

Spanning Tree Protocol

Slide Set 8.5

Overview

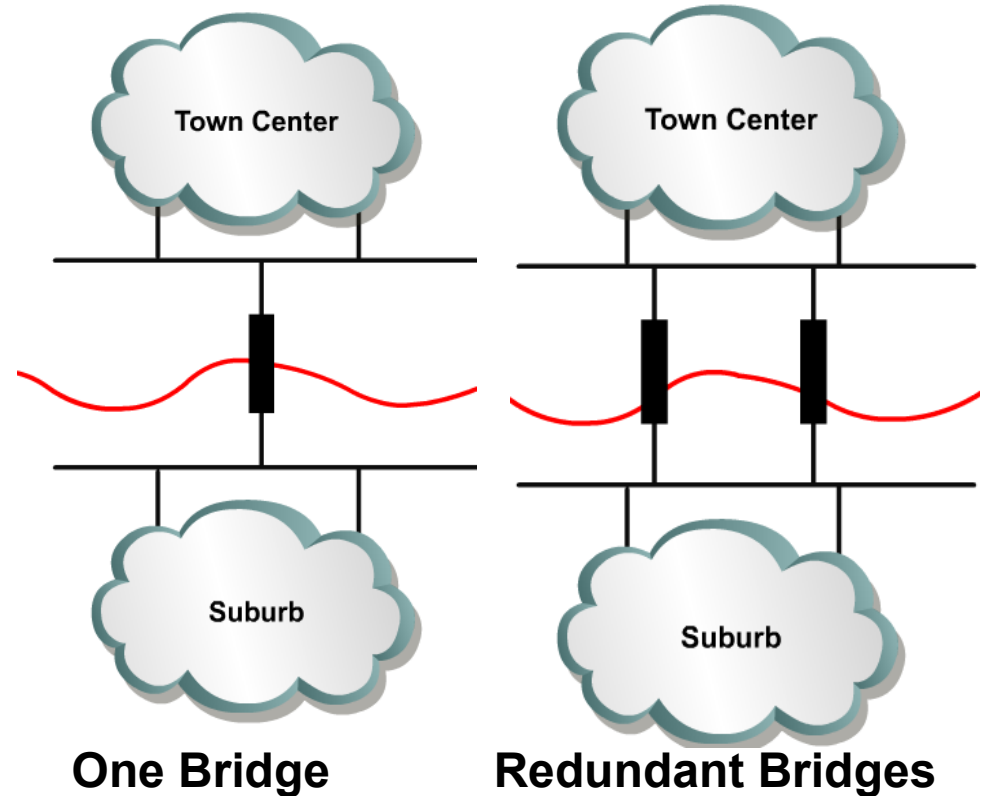
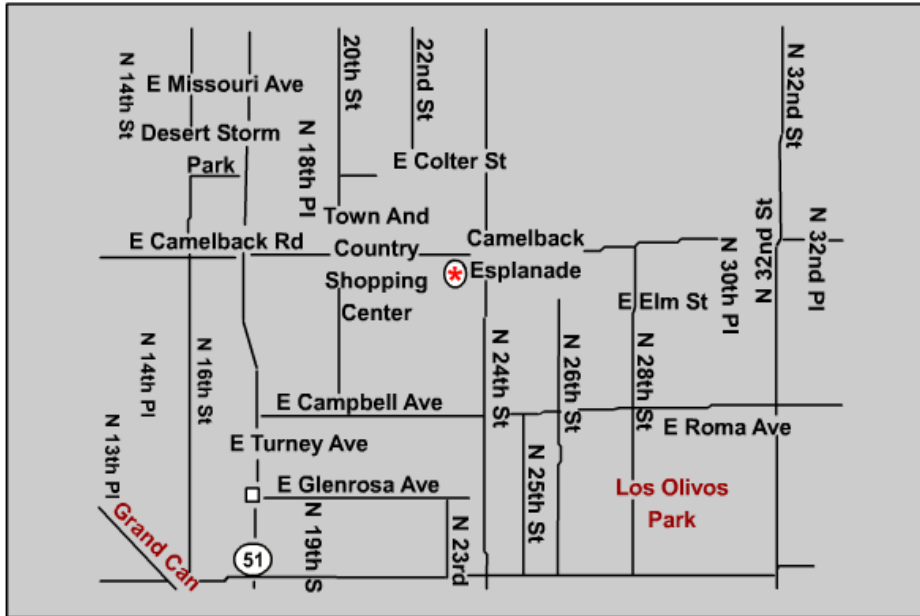
- Define redundancy and its importance in networking
- Describe the key elements of a redundant networking topology
- Define broadcast storms and describe their impact on switched networks
- Define multiple frame transmissions and describe their impact on switched networks
- Identify causes and results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



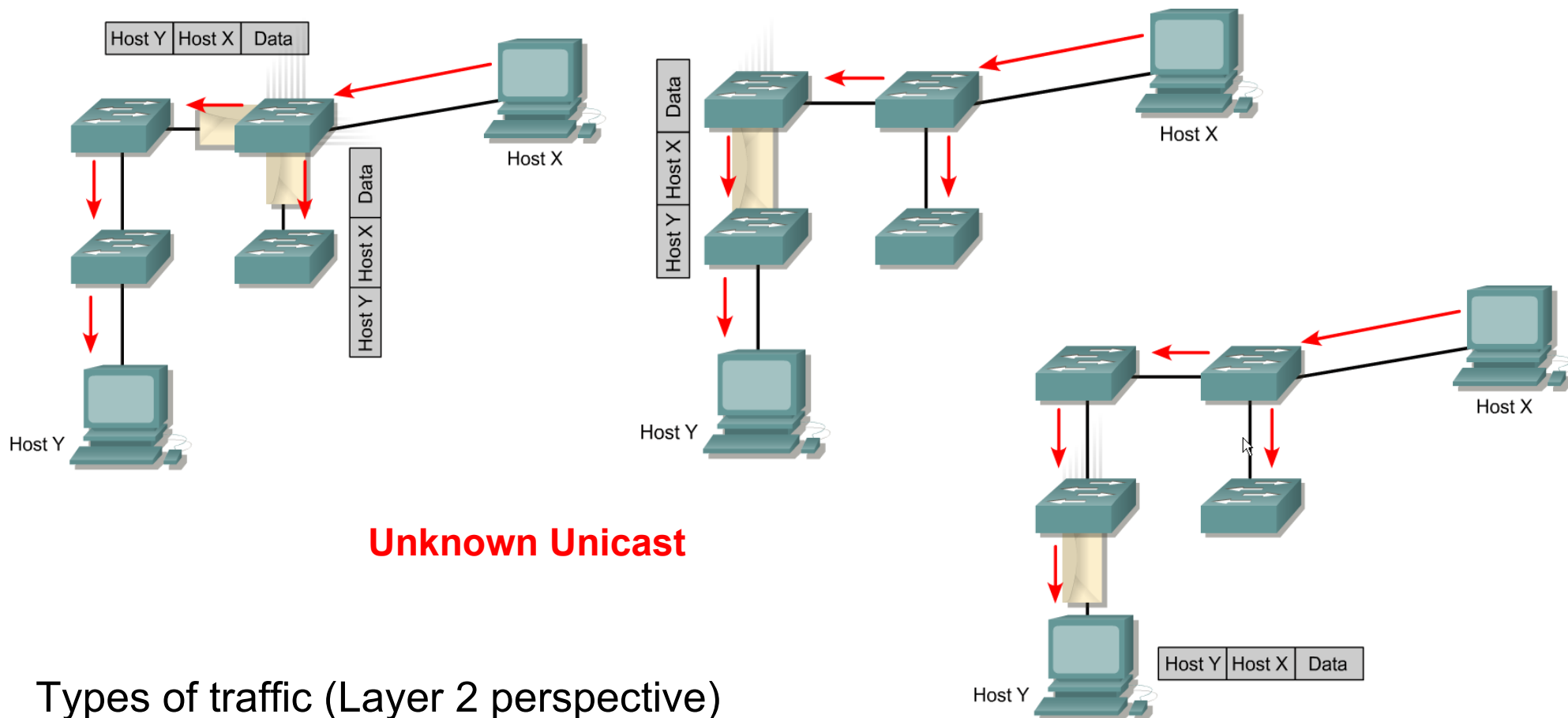
- Achieving such a goal requires extremely reliable networks.
- Reliability in networks is achieved by reliable equipment and by designing networks that are tolerant to failures and faults.
- The network is designed to reconverge rapidly so that the fault is bypassed.
- Fault tolerance is achieved by redundancy.
- Redundancy means to be in excess or exceeding what is usual and natural.

Redundant topologies



- A network of roads is a global example of a redundant topology.
- If one road is closed for repair there is likely an alternate route to the destination

Types of Traffic

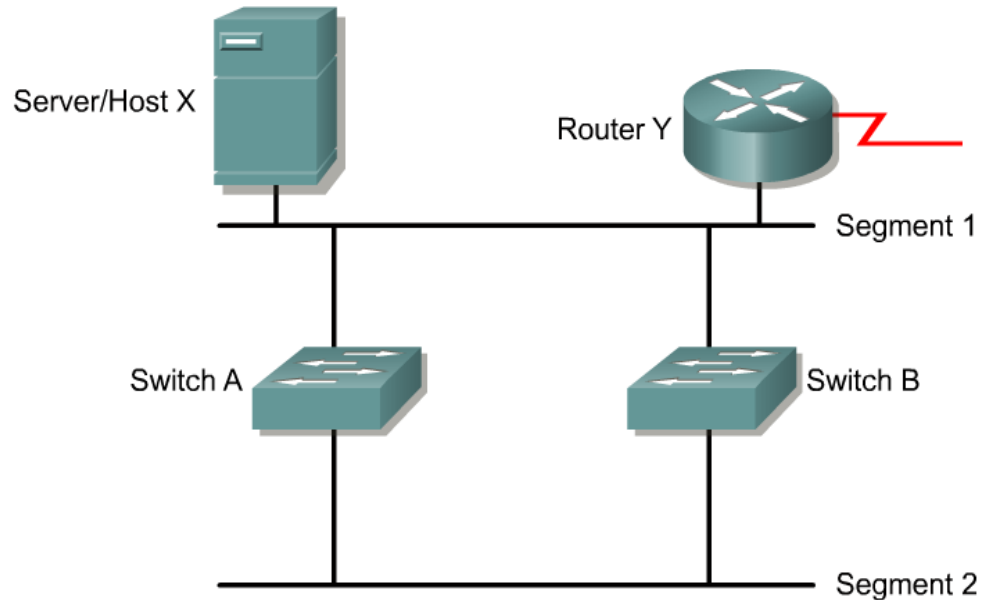


Unknown Unicast

Types of traffic (Layer 2 perspective)

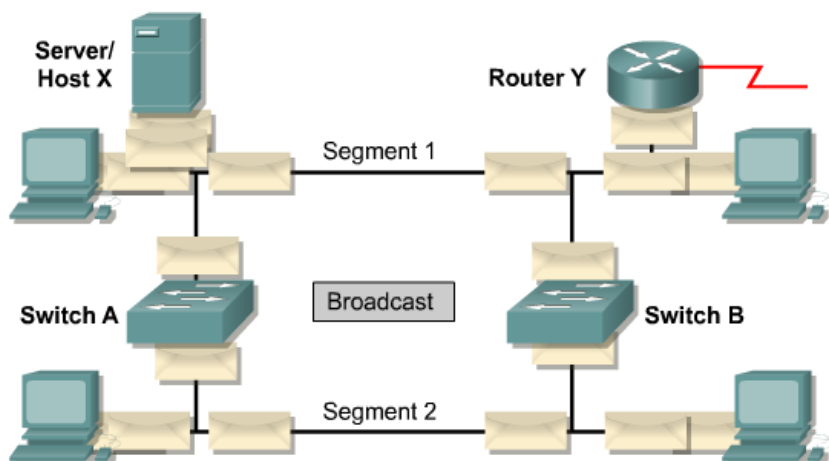
- Known Unicast: Destination addresses are in Switch Tables
- Unknown Unicast: Destination addresses are not in Switch Tables
- Multicast: Traffic sent to a group of addresses
- Broadcast: Traffic forwarded out all interfaces except incoming interface.

Redundant switched topologies



- 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. (Unless switch is doing Multicast Snooping or IGMP)
- A redundant switched topology *may* (STP disabled) cause broadcast storms, multiple frame copies, and MAC address table instability problems.

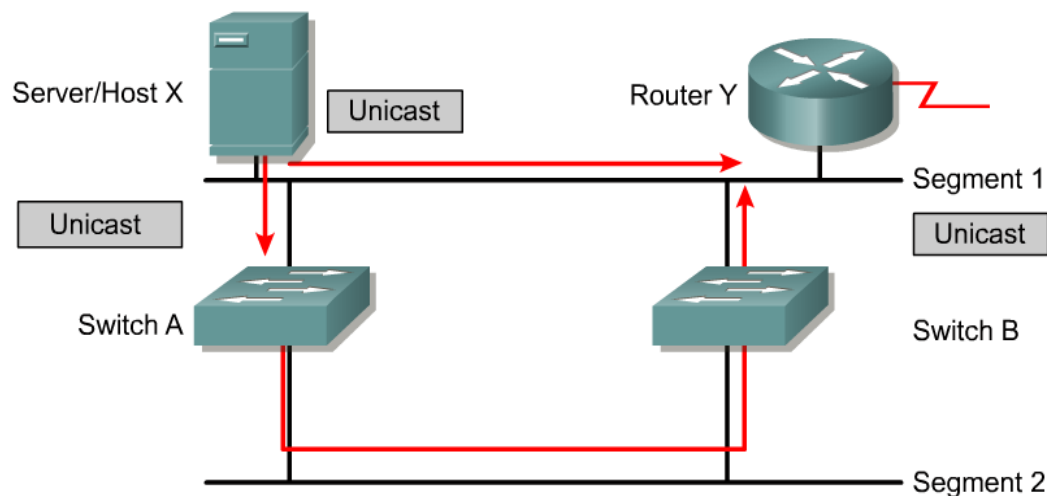
Broadcast Storm



A state in which a message that has been broadcast across a network results in even more responses, and each response results in still more responses in a snowball effect.
www.webopedia.com

- Broadcasts and multicasts can cause problems in a switched network.
- If Host X sends a broadcast, like an ARP request for the Layer 2 address of the router, then Switch A will forward the broadcast out all ports.
- Switch B, being on the same segment, also forwards all broadcasts.
- Switch B sees all the broadcasts that Switch A forwarded and Switch A sees all the broadcasts that Switch B forwarded.
- Switch A sees the broadcasts and forwards them.
- Switch B sees the broadcasts and forwards them.
- The switches continue to propagate broadcast traffic over and over.
- This is called a broadcast storm.

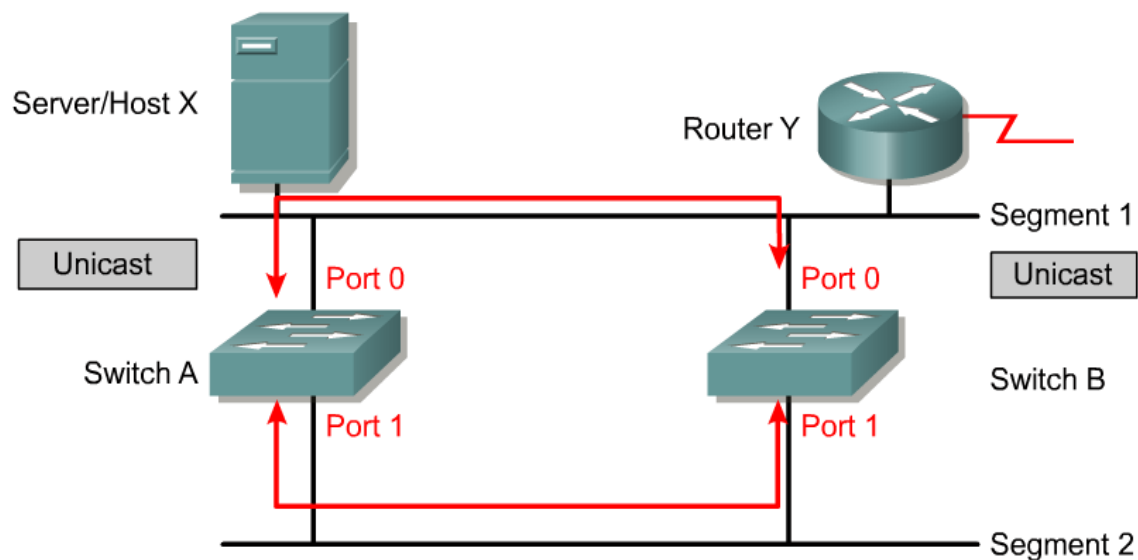
Multiple frame transmissions



(Some changes to curriculum)

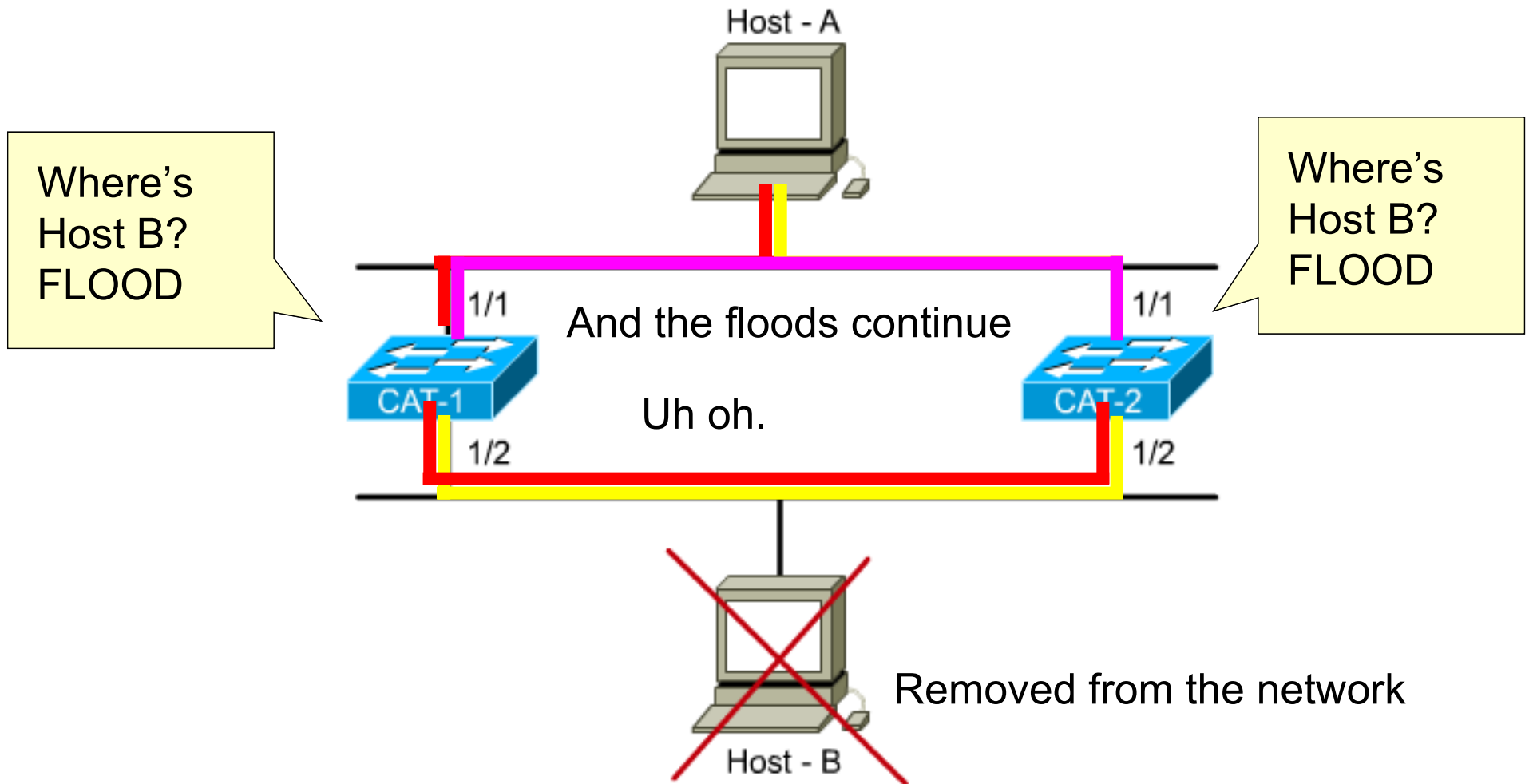
- The router receives the frame because it is on the same segment as Host X.
- Switch A does not have the MAC address of the Router Y and will therefore flood the frame out its ports. (Segment 2)
- Switch B also does not know which port Router Y is on.
- Note: Switch B will forward the the unicast onto Segment 2, creating multiple frames on that segment.
- After Switch B receives the frame from Switch A , it then floods the frame it received causing Router Y to receive multiple copies of the same frame.
- This is a causes of unnecessary processing in all devices.

Media access control database instability



- In a redundant switched network it is possible for switches to learn the wrong information.
- A switch can incorrectly learn that a MAC address is on one port, when it is actually on a different port.
- Host X sends a frame directed to Router Y.
- Switches A and B learn the MAC address of Host X on port 0.
- The frame to Router Y is flooded on port 1 of both switches.
- Switches A and B see this information on port 1 and incorrectly learn the MAC address of Host X on port 1.

Layer 2 Loops - Flooded unicast frames

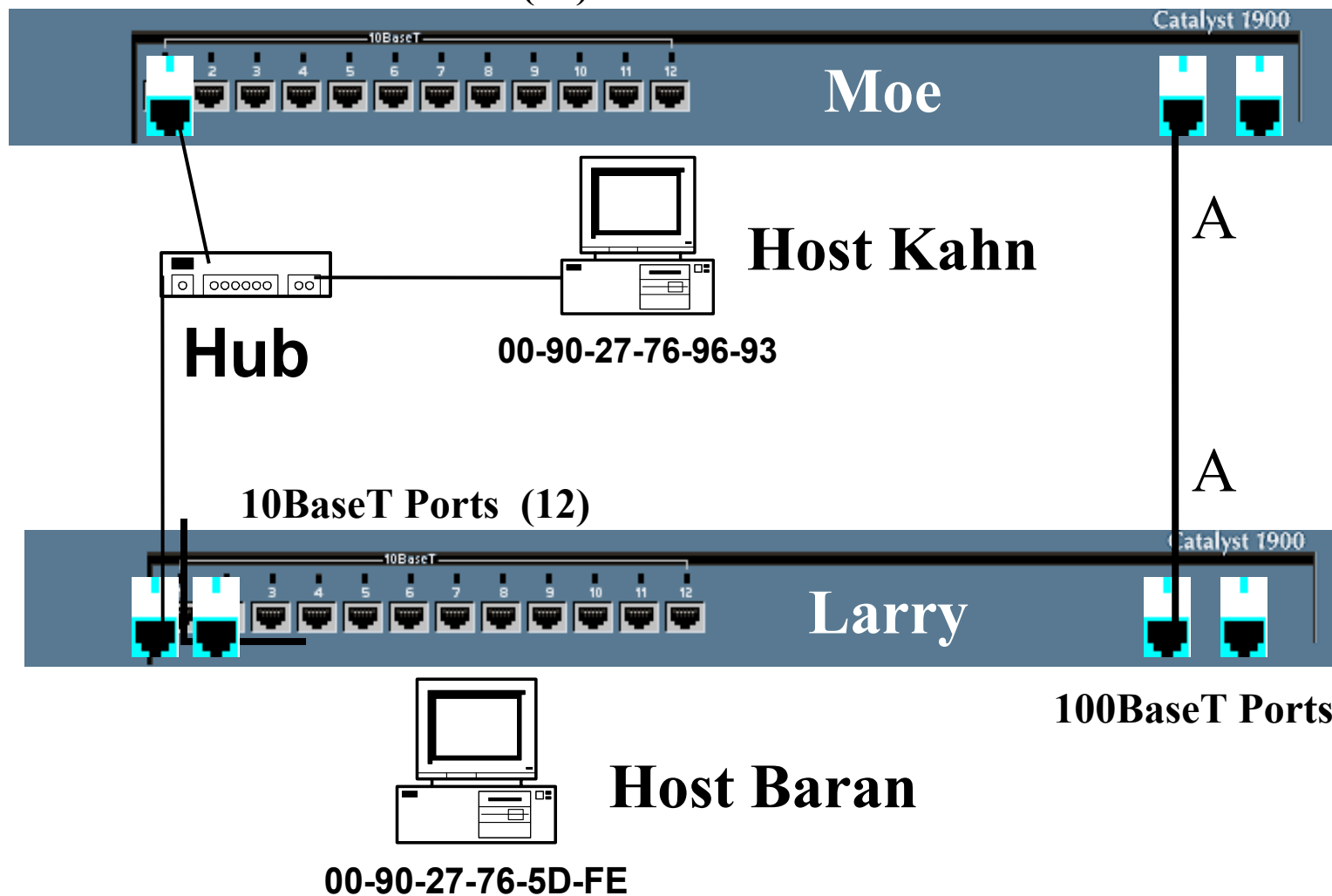


Redundant Paths and No Spanning Tree

Another problem, incorrect MAC Address Tables

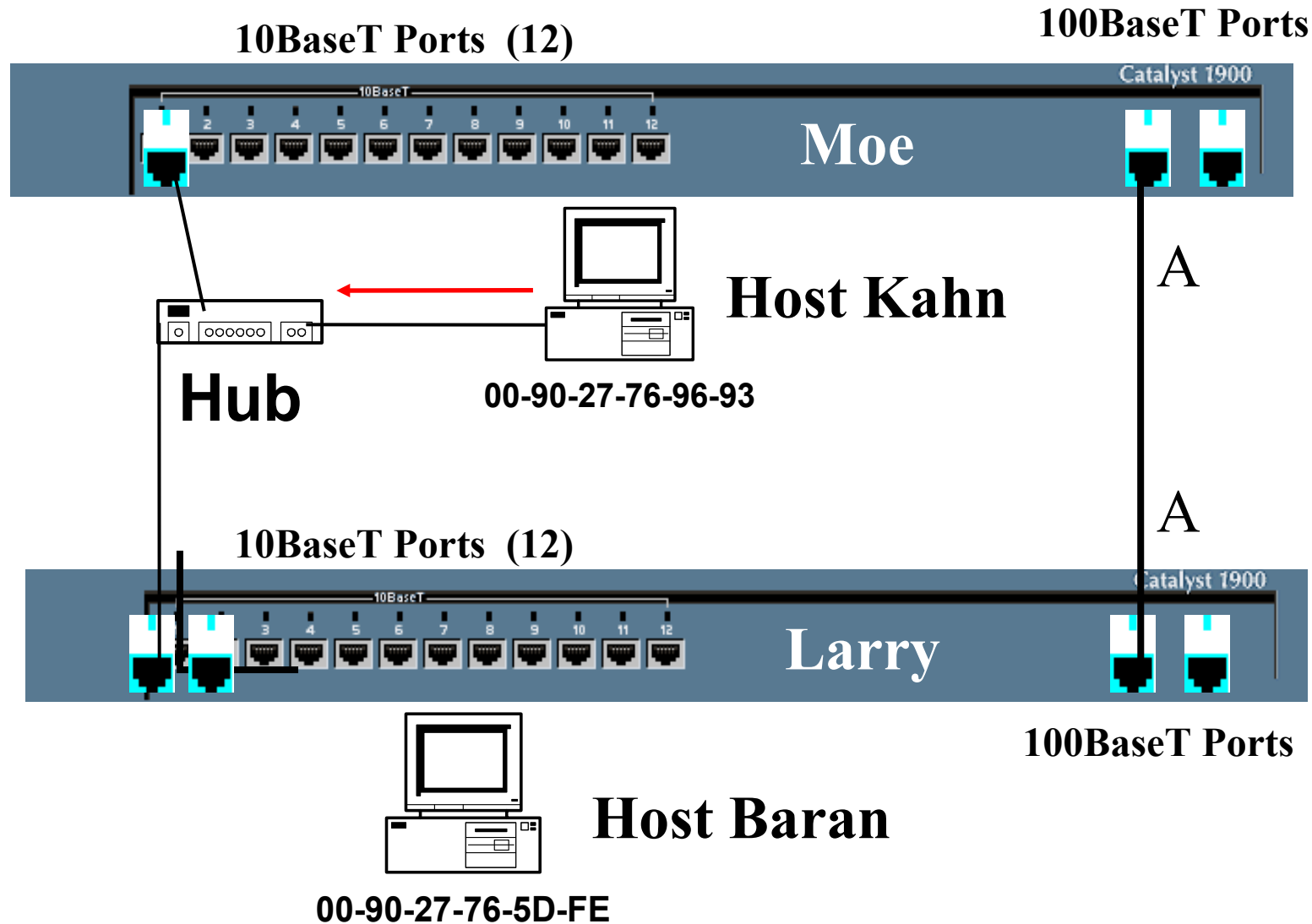
10BaseT Ports (12)

100BaseT Ports



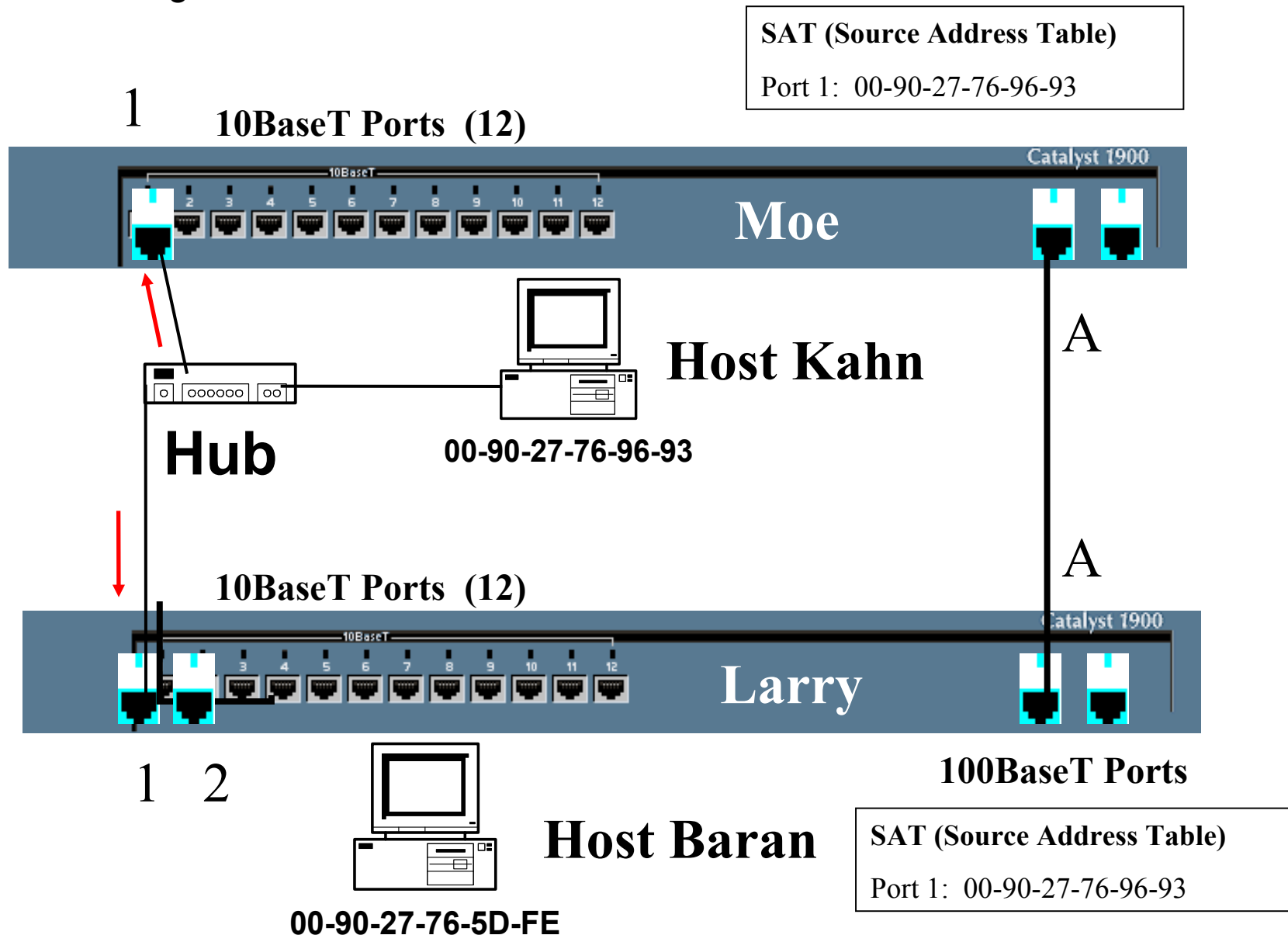
Redundant Paths and No Spanning Tree

Host Kahn sends an Ethernet frame to Host Baran. Both Switch Moe and Switch Larry see the frame and record Host Kahn's Mac Address in their switching tables.



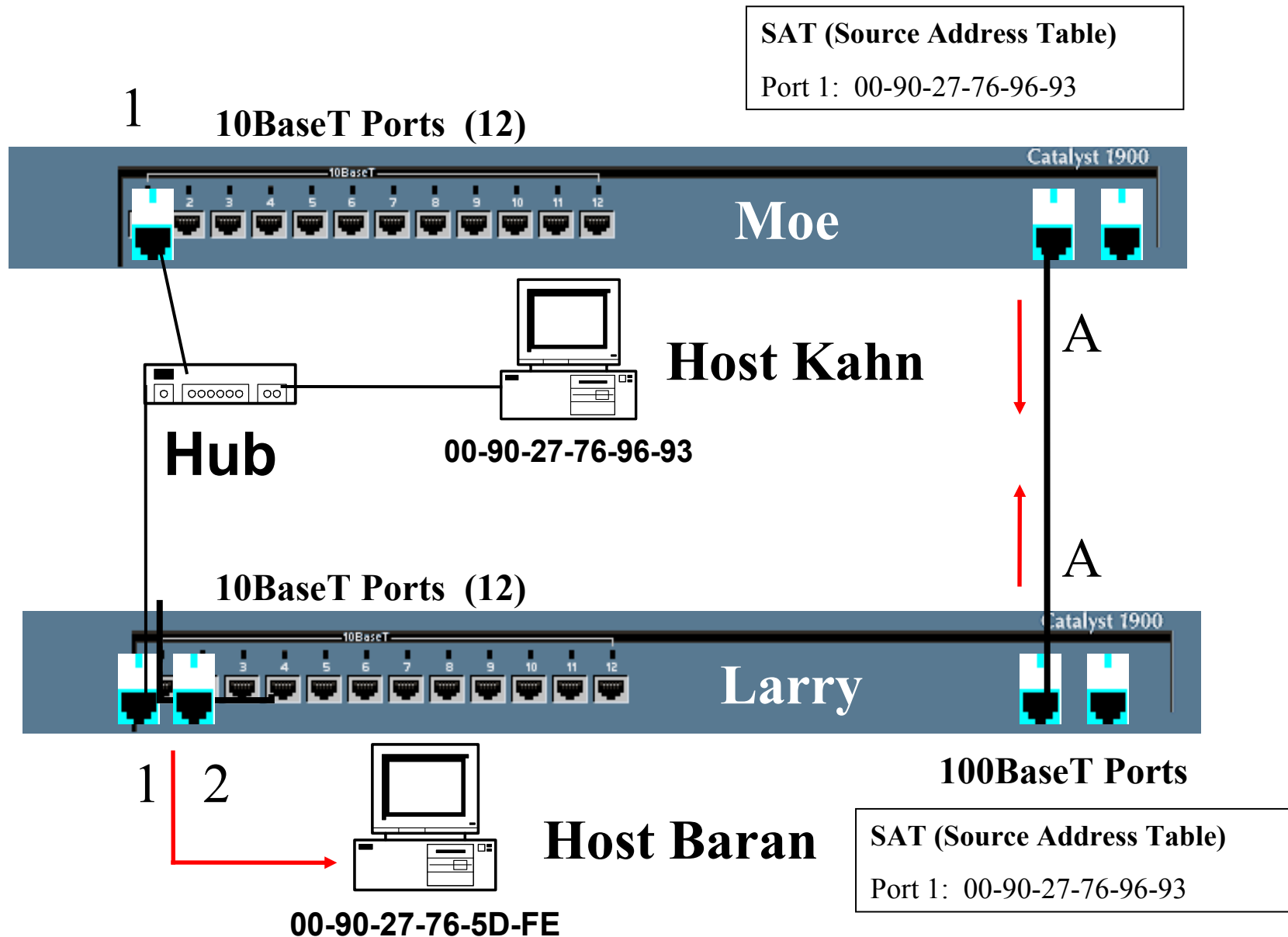
Redundant Paths and No Spanning Tree

Both Switch Moe and Switch Larry see the frame and record Host Kahn's Mac Address in their switching tables.



Redundant Paths and No Spanning Tree

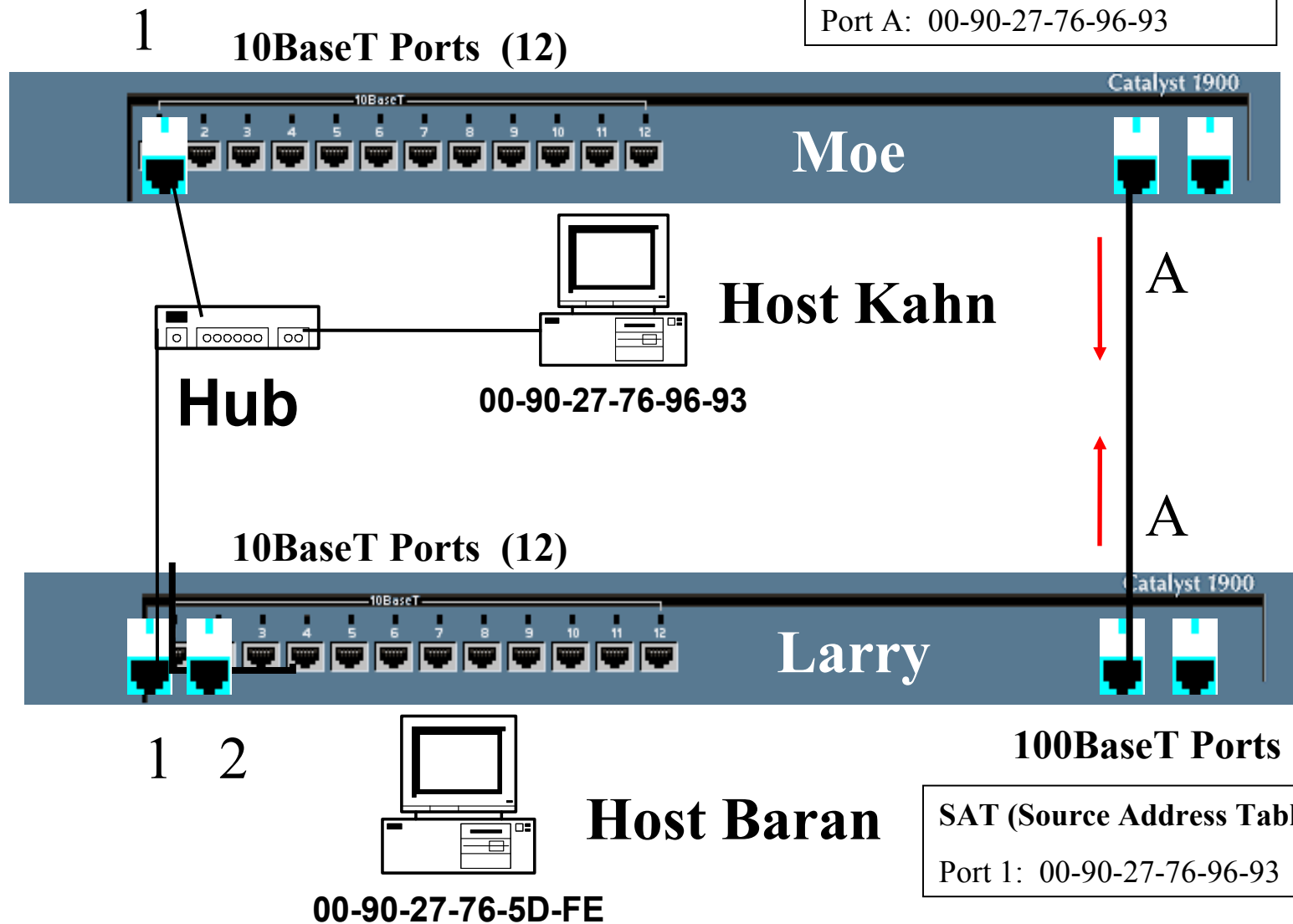
Both Switches do not have the **destination MAC address** in their table so they **both flood** it out all ports. Host Baran receives the frame.)



Redundant Paths and No Spanning Tree

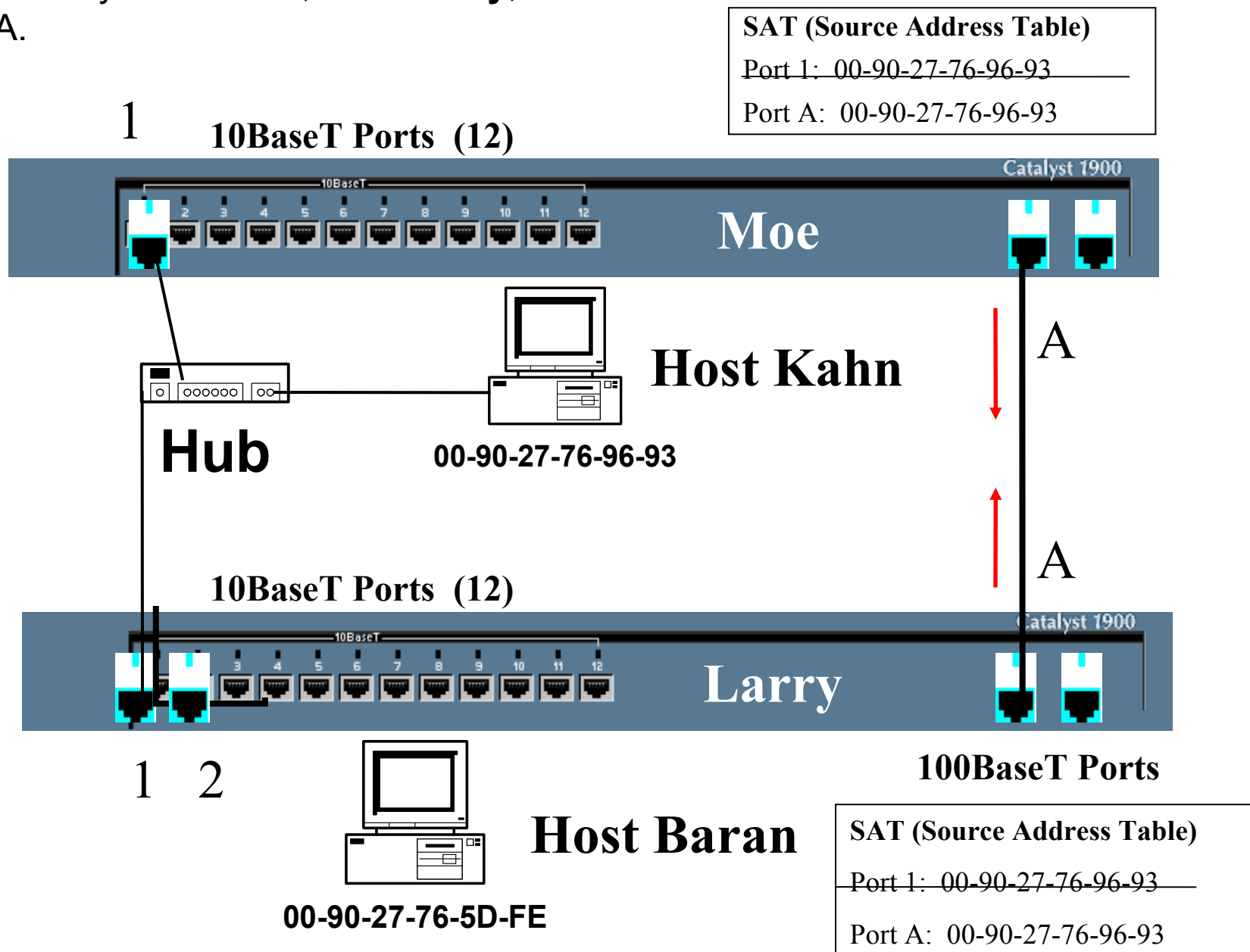
Switch Moe now learns, **incorrectly**, that the Source Address 00-90-27-76-96-93 is on Port A.

SAT (Source Address Table)	
Port 1:	00-90-27-76-96-93
Port A:	00-90-27-76-96-93



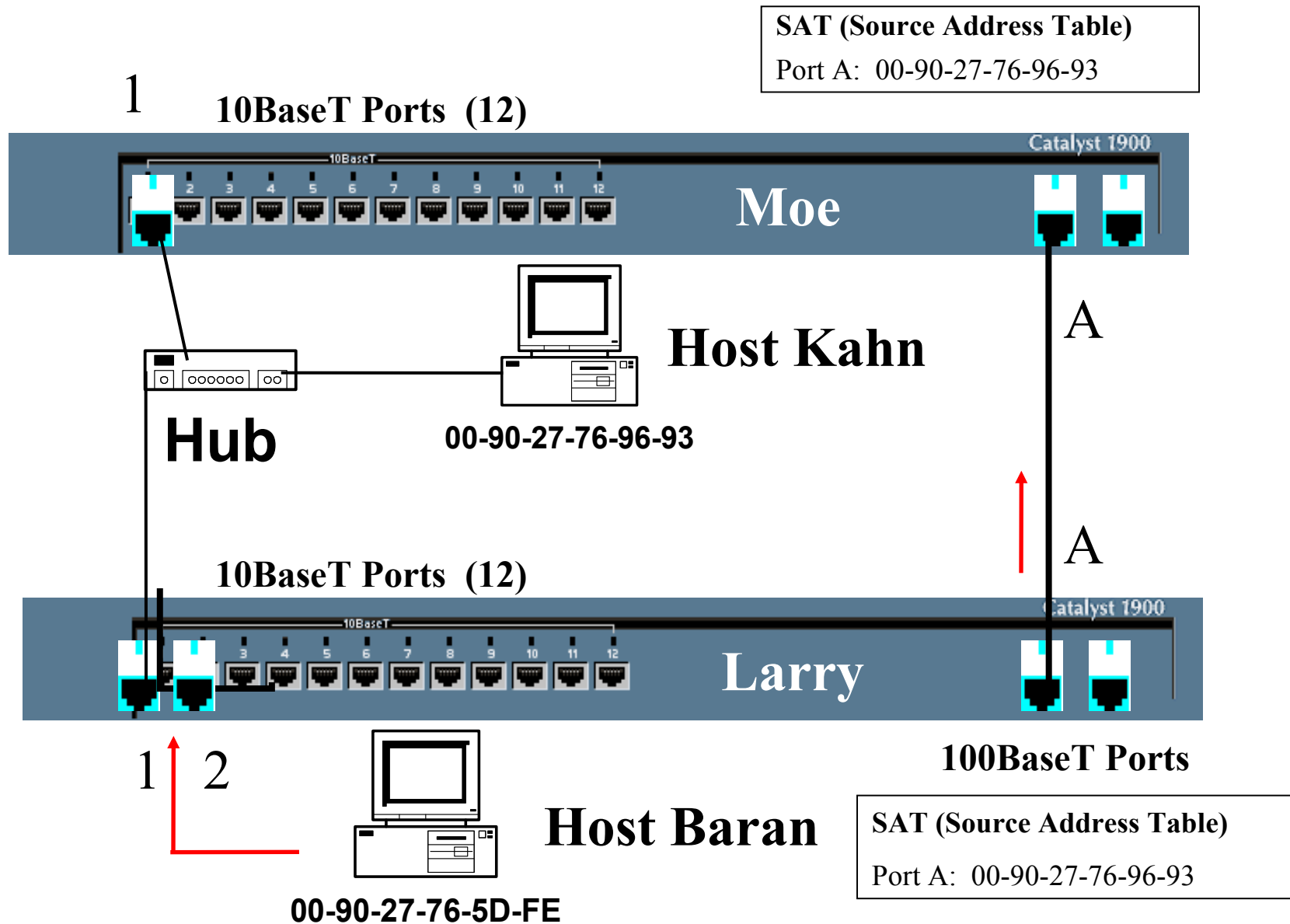
Redundant Paths and No Spanning Tree

Switch Larry also learns, **incorrectly**, that the Source Address 00-90-27-76-96-93 is on Port A.



Redundant Paths and No Spanning Tree

Now, when Host Baran sends a frame to Host Kahn, it will be sent the longer way, through Switch Larry's port A.

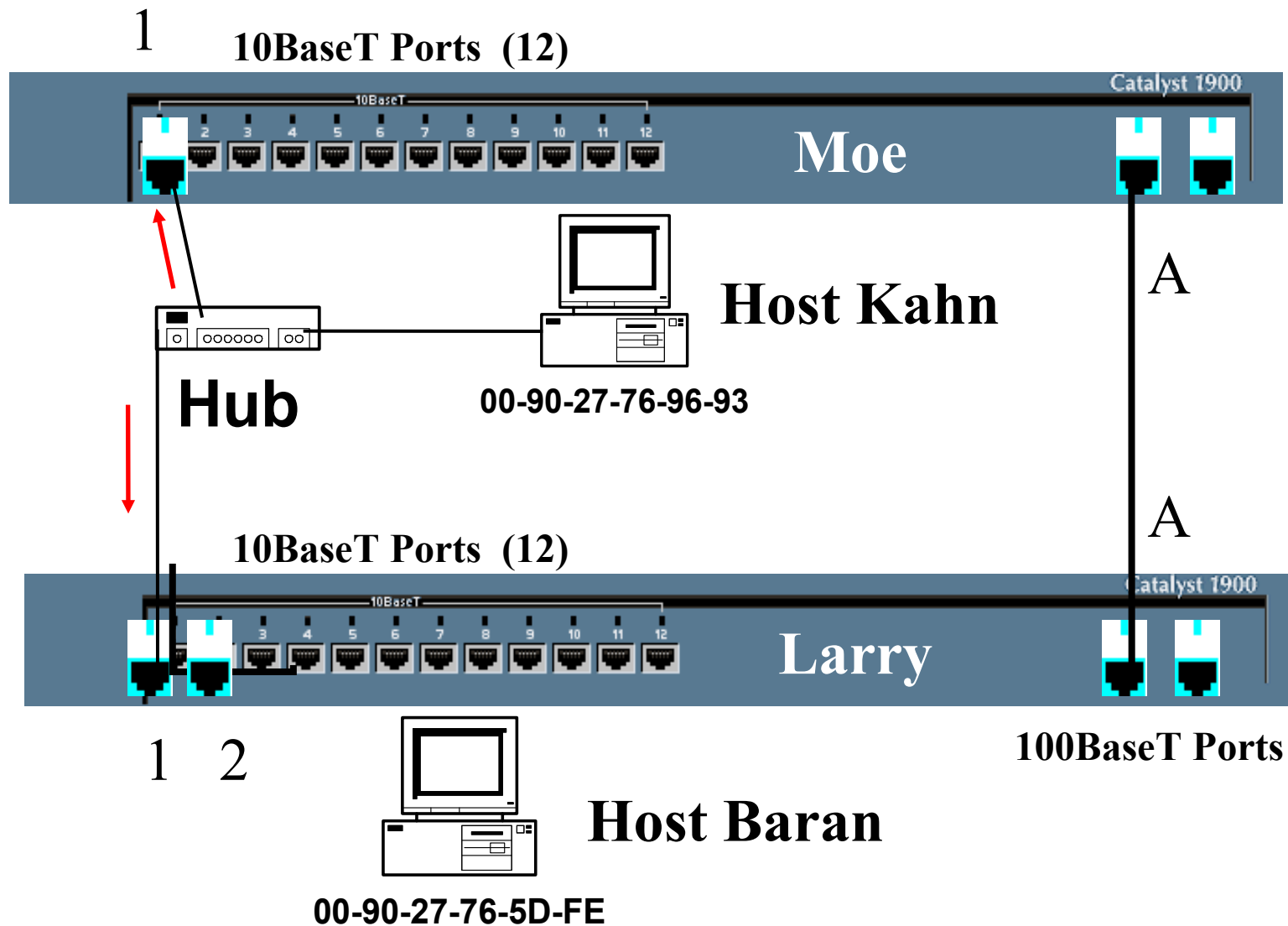


Redundant Paths and No Spanning Tree

- Then, the same confusion happens, but this time with Host Baran.
- Okay, maybe not the end of the world.
- Frames will just take a longer path and you may also see other “unexpected results.”
- But what about broadcast frames, like ARP Requests?

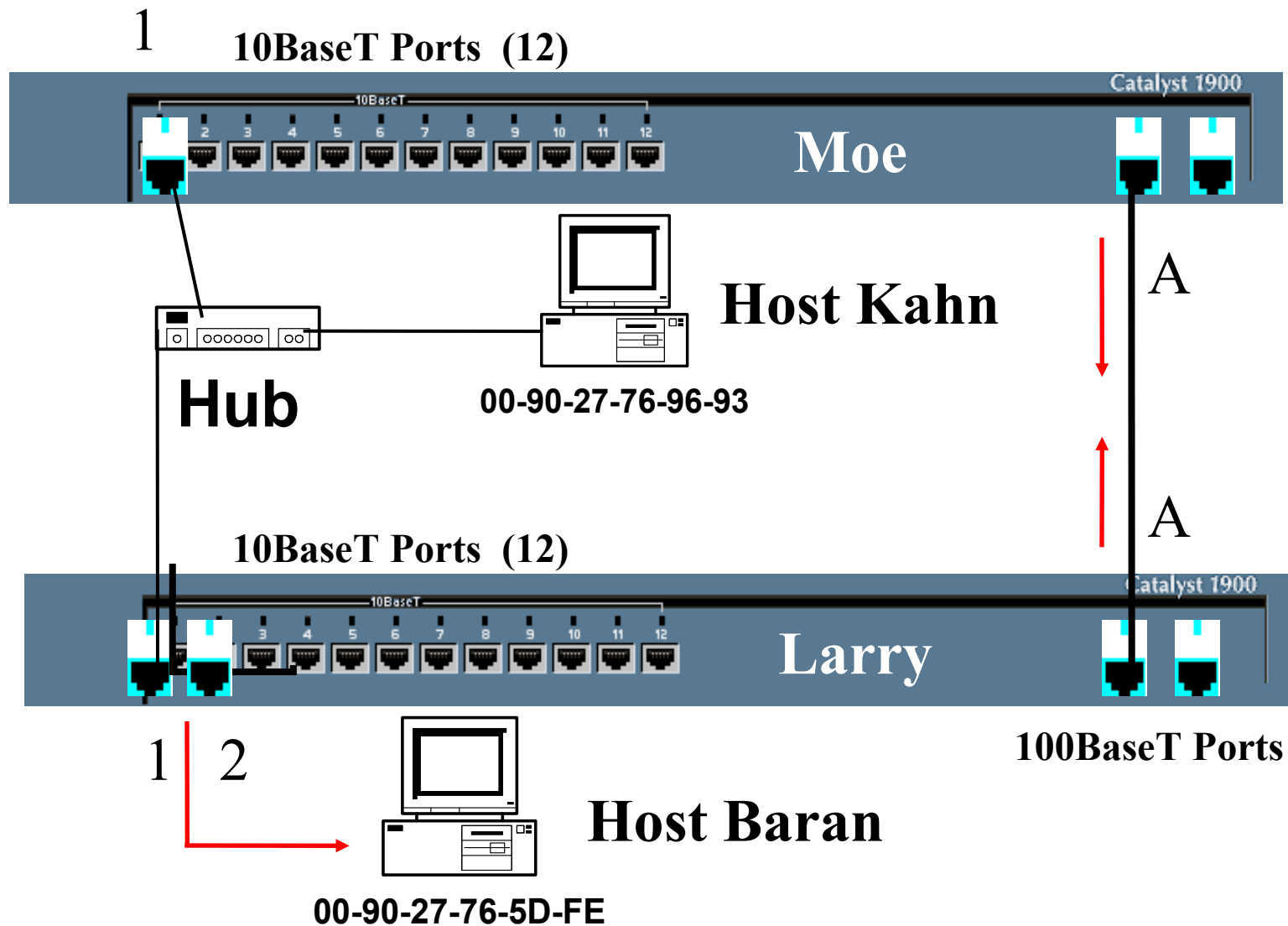
Broadcasts and No Spanning Tree

Lets, leave the switching tables alone and just look at what happens with the frames. Host Kahn sends out a layer 2 broadcast frame, like an ARP Request.



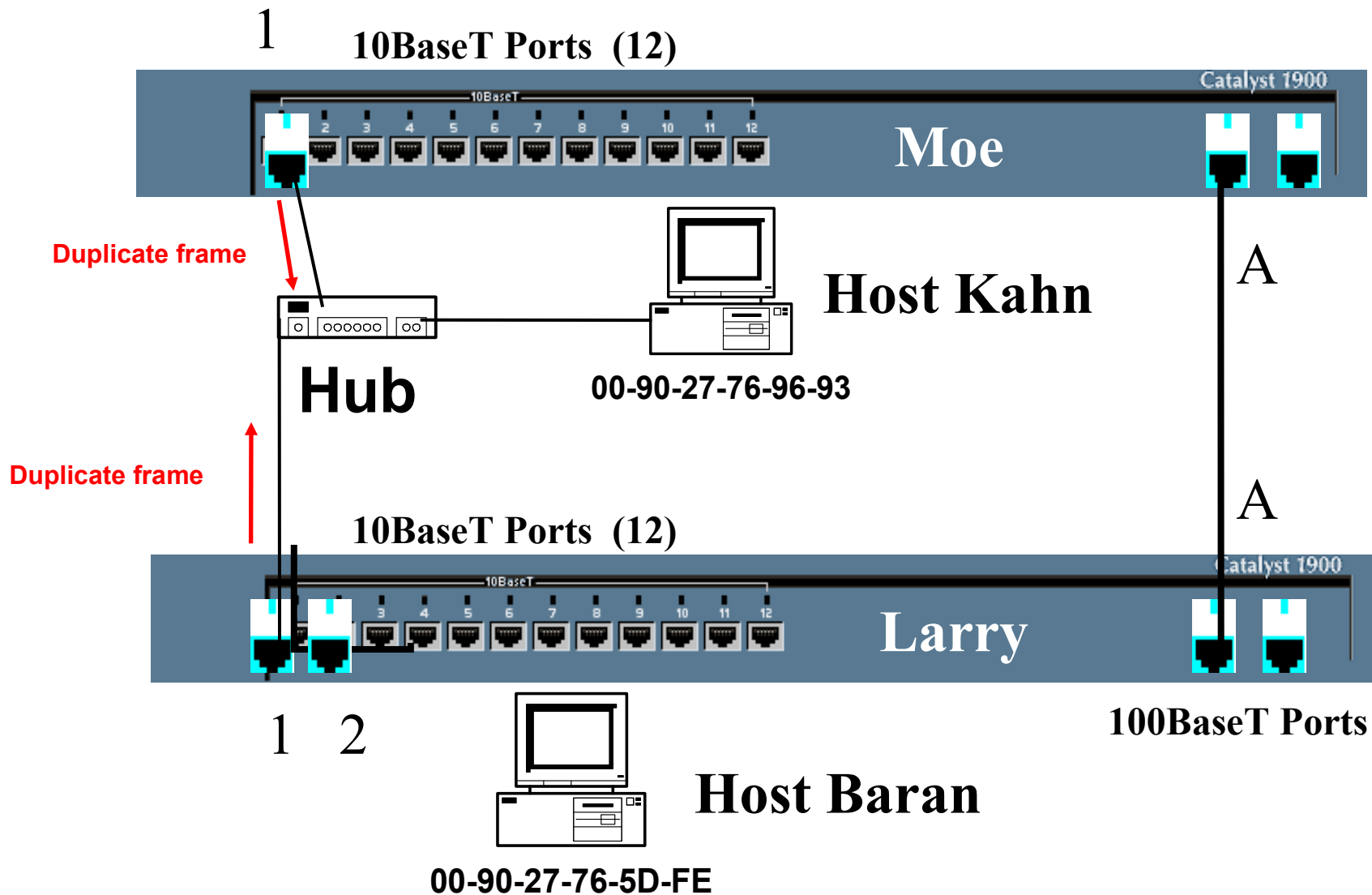
Broadcasts and No Spanning Tree

Because it is a layer 2 broadcast frame, both switches, Moe and Larry, **flood the frame out all ports**, including their port A's.



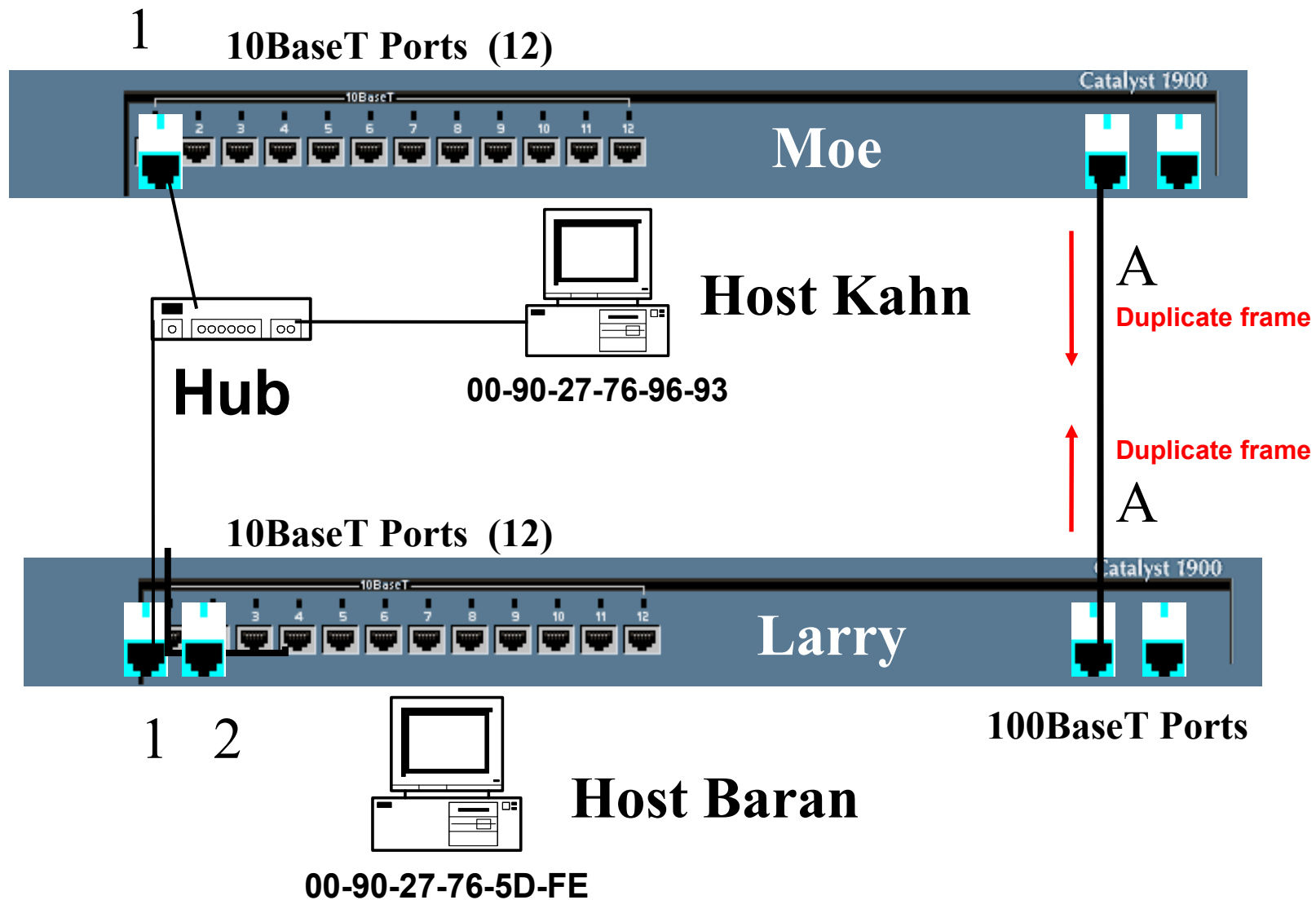
Broadcasts and No Spanning Tree

Both switches receive the same broadcast, but on a different port. Doing what switches do, **both switches flood the duplicate broadcast frame out their other ports.**



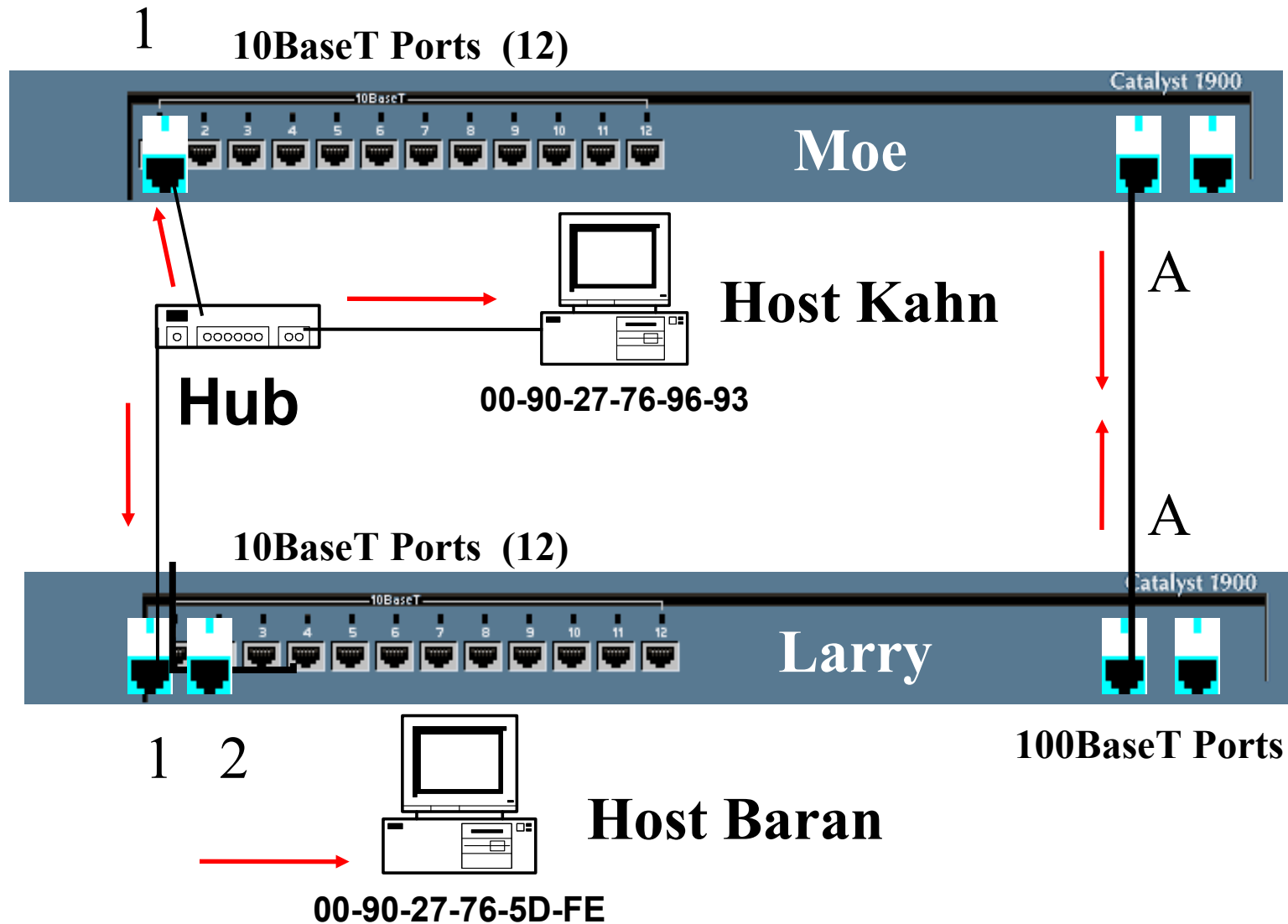
Broadcasts and No Spanning Tree

Here we go again, with the switches flooding the same broadcast again out its other ports. This results in duplicate frames, known as a ***broadcast storm!***

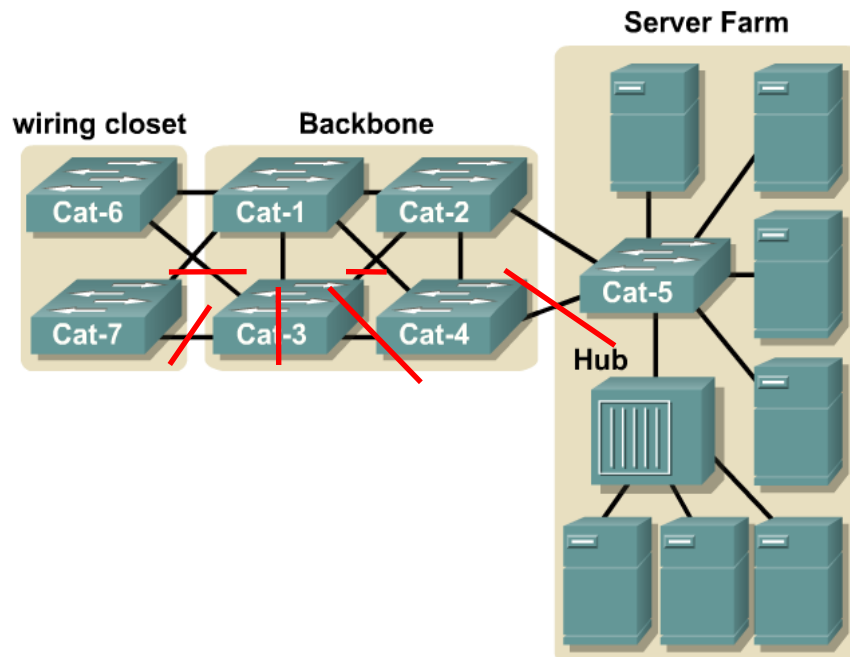
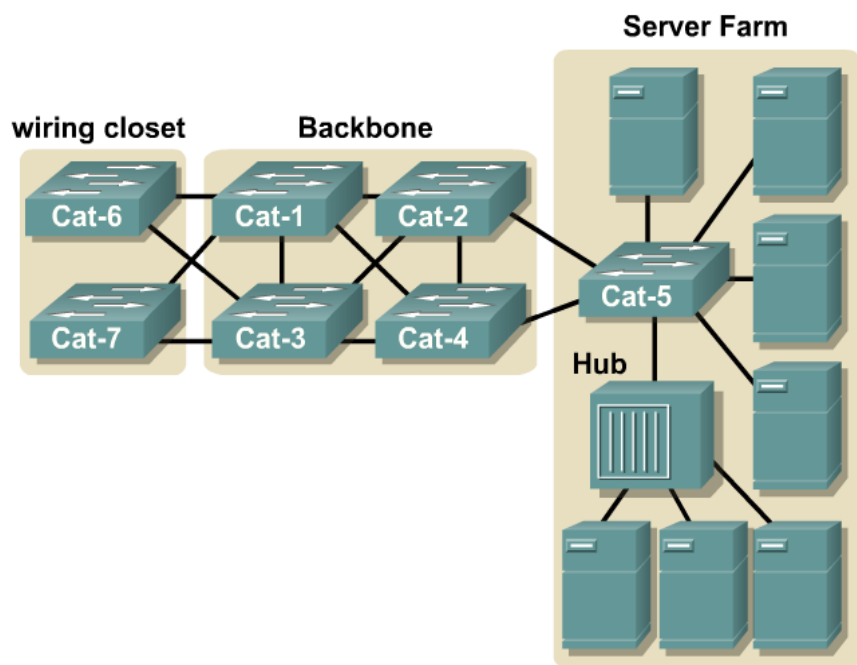


Broadcasts and No Spanning Tree

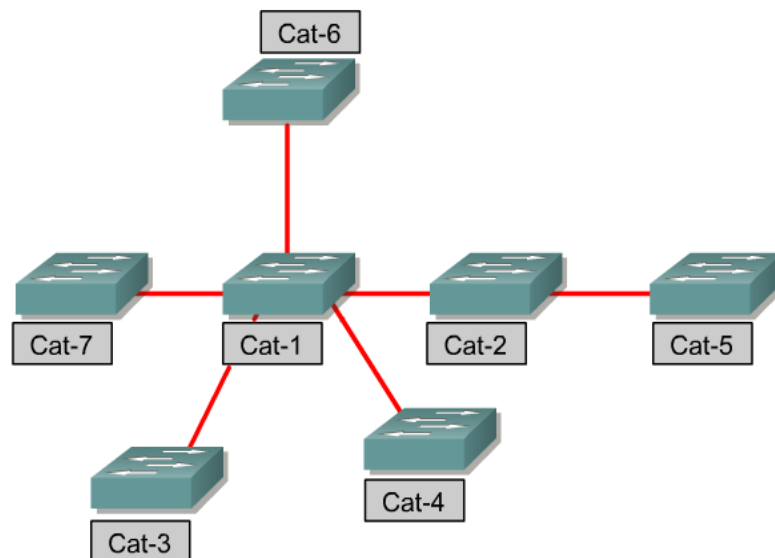
Remember, that layer 2 broadcasts not only take up network bandwidth, but must be processed by each host. This can severely impact a network, to the point of making it unusable.



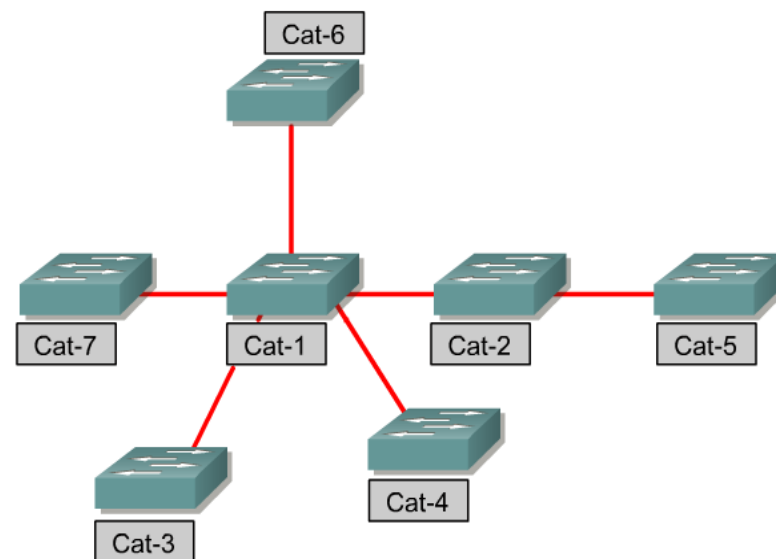
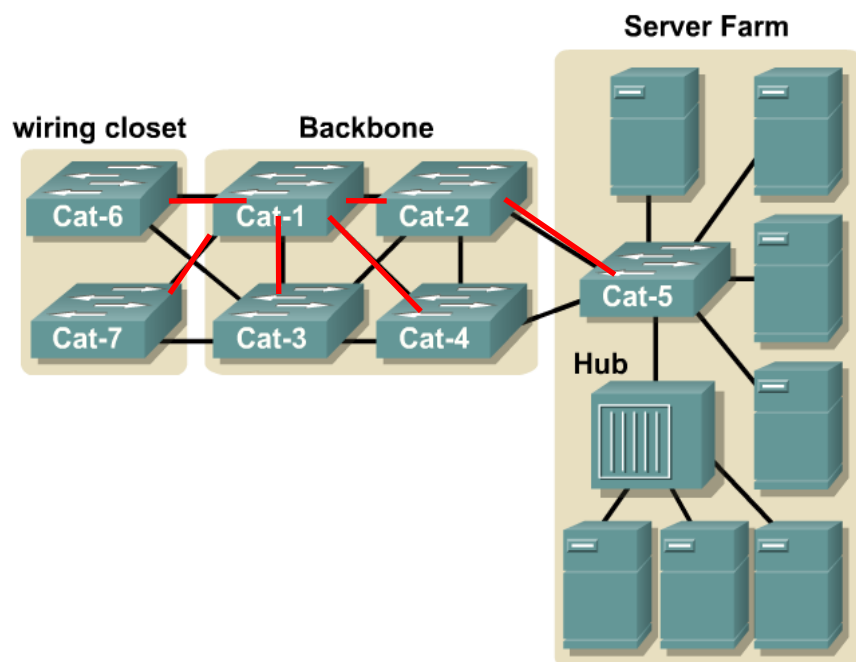
Redundant topology and spanning tree



- Unlike IP, in the Layer 2 header there is no Time To Live (TTL).
- The solution is to allow physical loops, but create a loop free logical topology.
- 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.



Redundant topology and spanning tree

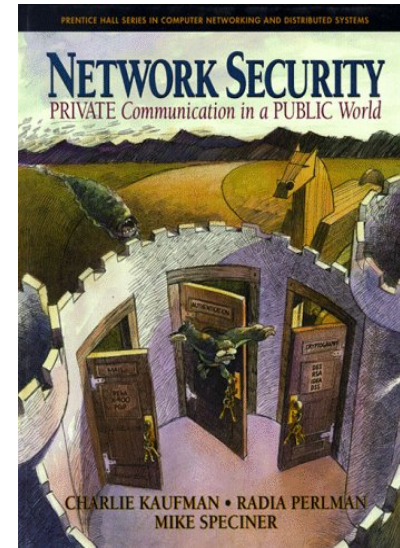
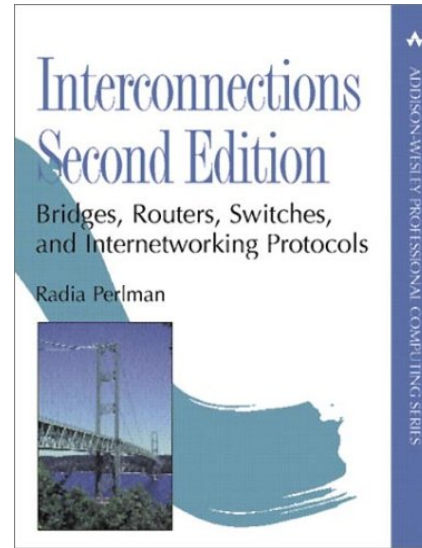


- It is a spanning tree because all devices in the network are reachable or spanned.
- The algorithm used to create this loop free logical topology is the **spanning-tree algorithm**.
- This algorithm can take a relatively long time to converge.
- A new algorithm called the **rapid spanning-tree algorithm** is being introduced to reduce the time for a network to compute a loop free logical topology. (later)

Spanning-Tree Protocol (STP)

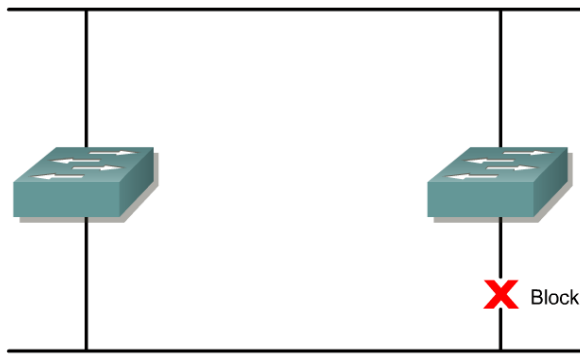


**Radia Perlman,
networking hero!**

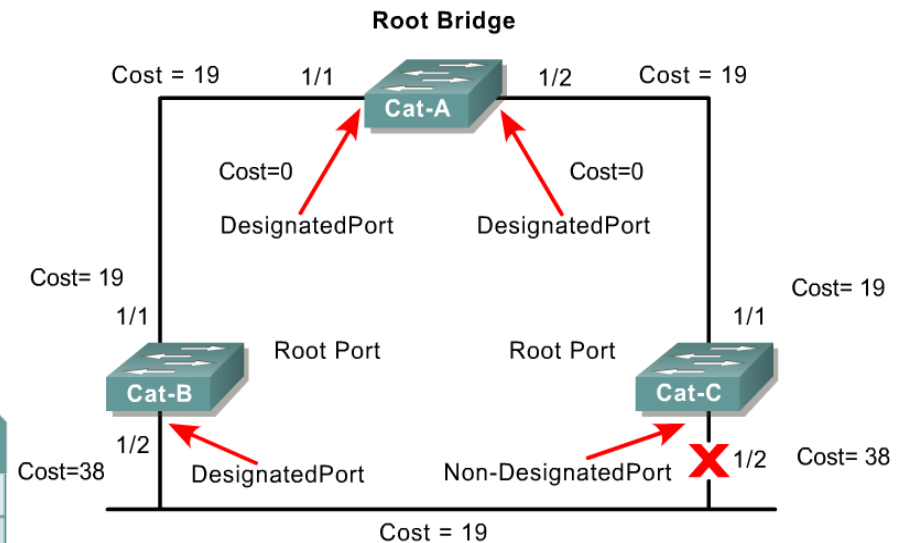


- Ethernet bridges and switches can implement the **IEEE 802.1D Spanning-Tree Protocol** and use the spanning-tree algorithm to **construct a loop free shortest path network**.
- Radia Perlman “is the inventor of the spanning tree algorithm used by bridges (switches), and the mechanisms that make link state routing protocols such as IS-IS (which she designed) and OSPF (which adopted many of the ideas) stable and efficient. Her thesis on sabotage-proof networks is well-known in the security community.”
<http://www.equipecom.com/radia.html>

Spanning-Tree Protocol (STP)



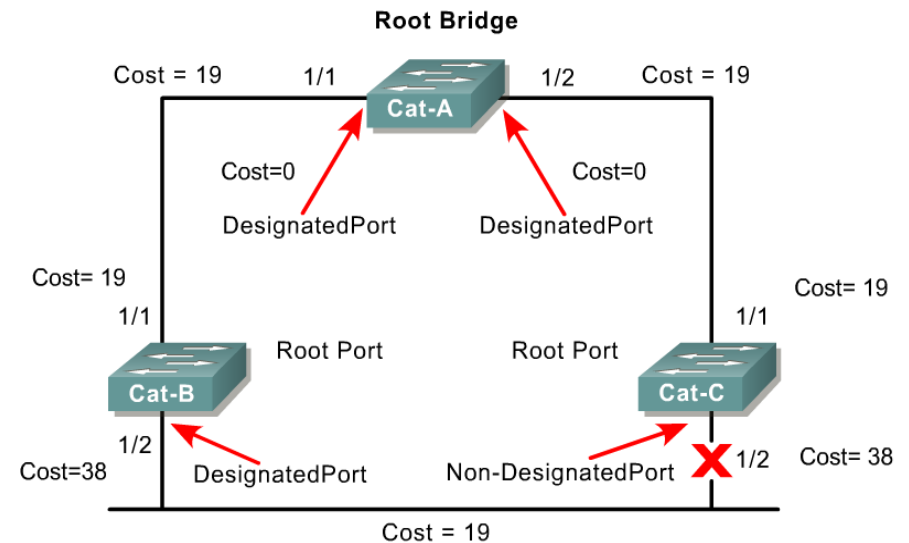
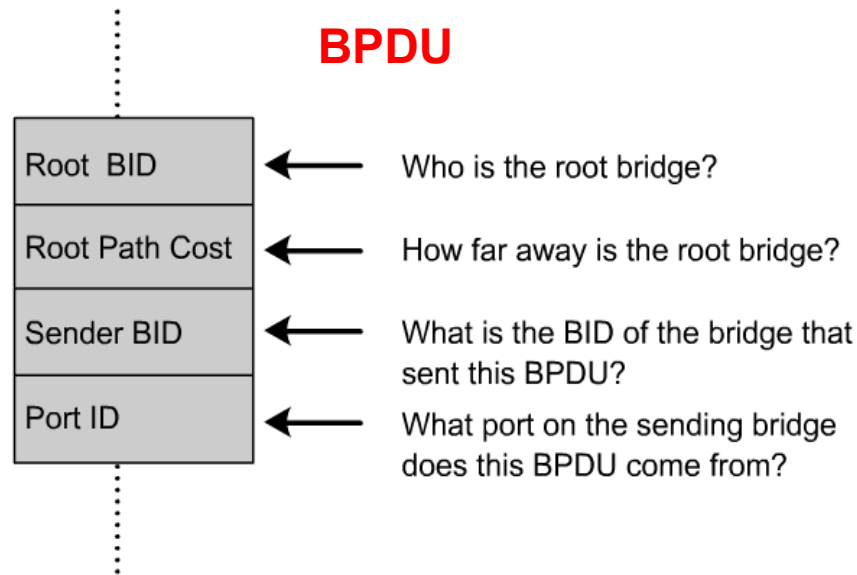
Link Speed	Cost(Revised IEEE Spec)	Cost (Previous IEEE Spec)
10 Gbps	2	1
1 Gbps	4	1
100 Mbps	19	10
10 Mbps	100	100



We will see how this works in a moment.

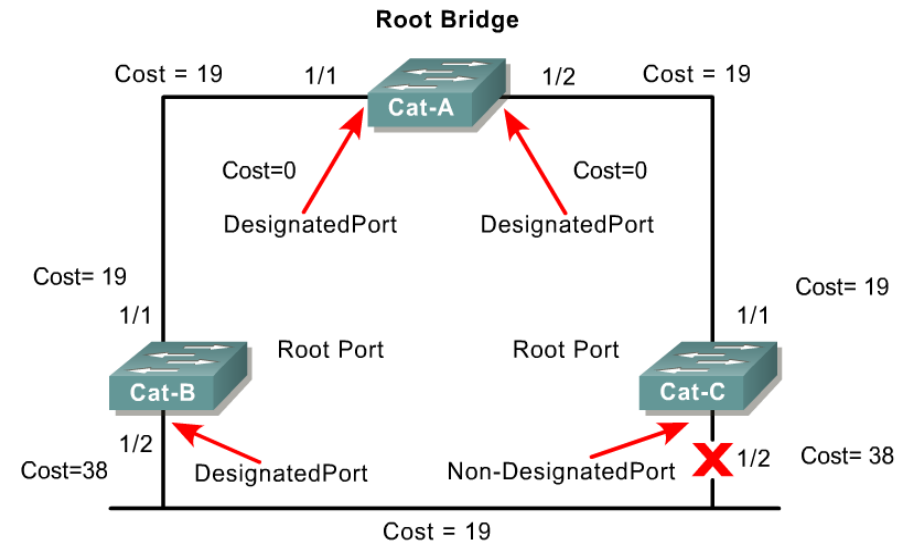
- Shortest path is based on cumulative link costs.
- Link costs are based on the speed of the link.
- The Spanning-Tree Protocol establishes a root node, called the root bridge.
- The Spanning-Tree Protocol constructs a topology that has one path for reaching every network node.
- The resulting tree originates from the **root bridge**.
- **Redundant links** that are not part of the shortest path tree are **blocked**.

Spanning-Tree Protocol (STP)



- It is because certain paths are blocked that a loop free topology is possible.
- Data frames received on blocked links are dropped.
- 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.
- 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.

Spanning-Tree Protocol (STP)



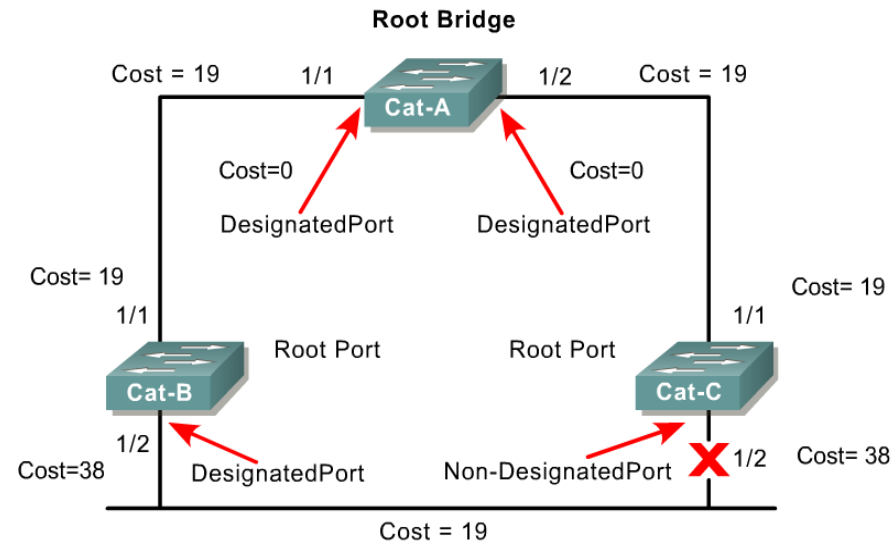
BPDUs contain enough information so that all switches can do the following:

- 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.

Let's see how this is done!

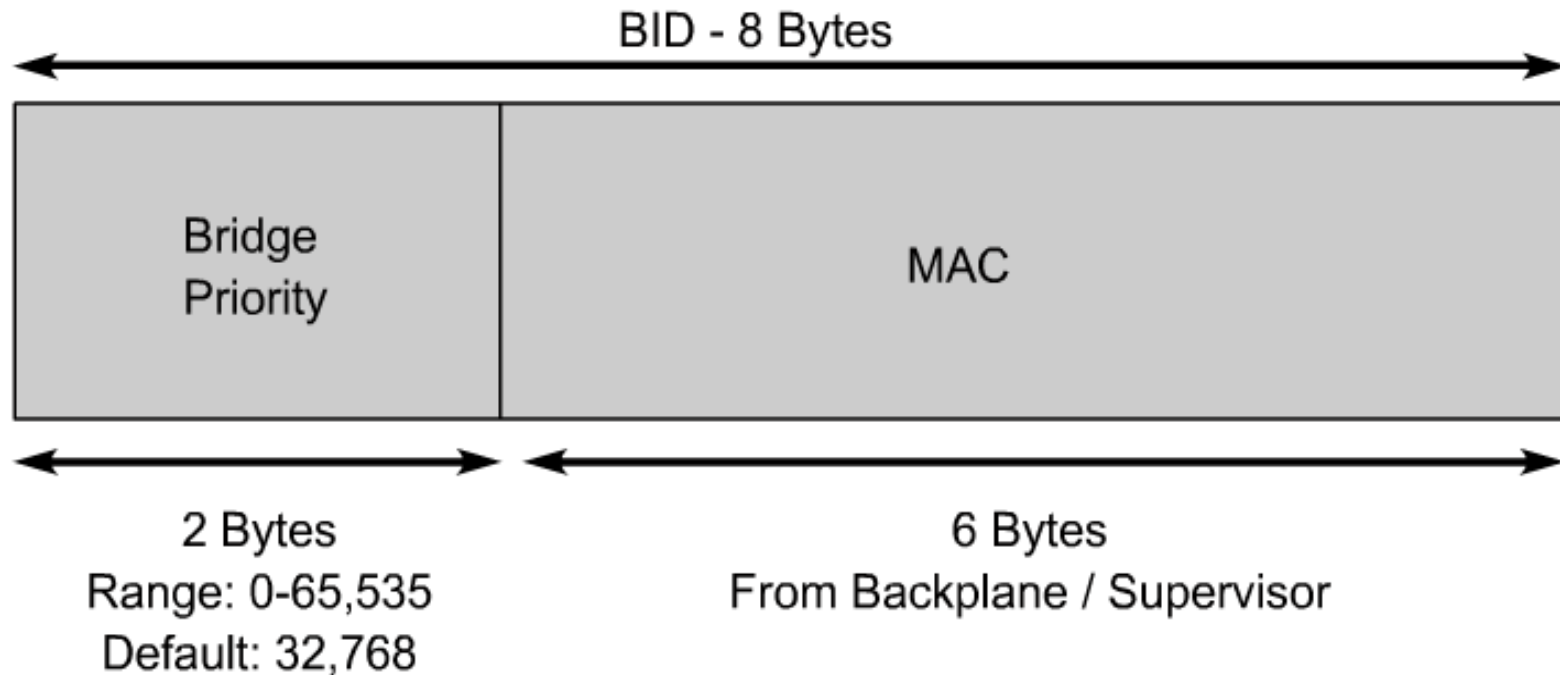
Some of this is extra information or information explained that is not explained fully in the curriculum.

Two Key Concepts: BID and Path Cost



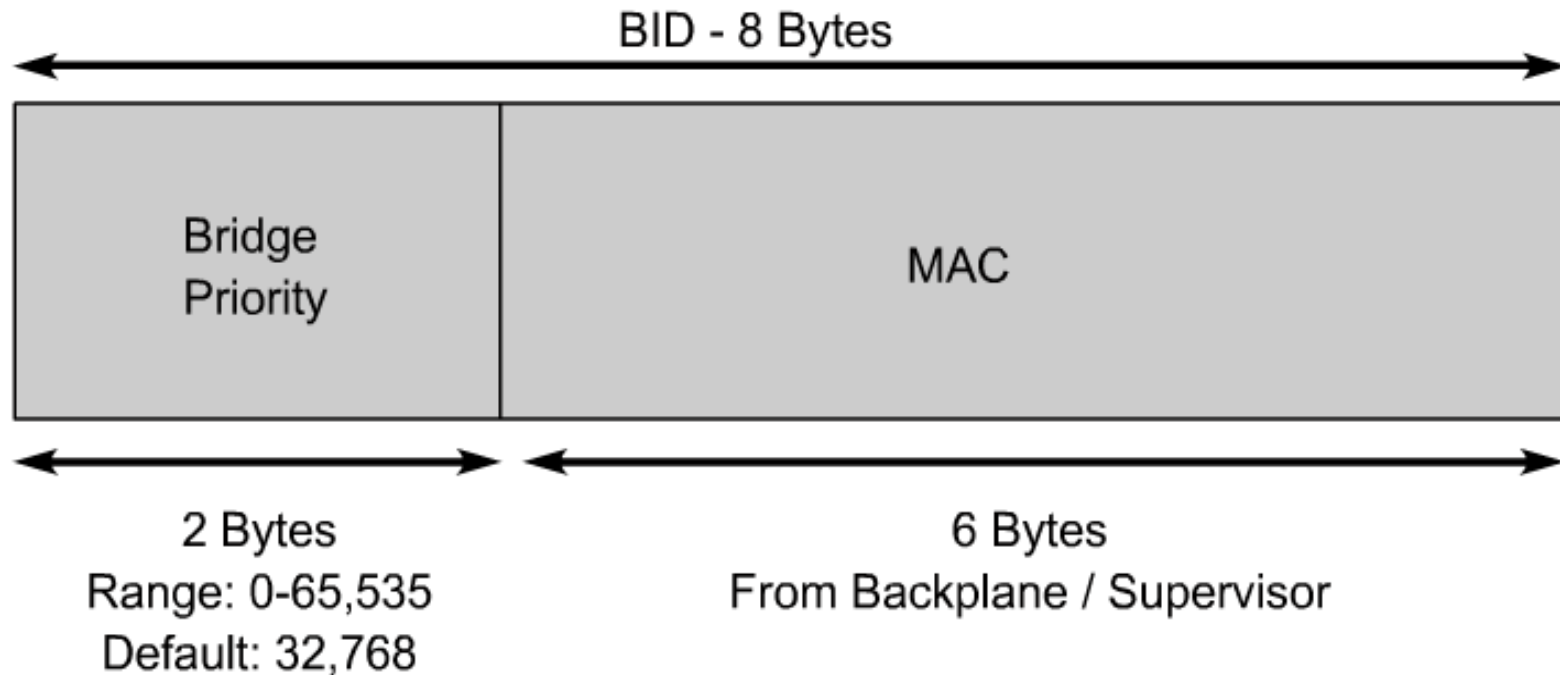
- STP executes an algorithm called Spanning Tree Algorithm (STA).
- STA chooses a reference point, called a root bridge, and then determines the available paths to that reference point.
 - If more than two paths exists, STA picks the best path and blocks the rest
- STP calculations make extensive use of two key concepts in creating a loop-free topology:
 - **Bridge ID**
 - **Path Cost**

Bridge ID (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.
- Consists of two components:
 - **A 2-byte Bridge Priority:** Cisco switch defaults to **32,768** or 0x8000.
 - **A 6-byte MAC address**

Bridge ID (BID)



- **Bridge Priority** is usually expressed in **decimal format** and the **MAC address** in the BID is usually expressed in **hexadecimal format**.
- BID is used to elect a root bridge (coming)
- **Lowest Bridge ID is the root.**
- If all devices have the same priority, the bridge with the lowest MAC address becomes the root bridge. (Yikes!)

Path Cost

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

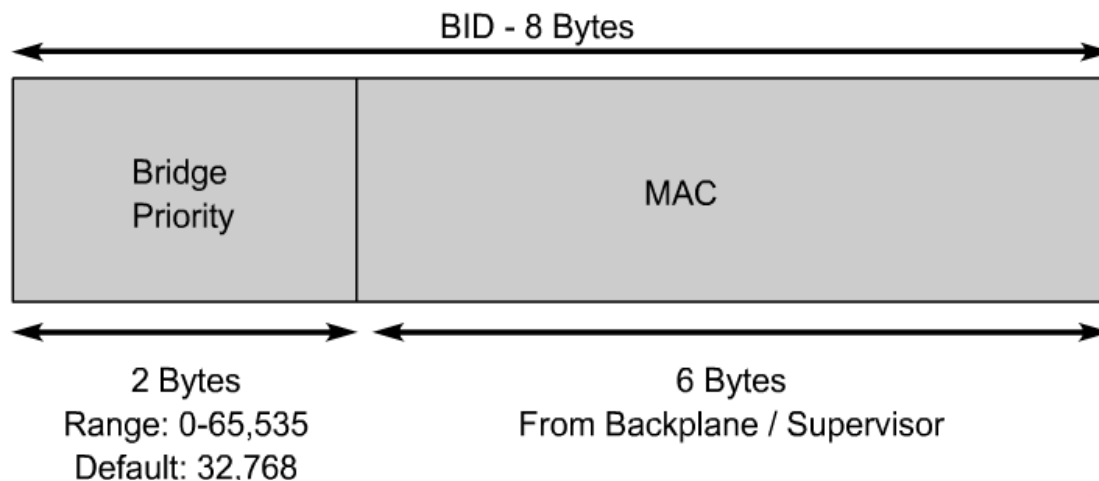
- Bridges use the concept of cost to evaluate how close they are to other bridges.
- This will be used in the STP development of a loop-free topology .
- **Originally, 802.1d** defined cost as $1000/\text{bandwidth of the link in Mbps}$.
 - Cost of 10Mbps link = 100 or $1000/10$
 - Cost of 100Mbps link = 10 or $1000/100$
 - Cost of 1Gbps link = 1 or $1000/1000$
- Running out of room for faster switches including 10 Gbps Ethernet.

Path Cost

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

- IEEE modified the most to use a non-linear scale with the new values of:
 - 4 Mbps 250 (cost)
 - 10 Mbps 100 (cost)
 - 16 Mbps 62 (cost)
 - 45 Mbps 39 (cost)
 - 100 Mbps 19 (cost)
 - 155 Mbps 14 (cost)
 - 622 Mbps 6 (cost)
 - 1 Gbps 4 (cost)
 - 10 Gbps 2 (cost)

Path Cost



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

- You can modify the path cost by modifying the cost of a port.
 - Exercise caution when you do this!
- BID and Path Cost are used to develop a loop-free topology .
- Coming very soon!
- But first the **Four-Step STP Decision Sequence**

Four-Step STP Decision Sequence

- When creating a loop-free topology, STP always uses the same four-step decision sequence:

Four-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID

- Bridges use Configuration BPDUs during this four-step process.
 - There is another type of BPDU known as Topology Change Notification (TCN) BPDU.

Four-Step STP Decision Sequence

BPDU key concepts:

- Bridges save a copy of only the best BPDU seen on every port.
- When making this evaluation, it considers all of the BPDUs received on the port, as well as the BPDU that would be sent on that port.
- As every BPDU arrives, it is checked against this four-step sequence to see if it is more attractive (lower in value) than the existing BPDU saved for that port.
- Only the lowest value BPDU is saved.
- Bridges send configuration BPDUs until a more attractive BPDU is received.
- Okay, lets see how this is used...

Four-Step STP Decision Sequence

BPDU key concepts:

- Bridges save a copy of only the best BPDU seen on every port.
- When making this evaluation, it considers all of the BPDUs received on the port, as well as the BPDU that would be sent on that port.
- As every BPDU arrives, it is checked against this four-step sequence to see if it is more attractive (lower in value) than the existing BPDU saved for that port.
- Only the lowest value BPDU is saved.
- Bridges send configuration BPDUs until a more attractive BPDU is received.
- Okay, lets see how this is used...

Three Steps of Initial STP Convergence

- The STP algorithm uses three simple steps to converge on a loop-free topology.
- Switches go through three steps for their initial convergence:

STP Convergence

Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

- All STP decisions are based on a the following predetermined sequence:

Four-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID

Three Steps of Initial STP Convergence

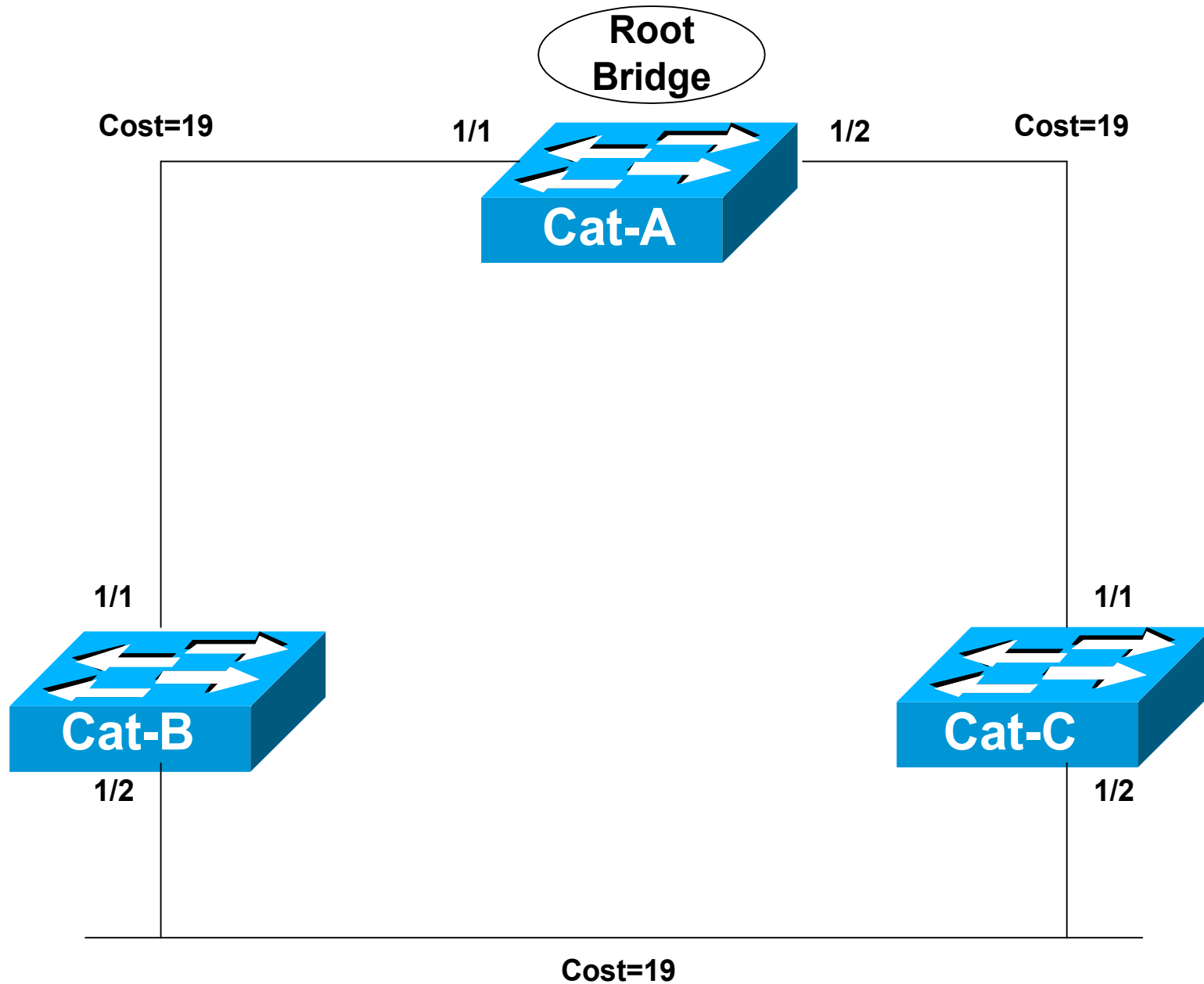
STP Convergence

Step 1 Elect one Root Bridge

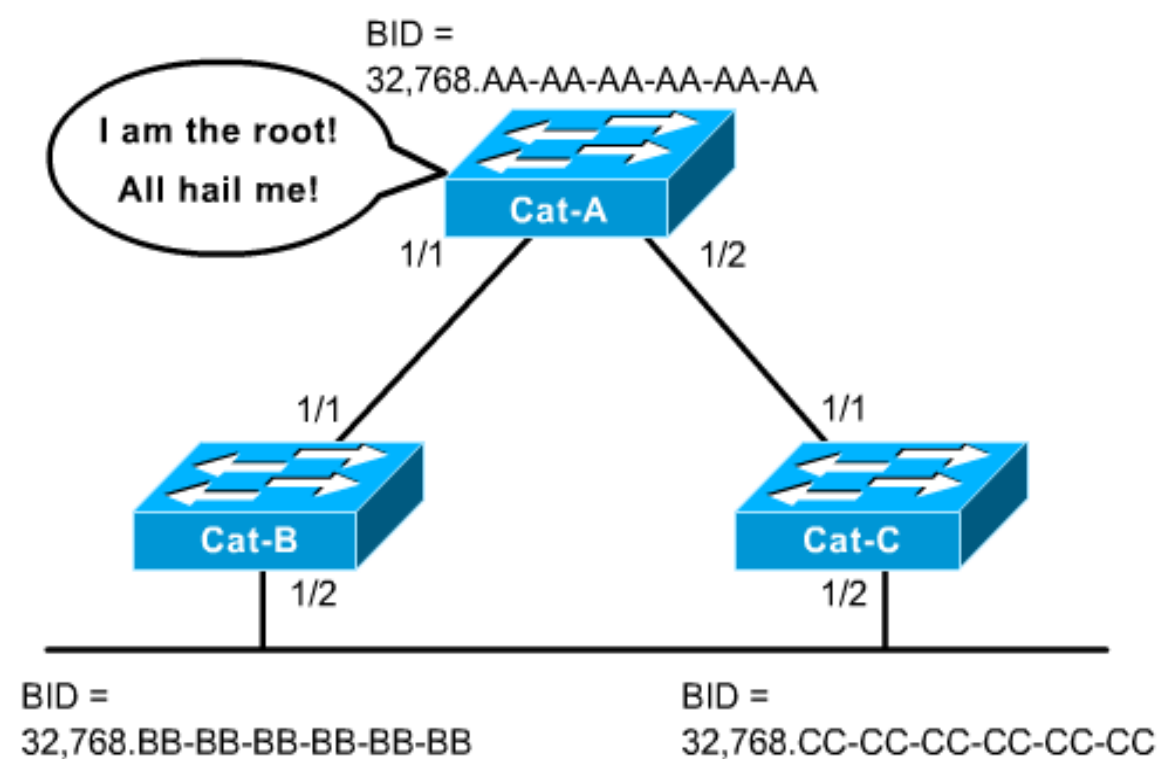
Step 2 Elect Root Ports

Step 3 Elect Designated Ports

Step 1 Elect one Root Bridge



Step 1 Elect one Root Bridge



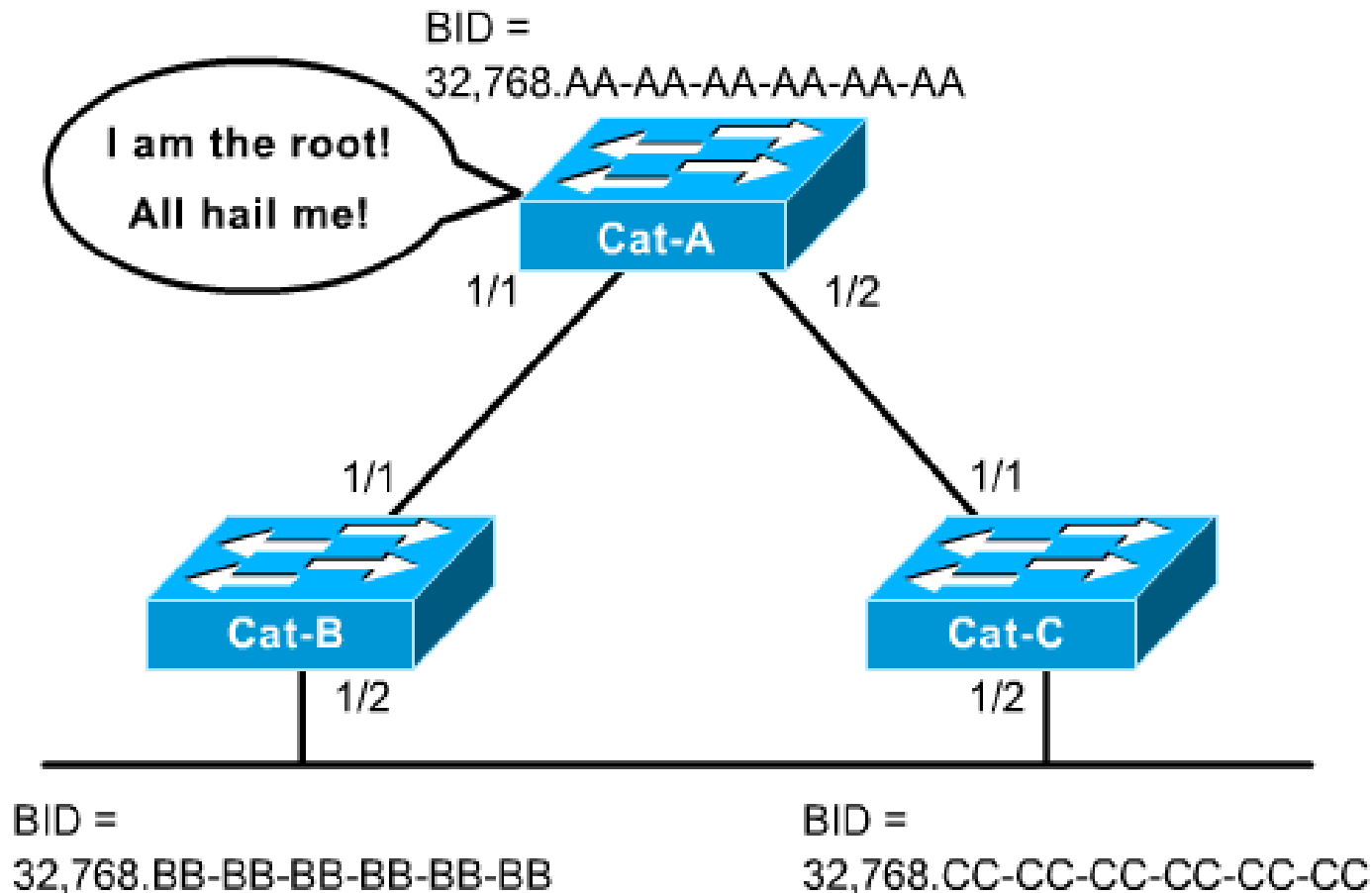
- When the network first starts, all bridges are announcing a chaotic mix of BPDUs.
- All bridges immediately begin applying the four-step sequence decision process.
- Switches need to elect a single Root Bridge.
- Switch with the **lowest BID** wins!

Note: Many texts refer to the term “highest priority” which is the “lowest” BID value.

- This is known as the “Root War.”

Step 1 Elect one Root Bridge

Cat-A has the lowest Bridge MAC Address, so it wins the Root War!



All 3 switches have the same default Bridge Priority value of 32,768

Step 1 Elect one Root Bridge

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

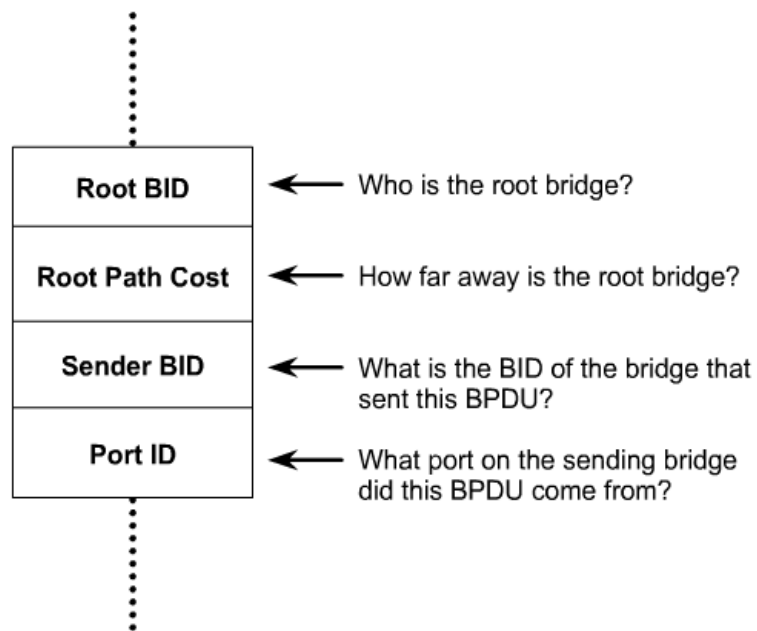
802.2 Logical Link Control (LLC) Header

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

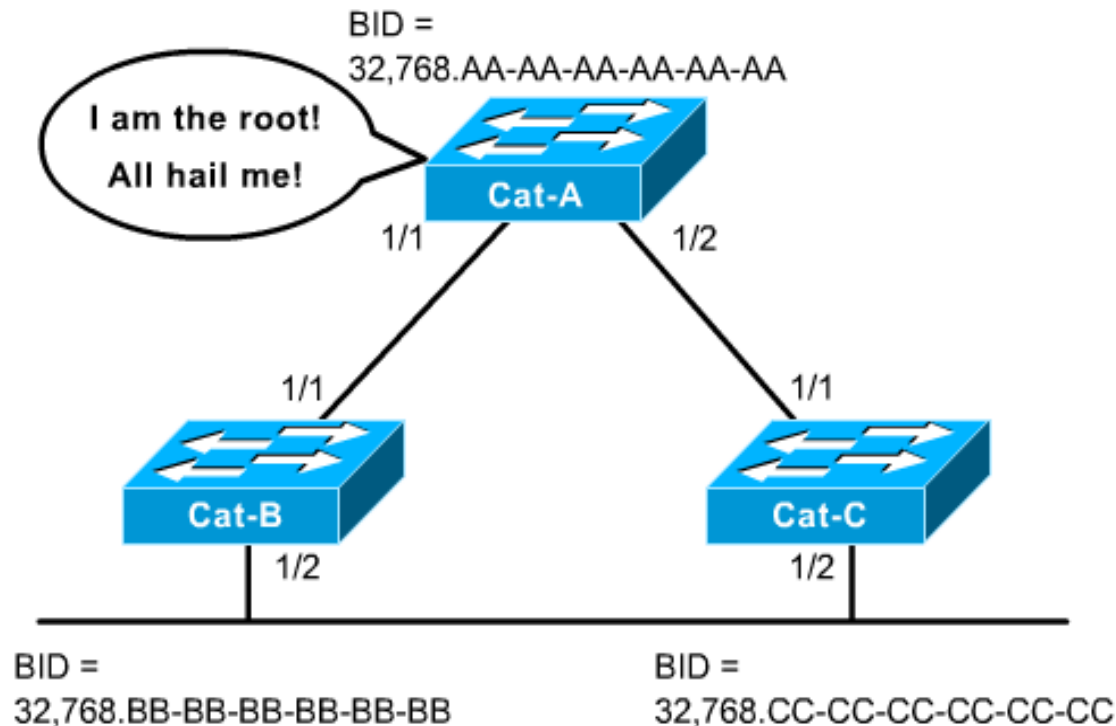
Protocol Identifier: 0
Protocol Version ID: 0
Message Type: 0 *Configuration Message*
Flags: %00000000
Root Priority/ID: 0x8000/ 00:D0:C0:F5:18:C0
Cost Of Path To Root: 0x00000000 (0)
Bridge Priority/ID: 0x8000/ 00:D0:C0:F5:18:C0
Port Priority/ID: 0x80/ 0x1D
Message Age: 0/256 seconds (*exactly 0 seconds*)
Maximum Age: 5120/256 seconds (*exactly 20 seconds*)
Hello Time: 512/256 seconds (*exactly 2 seconds*)
Forward Delay: 3840/256 seconds (*exactly 15 seconds*)

Its all done with BPDUs!



Configuration BPDUs are sent every 2 seconds by default.

Step 1 Elect one Root Bridge



- At the beginning, all bridges assume they are the center of the universe and declare themselves as the Root Bridge, by placing its own BID in the Root BID field of the BPDU.
- Once all of the switches see that Cat-A has the lowest BID, they are all in agreement that Cat-A is the Root Bridge.

Three Steps of Initial STP Convergence

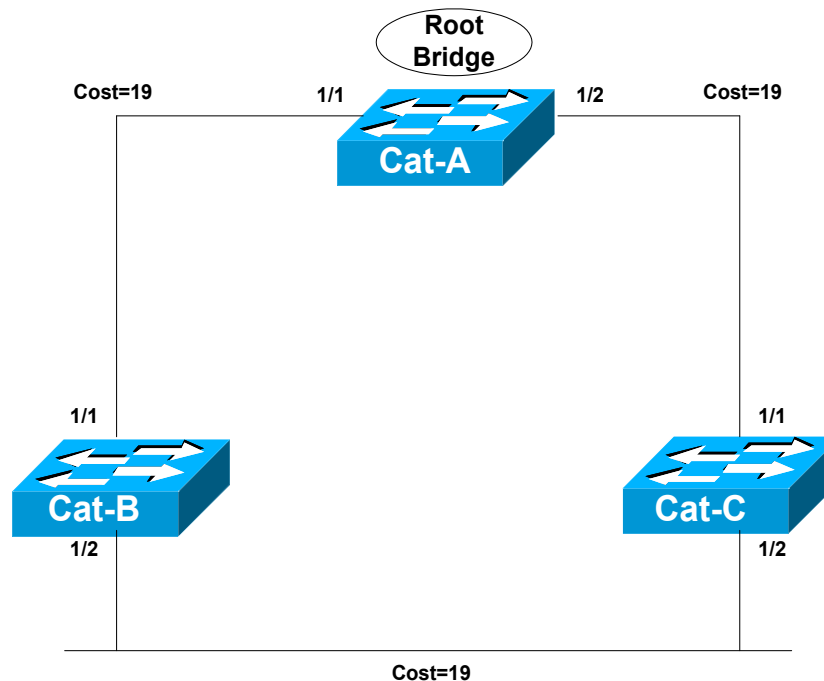
STP Convergence

Step 1 Elect one Root Bridge

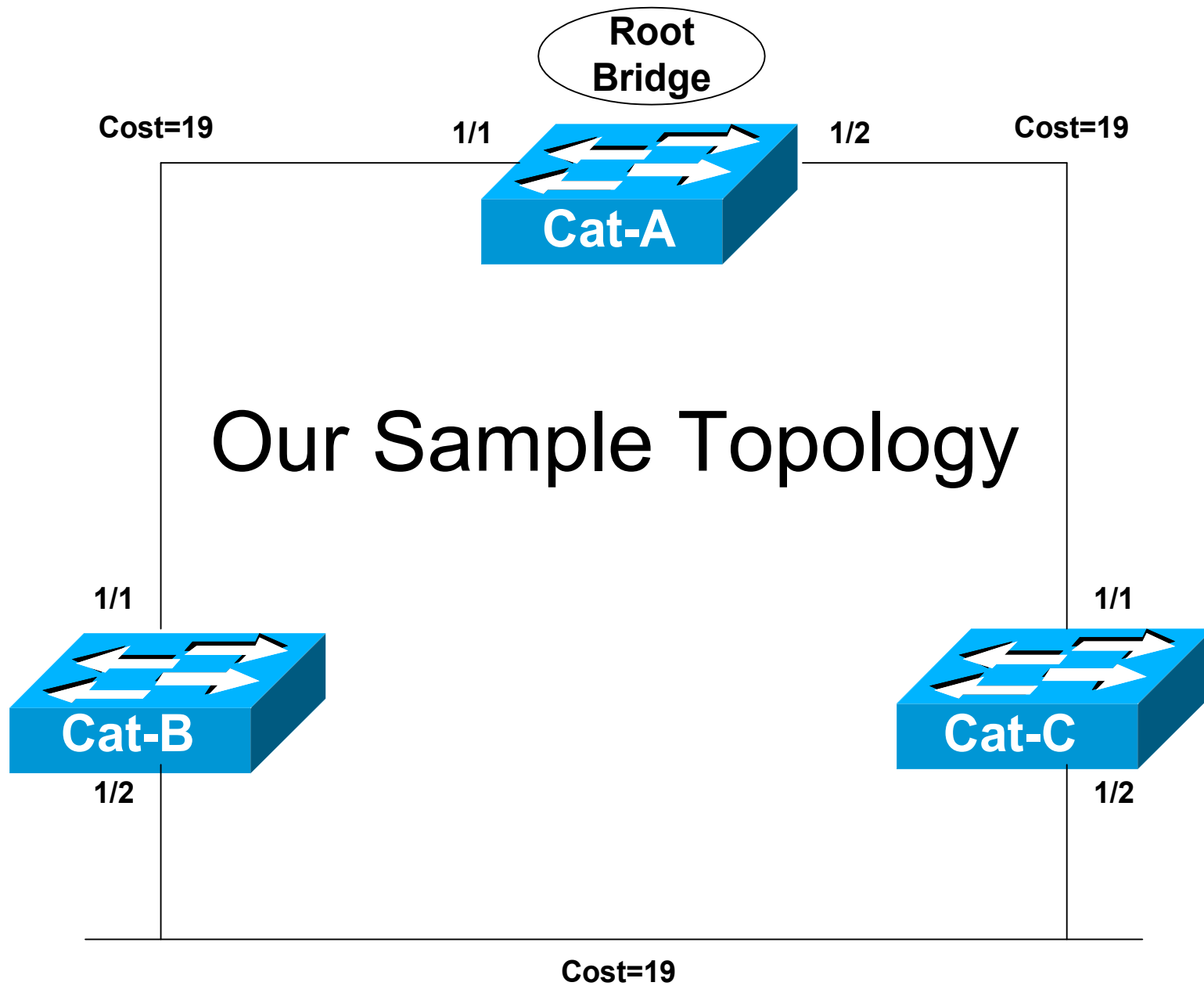
Step 2 Elect Root Ports

Step 3 Elect Designated Ports

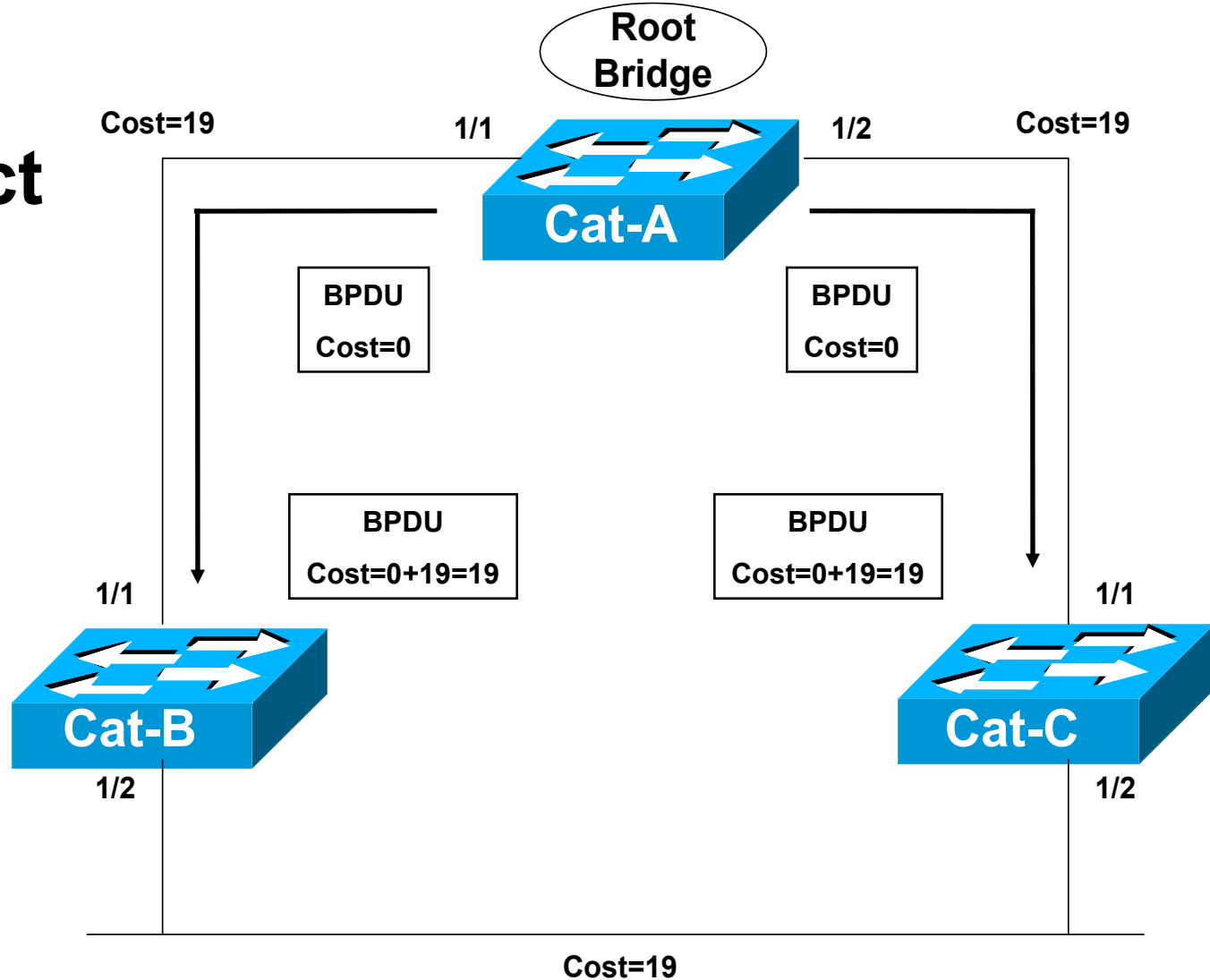
Step 2 Elect Root Ports



- Now that the Root War has been won, switches move on to selecting **Root Ports**.
- A bridge's **Root Port** is the *port closest to the Root Bridge*.
- Bridges use the **cost** to determine closeness.
- **Every non-Root Bridge will select one Root Port!**
- Specifically, bridges track the **Root Path Cost**, the cumulative cost of all links to the Root Bridge.



Step 2 Elect Root Ports



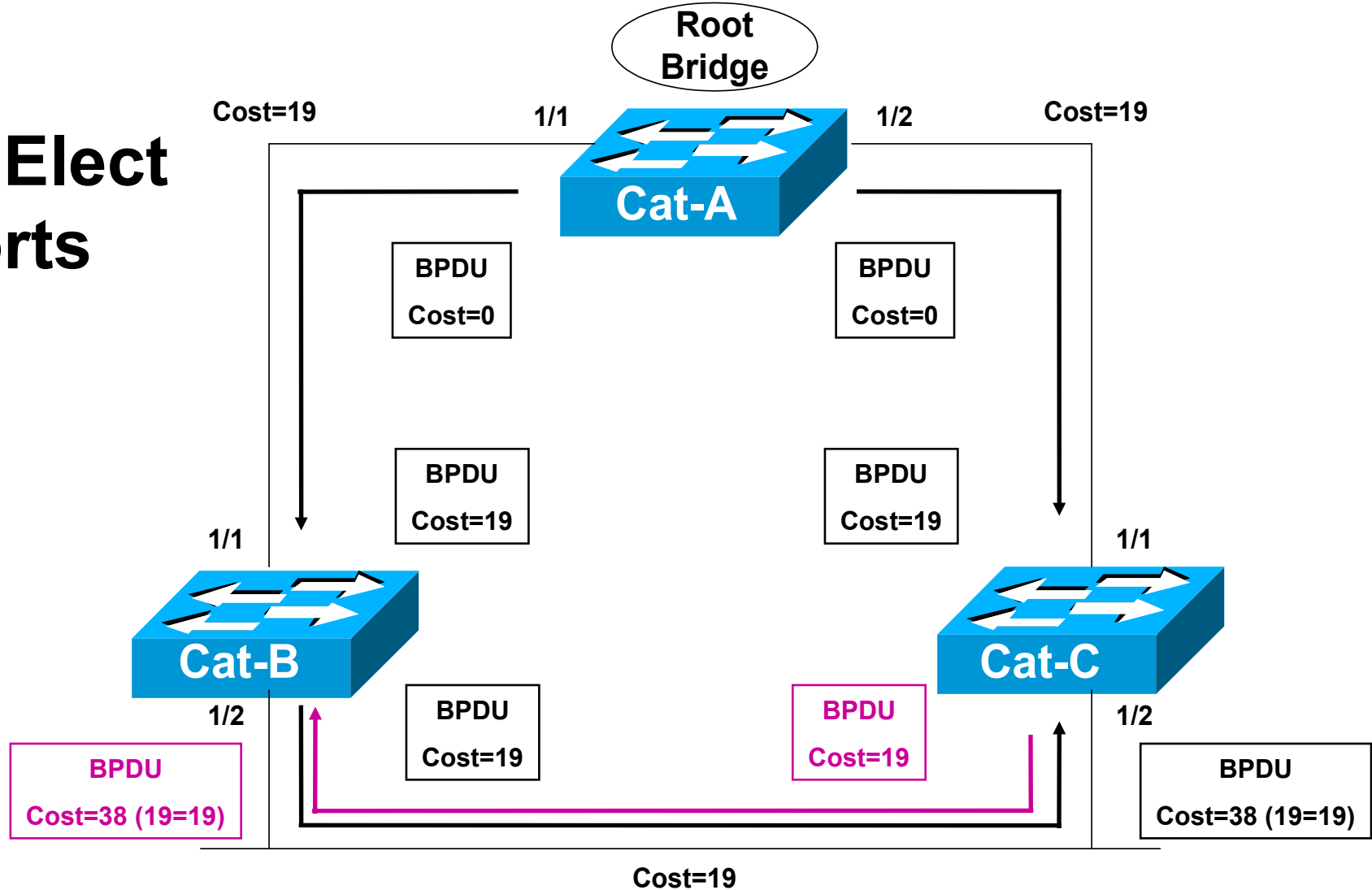
Step 1

- Cat-A sends out BPDUs, containing a Root Path Cost of 0.
- Cat-B receives these BPDUs and adds the Path Cost of Port 1/1 to the Root Path Cost contained in the BPDU.

Step 2

- Cat-B adds Root Path Cost 0 PLUS its Port 1/1 cost of 19 = 19

Step 2 Elect Root Ports



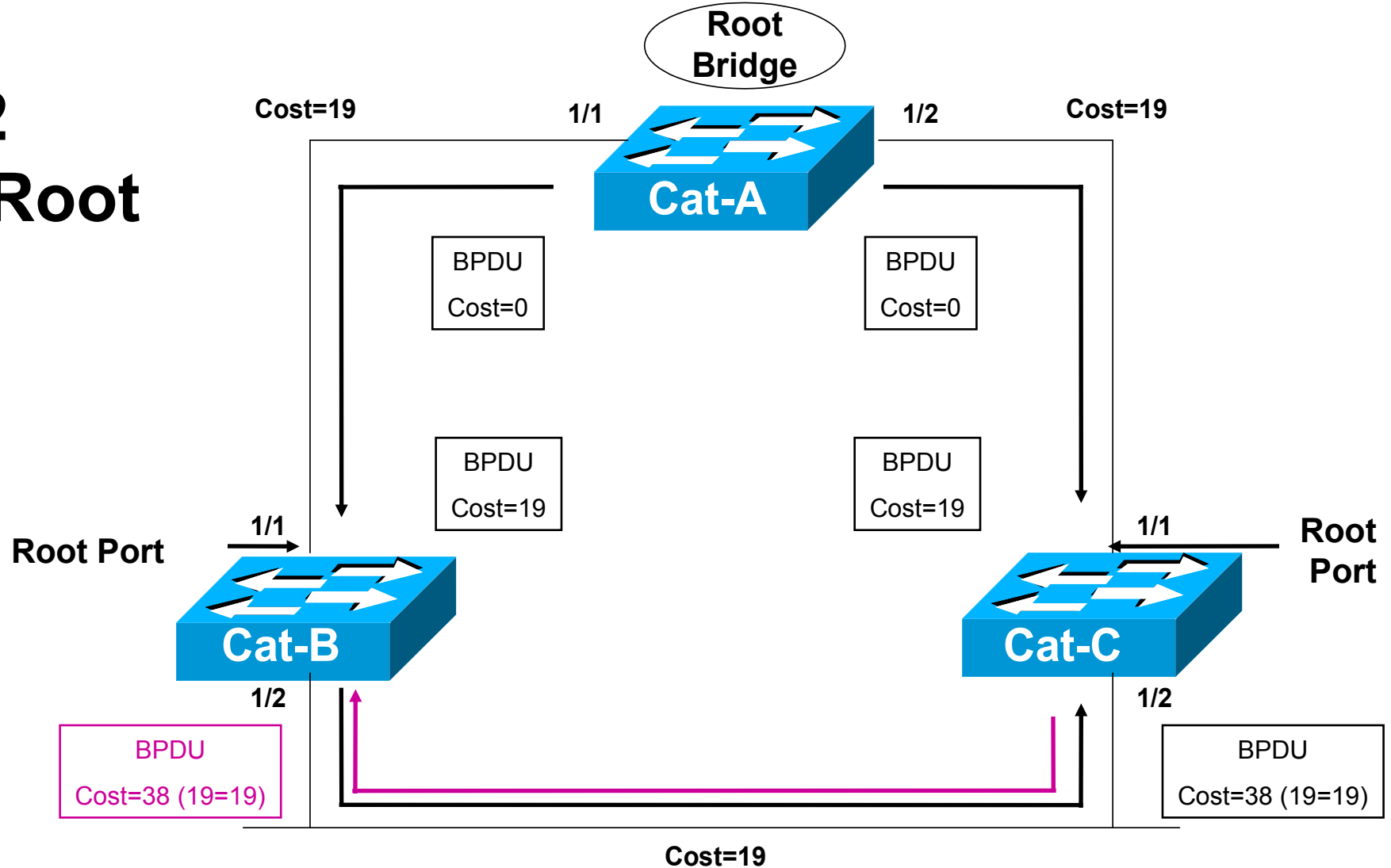
Step 3

- Cat-B uses this value of 19 internally and sends BPDUs with a Root Path Cost of 19 out Port 1/2.

Step 4

- Cat-C receives the BPDU from Cat-B, and increased the Root Path Cost to 38 (19+19). (Same with Cat-C sending to Cat-B.)

Step 2 Elect Root Ports



Step 5

- Cat-B calculates that it can reach the Root Bridge at a cost of 19 via Port 1/1 as opposed to a cost of 38 via Port 1/2.
- Port 1/1 becomes the Root Port for Cat-B, the port closest to the Root Bridge.
- Cat-C goes through a similar calculation. Note: Both Cat-B:1/2 and Cat-C:1/2 save the best BPDU of 19 (its own).

Three Steps of Initial STP Convergence

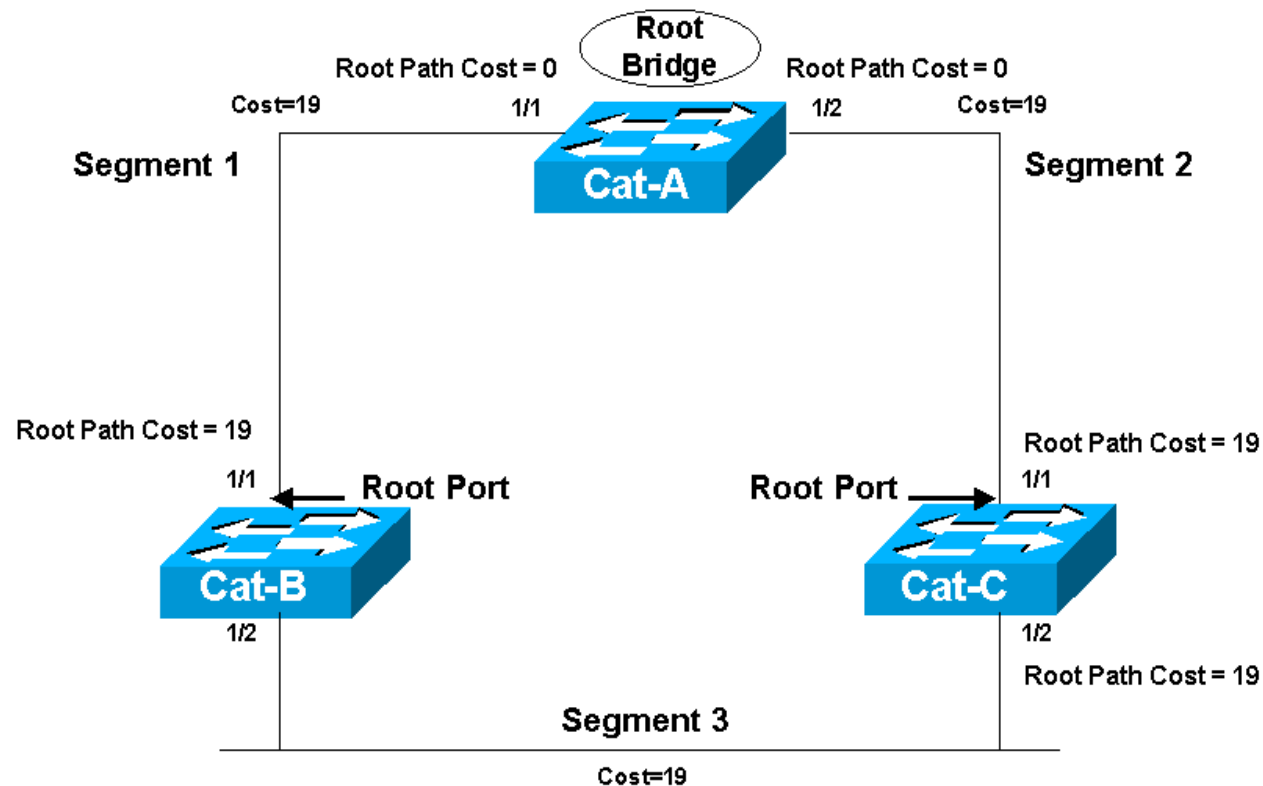
STP Convergence

Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

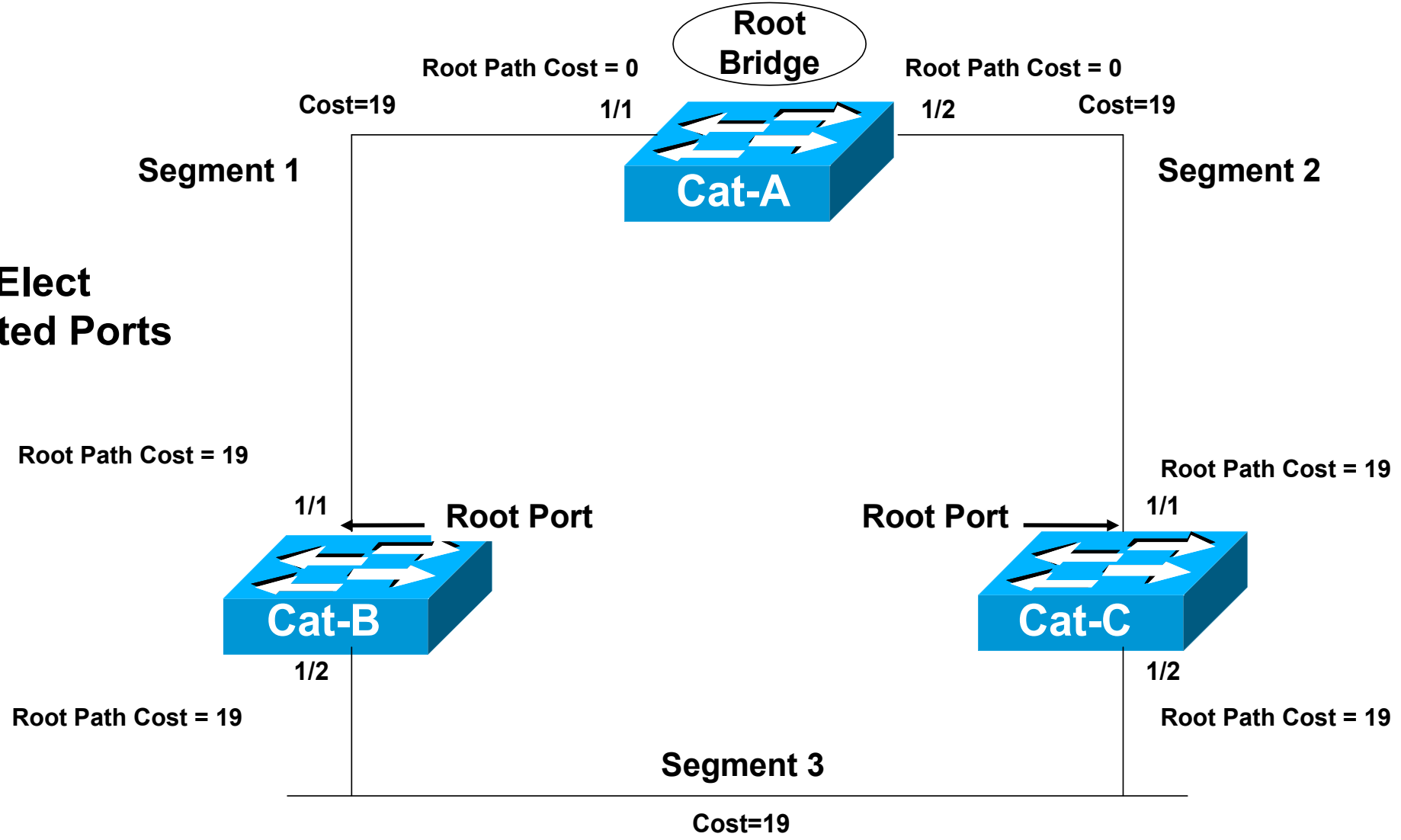
Step 3 Elect Designated Ports

Step 3 Elect Designated Ports

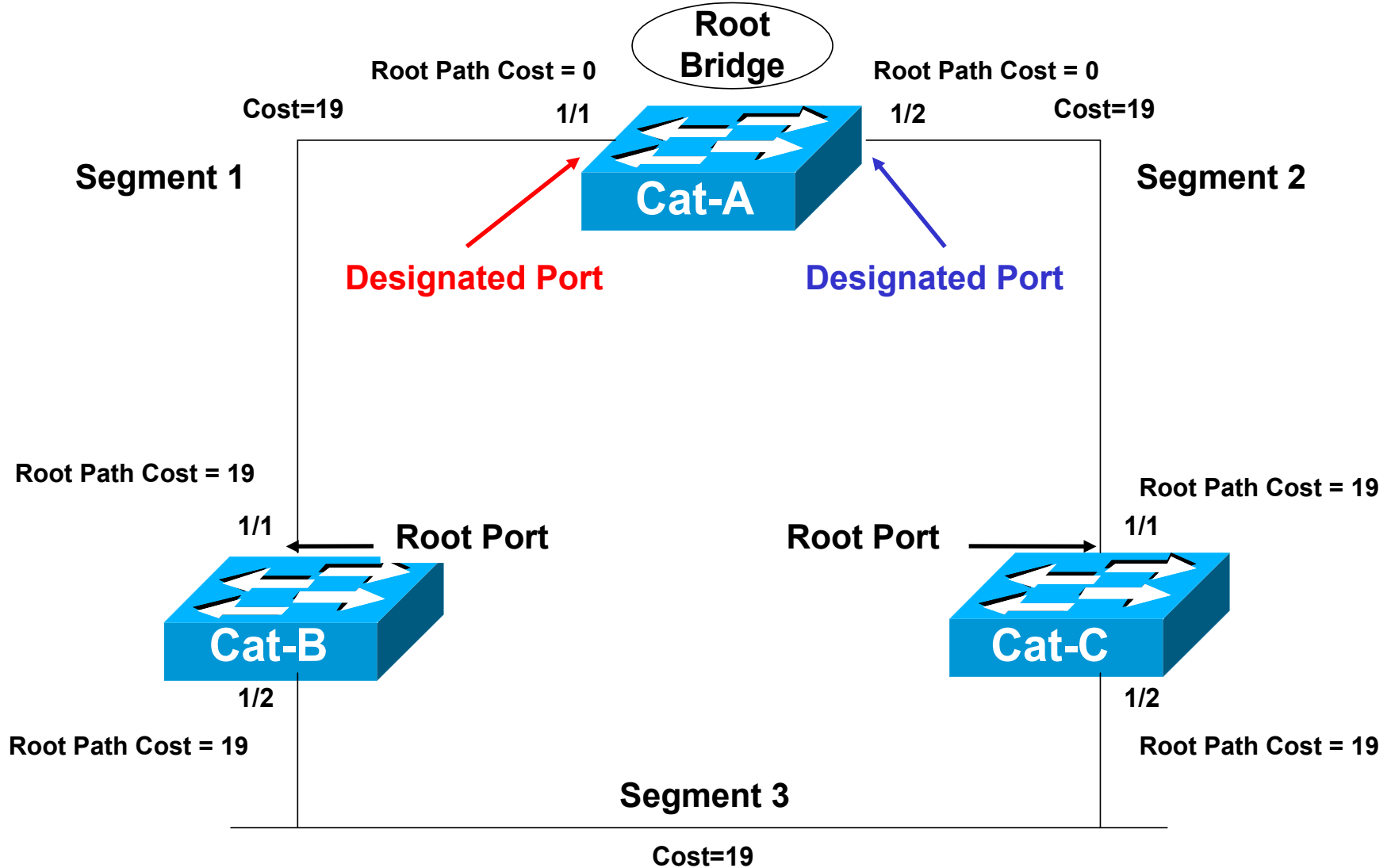


- The loop prevention part of STP becomes evident during this step, electing designated ports.
- A **Designated Port** functions as *the single bridge port that both sends and receives traffic to and from that segment and the Root Bridge*.
- **Each segment in a bridged network has one Designated Port, chosen based on cumulative Root Path Cost to the Root Bridge.**
- The switch containing the Designated Port is referred to as the **Designated Bridge** for that segment.
- To locate Designated Ports, let's take a look at each segment.
- **Root Path Cost**, the cumulative cost of all links to the Root Bridge.

Step 3 Elect Designated Ports

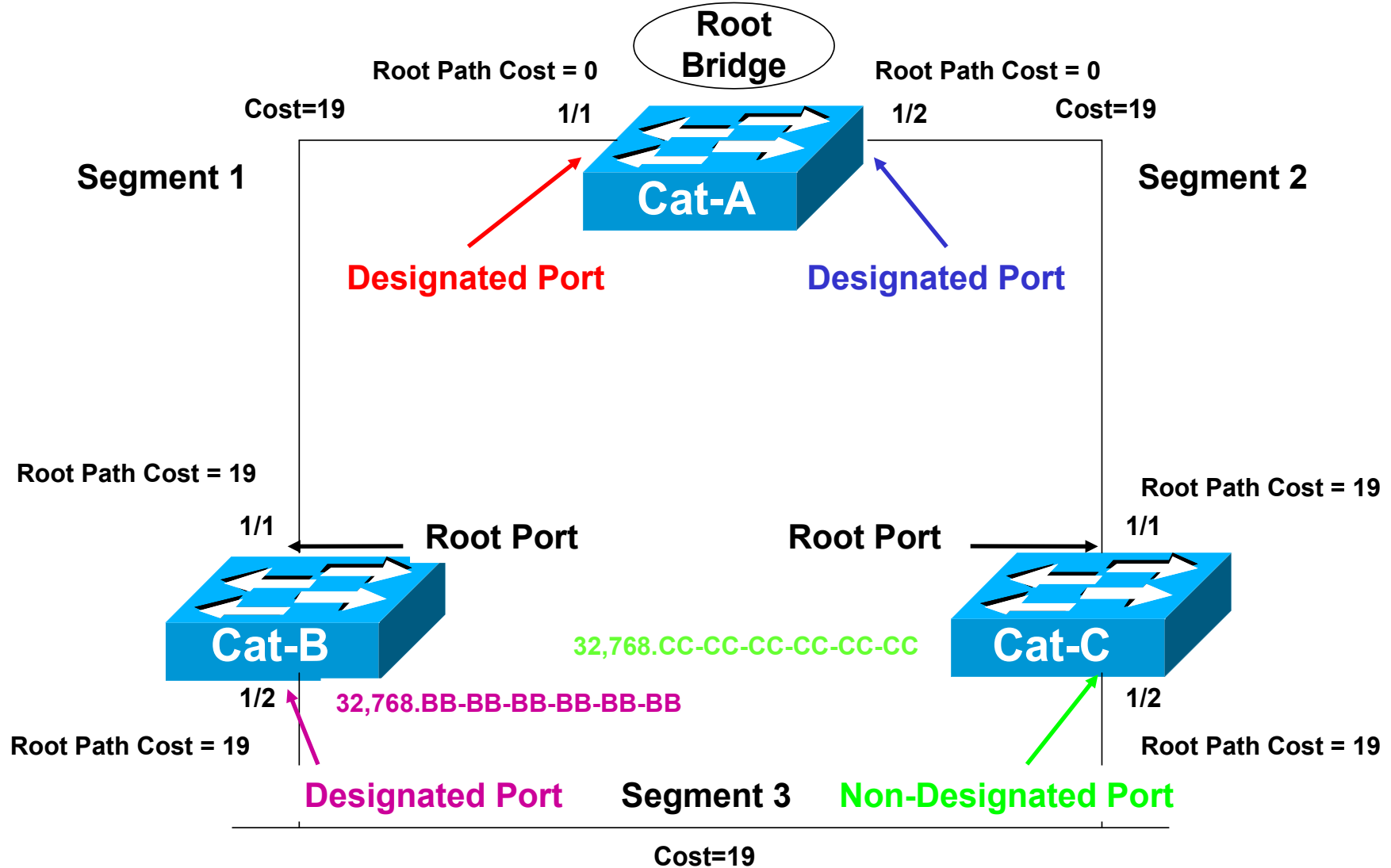


- **Segment 1:** Cat-A:1/1 has a Root Path Cost = 0 (after all it has the Root Bridge) and Cat-B:1/1 has a Root Path Cost = 19.
- **Segment 2:** Cat-A:1/2 has a Root Path Cost = 0 (after all it has the Root Bridge) and Cat-C:1/1 has a Root Path Cost = 19.
- **Segment 3:** **Cat-B:1/2** has a **Root Path Cost = 19** and **Cat-C:1/2** has a **Root Path Cost = 19**. *It's a tie!*



Segment 3

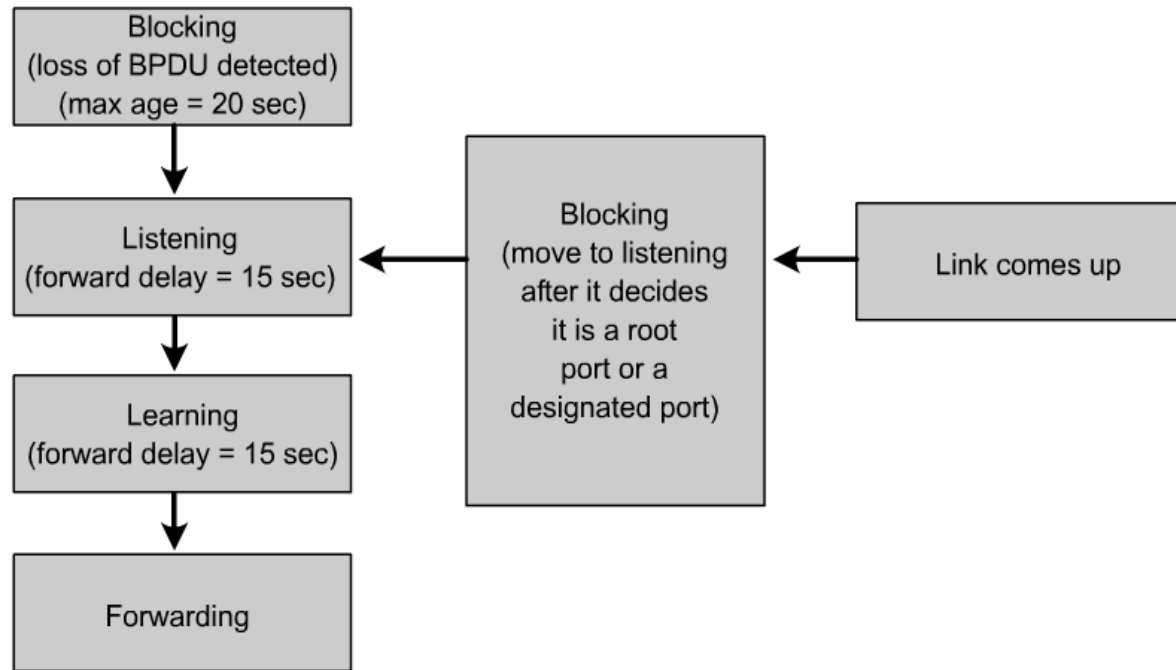
- **Both Cat-B and Cat-C have a Root Path Cost of 19, a tie!**
- When faced with a tie (or any other determination) STP always uses the four-step decision process:
 1. Lowest Root BID;
 2. Lowest Path Cost to Root Bridge;
 3. Lowest Sender BID;
 4. Lowest Port ID



Segment 3 (continued)

- 1) All three switches agree that Cat-A is the Root Bridge, so this is a tie.
- 2) Root Path Cost for both is 19, also a tie.
- 3) The sender's BID is lower on Cat-B, than Cat-C, so Cat-B:1/2 becomes the **Designated Port for Segment 3**.
- Cat-C:1/2 therefore becomes the **non-Designated Port for Segment 3**.

Stages of spanning-tree port states

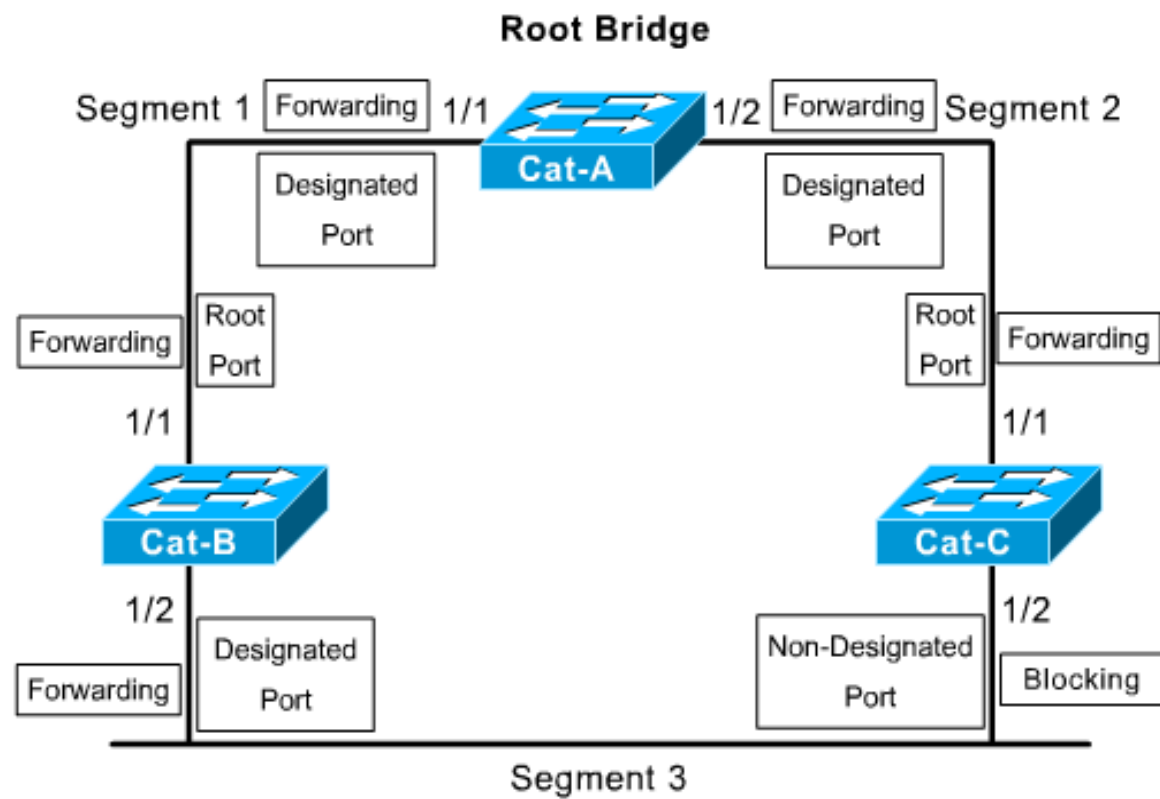


- Time is required for (BPDU) protocol information to propagate throughout a switched network.
- Topology changes in one part of a network are not instantly known in other parts of the network.
- There is propagation delay.
- A switch should not change a port state from inactive (Blocking) to active (Forwarding) immediately, as this may cause data loops.
- Each port on a switch that is using the Spanning-Tree Protocol has one of five states,

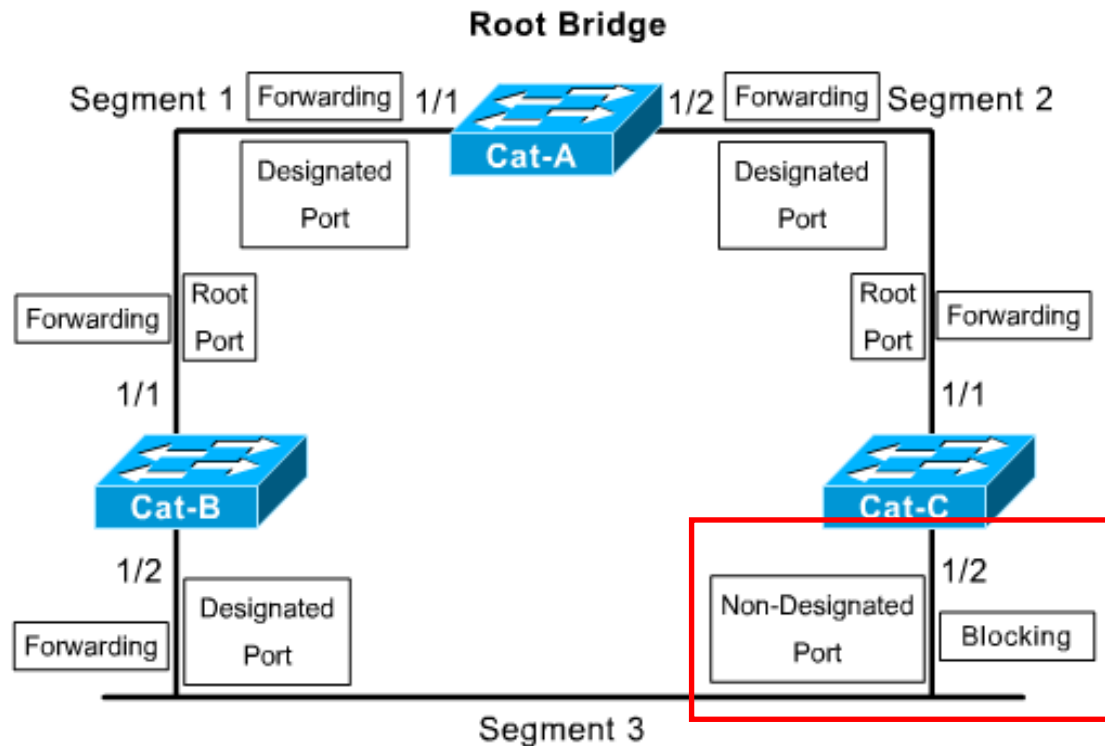
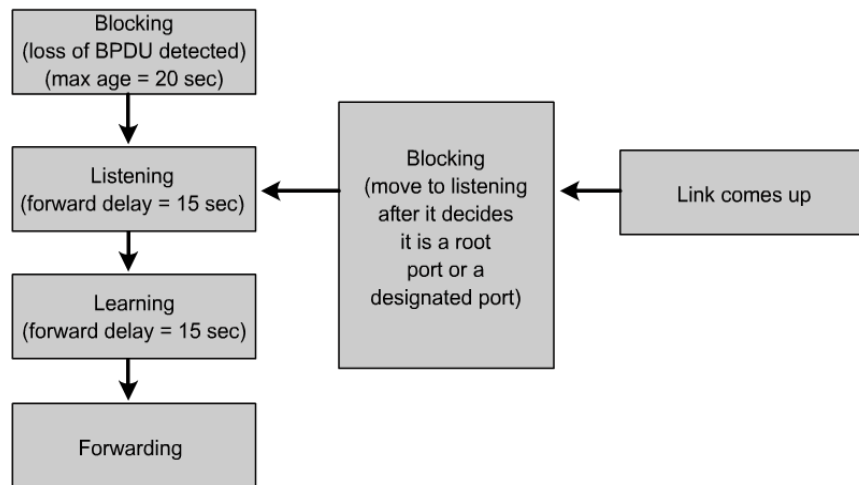
State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Designated Ports & Root Ports

Non-Designated Ports

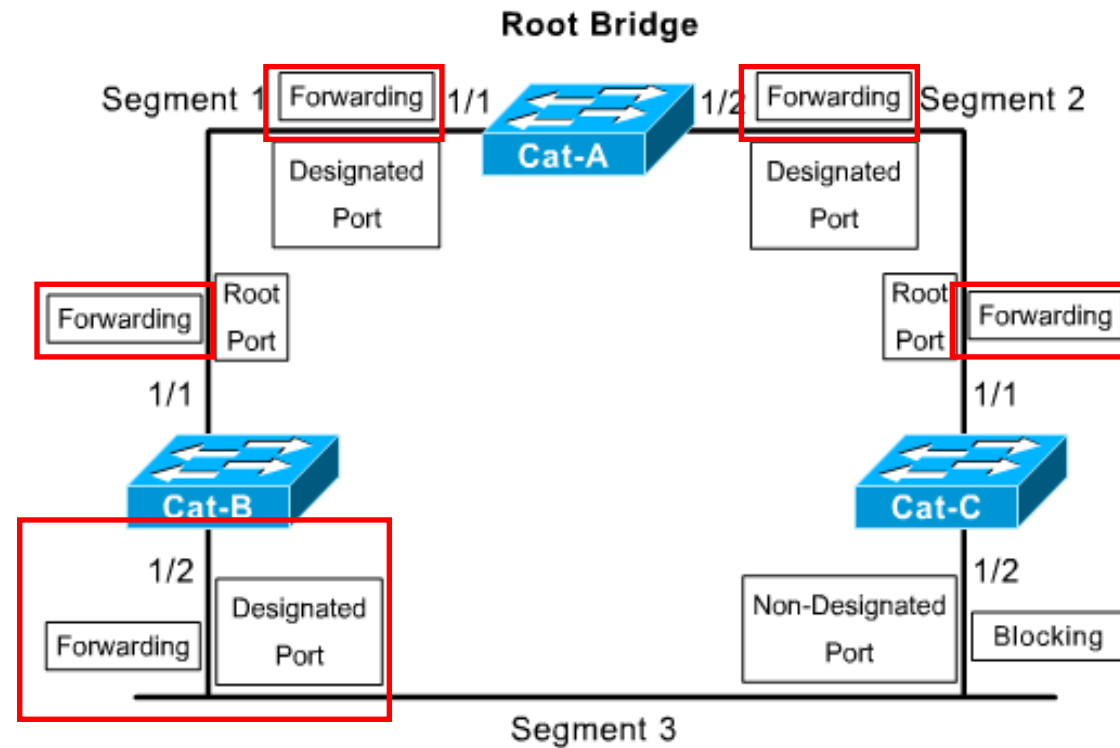
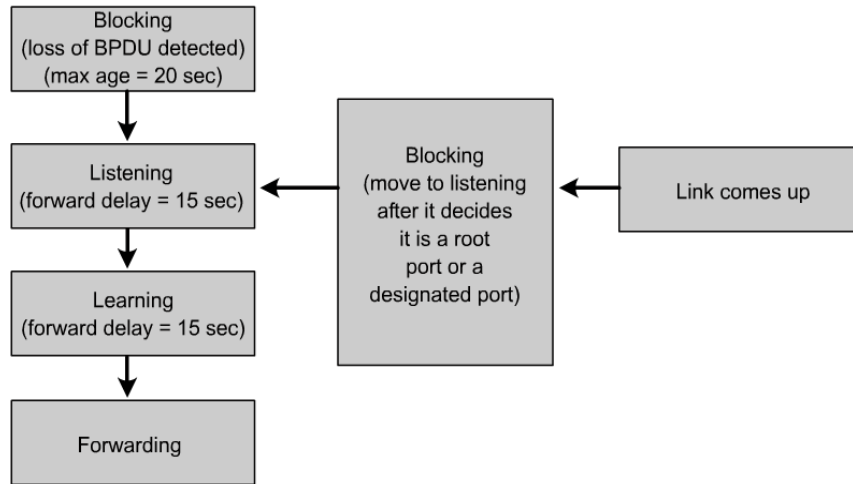


STP Port States



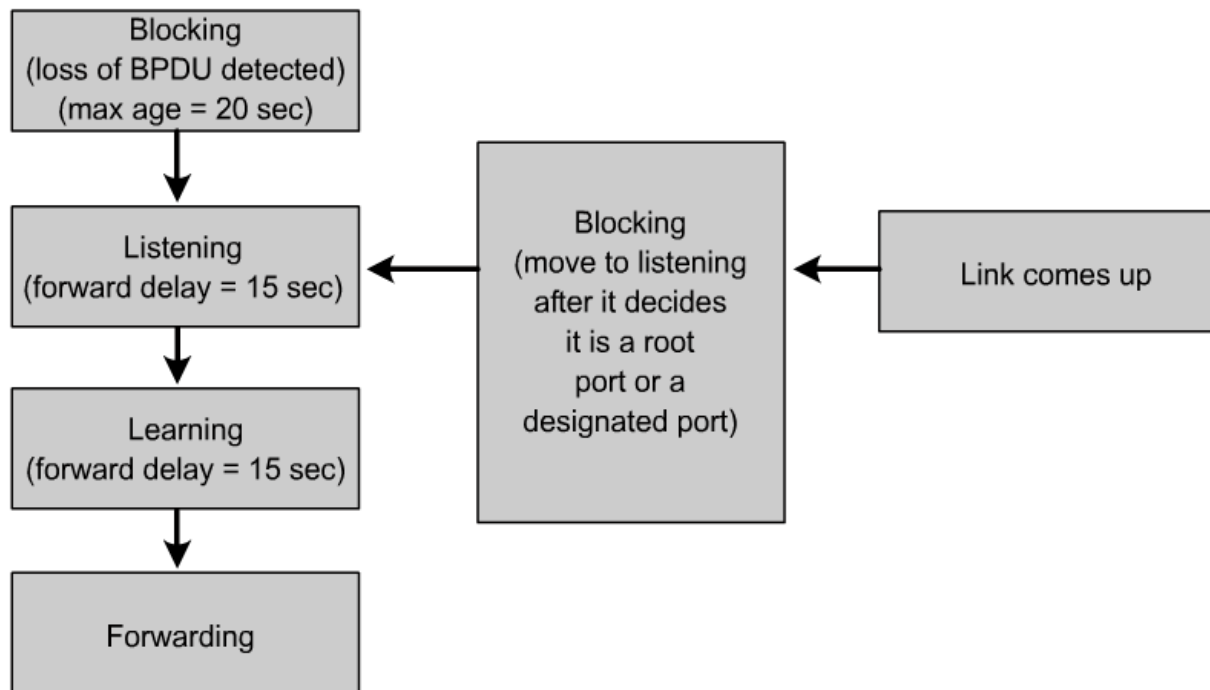
- In the **blocking state**, ports can only receive BPDUs.
 - Data frames are discarded and no addresses can be learned.
 - It may take up to 20 seconds to change from this state.
- Ports go from the blocked state to the **listening state**.
 - Switch **determines if there are any other paths to the root bridge**.
 - The **path that is not the least cost path to the root bridge goes back to the blocked state**.
 - The listening period is called the forward delay and lasts for 15 seconds.
 - In the listening state, user data is not being forwarded and MAC addresses are not being learned.
 - BPDUs are still processed.

STP Port States

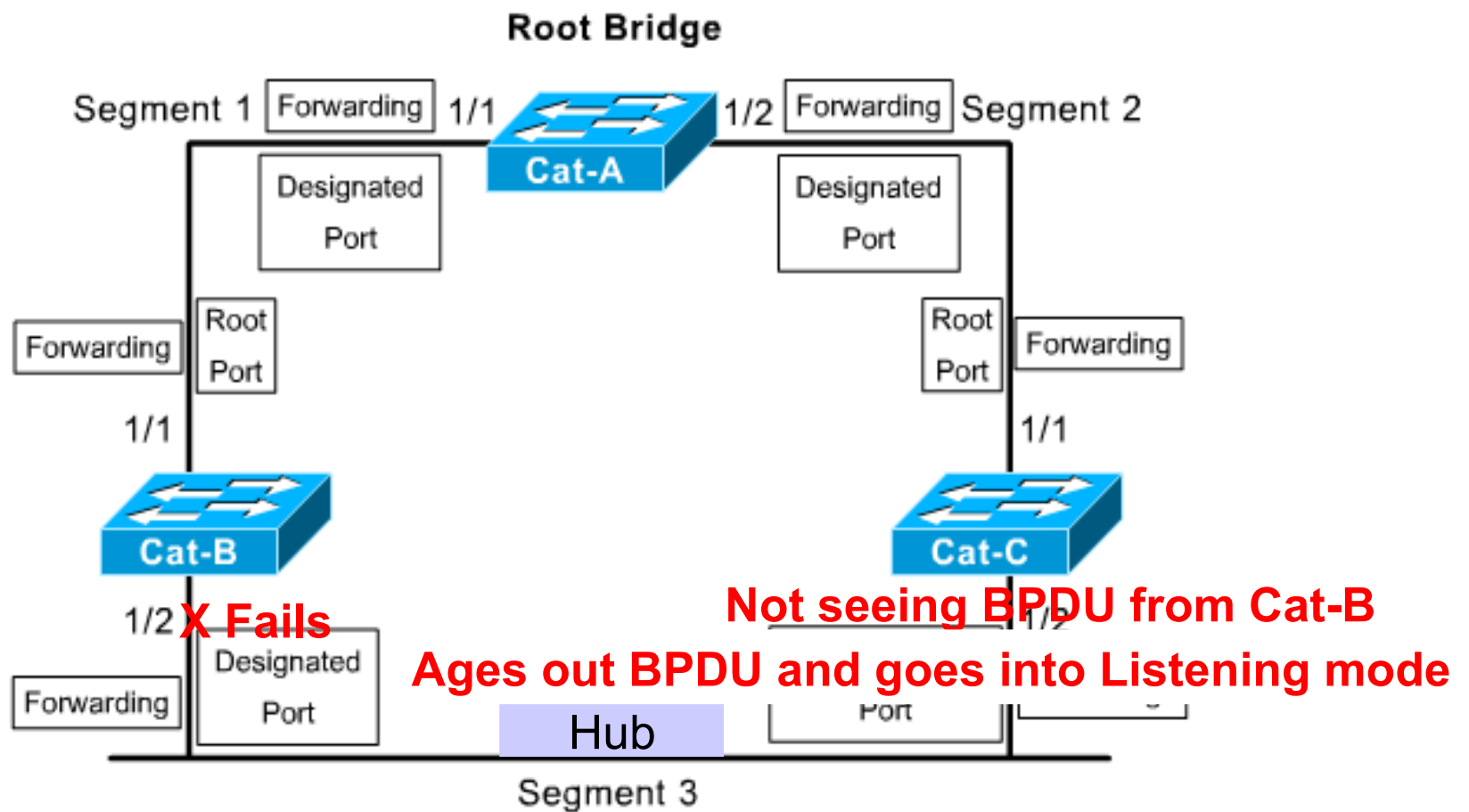


- Ports transition from the listening to the **learning state**.
 - In this state **user data is not forwarded, but MAC addresses are learned** from any traffic that is seen.
 - The learning state lasts for 15 seconds and is also called the forward delay.
 - BPDUs are still processed.
- A port goes from the learning state to the **forwarding state**.
 - In this state **user data is forwarded and MAC addresses continue to be learned**.
 - BPDUs are still processed.

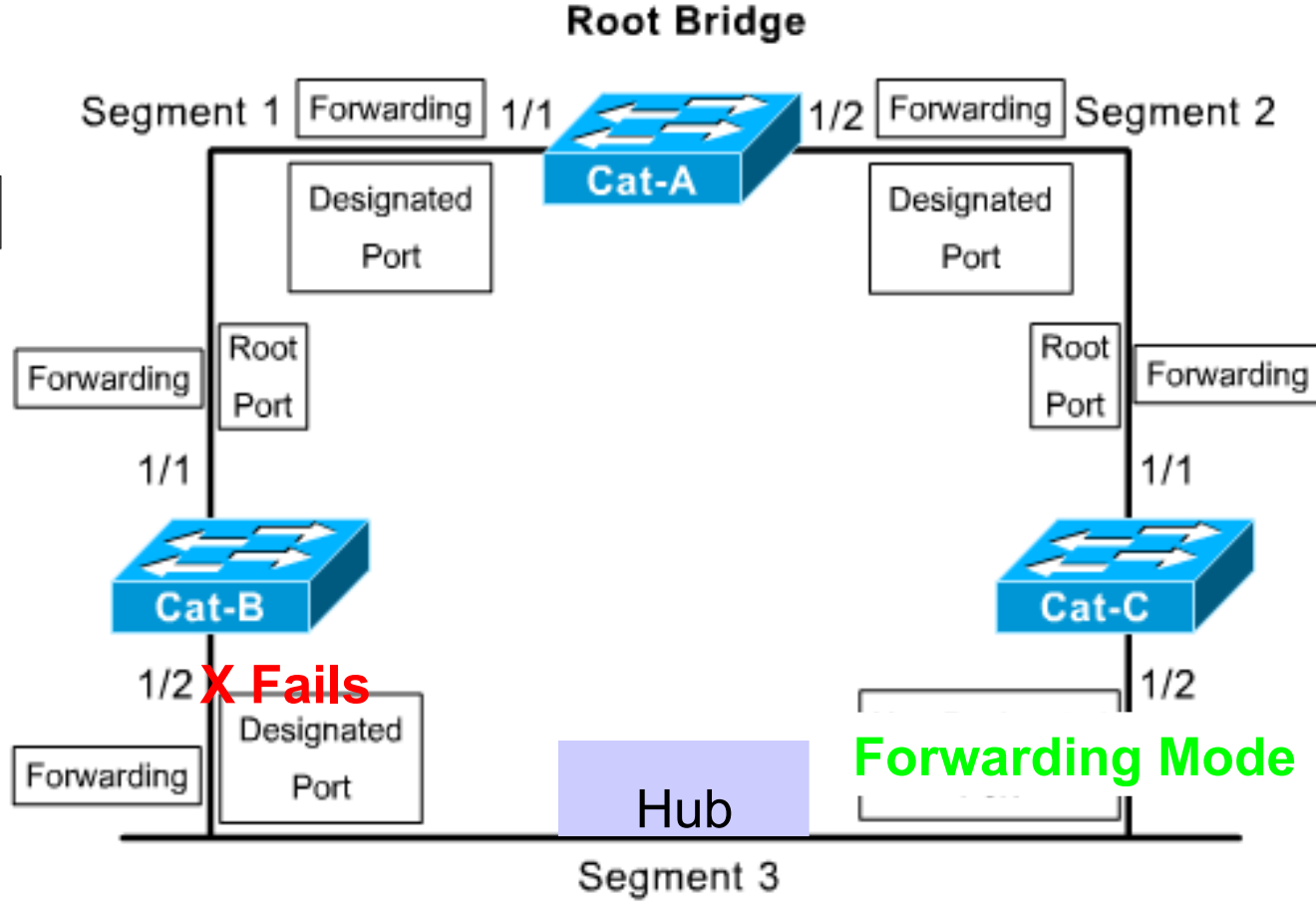
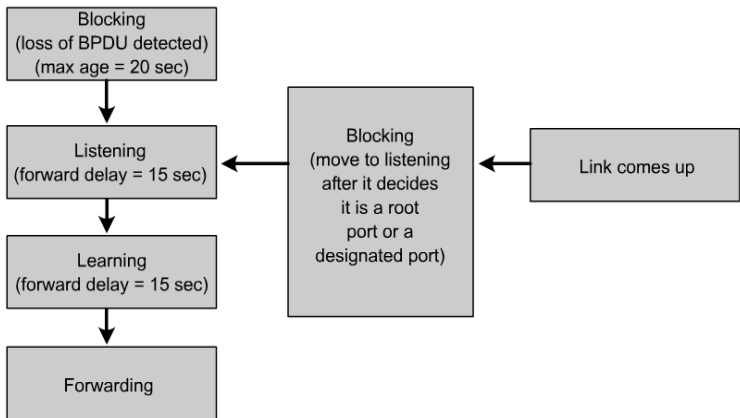
STP Timers



- Some details have been left out, such as timers, STP FSM, etc.
- The time values given for each state are the default values.
- These values have been calculated on an assumption that there will be a maximum of seven switches in any branch of the spanning tree from the root bridge.

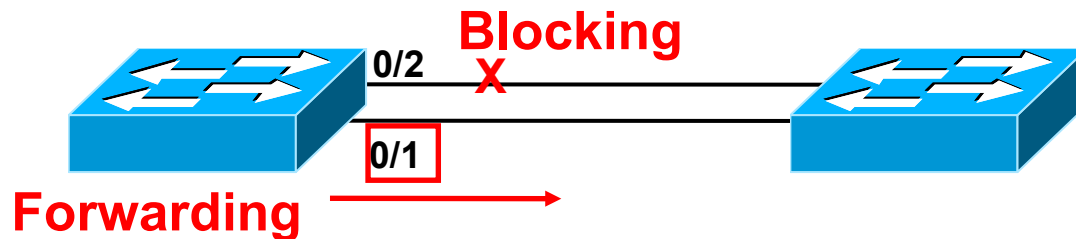


- Cat-B:1/2 fails.
- Cat-C has no immediate notification because it's still receiving a link from the hub.
- Cat-C notices it is not receiving BPDUs from Cat-B.
- **20 seconds (max age)** after the failure, Cat-C ages out the BPDUs that lists Cat-B as having the DP for segment 3.
- This causes **Cat-C:1/2 to transition into the Listening state (15 seconds)** in an effort to become the DP.



- Because Cat-C:1/2 now offers the most attractive access from the Root Bridge to this link, it **eventually transitions to Learning State (15 seconds), then all the way into Forwarding mode.**
- In practice this will take **50 seconds (20 max age + 15 Listening + 15 Learning)** for Cat-C:1/2 to take over after the failure of Cat-B:1/2.

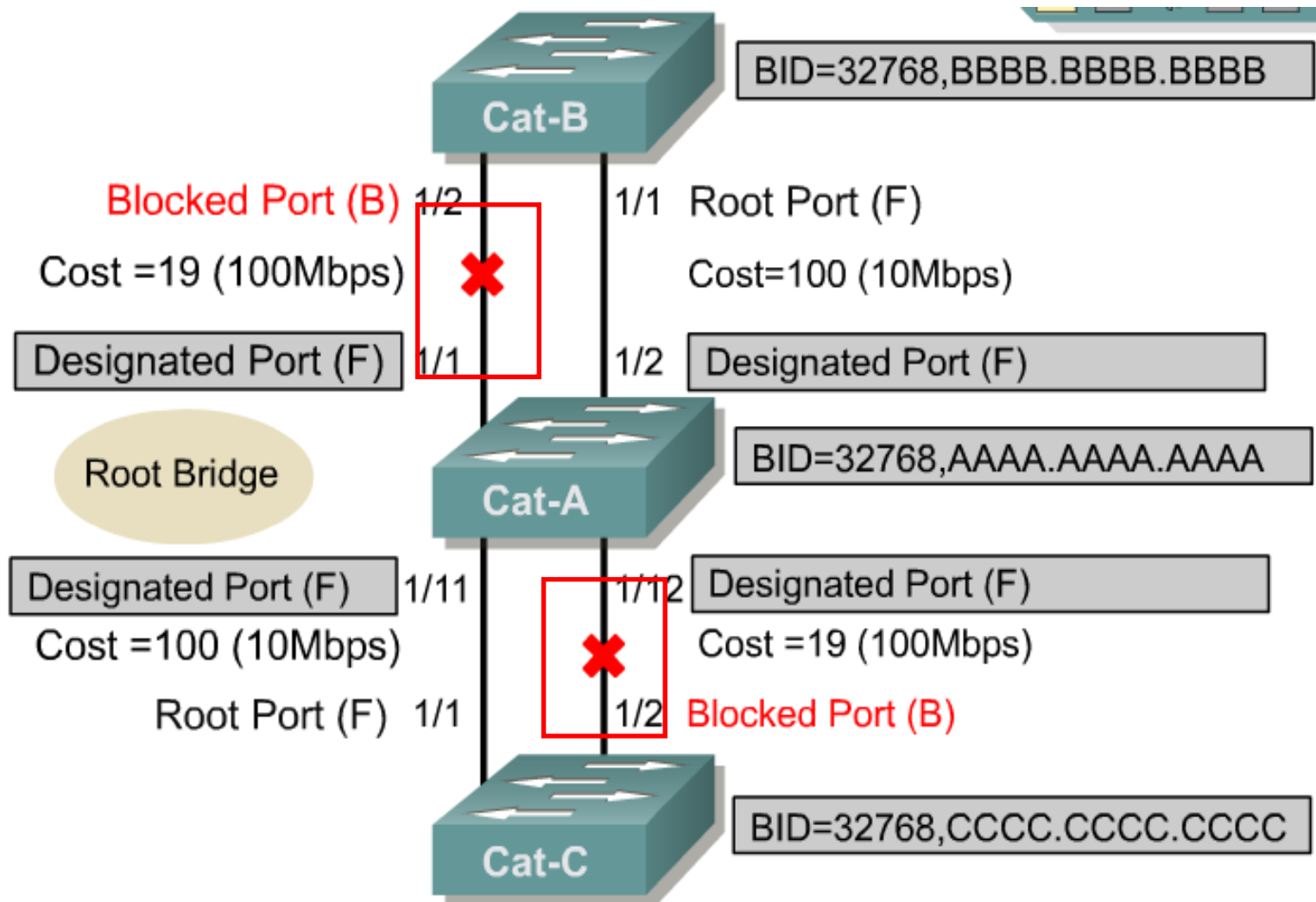
Port Cost/Port ID



Assume path cost and port priorities are default (32). Port ID used in this case. Port 0/1 would forward because it's the lower than Port 0/2.

- 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 (all ports set to 32).
- You can set the priority from 0 – 63.
- If all ports have the same priority, the port with the lowest port number forwards frames.

Port Cost/Port ID



- If all ports have the same priority, the port with the lowest port number forwards frames.

STP Convergence Recap

- Recall that switches go through three steps for their initial convergence:

STP Convergence

Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

- Also, all STP decisions are based on a the following predetermined sequence:

Four-Step decision Sequence

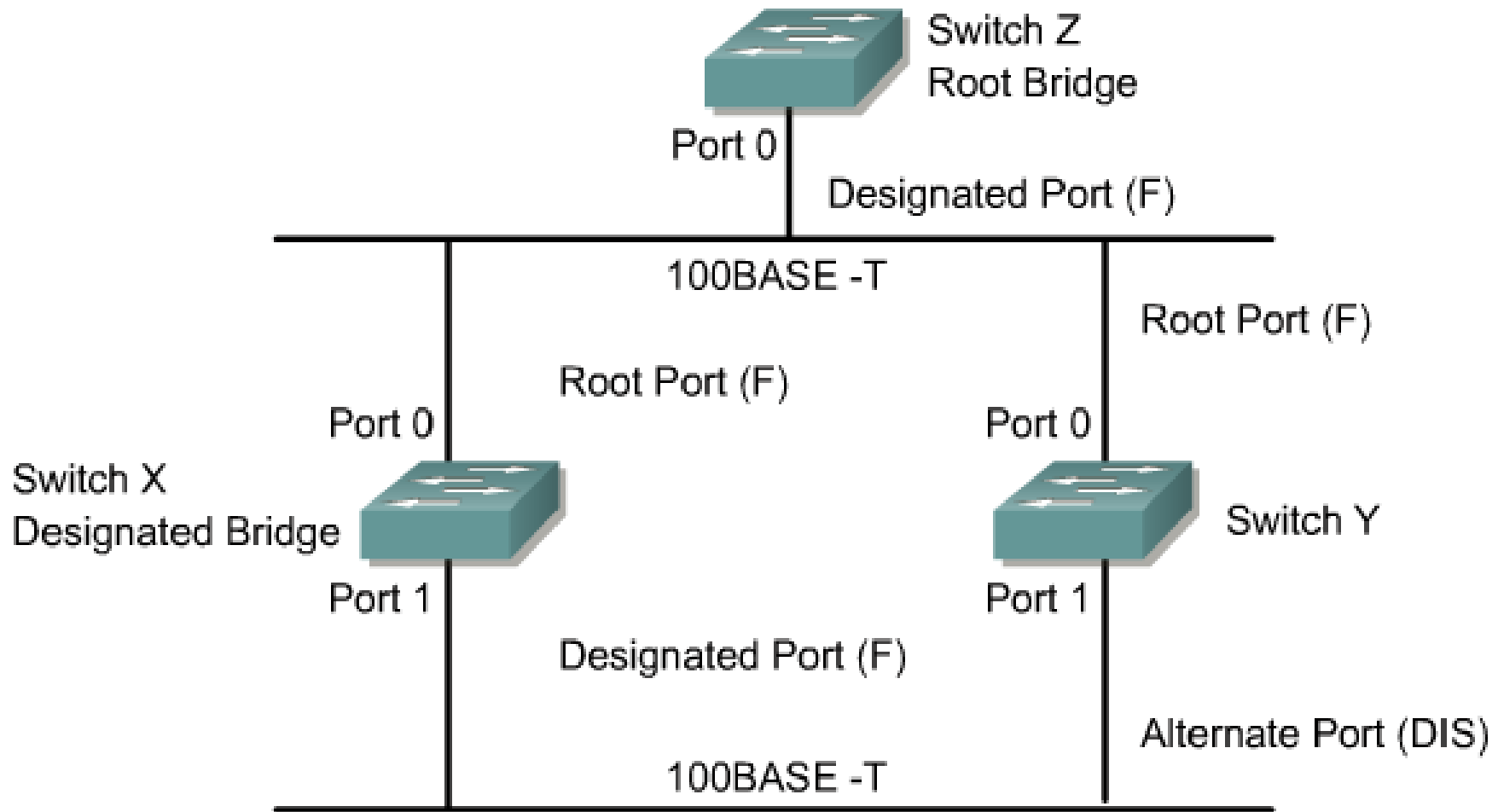
Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

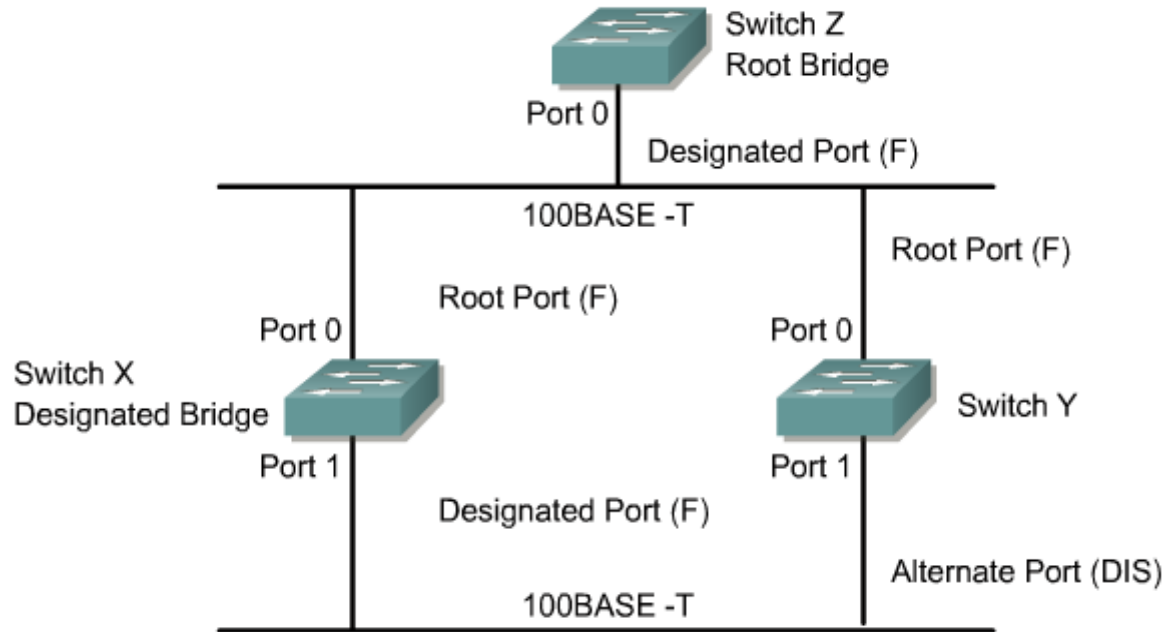
Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID

Rapid Spanning Tree Protocol (RSTP)

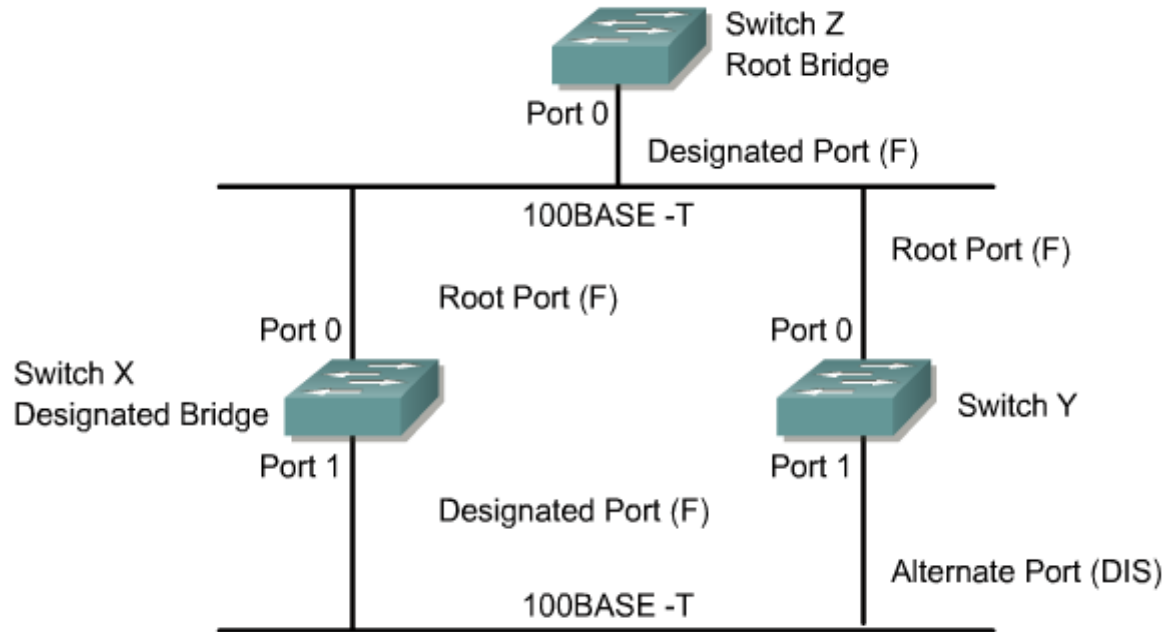


Rapid Spanning Tree Protocol (RSTP)



- 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

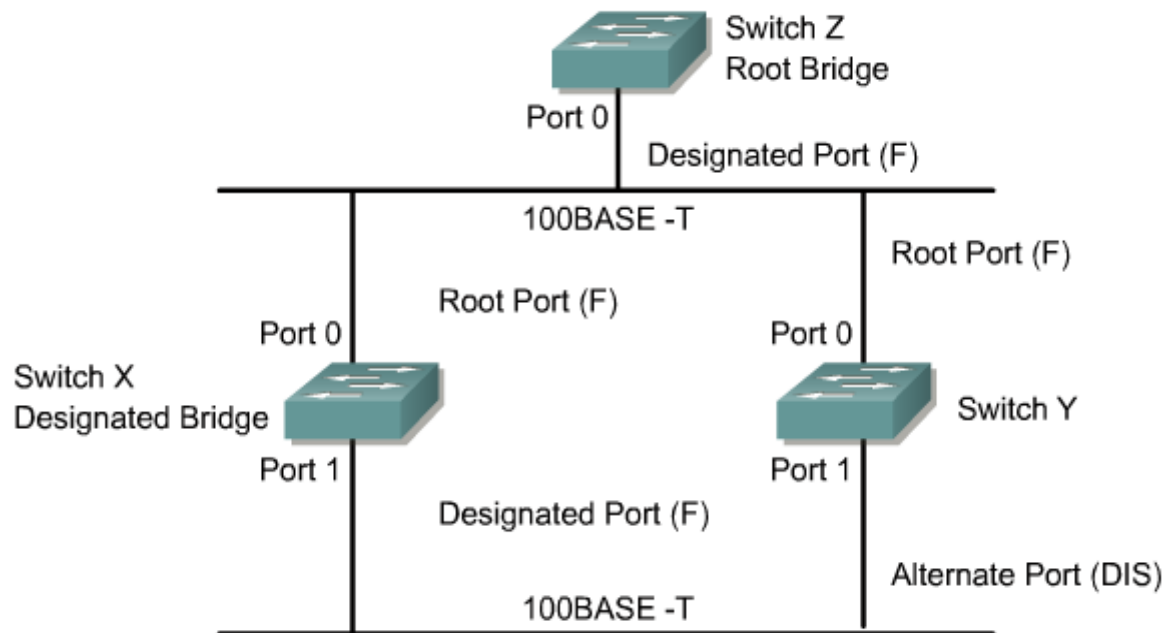


- Link types have been defined as point-to-point, edge-type, and shared.
- These changes allow failure of links in switched network to be learned rapidly.
- Point-to-point links and edge-type links can go to the forwarding state immediately.
- Network convergence does not need to be any longer than 15 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

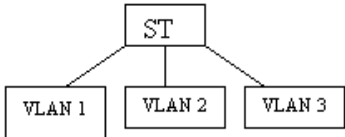
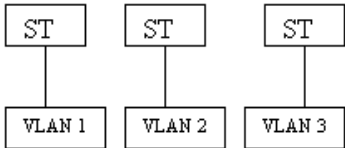
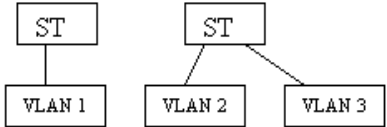
RSTP Port Roles



- 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.

Rapid Spanning Tree Protocol (RSTP)

Evolution of STP

Cisco's Implementation	Spanning Tree Protocol Process	IEEE Standard
Spanning Tree Protocol (STP): <ul style="list-style-type: none"> • 802.1D • Common Spanning Tree (CST) • Mono Spanning Tree (MST) 	ST = Spanning Tree 	
Cisco Enhancements (First evolution): <ul style="list-style-type: none"> • Portfast • Uplinkfast • Backbonefast 		RSTP: <ul style="list-style-type: none"> • 802.1w • Edge Fast (Cisco Port Fast) • Uplink Fast RSTP (Cisco Uplink Fast) • Backbone Fast Engine (Cisco Backbone Fast)
Cisco Enhancements (Second Evolution): <ul style="list-style-type: none"> • PVST: ISL • PVST+: ISL & 802.1Q • Includes previous enhancements • Additional enhancements: <ul style="list-style-type: none"> ○ BPDU Guard ○ Root Guard 		
Cisco MISTP: <ul style="list-style-type: none"> • Uses PVST+ • Includes previous enhancements • Catalyst 4000/6000 		MST (Multiple Spanning Tree): <ul style="list-style-type: none"> • 802.1s • Uses RSTP

- RSTP adds features to the standard similar to vendor proprietary features including Cisco's Port Fast, Uplink Fast and Backbone Fast.
- Cisco recommends that administrators upgrade to the IEEE 802.1w standard when possible.

Cisco's Port Fast and RSTP's Edge Fast

- A common problem is with DHCP and STP Port States.
- The workstation will power up and start looking for a DHCP servers before its port has transitioned to Forwarding State.
- The workstation will not be able to get a valid IP address, and may default to an IP address such as 169.x.x.x.
- Spanning-tree PortFast causes a port to enter the spanning-tree forwarding state immediately, bypassing the listening and learning states.
- You can use PortFast on switch ports connected to a single workstation or server to allow those devices to connect to the network immediately, instead of waiting for the port to transition from the listening and learning states to the forwarding state.
- **Caution** PortFast should be used *only* when connecting a single end station to a switch port.
 - If you enable PortFast on a port connected to another networking device, such as a switch, you can create network loops.

Algorithm 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 is 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.

