

# CCNA 3 Module 7



Cabrillo College

## Spanning Tree Protocol

# Learner Outcomes

After completing this module, students will be able to:

- Define redundancy and its importance in networking
- Describe the key elements of a redundant networking topology
- Define broadcast storms and describe their impact on switched networks
- Define multiple frame transmissions and describe their impact on switched networks
- Identify causes and results of MAC address database instability

# Learner Outcomes Continued

- Identify the benefits and risks of a redundant topology
- Describe the role of spanning tree in a redundant-path switched network
- Identify the key elements of spanning tree operation
- Describe the process for root bridge election
- List the spanning-tree states in order
- Compare Spanning-Tree Protocol and Rapid Spanning-Tree Protocol

# Redundancy

- Redundancy in a network is extremely important because redundancy allows networks to be fault tolerant. Redundant topologies protect against network downtime due to a failure of a single link, port, or networking device. Network engineers are often required to make difficult decisions, balancing the cost of redundancy with the need for network availability.

# Downtime not an Option

- Many companies and organizations increasingly rely on computer networks for their operations.
- Access to file servers, databases, the Internet, intranets, and extranets is critical for successful businesses.
- If the network is down, productivity is lost and customers are dissatisfied.
- Reliability
- Fault Tolerance

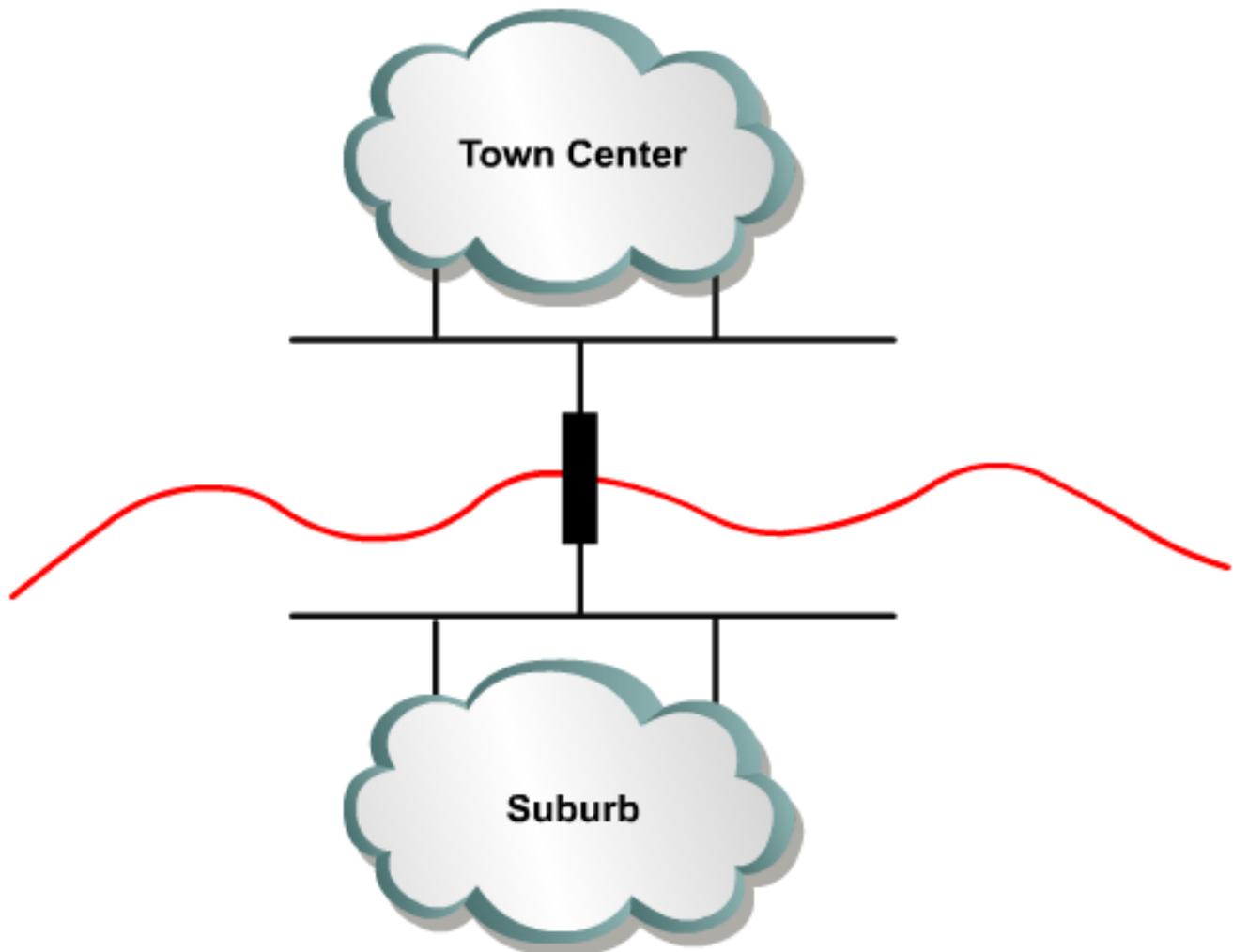
# Redundancy

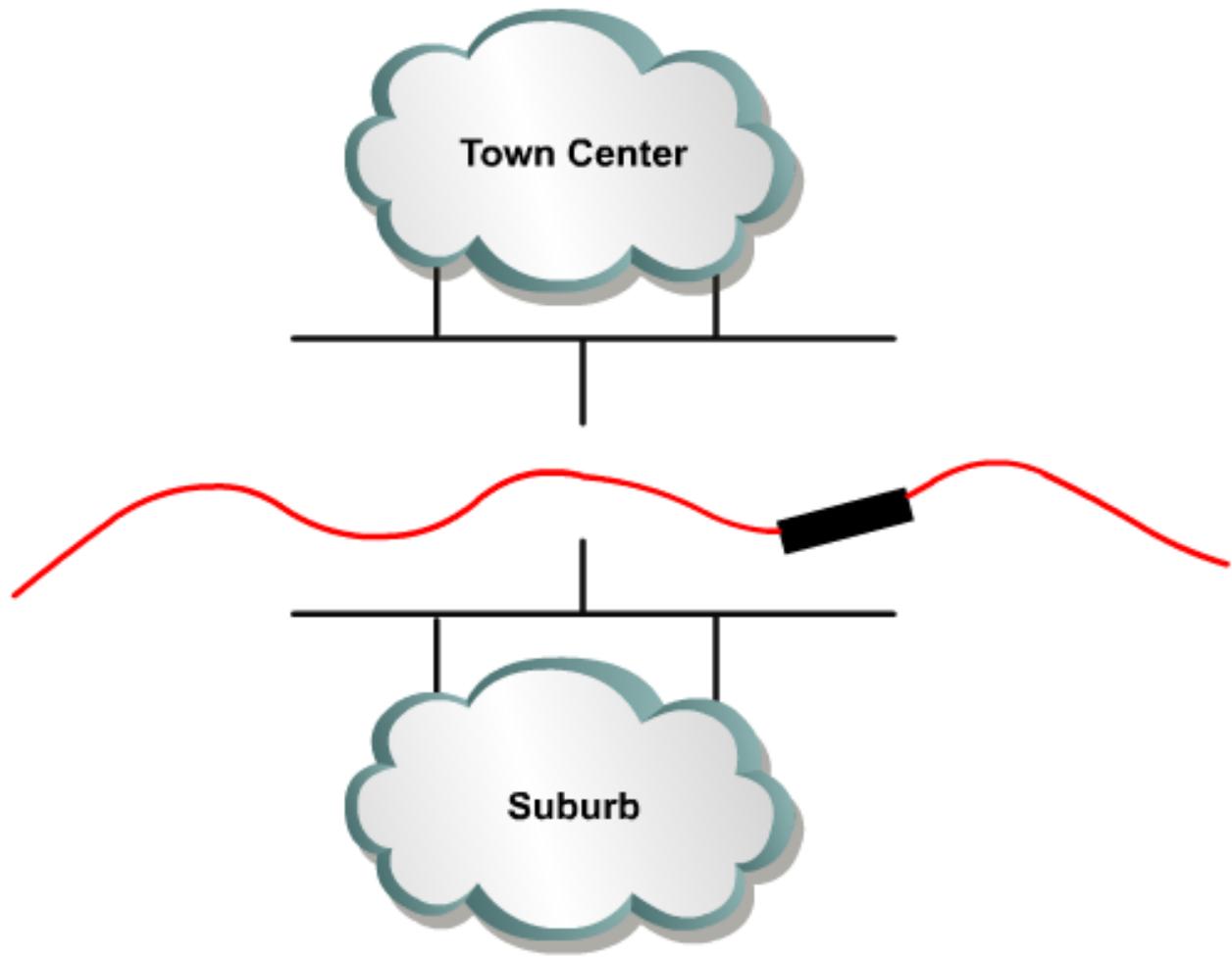


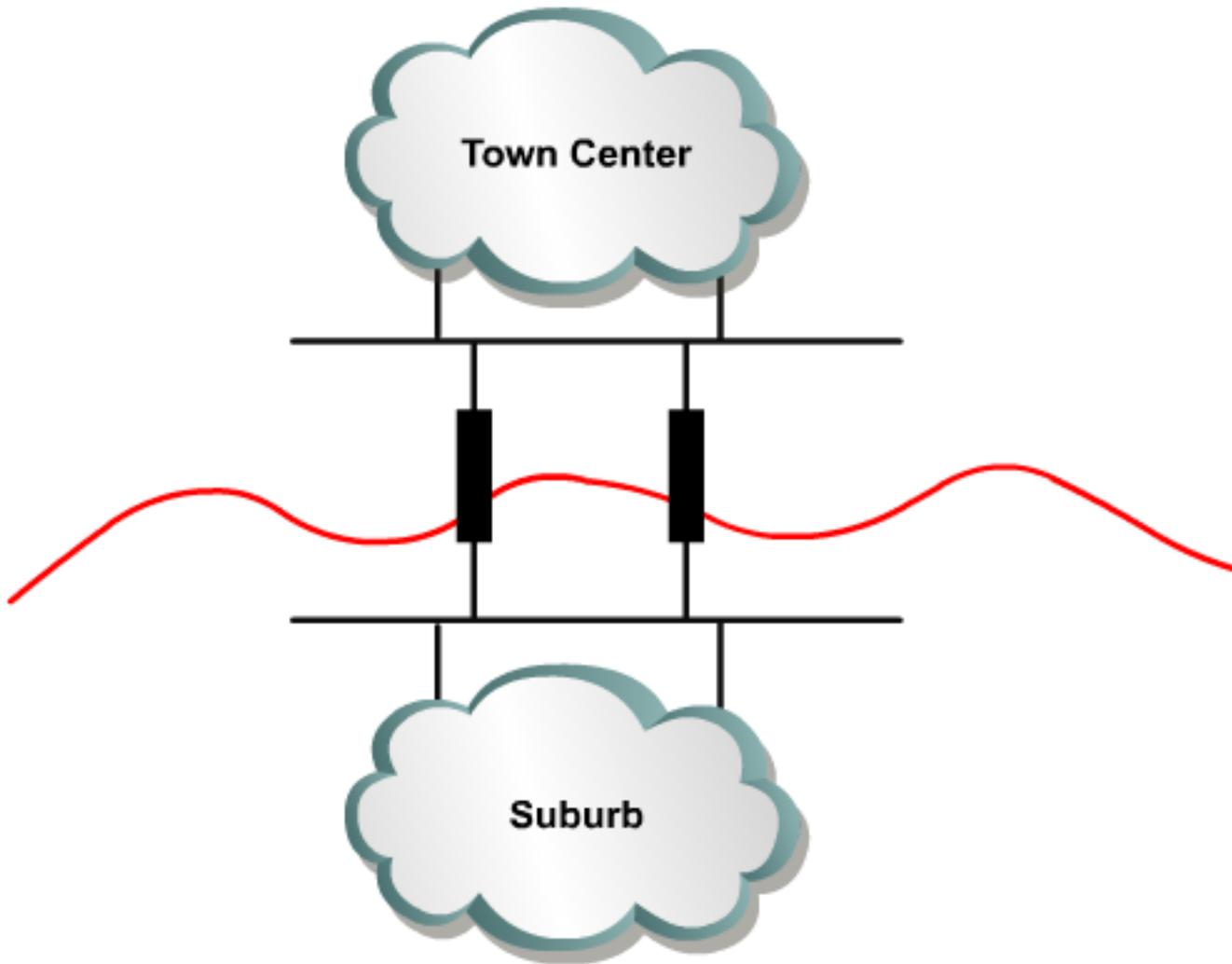
**Kaputt cannot go to work**

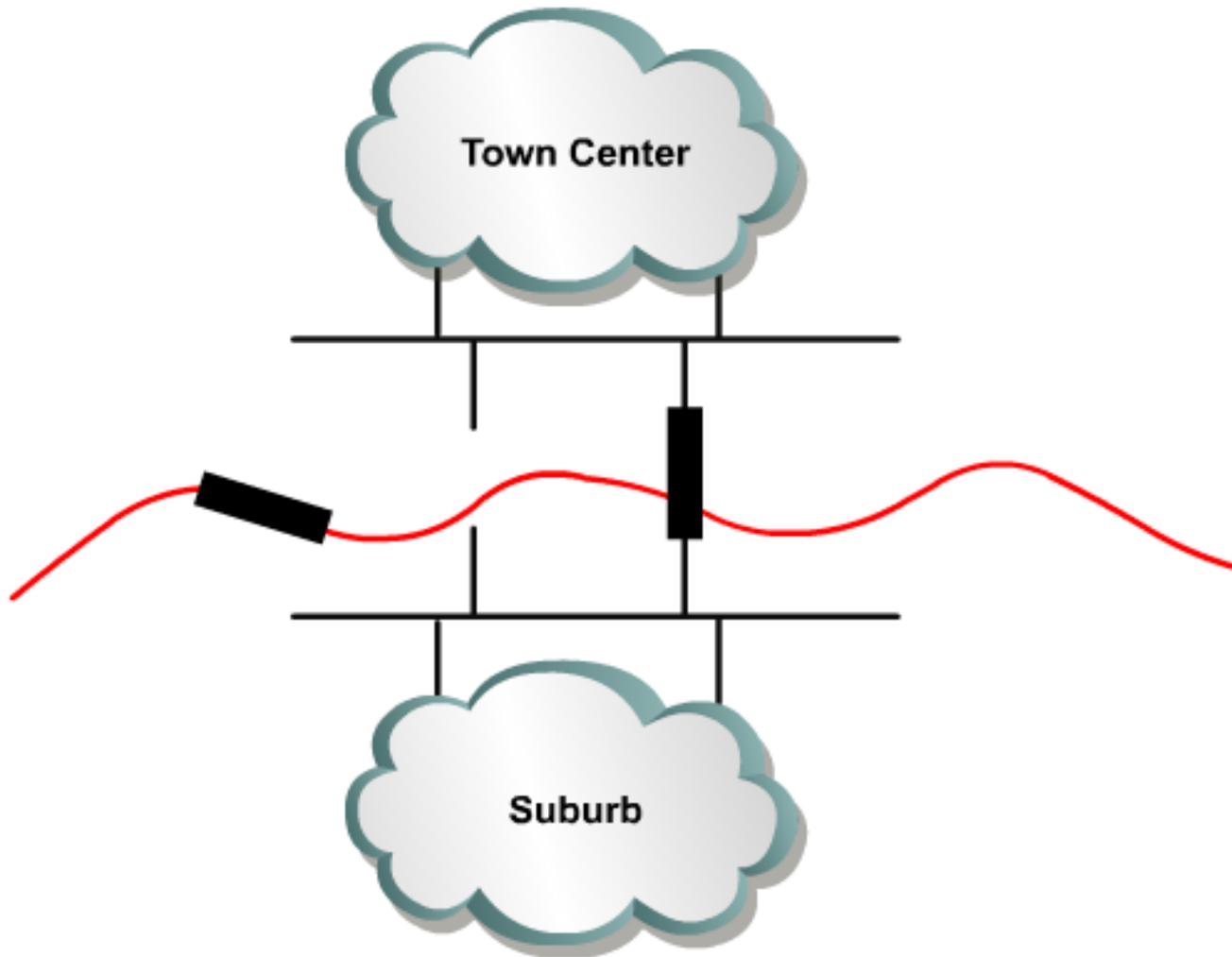
# Plenty of Redundancy



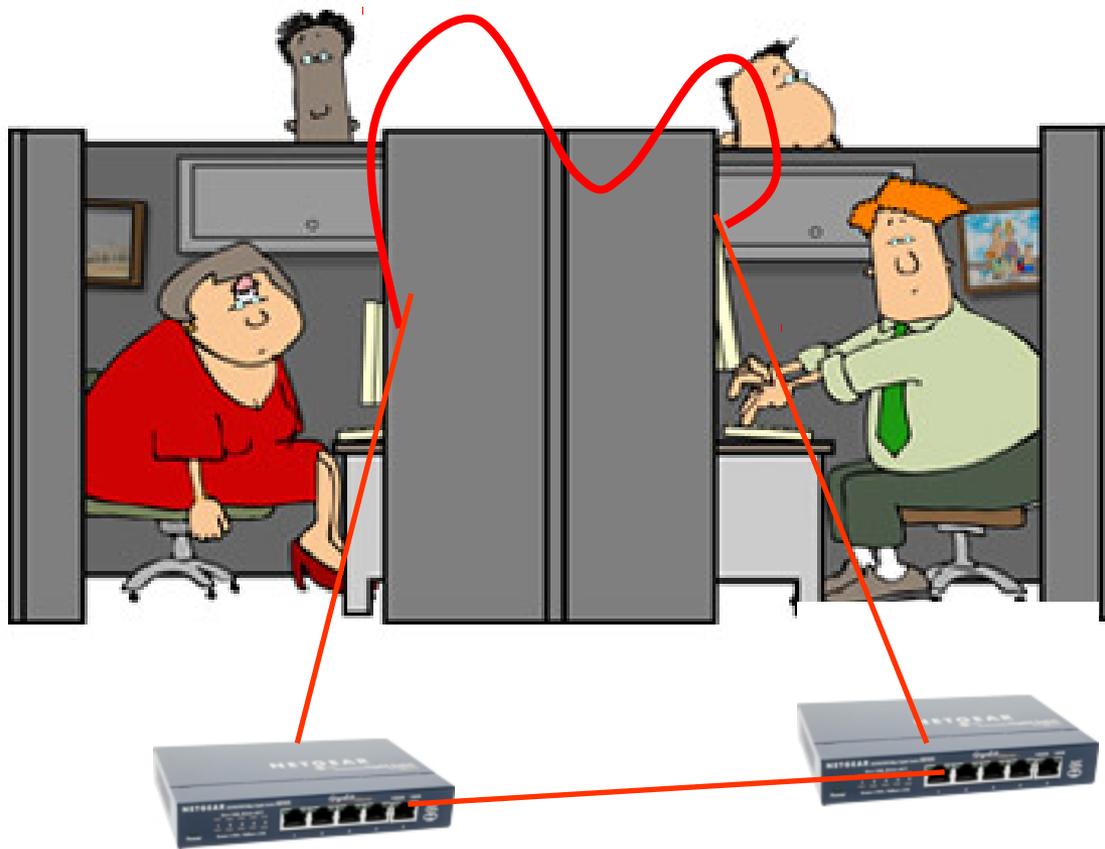




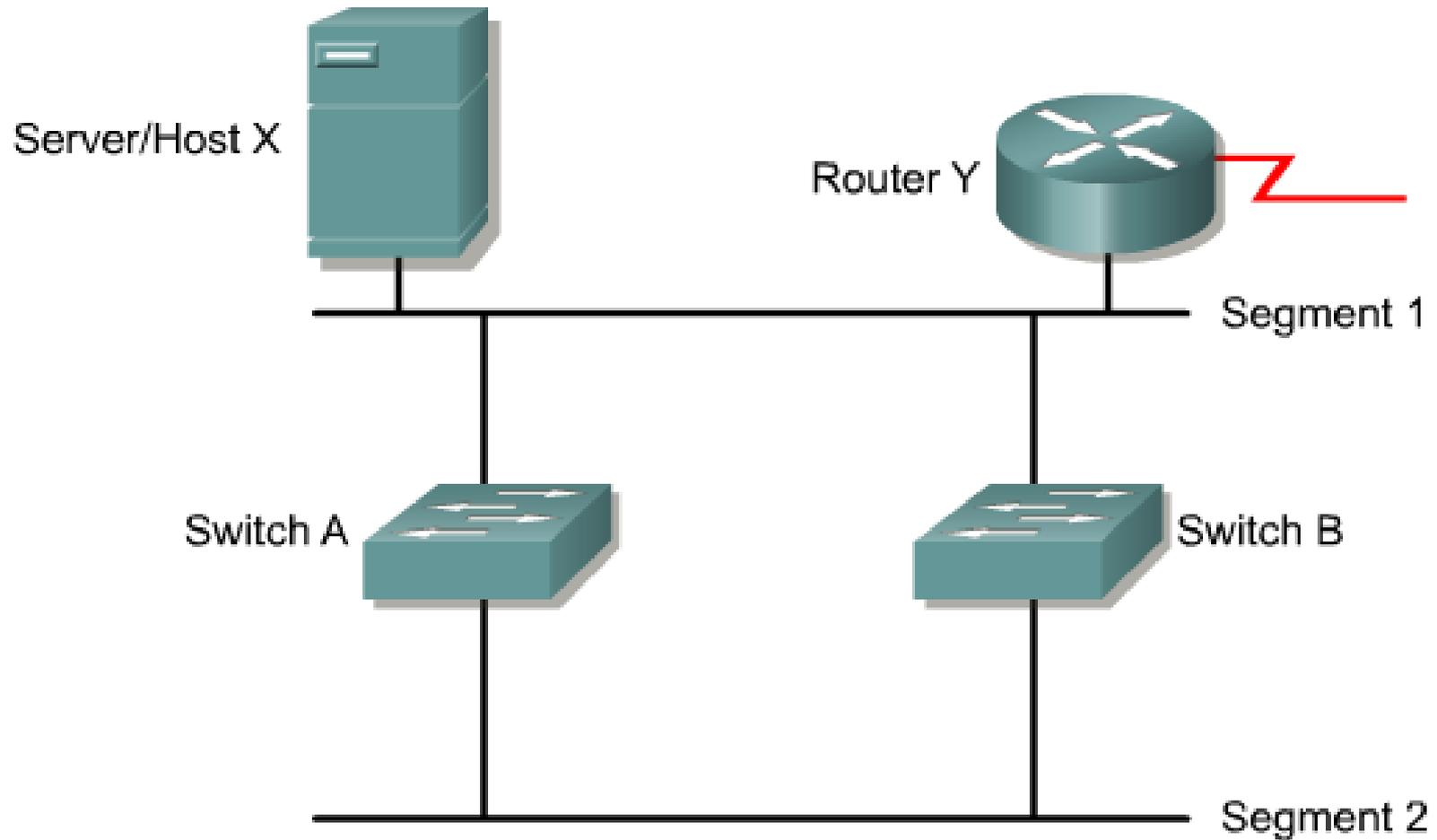




# Switching Loop – User Error

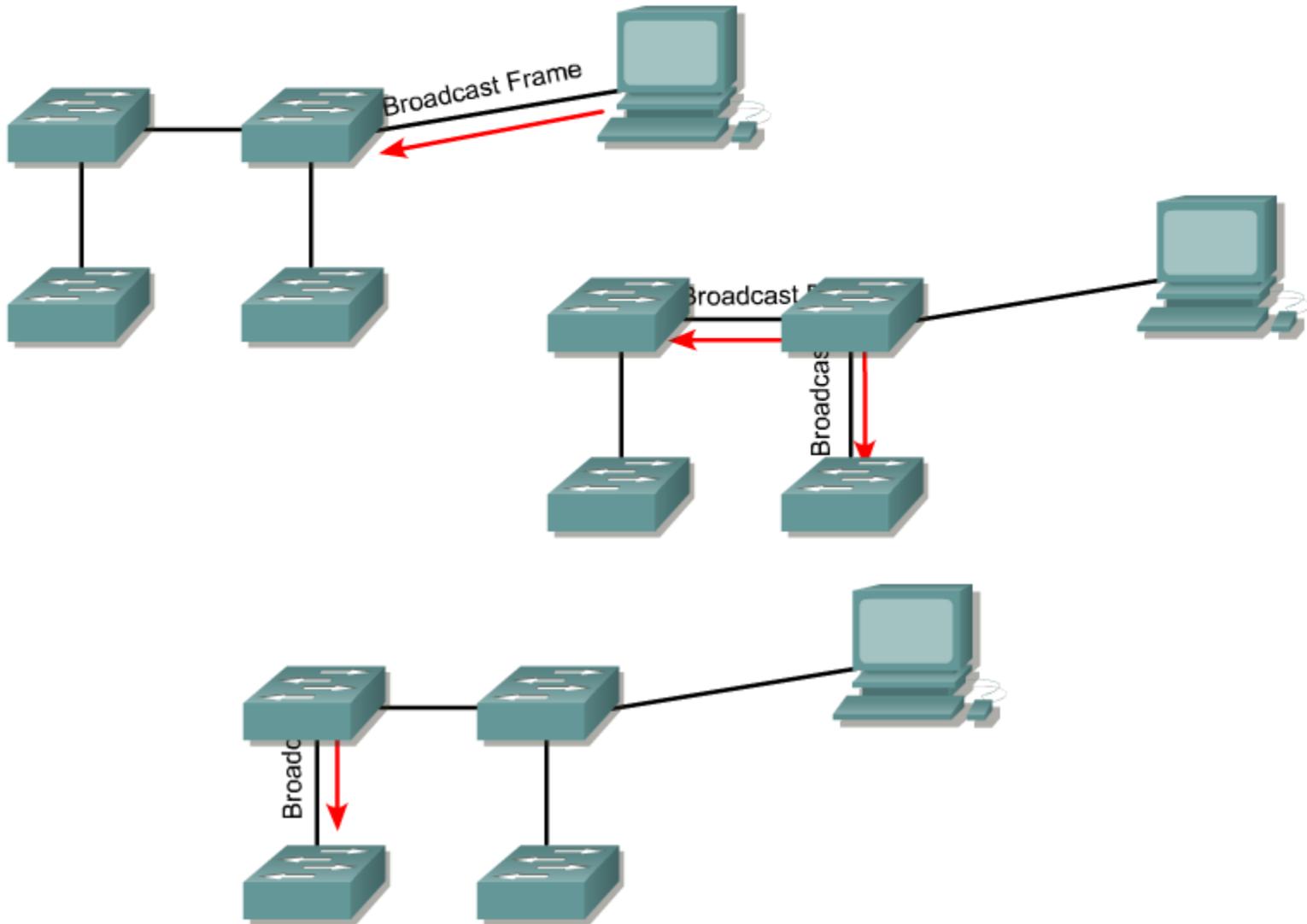


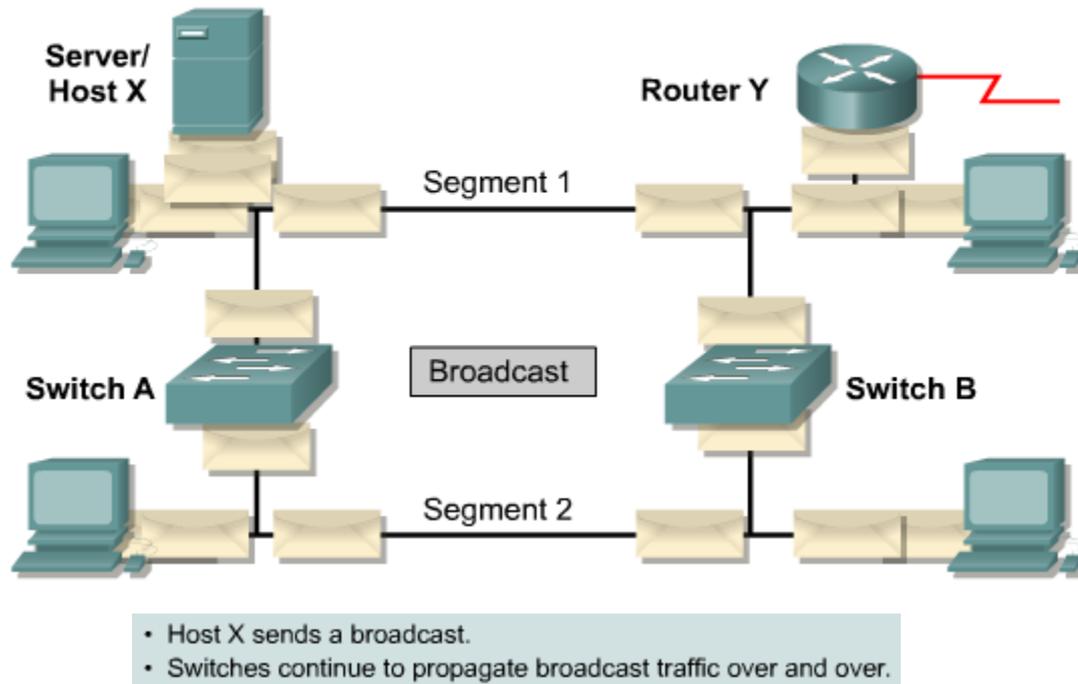
# Redundant Switched Topologies



# Redundancy in Switched Networks

- Switches learn the MAC addresses of devices on their ports so that data can be properly forwarded to the destination. Switches will flood frames for unknown destinations until they learn the MAC addresses of the devices.
- Broadcasts and multicasts are also flooded.
- A redundant switched topology may cause broadcast storms, multiple frame copies, and MAC address table instability problems.

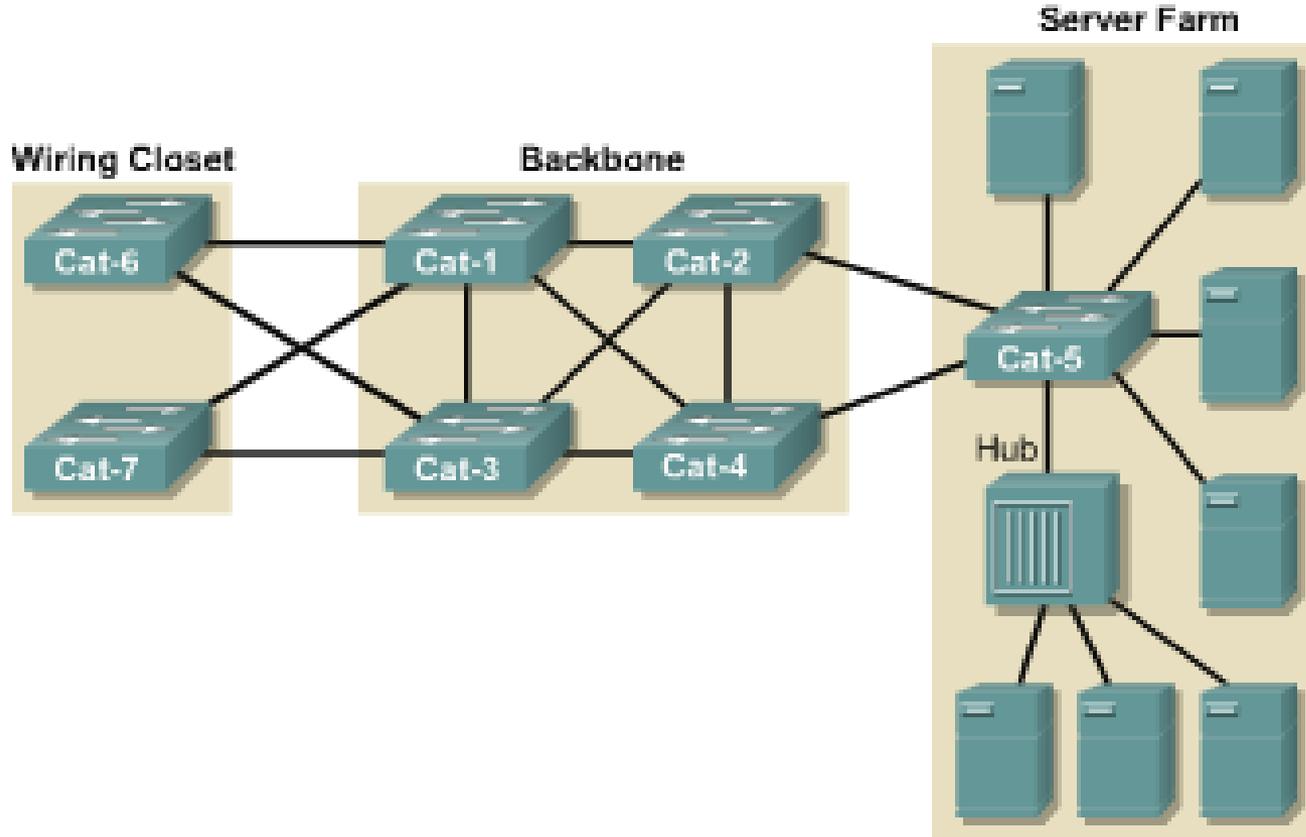




**Broadcast frames causing broadcast storms.**

**WHY?**

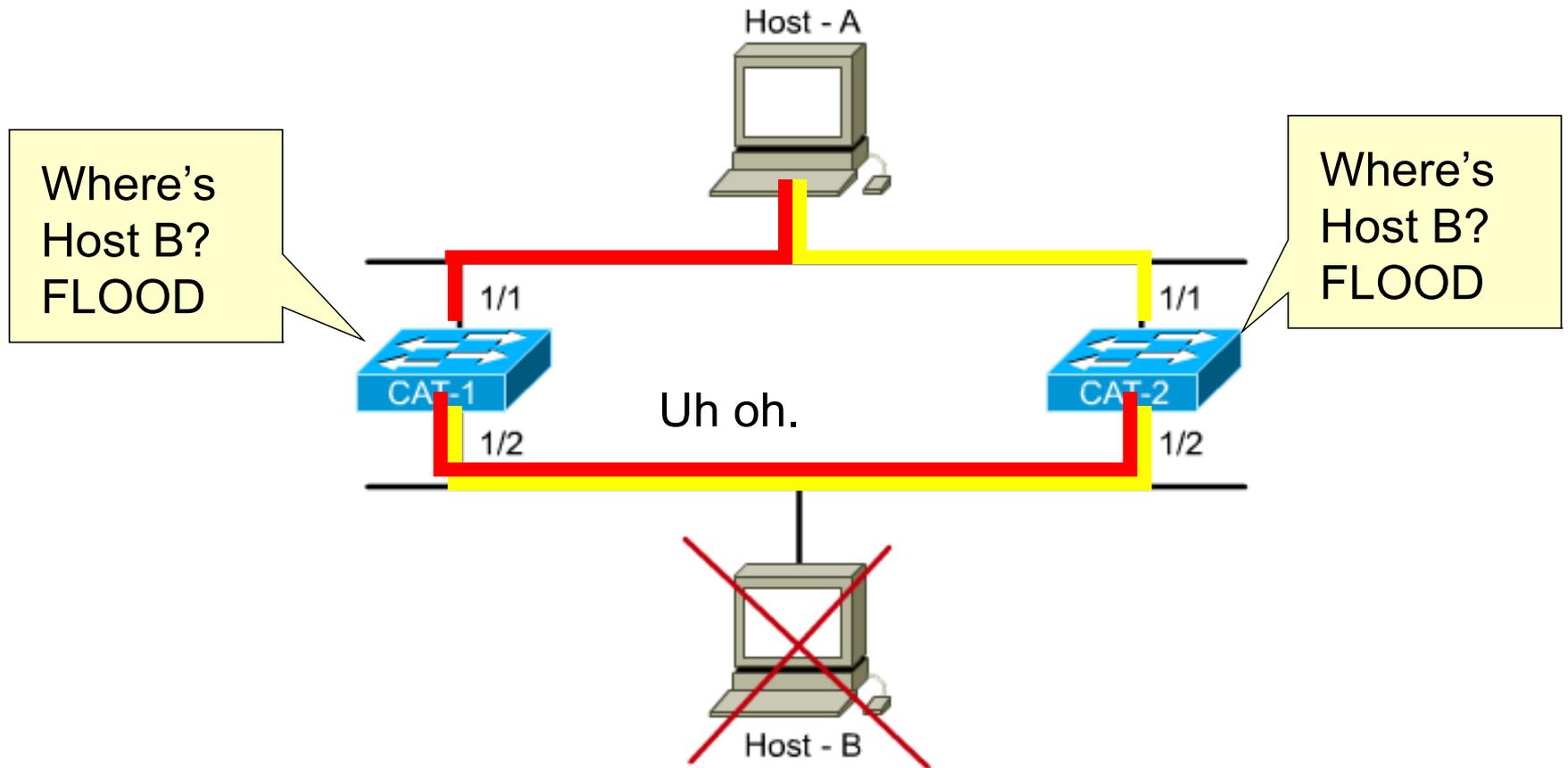
# Redundancy Creates Loops



# L2 Loops

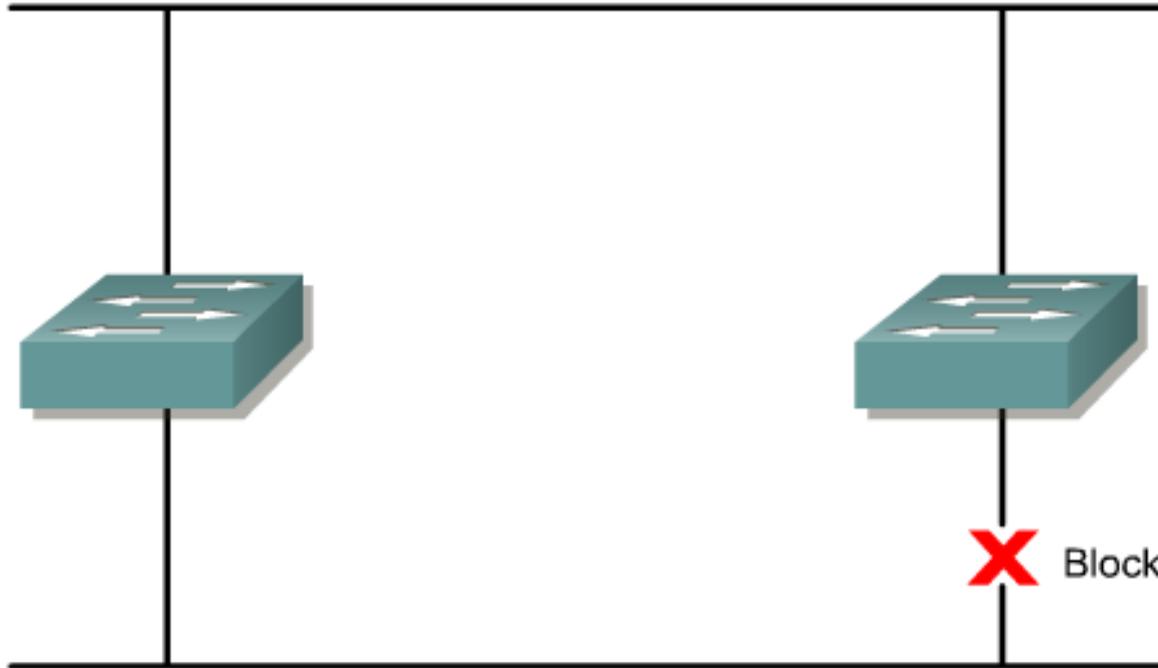
- Broadcasts and Layer 2 loops can be a dangerous combination.
- Ethernet frames have no TTL field
- After an Ethernet frame starts to loop, it will probably continue until someone shuts off one of the switches or breaks a link.
- Physical loops without STP can be disastrous

# L2 Loops





# STP: Redundancy Without Loops



Provides a loop-free, redundant network topology by placing certain ports in the blocking state.

# Spanning Tree Protocol (STP)

- The algorithm used to create this loop free logical topology is the spanning-tree algorithm.
- Allows redundancy without loops
- The loop free logical topology created is called a tree.
- This topology is a star or extended star logical topology, the spanning tree of the network.
- It is a spanning tree because all devices in the network are reachable or spanned.
- Slow to converge, so now there is Rapid Spanning Tree

# STP

- Ethernet switches can implement the IEEE 802.1D Spanning-Tree Protocol and use the spanning-tree algorithm to construct a loop free shortest path network.
- Shortest path is based on cumulative link costs. Link costs are based on the speed of the link.
- STP establishes a root node, called the root bridge. The Spanning-Tree Protocol constructs a topology that has one path for reaching every network node.
- The resulting tree originates from the root bridge.
- Redundant links that are not part of the shortest path tree are blocked, causing a loop free topology

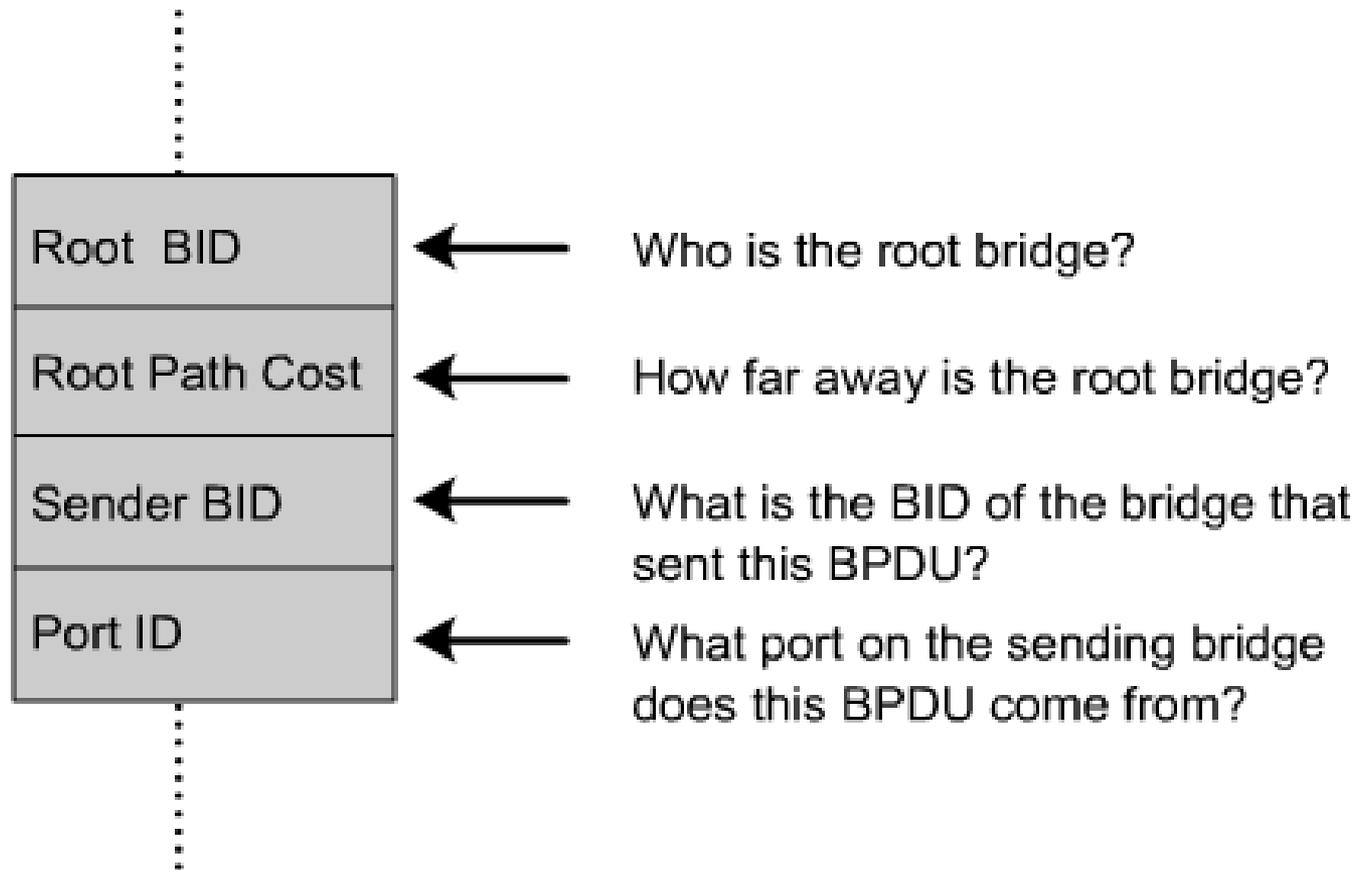
# Spanning Tree Link Costs

Link Speed	Cost(Revised IEEE Spec)	Cost (Previous IEEE Spec)
10 Gbps	2	1
1 Gbps	4	1
100 Mbps	19	10
10 Mbps	100	100

# BPDUs

- The Spanning-Tree Protocol requires network devices to exchange messages to detect bridging loops.
- Links that will cause a loop are put into a blocking state.
- The message that a switch sends, allowing the formation of a loop free logical topology, is called a Bridge Protocol Data Unit (BPDU).
- BPDUs continue to be received on blocked ports. This ensures that if an active path or device fails, a new spanning tree can be calculated.

# What's in a BPDU?



# BPDUs

## 802.3 Header

**Destination:** 01:80:C2:00:00:00 *Mcast 802.1d Bridge group*  
**Source:** 00:D0:C0:F5:18:D1  
**LLC Length:** 38

## 802.2 Logical Link Control (LLC) Header

**Dest. SAP:** 0x42 *802.1 Bridge Spanning Tree*  
**Source SAP:** 0x42 *802.1 Bridge Spanning Tree*  
**Command:** 0x03 *Unnumbered Information*

## 802.1 - Bridge Spanning Tree

**Protocol Identifier:** 0  
**Protocol Version ID:** 0  
**Message Type:** 0 *Configuration Message*

**Flags:** %00000000

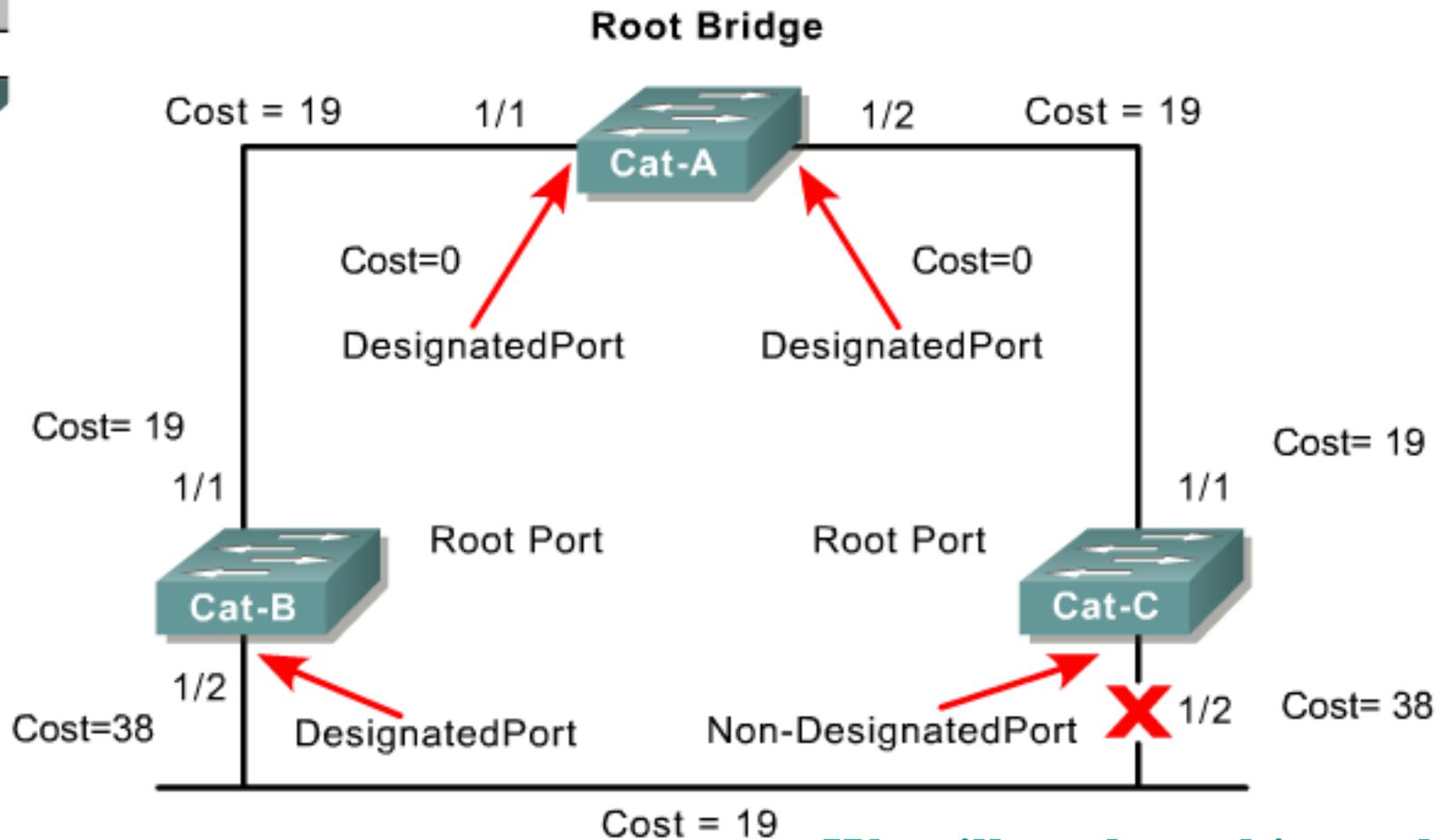
**Root Priority/ID:** 0x8000/ 00:D0:C0:F5:18:C0  
**Cost Of Path To Root:** 0x00000000 (0)  
**Bridge Priority/ID:** 0x8000/ 00:D0:C0:F5:18:C0  
**Port Priority/ID:** 0x80/ 0x1D

**Message Age:** 0/256 seconds (*exactly 0 seconds*)  
**Maximum Age:** 5120/256 seconds (*exactly 20 seconds*)  
**Hello Time:** 512/256 seconds (*exactly 2 seconds*)  
**Forward Delay:** 3840/256 seconds (*exactly 15 seconds*)

# Switches use BPDUs to:

- Select a single switch that will act as the **root** of the spanning tree
- Calculate the shortest path from itself to the root switch
- Designate one of the switches as the closest one to the root, for each LAN segment. This bridge is called the **“designated switch”**. The designated switch handles all communication from that LAN towards the root bridge.
- Choose one of its ports as its **root port**, for each non-root switch. This is the interface that gives the best path to the root switch.
- Select ports that are part of the spanning tree, the **designated ports**. Non-designated ports are blocked.

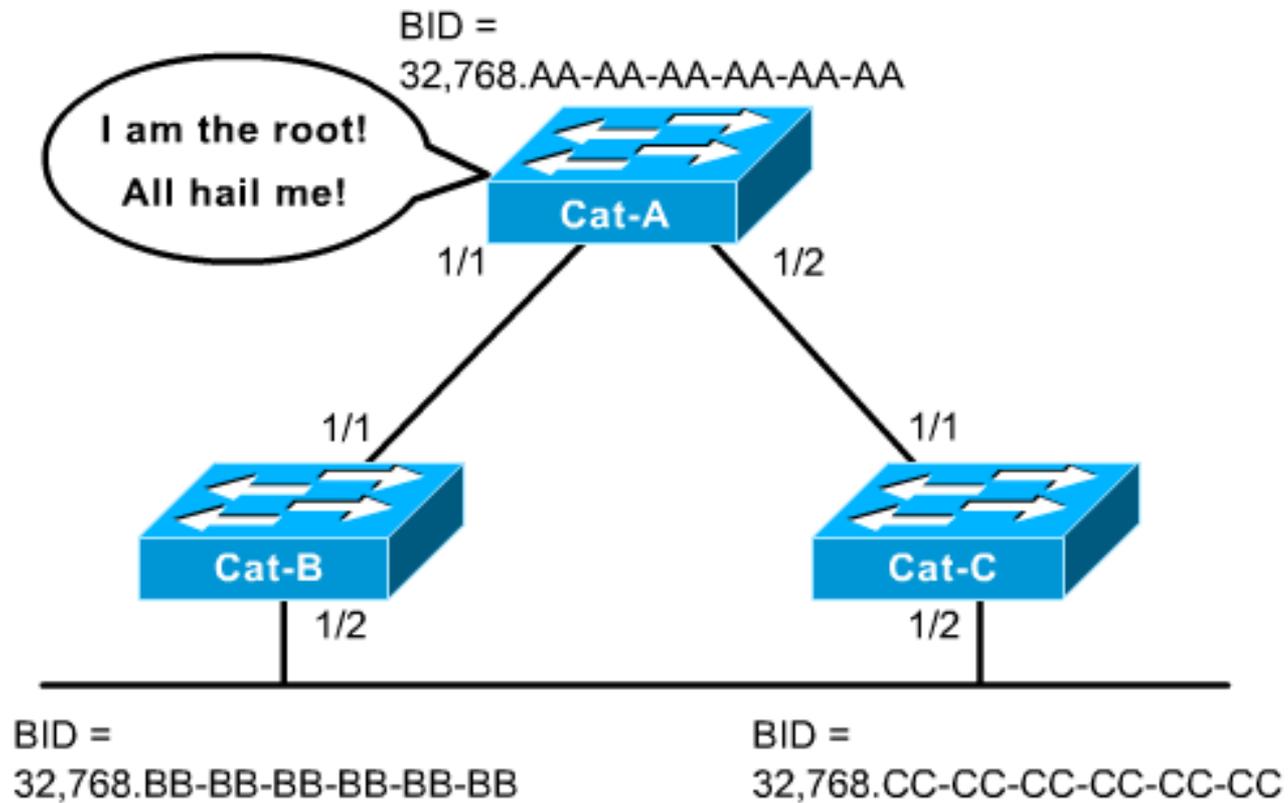
# Spanning Tree Operation



We will see how this works

# Electing a Root Bridge

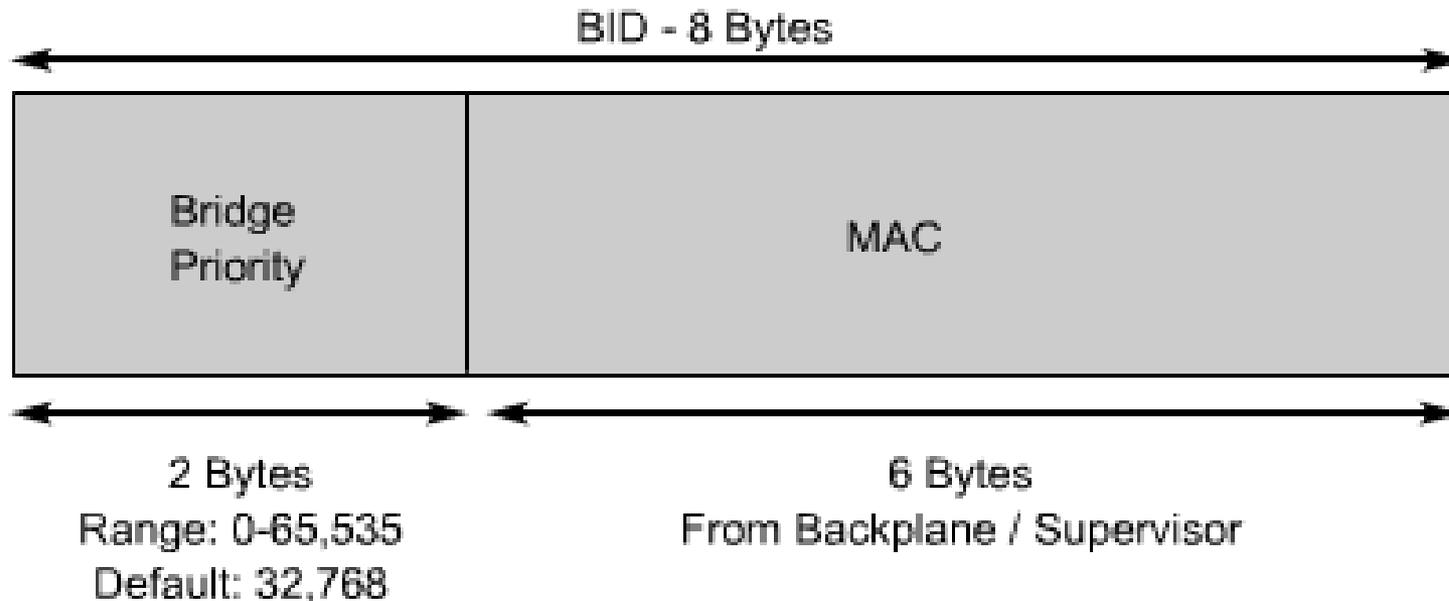
**BID = Bridge ID (Priority/MAC Address) the lower, the better**



# Electing the Root Bridge

- When a switch is turned on, the spanning-tree algorithm is used to identify the root bridge. BPDUs are sent out with the Bridge ID (BID).
- The BID consists of a bridge priority that defaults to 32768 and the switch base MAC address (from PROM). By default BPDUs are sent every two seconds.
- When a switch first starts up, it assumes it is the root switch and sends “inferior” BPDUs. These BPDUs contain the switch MAC address in both the root and sender BID. All switches see the BIDs sent. As a switch receives a BPDU with a lower root BID it replaces that in the BPDUs that are sent out. All bridges see these and decide that the bridge with the smallest BID value will be the root bridge.

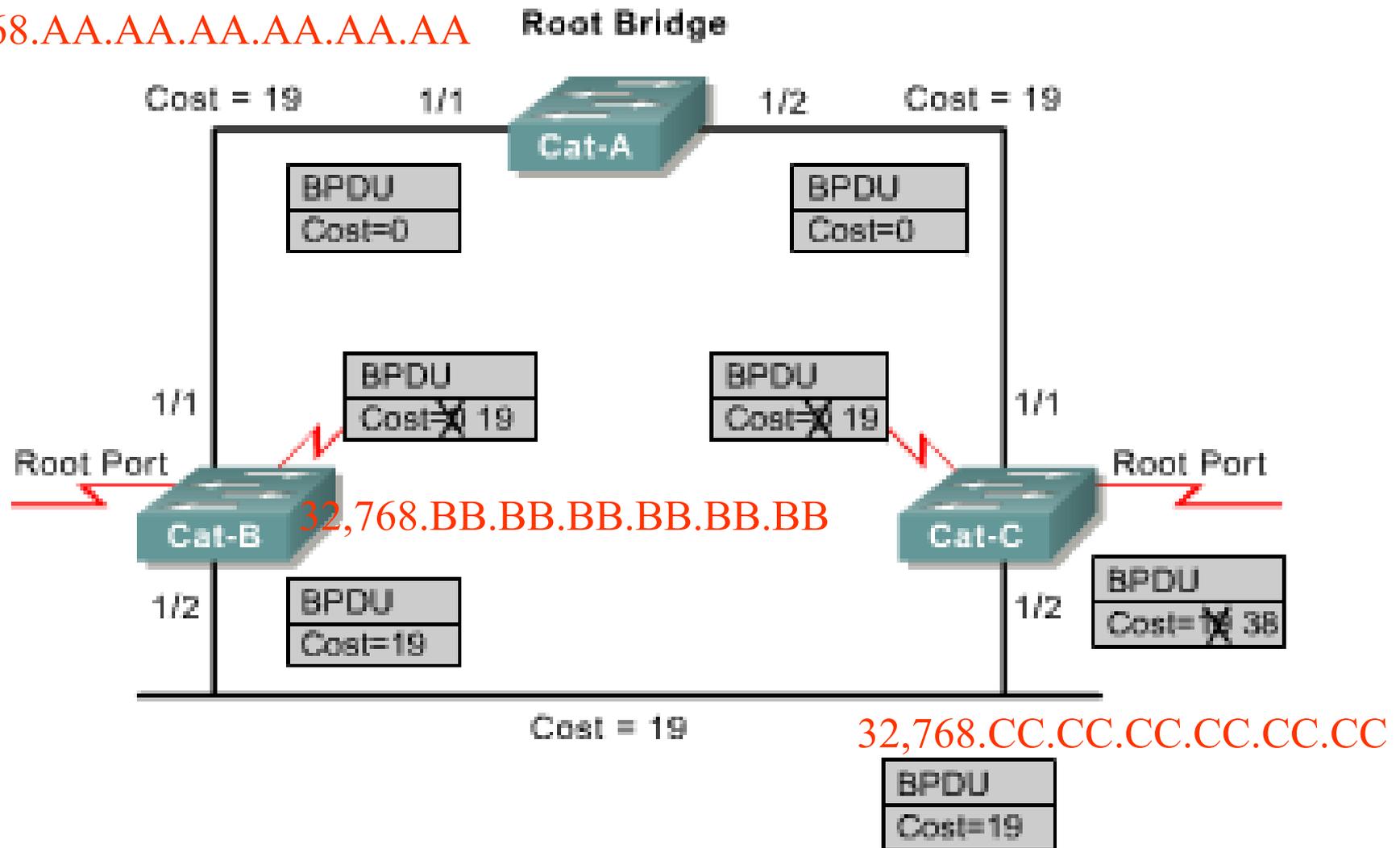
# BID



- Bridge ID (BID) is used to identify each bridge/switch.
- The BID is used in determining the center of the network, in respect to STP, known as the root bridge.

# Electing Root Ports

32,768.AA.AA.AA.AA.AA.AA



# Electing Root Ports

- Each switch must form an association with the root bridge.
- At the conclusion of the root war, the switches move on to selecting Root Ports.
- A bridge's Root Port is the port that is closest to the Root Bridge in terms of Path Cost. Every non-Root Bridge must select one Root Port. Again, bridges use the concept of cost to measure closeness.
- If a switch receives BPDUs on multiple ports, it has a redundant path to the root bridge (or it is the root bridge!)

# Electing Root Ports

- In order to choose which ports will forward data and which ports will block data, the switch looks at three components of the BPDU:
  - Lowest path cost to root bridge
  - Lowest sender Bridge ID
  - Lowest port priority/port ID

# Path Cost

- Lowest cost path to the root preferred
- Path cost calculated based on link speed and the number of links that the BPDU crossed downstream from the root.
- If one port has the lowest cost, it is placed in forwarding mode. All other ports receiving BPDUs are placed in blocking mode.

# Path Cost

Bandwidth	STP Cost
4 Mbps	250
10 Mbps	100
16 Mbps	62
45 Mbps	39
100 Mbps	19
155 Mbps	14
622 Mbps	6
1 Gbps	4
10 Gbps	2

# Path Cost

- You can tweak the path cost by modifying the cost of a port.
  - Exercise caution when you do this!

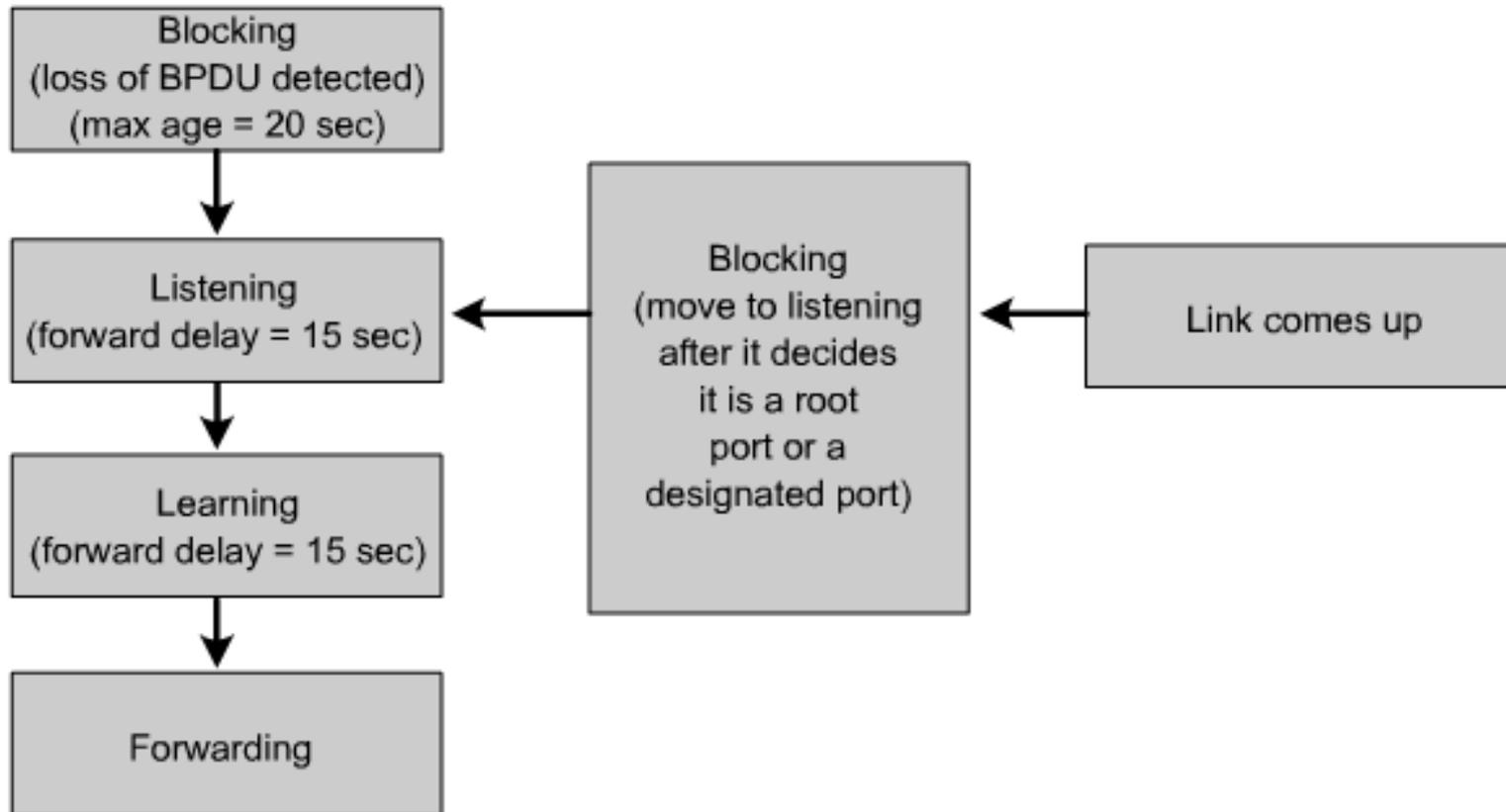
# Bridge IDs

- If the path costs of the received BPDUs are equal, the switch looks at the Bridge ID to determine which port should forward.
- The port receiving the lowest Bridge ID is chosen to forward, all others block.

# Port Cost/Port ID

- If the path cost and bridge IDs are equal (as in the case of parallel links), the switch goes to the port priority as a tiebreaker.
- Lowest port priority wins (default 128).
- You can set the priority from 0 – 255.
- If all ports have the same priority, the port with the lowest port number forwards frames.

# Spanning Tree Port States



Spanning tree transits each port through several different states.

# Spanning-Tree Port States

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

# Spanning-Tree Port States

- Blocked
  - All ports start in blocked mode in order to prevent the bridge from creating a bridging loop. The port stays in a blocked state if Spanning Tree determines that there is a better path to the root bridge.

# Blocked State

- Discards frames received from the attached segment or internally forwarded through switching
- Receives BPDUs and directs them to the system module
- Has no address database
- Does not transmit BPDUs received from the system module
- Receives and responds to network management messages but does not transmit them

# Listening State

- Discards frames received from the attached segment or frames switched from another port
- Has no address database
- Receives BPDUs and directs them to the system module
- Processes BPDUs received from the system module Receives and responds to network management messages
- It is during the Listening state that the three initial convergence steps take place – elect a Root Bridge, elect Root Ports, and elect Designated Ports.

# Spanning-Tree Port States

- **Learn:** The learn state is very similar to the listen state, except that the port can add information it has learned to its address table.
  - Still not allowed to send or receive data
  - Learns for a period of time called the fwd delay

# Learning State

- Discards frames received from the attached segment
- Discards frames switched from another port for forwarding
- Receives and responds to network management messages
- Builds the bridging table

# Spanning-Tree Port States

- **Forward**: The port can send and receive data.
  - A port is not placed in the forwarding state unless there are no redundant links or it is determined that it has the best path to the root.

# Learning State

- As the bridge receives a frame, it places the source MAC address and port into the bridging table. Discards frames switched from another port for forwarding
- Incorporates station location into its address database
- Receives BPDUs and directs them to the system module
- Receives, processes, and transmits BPDUs received from the system module
- Receives and responds to network management messages

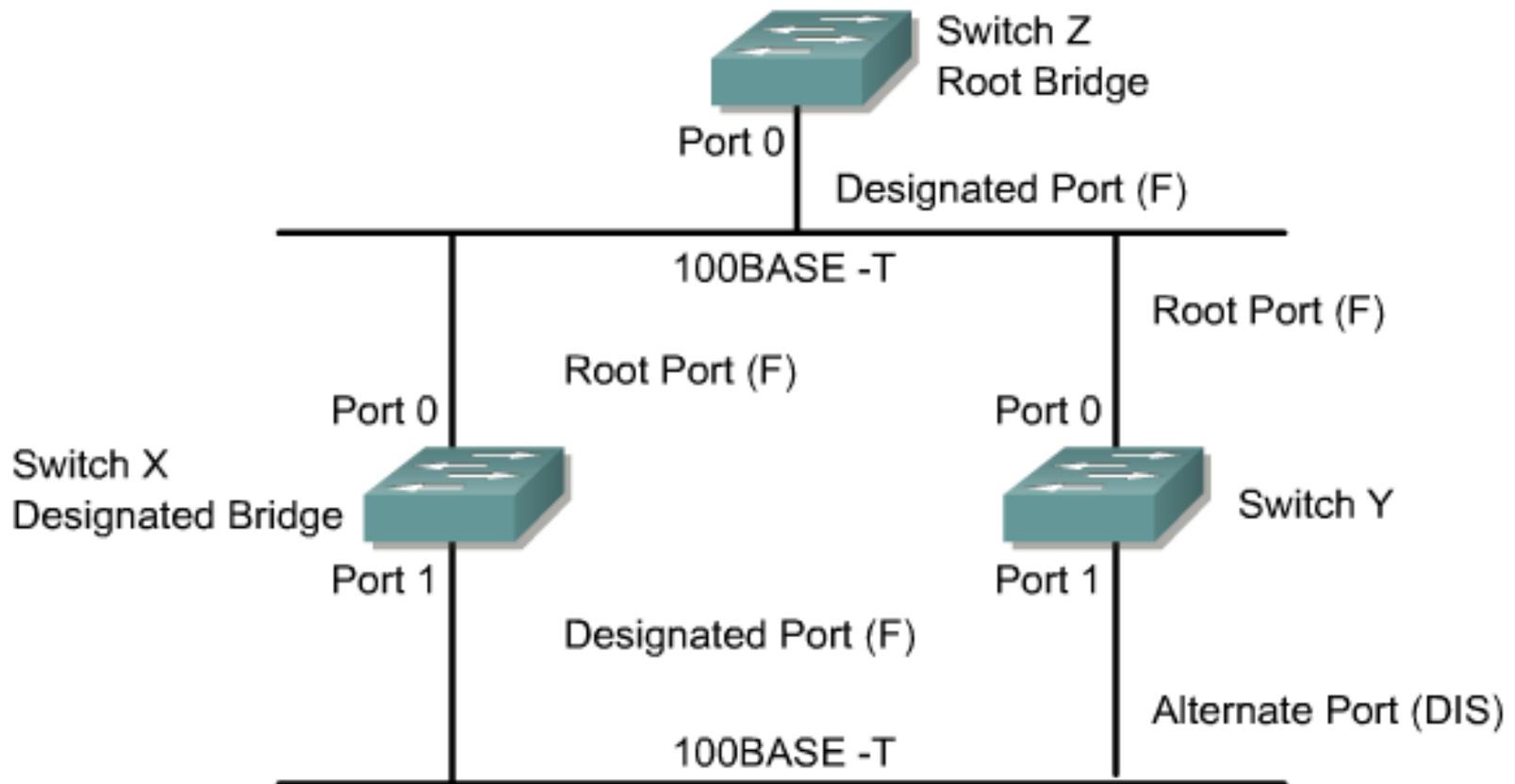
# Spanning-Tree Port States

- **Disabled**: The port is shutdown.

# Spanning Tree Recalculation

- A switched internetwork has converged when all the switch and bridge ports are in either the forwarding or blocked state. Forwarding ports send and receive data traffic and BPDUs. Blocked ports will only receive BPDUs. When the network topology changes, switches and bridges recompute the Spanning Tree and cause a disruption of user traffic.
- Convergence on a new spanning-tree topology using the IEEE 802.1D standard can take up to 50 seconds. This convergence is made up of the max-age of 20 seconds, plus the listening forward delay of 15 seconds, and the learning forward delay of 15 seconds.

# Rapid Spanning Tree



# RSTP

- The Rapid Spanning-Tree Protocol is defined in the IEEE 802.1w LAN standard. The standard and protocol introduce the following:
- Clarification of port states and roles
- Definition of a set of link types that can go to forwarding state rapidly
- Concept of allowing switches, in a converged network, to generate their own BPDUs rather than relaying root bridge BPDUs
- The “blocked” state of a port has been renamed as the “discarding” state.

# RSTP Link Types

- Link types have been defined as point-to-point, edge-type, and shared. These changes allow failure of links in switched network to be learned rapidly.
- Point-to-point links and edge-type links can go to the forwarding state immediately.
- Network convergence does not need to be any longer than 15 seconds with these changes.
- The Rapid Spanning-Tree Protocol, IEEE 802.1w, will eventually replace the Spanning-Tree Protocol, IEEE 802.1D

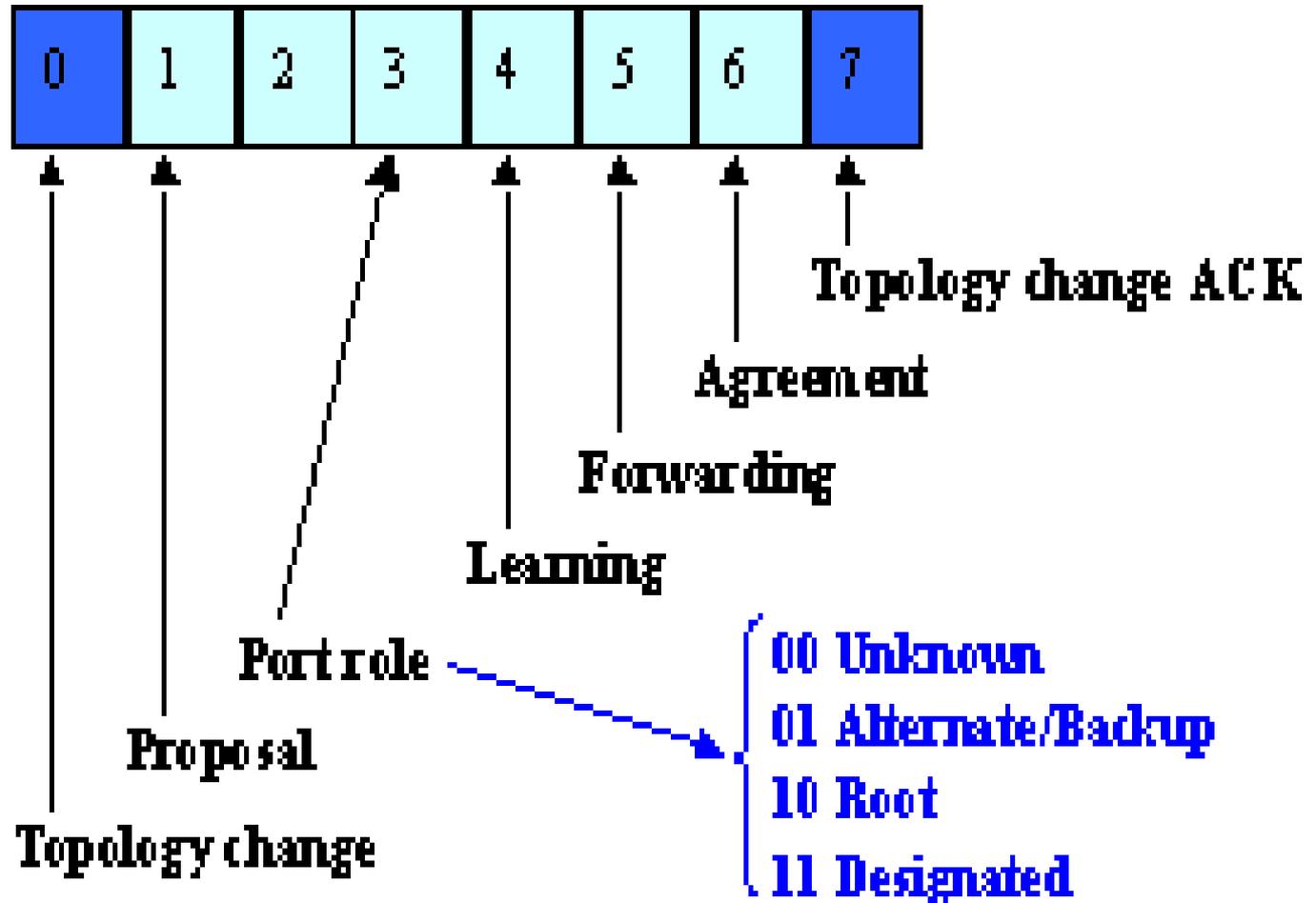
# RSTP Port States

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

# RSTP Port Roles

- The role is now a variable assigned to a given port.
- The root port and designated port roles remain.
- The blocking port role is now split into the **backup** and **alternate** port roles.
- The Spanning Tree Algorithm (STA) determines the role of a port based on Bridge Protocol Data Units (BPDUs).
- To keep things simple, the thing to remember about a BPDU is that there is always a way of comparing any two of them and deciding whether one is more useful than the other. This is based on the value stored in the BPDU and occasionally on the port on which they are received.

# 802.1w BPDUs



# Designated

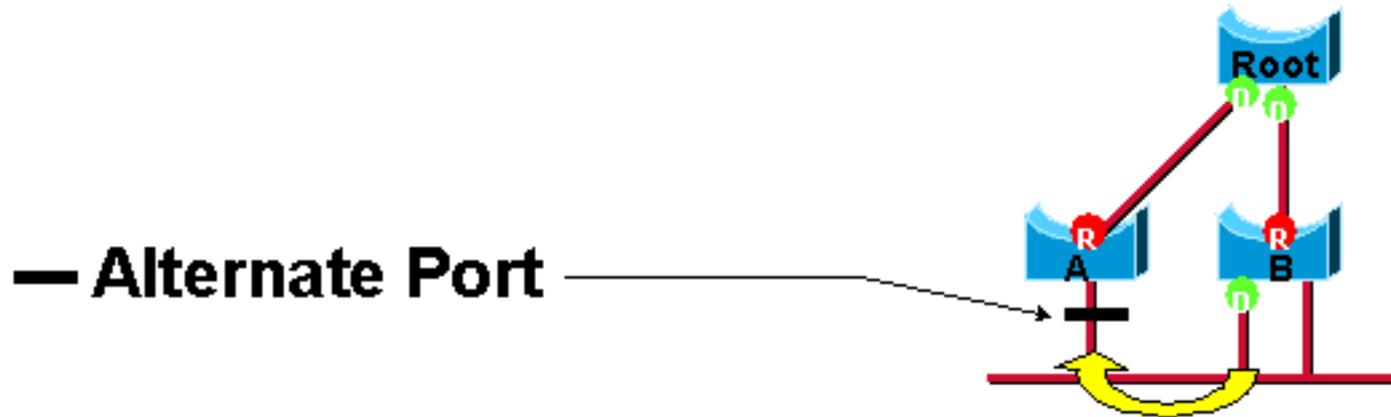
- A port is designated if it can send the best BPDU on the segment to which it is connected.
- 802.1D bridges link together different segments, such as Ethernet segments, to create a bridged domain.
- On a given segment, there can only be one path toward the root bridge.
- All bridges connected to a given segment listen to the BPDUs of each and agree on the bridge that sends the best BPDU as the designated bridge for the segment. The port on that bridge that corresponds is designated.

# Alternate and Backup

- These two port roles correspond to the blocking state of 802.1D.
- A blocked port is defined as not being the designated or root port.
- A blocked port receives a more useful BPDU than the one it sends out on its segment.
- Remember that a port absolutely needs to receive BPDUs in order to stay blocked. RSTP introduces these two roles for this purpose.

# Alternate

- an alternate port receives more useful BPDUs from another bridge and is a port blocked



# Backup

- A backup port receives more useful BPDUs from the same bridge it is on and is a port blocked



# Algorhyme by Radia Perlman

- I think I shall never see
- A graph more lovely than a tree.
- A tree whose crucial property
- Is loop-free connectivity
- A tree that must be sure to span
- So packets can reach every LAN.
- First the root must be elected.
- By ID it is elected.
- Least-cost paths from root are traced.
- In the tree, these paths are placed.
- A mesh is made by folks like me,
- Then bridges find a spanning tree.