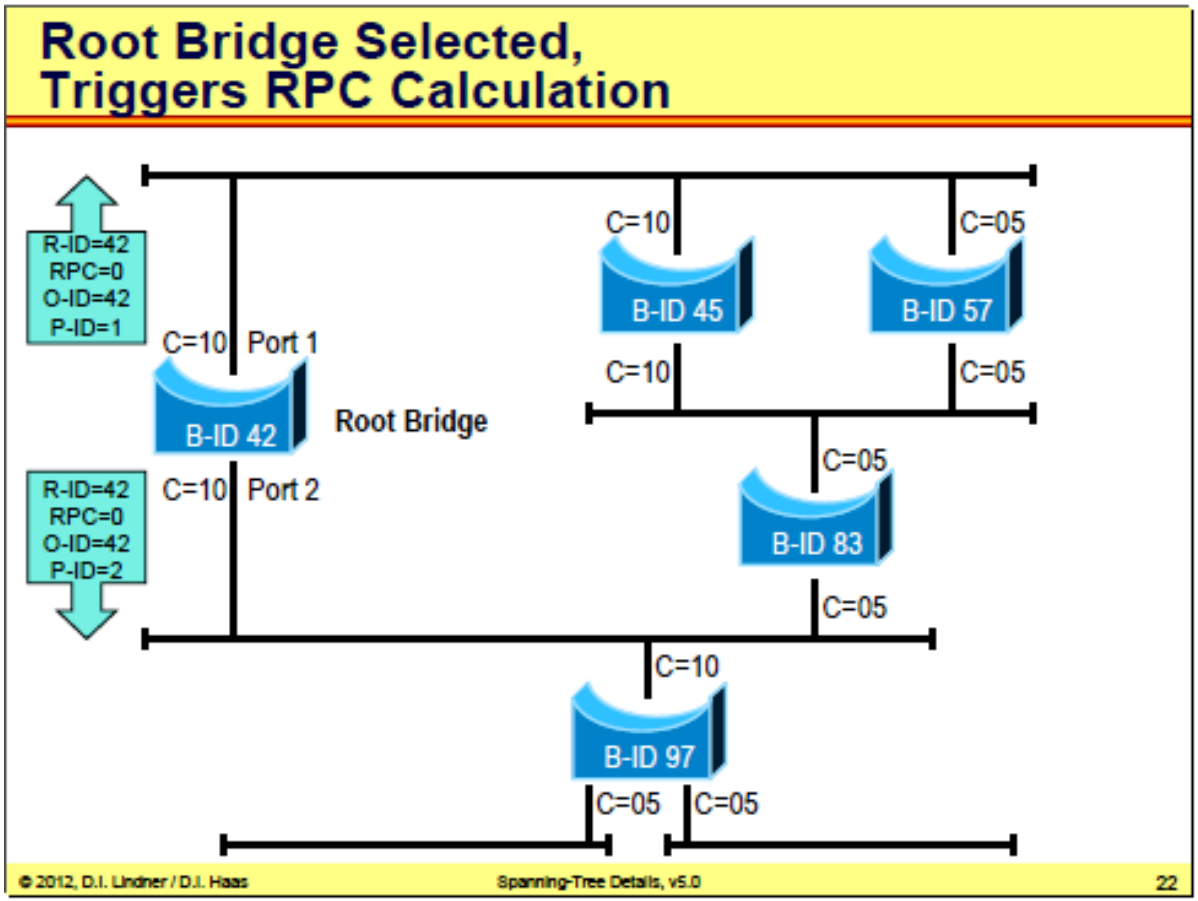
After power up all ports are set in a Blocking State and every bridge tries to become the Root Bridge (RB) of the Spanning Tree by sending Configuration BPDUs.

Blocking state means: End station Ethernet frames are not received and forwarded on such a port but BDPU frames can still be received, manipulated by the bridge and transmitted on such port. BPDU frames are actually filtered based on the well-known multicast address and are given to the CPU of the bridge.

Using such Configuration BPDUs, a bridge tells, which bridge actually is seen as RB, which path costs exist to this RB (Root Path Cost) and its own Bridge ID and Port ID.

Βήματα του Αλγορίθμου		Γέφυρα 42		Γέφυρα 45		Γέφυρα 57		Γέφυρα 83		Γέφυρα 97		
		Port 1	Port 2	Port 1	Port 2	Port 1	Port 2	Port 1	Port 2	Port 1	Port 2	Port 3
0	Αποστολή	42,42,0,1	42,42,0,2	45,45,0,1	45,45,0,2	57,57,0,1	57,57,0,2	83,83,0,1	83,83,0,2	97,97,0,1	97,97,0,2	97,97,0,3
	Λήψη	45,45,0,1 57,57,0,1	83,83,0,2 97,97,0,1	57,57,0,1 42,42,0,1	57,57,0,2 83,83,0,1	42,42,0,1 45,45,0,1	45,45,0,2 83,83,0,1	45,45,0,2 57,57,0,2	42,42,0,2 97,97,0,1	42,42,0,2 83,83,0,2		
	Απόφαση	**Είμαι η ρίζα**		Ρίζα είναι η **42**		Ρίζα είναι η **42**		Ρίζα είναι η **42**		Ρίζα είναι η **42**		
1	Αποστολή	42,42,0,1	42,42,0,2	---	45,**42**,10,2	---	57,**42**,05,2	83,**42**,05,1	---	---	97,**42**,10,2	97,**42**,10,3
	Λήψη	---	---	---	57,42,05,2 83,42,05,1	---	45,42,10,2 83,42,05,1	45,42,10,2 57,42,05,2	---	---	---	---
	Απόφαση			Root Port 1 Blocked Port 2		Root Port 1 Designated Port 2		Root Port 2 Blocked Port 1		Root Port 1 Designated Port 2 Designated Port 3		
2	Αποστολή	42,42,0,1	42,42,0,2	---	---	---	---	---	---	---	---	---

Own bridge-ID, Root bridge-ID, Root Path Cost, Port Number
e.g., 42, 42, 0, 1

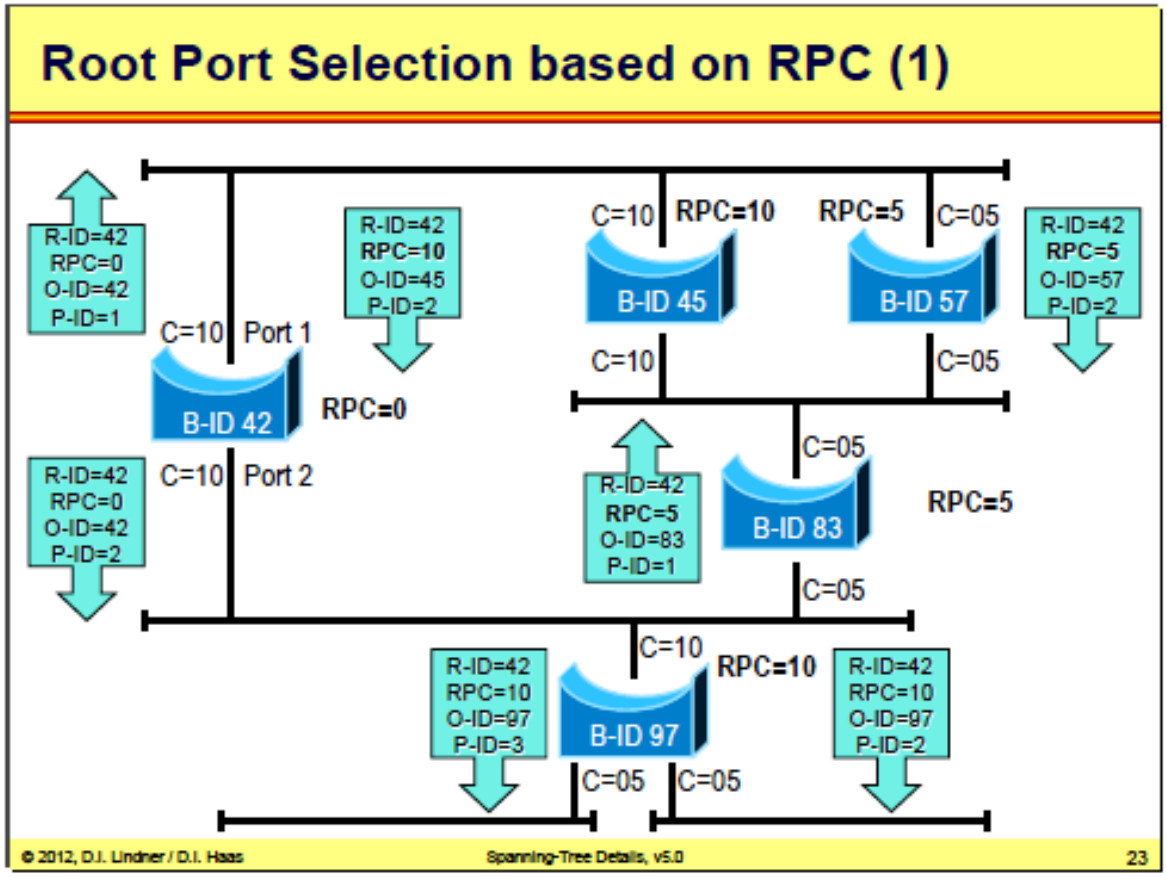


Bridge with the lowest Bridge ID becomes RB. after selection of the RB all sending of Configuration BPDUs are exclusively triggered by the RB. Other bridges just move such BPDUs on after actualizing the corresponding BPDU fileds.

Strategy to determinate the RB :

If bridge receives a Configuration BPDU with *lower* Root Bridge ID as own Bridge ID the bridge stops sending Configuration BPDUs on this port and the received and adapted Configuration BPDU is forwarded to all other ports.

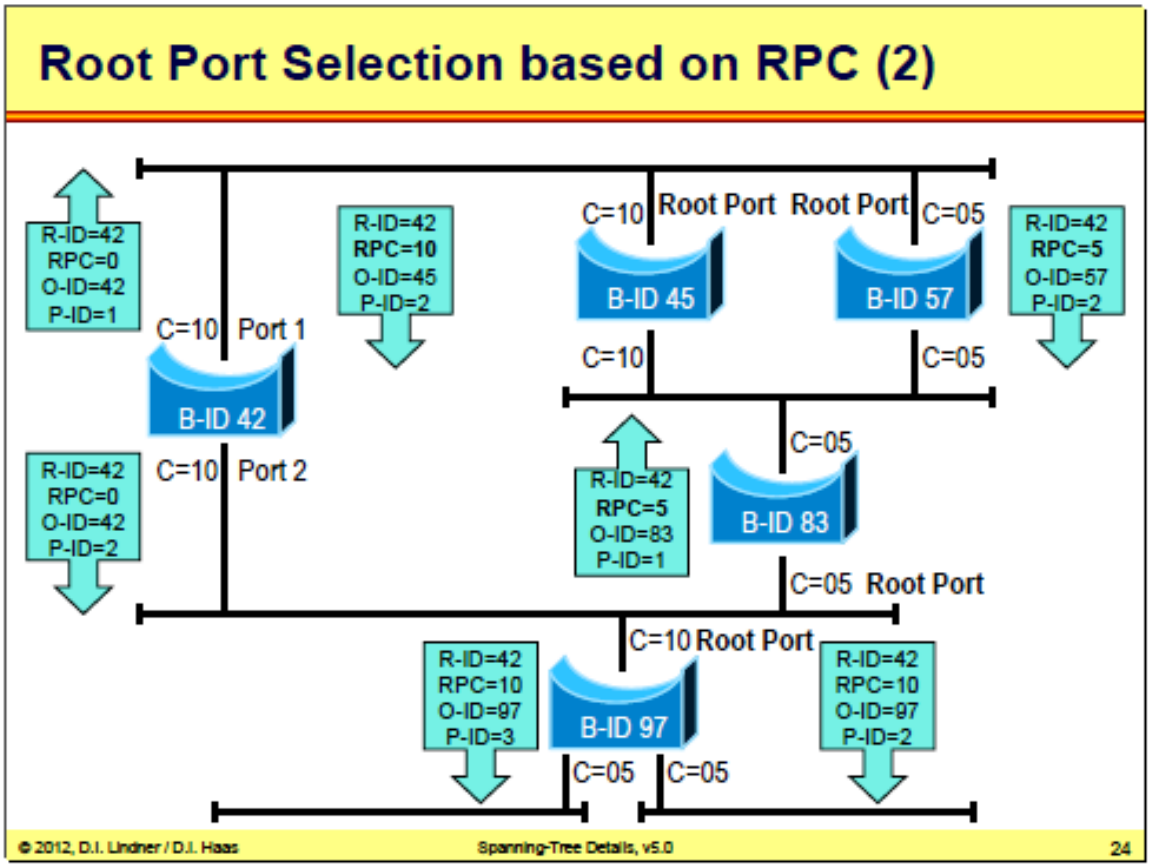
If bridge receives Configuration BPDU with *higher* Root Bridge ID as own Bridge ID the bridge continues sending Configuration BPDUs with own Bridge ID as proposed Root Bridge ID on all ports, the other bridges should give up.



Now, every bridge determines which of its ports has the lowest Root Path Cost. Root Path Cost = sum of all port costs from this bridge to the RB, including port costs of all intermediate bridges. This port becomes the Root Port. In case of equal costs the port ID decides (lower means better).

The principle calculation method: Root Path Cost received in BPDU + port cost of the local port receiving that BPDU.

Similar to Root Bridge selection, a Designated Bridge (DB) is selected for each LAN-segment which is the bridge with the lowest Root Path Cost on its Root Port. In case of equal costs the bridge with the lowest Bridge ID wins again.

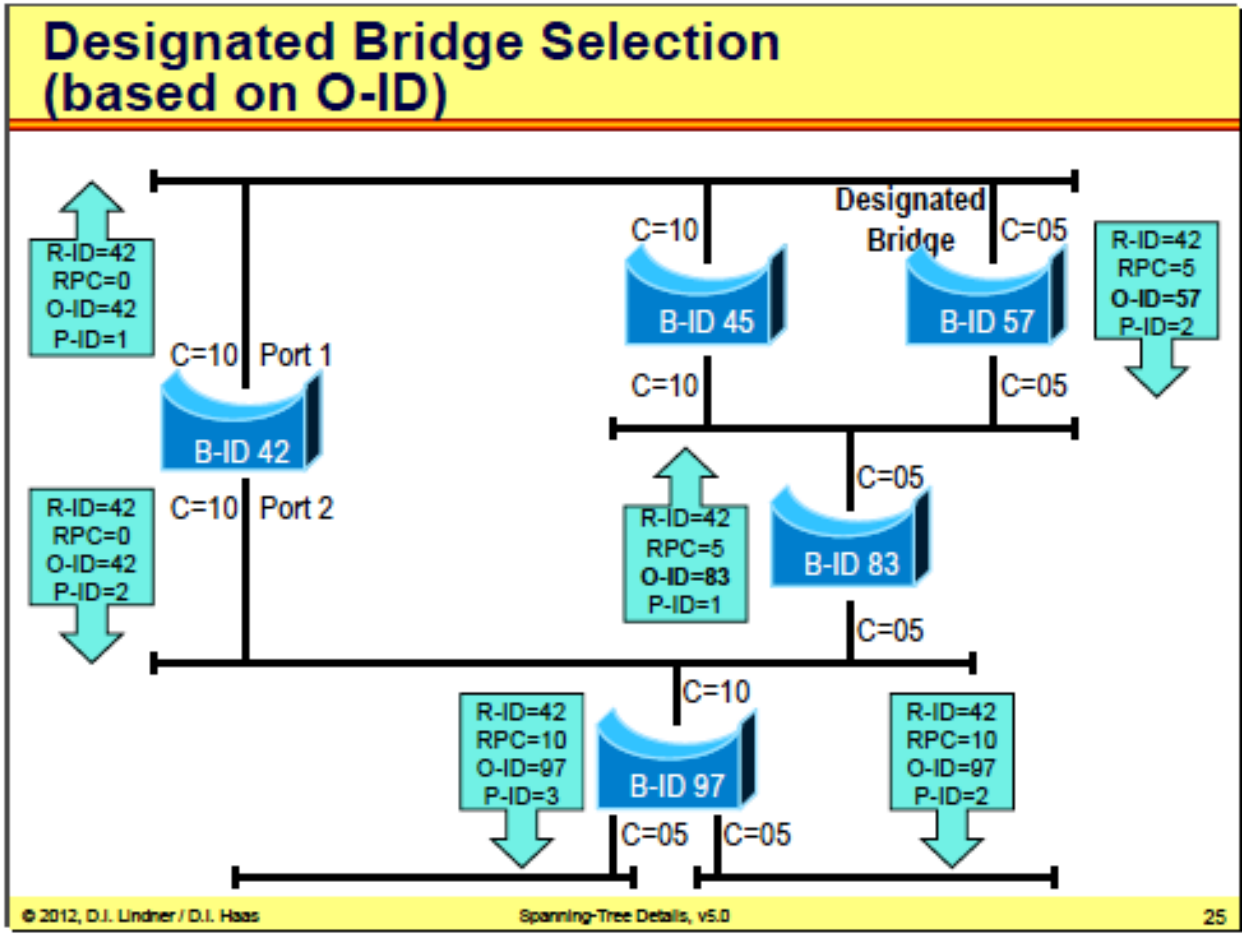


Using the Root Path Cost field in the Configuration BPDU, a bridge indicates its distance to the RB.

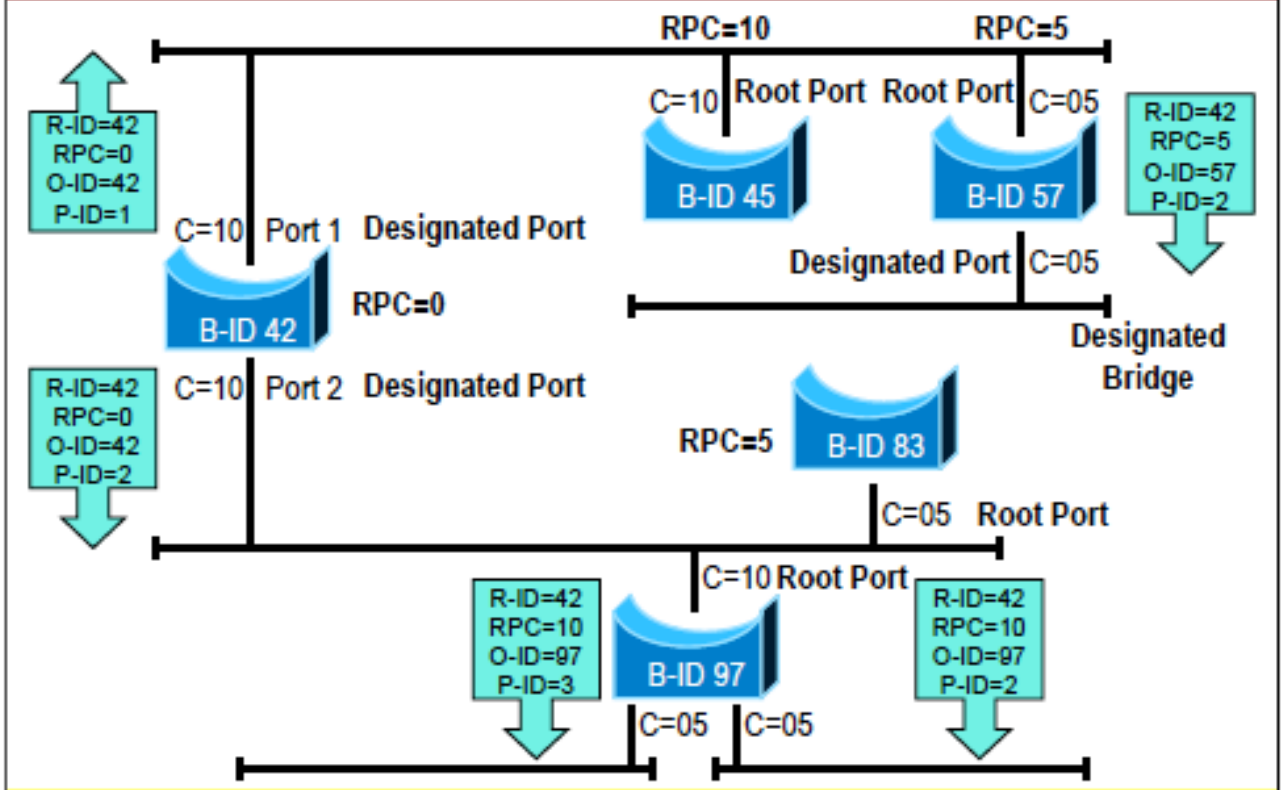
Strategy for decision:

If a bridge receives a Configuration BPDU from a bridge which is closer to the RB, the receiving bridge adds its own port costs to the Configuration BPDU and forwards this message to all other ports.

If a bridge receives a Configuration BPDU from a bridge which is more distant to the RB, the receiving bridge drops the message and sends its own Configuration BPDU on this port containing its own Root Path Cost.


Designated Bridge Selection (based on O-ID)
R-ID=42
RPC=0
O-ID=42
P-ID=1
C=10
Port 1
B-ID 42
C=10
Port 2
R-ID=42
RPC=0
O-ID=42
P-ID=2
C=10
B-ID 45
Designated Bridge
C=05
B-ID 57
R-ID=42
RPC=5
O-ID=57
P-ID=2
C=10
C=05
R-ID=42
RPC=5
O-ID=83
P-ID=1
C=05
B-ID 83
C=05
C=10
R-ID=42
RPC=10
O-ID=97
P-ID=3
B-ID 97
R-ID=42
RPC=10
O-ID=97
P-ID=2
C=05
C=05
© 2012, D.I. Lindner / D.I. Haas
Spanning-Tree Details, v5.0
25

Final Topology



Procedure Parameters Summary:

Root Bridge -> lowest Bridge ID.

Root Ports via Root Path Costs -> which sum of costs contained in the Configuration BPDU and the receiving interface Port Costs.

Designated Bridge -> lowest Root Path Costs for a given LAN segment.

Root switch has only Designated Ports, all of them are in forwarding state.

Other switches have exactly one Root Port (RP) upstream, zero or more Designated Ports (DP) downstream and zero or more Nondesignated Ports (blocked).

Now every designated bridge declares its ports as designated ports and puts them (together with the Root Port) in the Forwarding State.

All other bridges keep their non-RP and non-DP ports in the Blocking State.

From this moment on, the normal network operation is possible and there is only one path between any two arbitrary end systems.

Redundant links remain in active stand-by mode. If root port fails, other root port becomes active. Still it is reasonable to establish parallel paths in a switched network in order to utilize this redundancy in an event of failure. The STP automatically activates redundant paths if the active path is broken. Note that BPDUs are always sent or received on blocking ports. Note that (very-) low price switches might not support the STP and should not be used in high performance and redundant configurations.