Spanning Tree Protocol (STP) – Comprehensive Study Notes

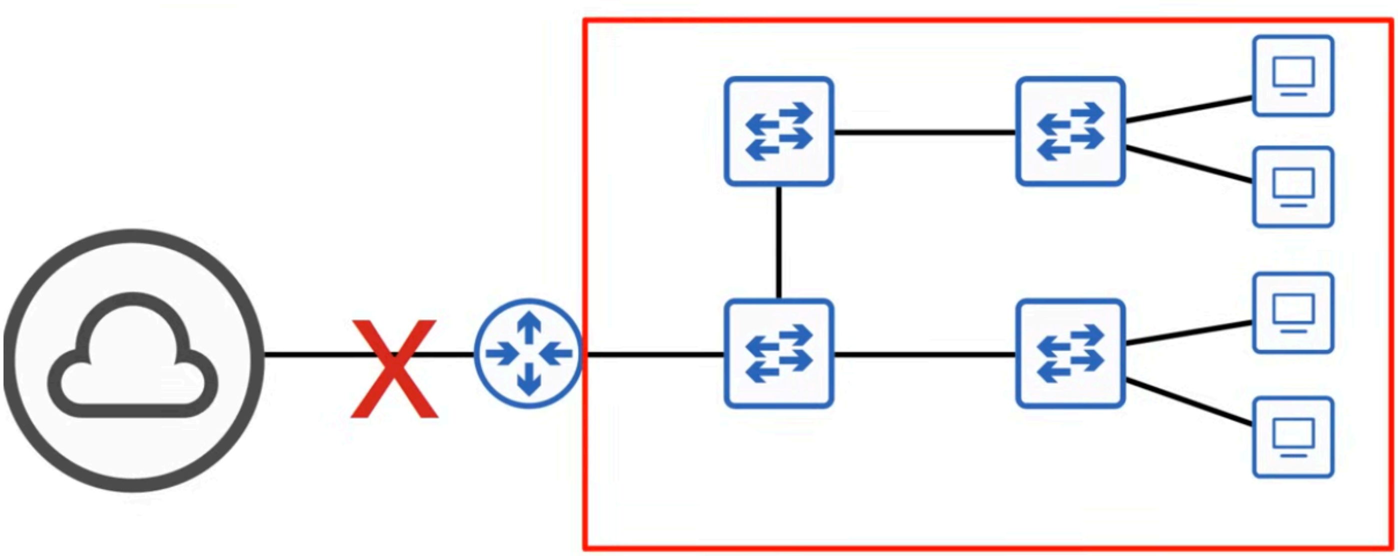
Redundancy in Network Design

Redundancy is an essential part of modern network design. Networks are expected to operate 24/7/365, and even short downtime can be disastrous for a business. A non-redundant network is simply unacceptable.

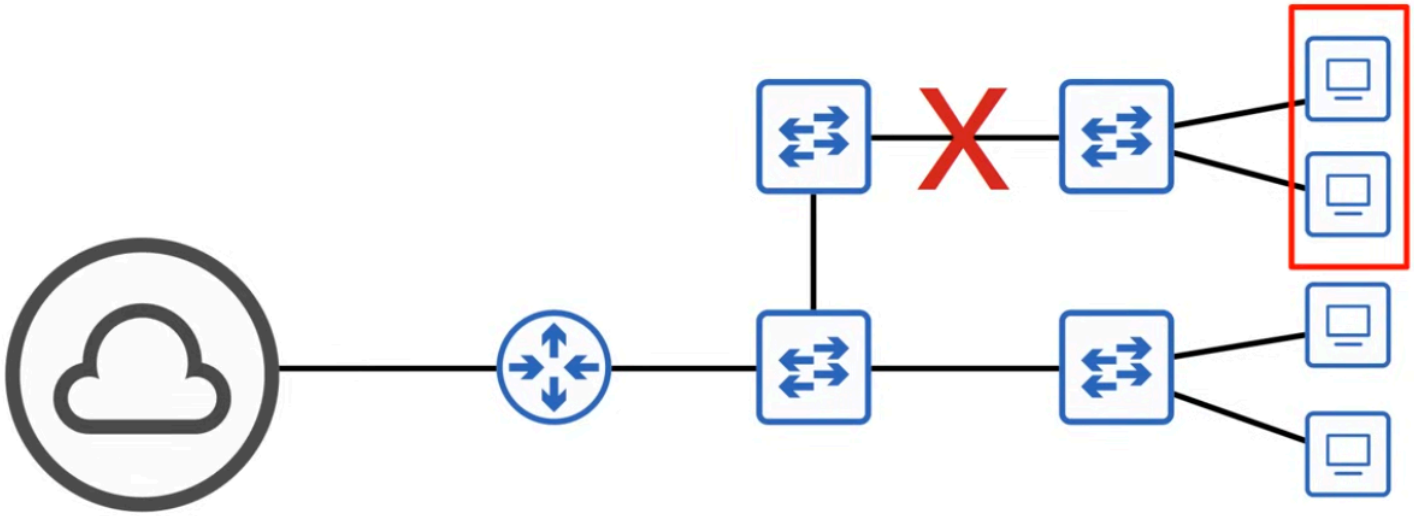
- **Single points of failure** must be eliminated by providing alternate paths.
- Redundancy should be implemented at every possible point in the network.
- **Network engineers** are responsible for business-critical infrastructure and must ensure resilience to failures.

Poor Network Design

Example 1:



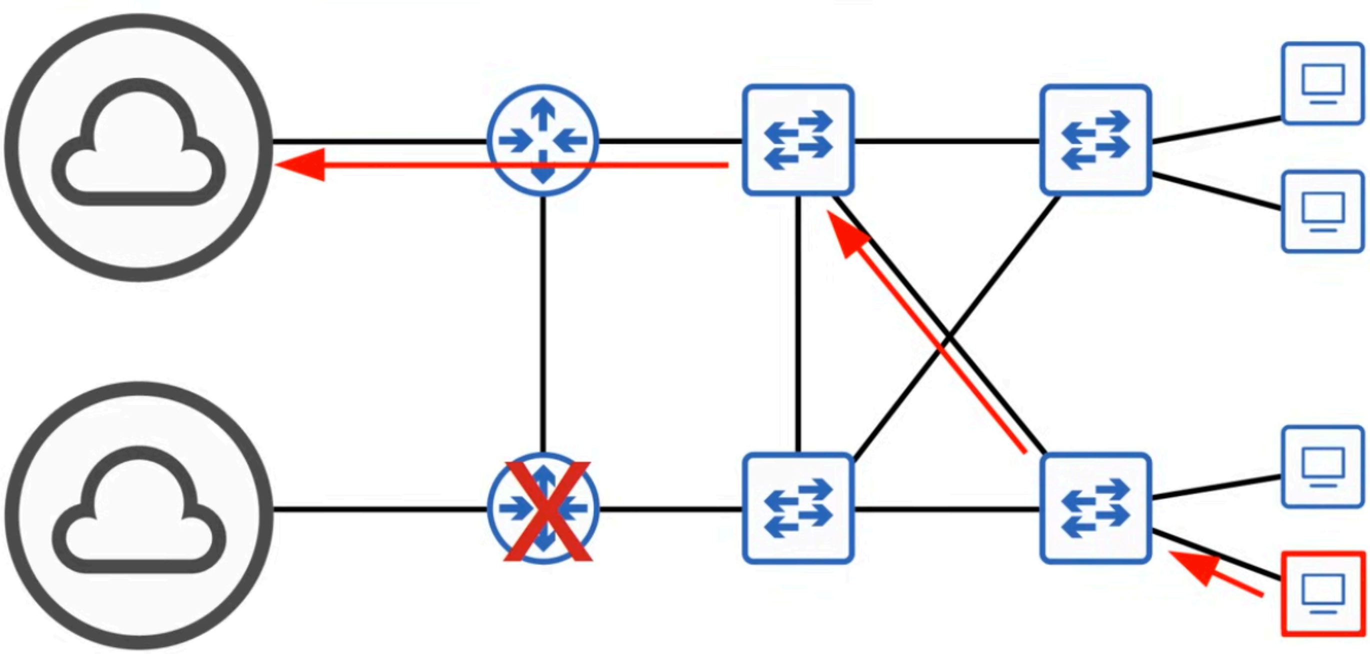
Example 2:



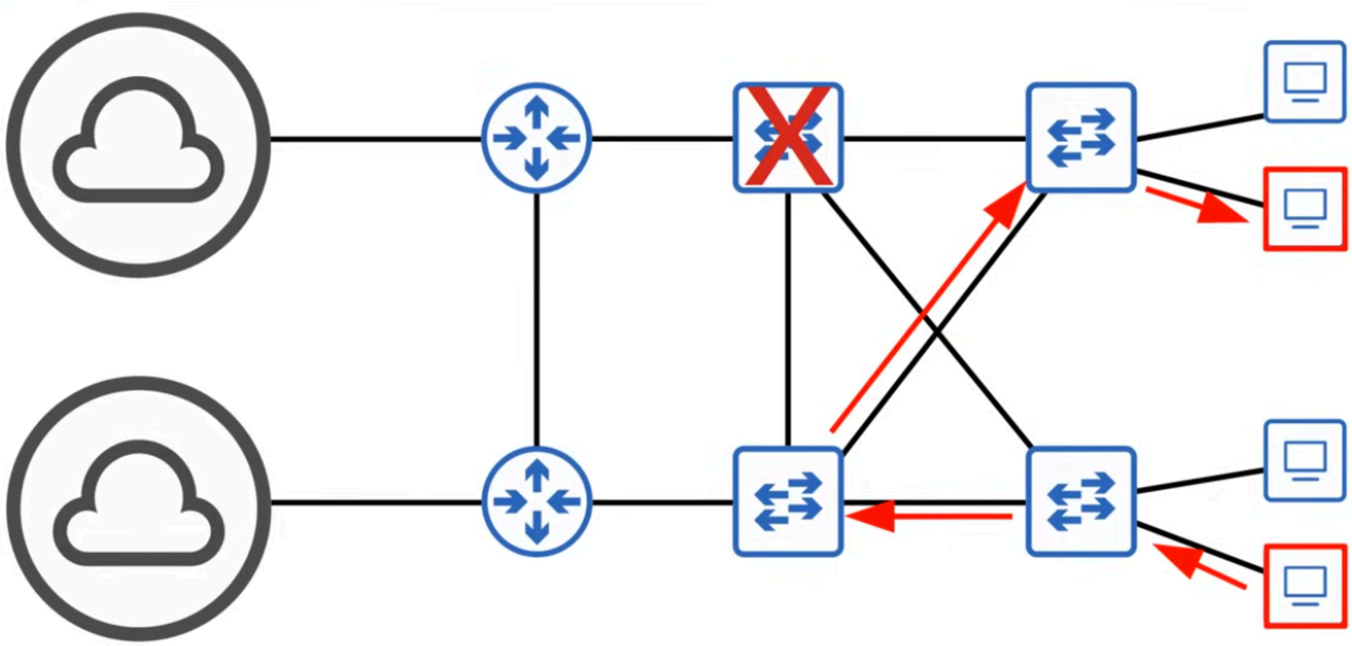
- A non-redundant network has many **single points of failure**.
- Example: If one link to the Internet is cut, the entire network loses connectivity.

Better Network Design

Example 1:

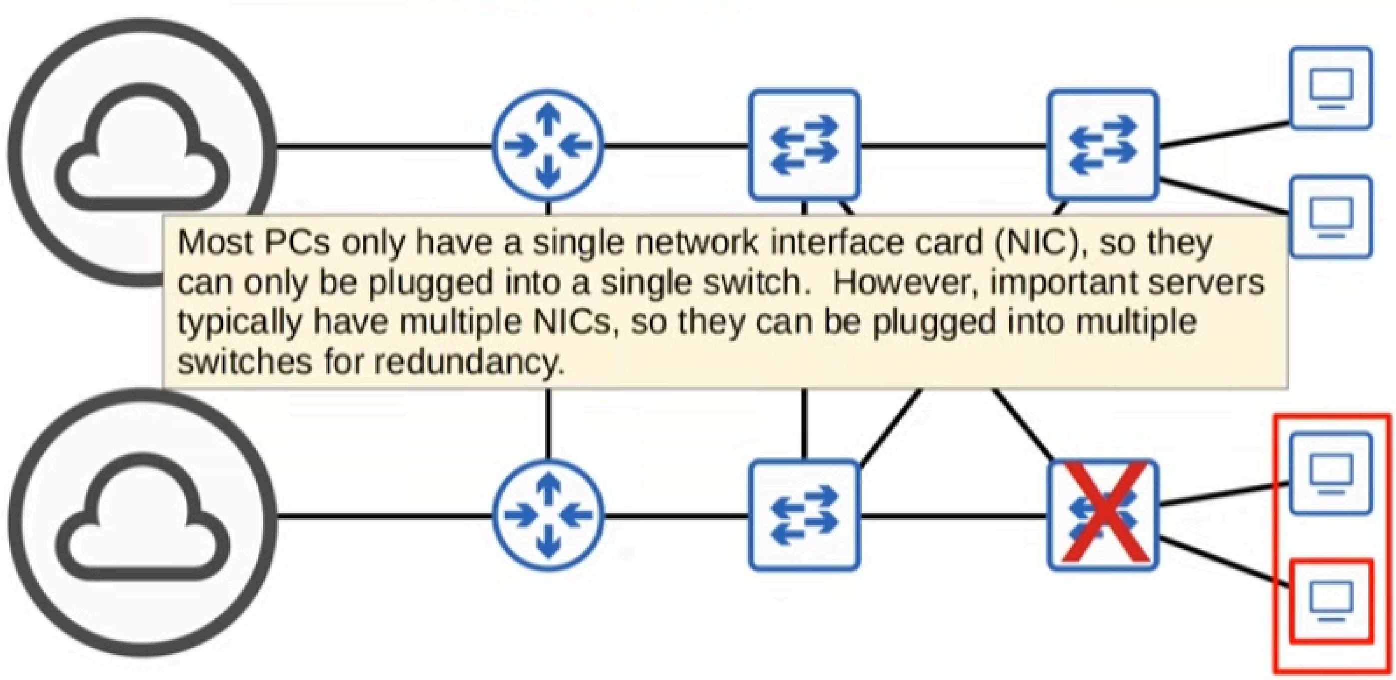


Example 2:



- A redundant network provides **alternate paths**.
- If a router or switch fails, traffic can still reach its destination via another path.

Key Limitation



- Most PCs have only **one NIC**, so they can only connect to a single switch.

- **Important servers** typically have multiple NICs and can be plugged into multiple switches for redundancy

Benefit for Redundancy

- Having multiple paths between these switches provides alternate paths if one connection fails.

Broadcast Storms and MAC Address Flapping

The Problem: Layer 2 Loops

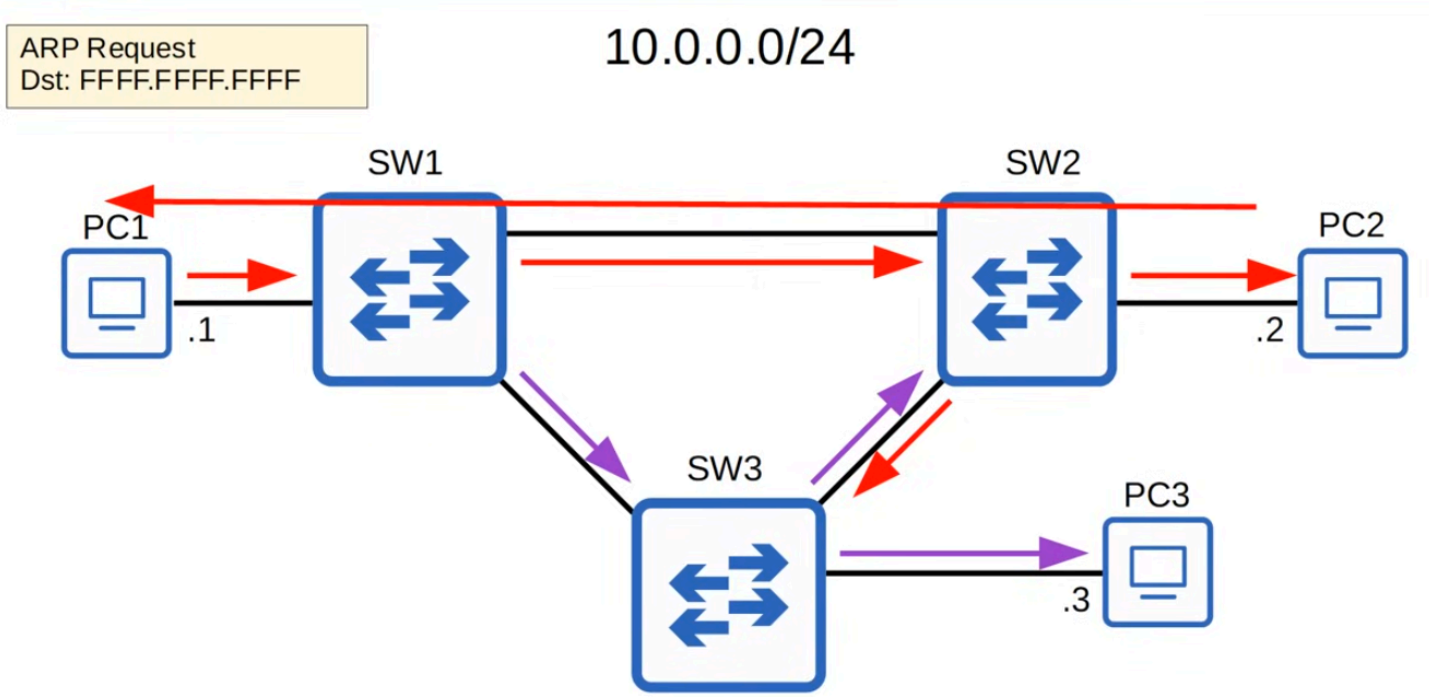
Without **Spanning Tree Protocol (STP)** , a redundant Layer 2 network can suffer from **broadcast storms**. The Ethernet header has no **TTL** (Time to Live) field, unlike IP headers. This means broadcast frames can loop indefinitely.

Broadcast Storm — Detailed Example

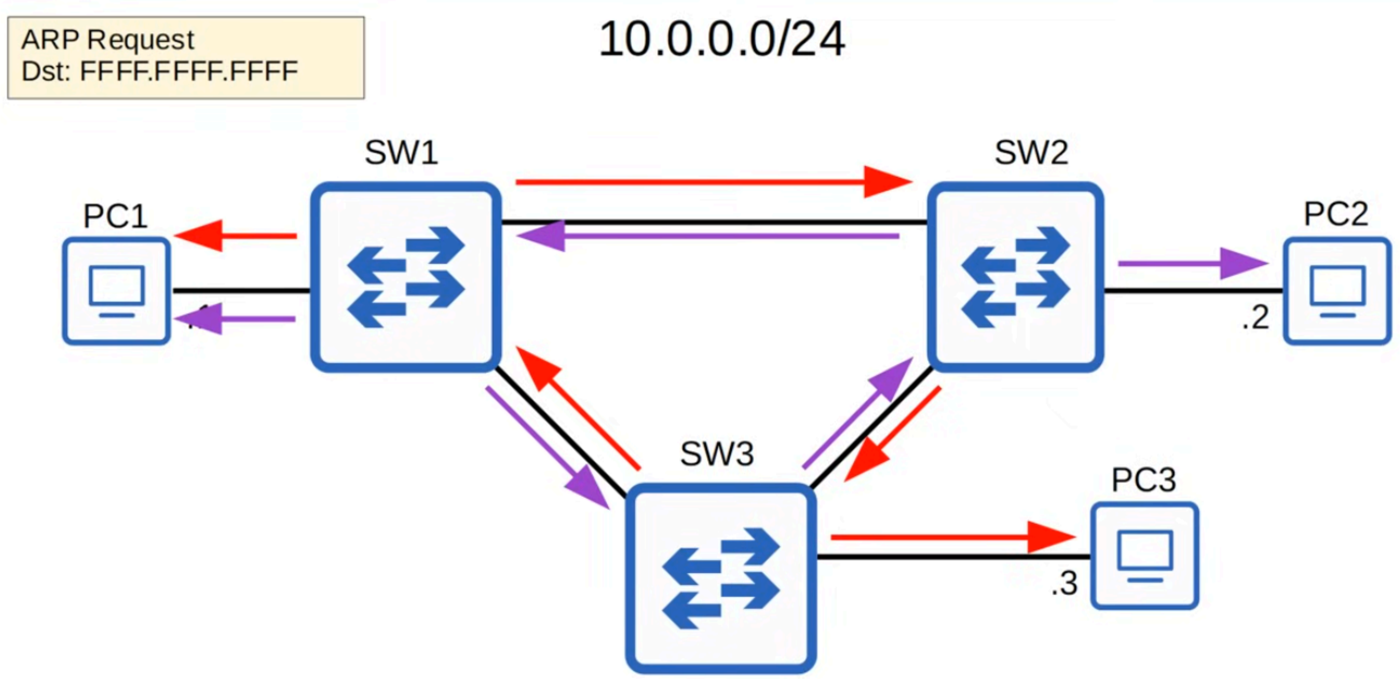
Topology:

- PC1: 10.0.0.1
- PC2: 10.0.0.2
- PC3: 10.0.0.3
- Switches: SW1, SW2, SW3 (three switches forming a triangle)

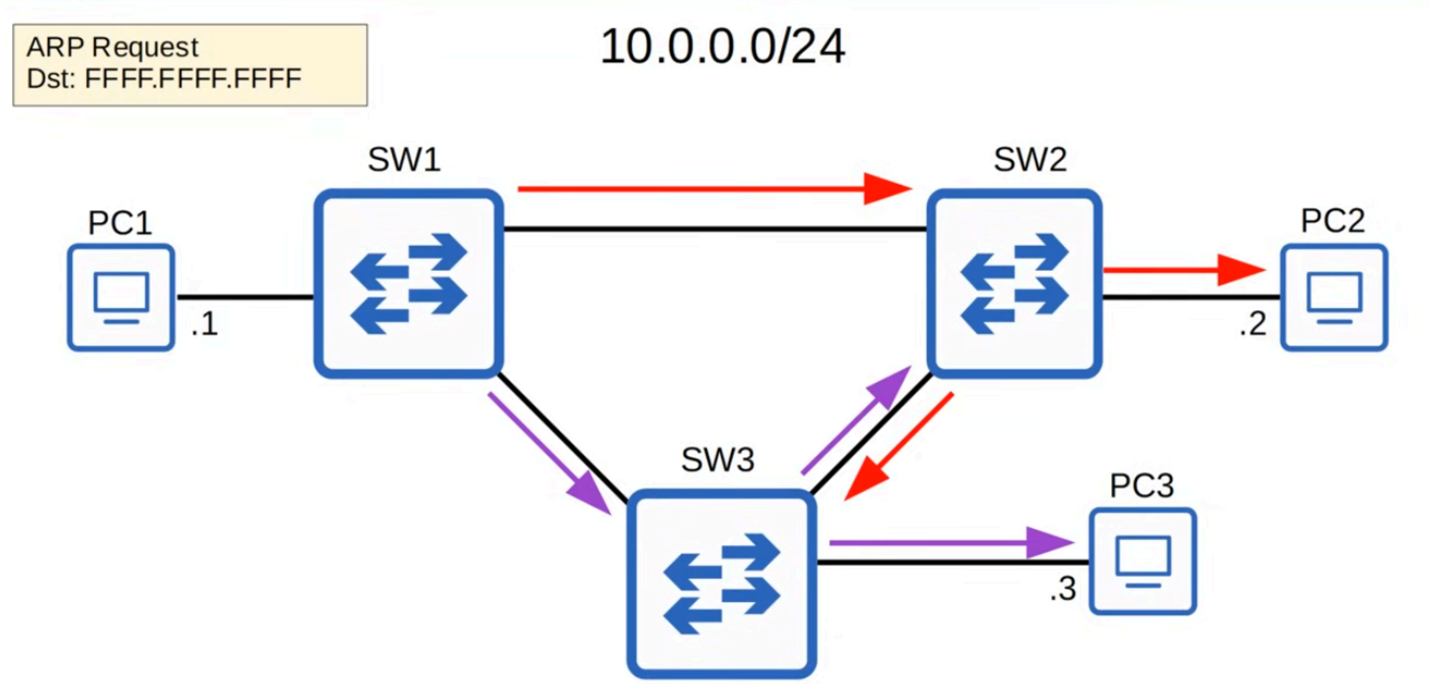
Step-by-step process:



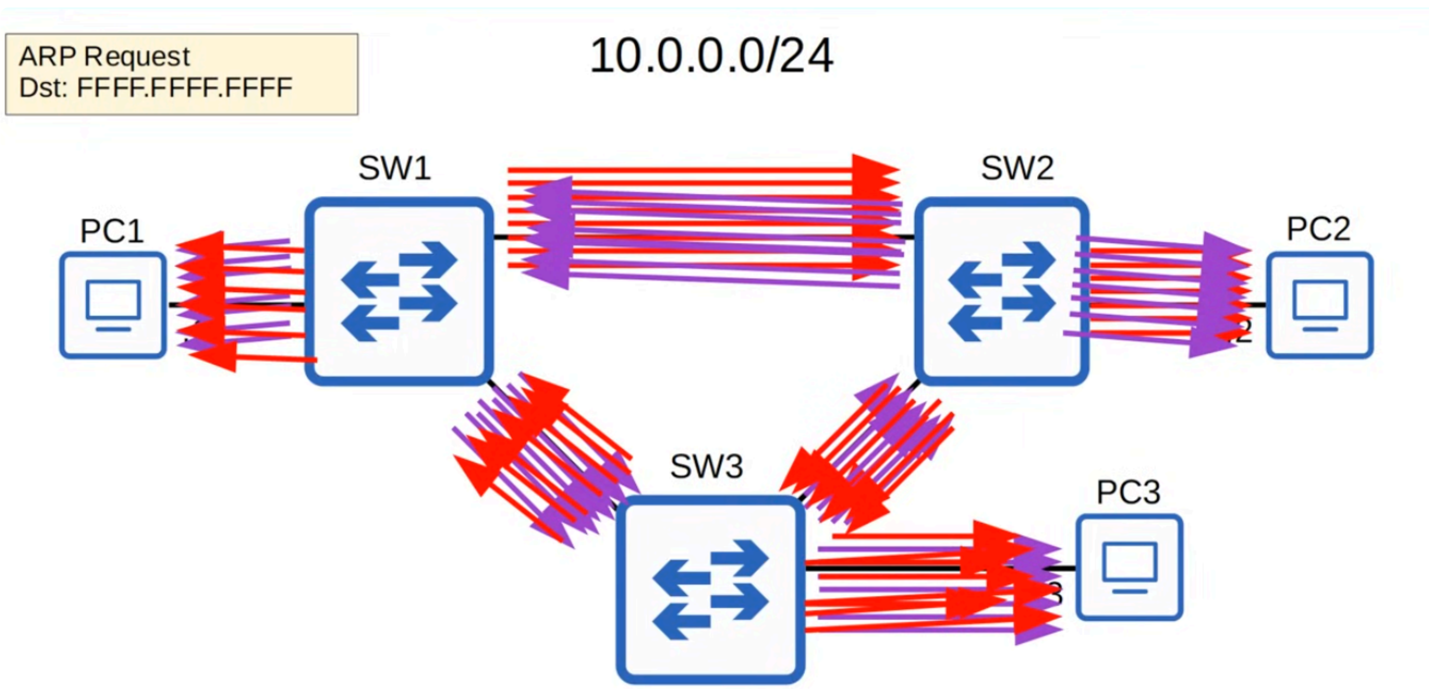
1. **PC1 sends an ARP request** to learn PC2's MAC address. This is a **broadcast frame** (destination MAC: `FFFF.FFFF.FFFF`).
2. **SW1 receives** the frame on one interface and **floods** it out all other interfaces. Both **SW2** and **SW3** receive a copy.
3. **SW2 and SW3** each receive the frame and **flood** it out all interfaces except the one it was received on. PC2 receives the ARP request and sends a unicast reply — but the broadcast frames are still in the network.



4. **SW1** now receives the broadcast frame back on **two different interfaces** (one from SW2, one from SW3). It floods both copies again out all other interfaces.



5. **SW2 and SW3** again receive broadcast frames and flood them again.



6. **This continues forever** — there is no TTL in the Ethernet header to stop the loop.

> ***Broadcast Storm: An accumulation of looped broadcast frames that congests the network, preventing legitimate traffic from passing.***

Visual representation:

- **Red arrows** = clockwise loop between the three switches
- **Purple arrows** = counter-clockwise loop

MAC Address Flapping

Each time a frame arrives on a switchport, the switch learns (or updates) the source MAC address in its MAC address table. When the same source MAC address repeatedly arrives on **different interfaces** due to loops, the switch continuously updates the interface entry. This is called **MAC Address Flapping**.

Example: If PC1's MAC address is learned on SW2's G0/0, then a moment later on SW2's G0/1, then back on G0/0 — the switch keeps overwriting the MAC table entry, wasting CPU resources and destabilizing the switch's forwarding decisions.

Key Takeaway

Redundant Layer 2 networks without STP cause:

1. **Broadcast storms** — indefinite looping of broadcast/unknown unicast frames
2. **MAC address flapping** — continuous MAC table updates destabilizing the switch

STP prevents these problems by placing redundant ports into a **blocking state**, creating a single loop-free path through the network.

Spanning Tree Protocol (STP) – IEEE 802.1D

STP is a Layer 2 protocol that prevents Layer 2 loops by placing redundant ports into a **blocking state**. These blocked ports act as backups; if an active (forwarding) port fails, a blocked port can transition to forwarding.

- STP is an industry standard (IEEE 802.1D).
- Switches from all vendors run STP by default.
- Interfaces in **forwarding state** behave normally (send/receive all traffic).

- Interfaces in **blocking state** only send/receive STP messages (BPDUs) and a few other special frames.

> ***BPDU (Bridge Protocol Data Unit):*** *STP messages exchanged between switches to elect the root bridge and determine port roles.*

Note: The term "bridge" is used in STP terminology, but in modern networks it refers to **switches**.

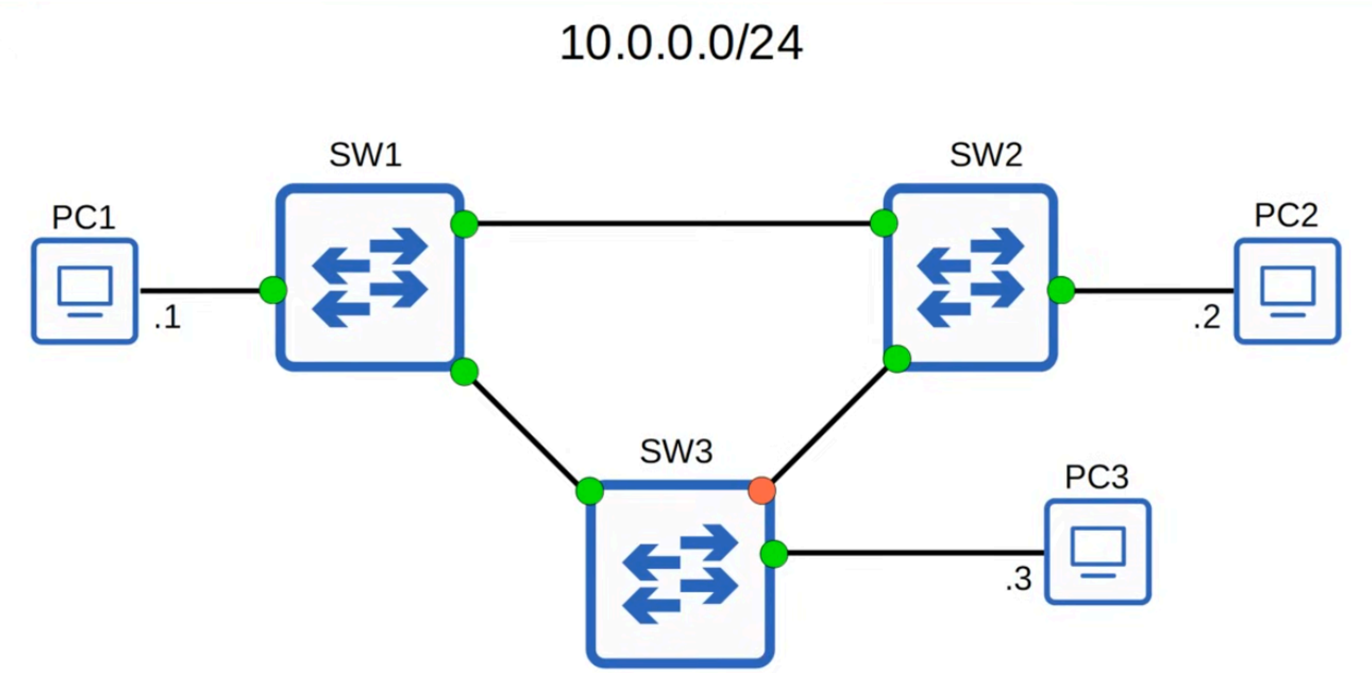
The Solution: Blocking Redundant Ports

To prevent Layer 2 loops, STP places some redundant ports into a **blocking state**. This effectively disables the link from the perspective of normal traffic.

Before STP (No Blocking)

- All three switches are connected in a triangle.
- Broadcast frames loop indefinitely (broadcast storm).

After STP (With Blocking)

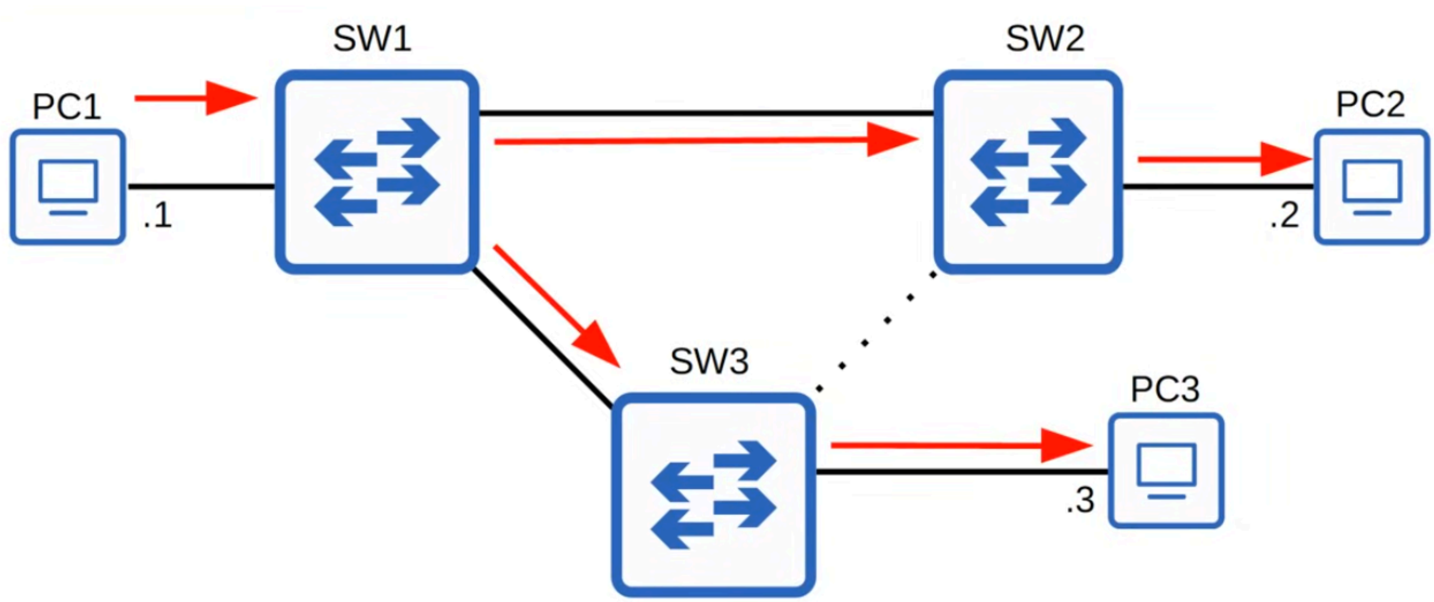


- **Example:** SW3's interface connecting to SW2 is placed in a **blocking state**.
- The link between SW2 and SW3 is effectively disabled — it is as if that link does not exist.

![Triangle topology with one link blocked — effectively a linear chain]

Result with blocked link:

When PC1 sends an ARP request broadcast frame:



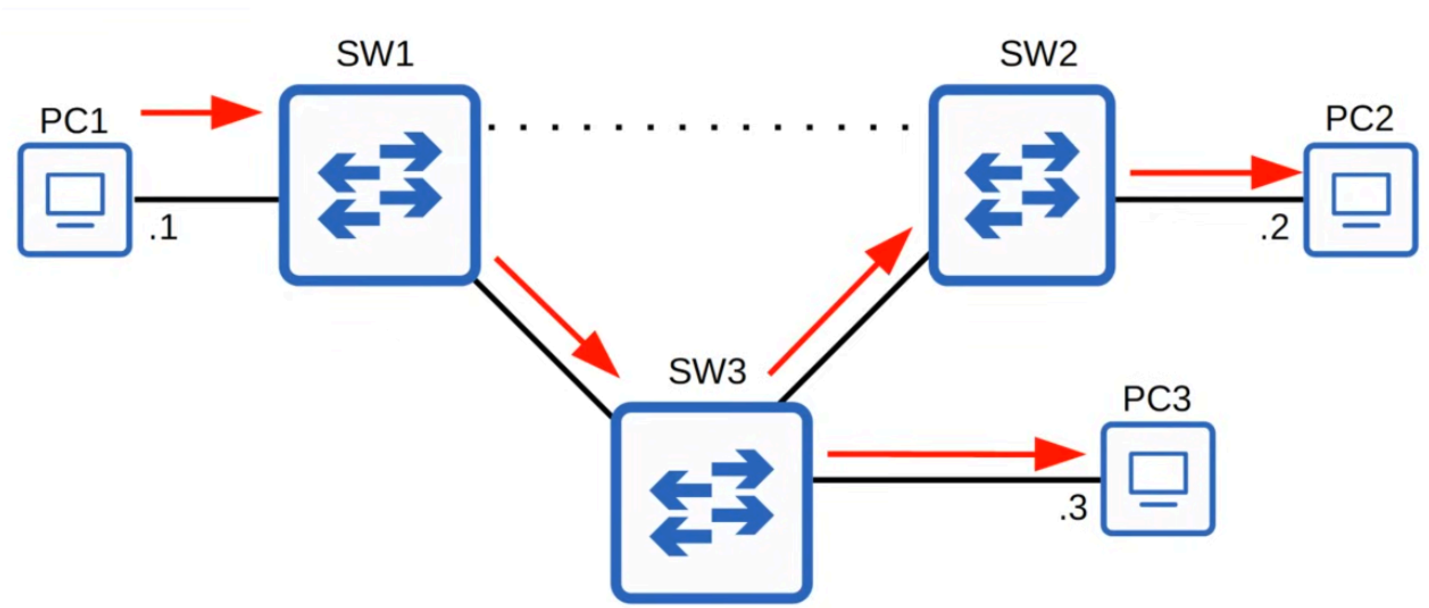
3. No loop exists → **broadcast storm is prevented**

The remaining path is loop-free and functions normally.

Automatic Recovery (Link Failure)

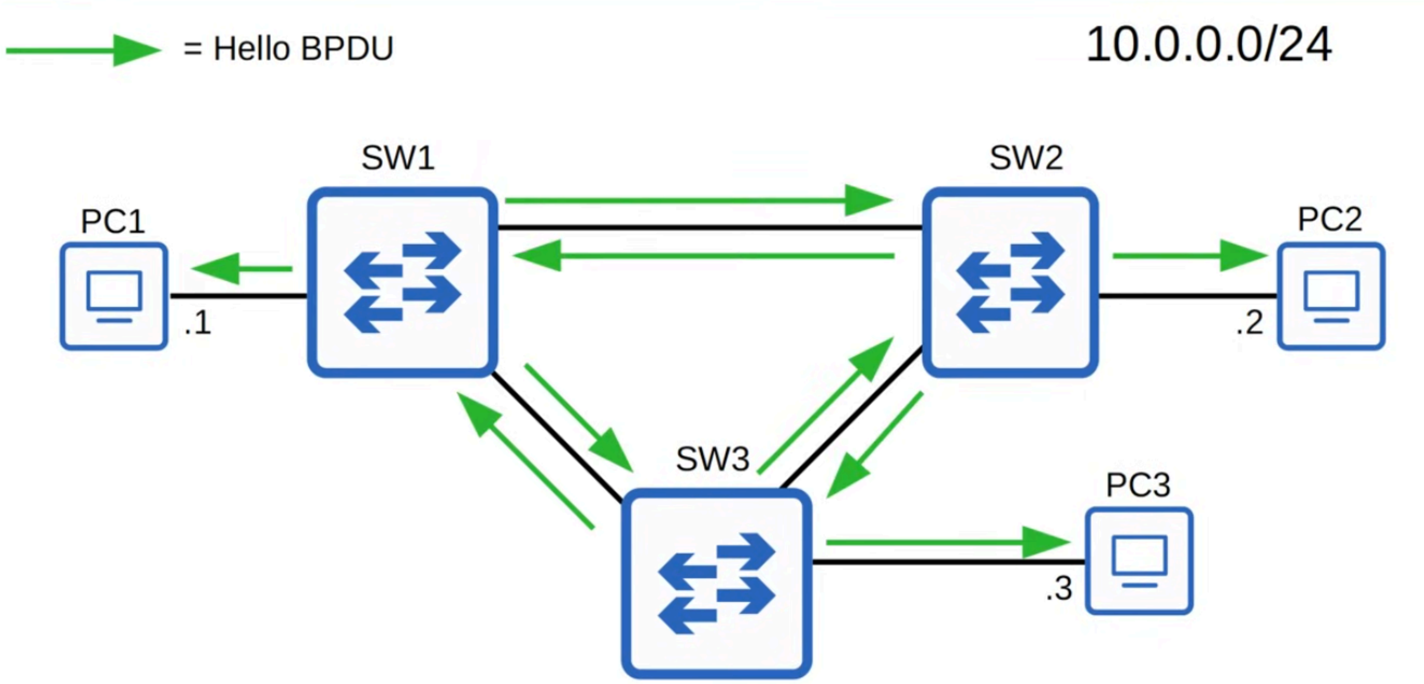
STP is dynamic. If an active (forwarding) link fails, STP **automatically recalculates** the topology and transitions a previously blocked port into a forwarding state.

Example of failure recovery:



- **Normal operation:** SW3's link to SW2 is blocked; traffic flows through SW1.
- **Failure:** SW1's link to SW3 fails.
- **STP reaction:** The blocked port on SW3 (to SW2) transitions to **forwarding**.
- **New path:** PC1 → SW1 → SW2 → SW3 — no loops, network remains operational.

> ***Key point: Blocked ports act as backups. They only send/receive STP messages (BPDUs) during normal operation, but can become active if a forwarding link fails.***
>
> ***These switches will send BPDU's out of each interface, like this:***



They use these BPDU'S to advertise themselves to other switches and to learn about other switches.

STP Process Overview

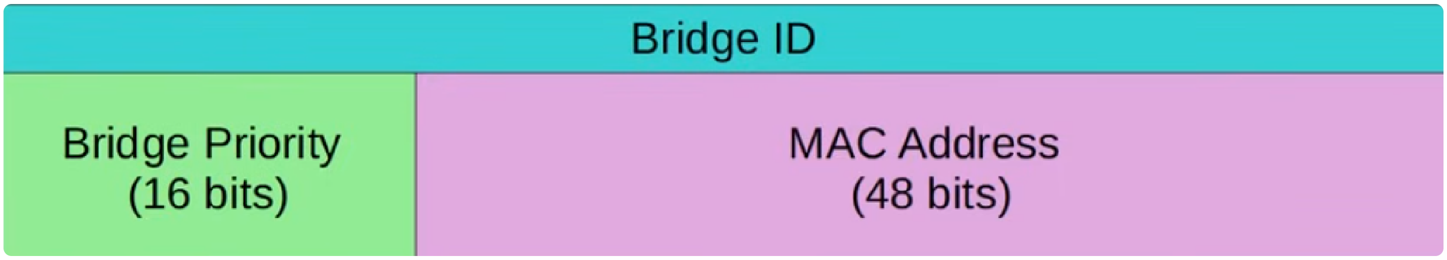
STP follows a defined sequence to determine which ports should forward and which should block:

1. Switches use one field in the STP BPDU, the **Bridge ID** field, to elect a root bridge.
2. Elect a **root bridge**.
3. Each non-root switch selects one **root port**.
4. Each remaining collision domain (link) selects one **designated port**; the other port becomes **non-designated** (blocking).

Step 1: Root Bridge Election

- Switches send **Hello BPDUs** out on all interfaces every 2 seconds by default.
- The **Bridge ID** field in BPDUs is used to elect the root bridge.
- The switch with the **lowest Bridge ID** becomes the root bridge.
- **All ports on the root bridge are designated ports (forwarding).**
  - And other switches in the topology must have a path to reach the root bridge.

The Bridge ID

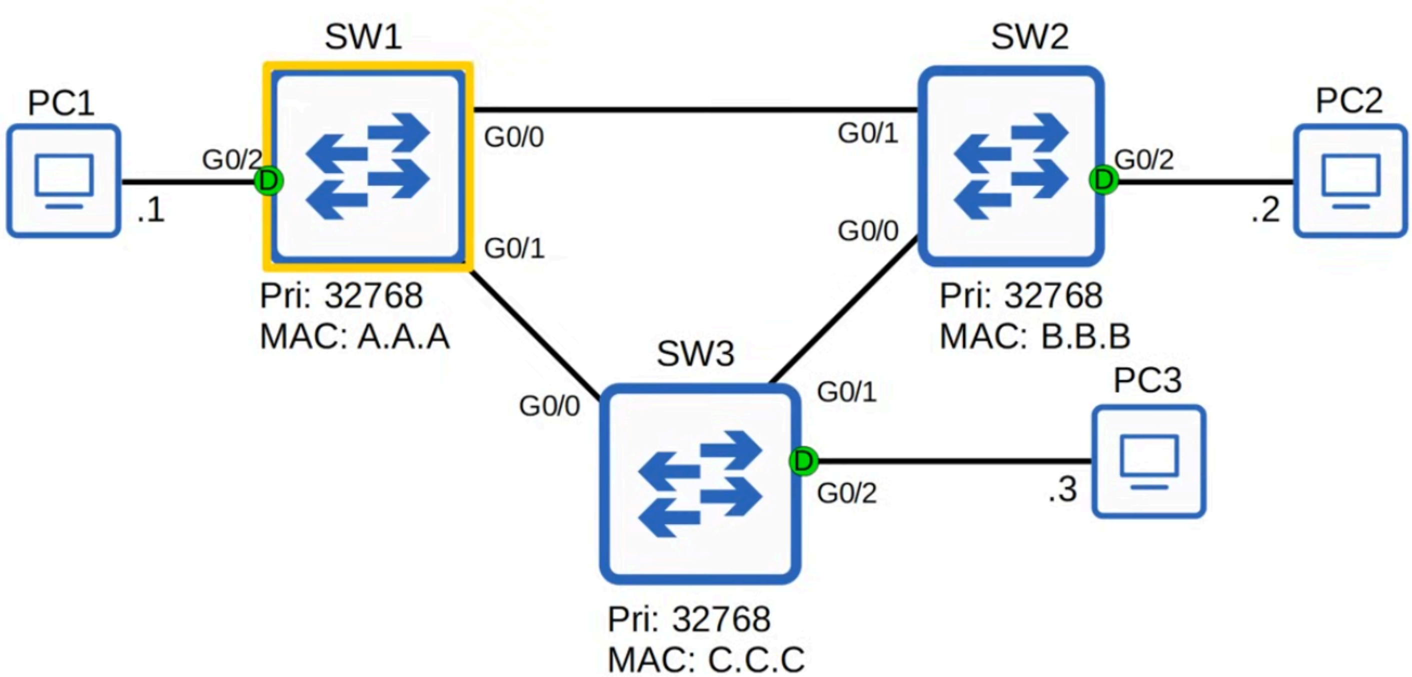


The Bridge ID consists of two parts:

- **Bridge Priority:** 16 bits (originally). The default priority is 32768.
- **MAC Address:** 48 bits (tiebreaker if priorities equal).
  - The bridge with the lowerst MAC address becomes the root bridge.

Topology Setup

Three switches with the **default bridge priority** of **32768**:



Switch	Bridge Priority	MAC Address (shortened)
SW1	32768	A
SW2	32768	B
SW3	32768	C

- All three have the same priority → **MAC address is the tiebreaker**
- Lowest MAC address wins: A (10) < B (11) < C (12)
- **Root bridge: SW1**

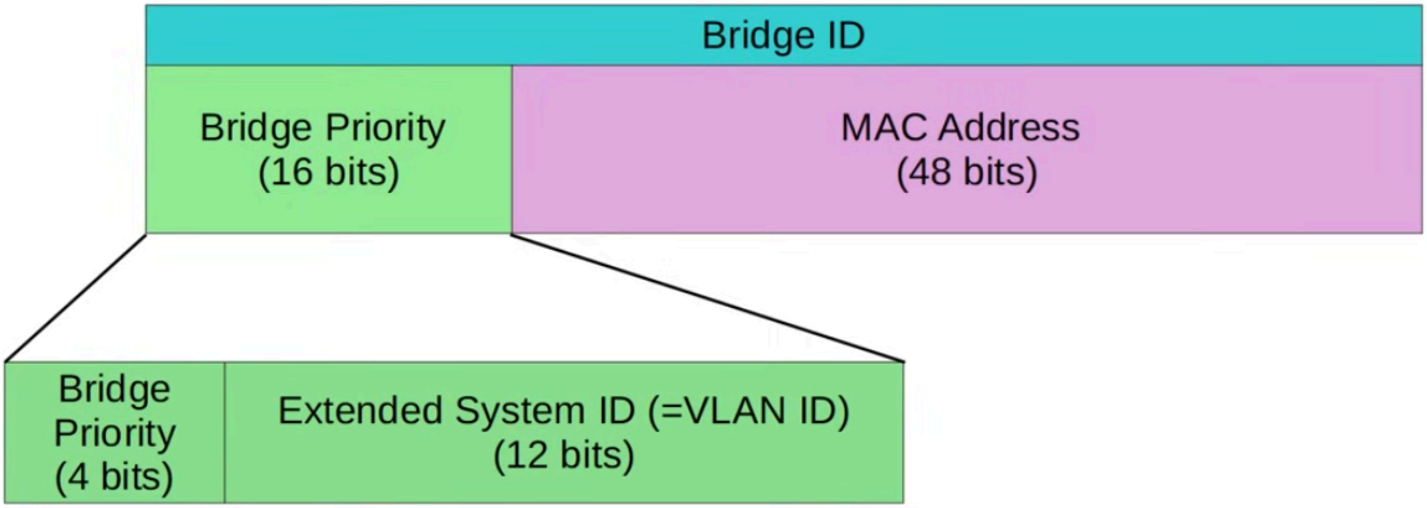
Port States

Port	State	Reason
SW1 G0/2	Forwarding (green)	Connected to PC — no BPDUs received, no loop risk
SW2 G0/2	Forwarding (green)	Connected to PC — same reason
SW3 G0/2	Forwarding (green)	Connected to PC — same reason
All ports on SW1 (root bridge)	**Designated (forwarding)**	All ports on root bridge are designated by default

Key Rule

> ***All ports on the root bridge are designated ports, placed in a forwarding state.***

However, Cisco switches use **PVST (Per-VLAN Spanning Tree)** , which runs a separate STP instance per VLAN. The Bridge ID was updated to include the **Extended System ID** (VLAN ID), 12 bits.



- The total field is still 16 bits: 4 bits for bridge priority + 12 bits for VLAN ID.
- The default priority becomes 32768 + VLAN ID (e.g., 32769 for VLAN 1).
- Bridge priority can only be changed in increments of **4096** (the value of the least significant bit of the 4-bit part).

Why 32768 is the Default

The original Bridge Priority field was **16 bits** in length. By default, the **most significant bit** is set to 1, giving a value of:

215=32768215=32768

Bit	15	14	13	12	11 - 0
Binary	1	0	0	0	0...0
Decimal	32768	-	-	-	-

Updated Bridge ID with Extended System ID

Cisco switches use **PVST (Per-VLAN Spanning Tree)** , which runs a separate STP instance per VLAN. The Bridge ID was updated to include the **Extended System ID** (VLAN ID), 12 bits.

The total field is still **16 bits**:

Component	Bits	Description
Bridge Priority	4 bits	Configurable (increments of 4096)
Extended System ID	12 bits	VLAN ID (automatically added)

Default calculation for VLAN 1:

$$\text{Total Bridge Priority} = \text{Bridge Priority} + \text{VLAN ID} = 32768 + 1 = 32769$$

Configurable Priority Values

The bridge priority portion is only **4 bits**, so it can only be changed in increments of **4096** (the value of the least significant bit of that 4-bit portion).

Valid Bridge Priority Values

Bridge Priority				Extended System ID (VLAN ID)											
32768	16384	8192	4096	2048	1024	512	256	128	64	32	16	8	4	2	1
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Example Calculation

To make a switch root bridge in VLAN 1, you might reduce its priority:

$$\text{Configured value} = 32768$$

$\text{Total in VLAN 1} = 32768 + 1 = 32769$

Field	Value
Bridge Priority (configured)	32768
Extended System ID (VLAN 1)	+1
Total Bridge Priority	32769

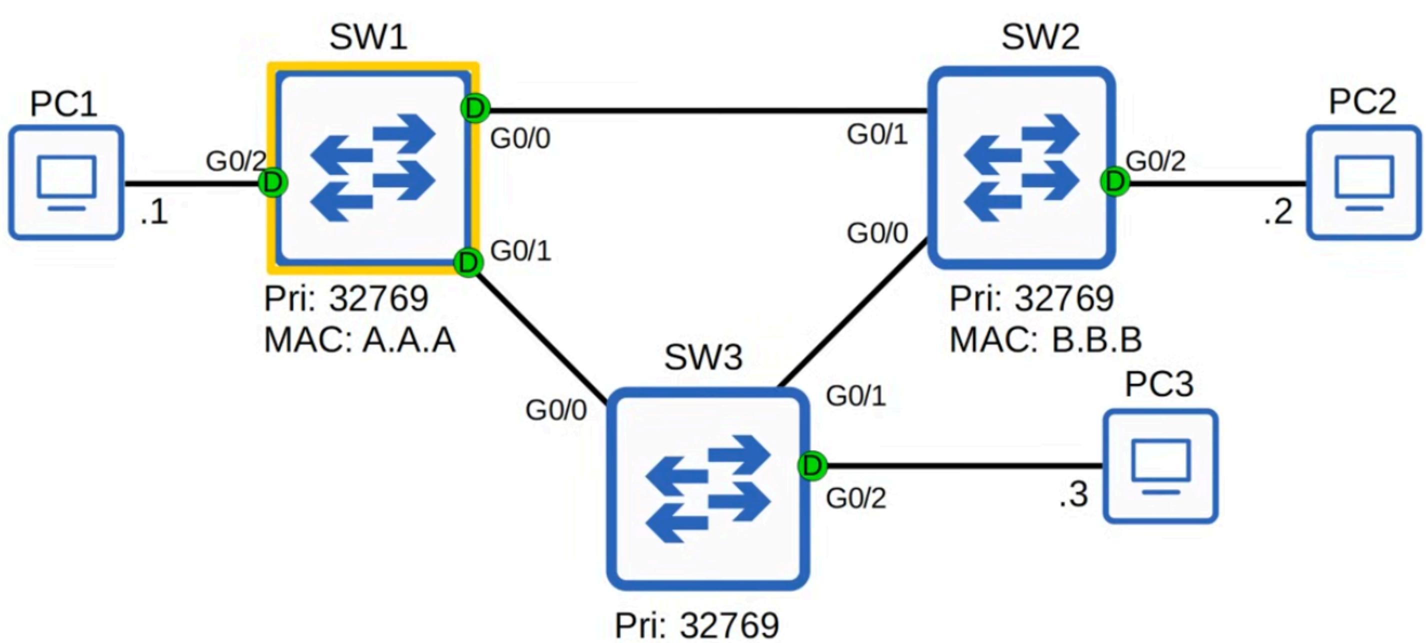
> ***Key Rule: The bridge priority can only be changed in increments of 4096:*** $0, 4096, 8192, 12288, 16384, 20480, 24576, 28672, 32768, \ldots$ ***The extended system ID (VLAN ID) is then added automatically.***

Per-VLAN Spanning Tree (PVST) – Multiple VLANs

Bridge Priority in Different VLANs

With PVST, a separate STP instance runs in **each VLAN**, meaning each VLAN has its own root bridge and port roles.

Default Priority per VLAN (All switches at default 32768)



VLAN	Extended System ID	Total Bridge Priority
1	+1	32769
2	+2	32770
3	+3	32771

Example: Multiple Root Bridges (One per VLAN)

You can configure different switches to be the root bridge in different VLANs:

Switch	VLAN 1 Priority	VLAN 2 Priority	VLAN 3 Priority
SW1	**32769** (root)	32770	32771
SW2	32769	**32770** (root)	32771
SW3	32769	32770	**32771** (root)

- **VLAN 1 root bridge:** SW1
- **VLAN 2 root bridge:** SW2
- **VLAN 3 root bridge:** SW3

> ***Benefit: This allows load balancing — traffic in different VLANs can use different paths, preventing one link from being overloaded while others sit idle.***

All interfaces on the root bridge are designated ports. Designated ports are in a forwarding state.

Few more points about the root bridge:

- When a switch is powered on, it assumes it is the root bridge.

- It will only give up its position if it receives a ‘superior’ BPDU, and superior means a BPDU from a switch with a lower bridge ID.
- Once the topology has converged and all switches agree on the root bridge, only the root bridge sends BPDUs. The reason all switches send BPDUs at first is because they all think they are the root bridge.
- Other switches in the network will forward BPDUs from the root bridge, but they will not generate their own original BPDUs.

Step 2: Root Port Selection

Each non-root switch selects **one** of its interfaces as its **root port**. The root port is the best path to the root bridge and is in a forwarding state.

Each remaining switch will select one of its interfaces to be its root port. The interface with the lowest root cost will be the root port. Root ports are also in the forwarding state.

Selection criteria (in order):

1. The interface with the **lowest root cost** (cumulative cost along the path to the root bridge).
2. If equal root cost, the interface connected to the neighbor with the **lowest Bridge ID**.
3. If still equal, the interface connected to the neighbor with the **lowest port ID**.

> ***Root Cost:*** *The sum of the costs of all outgoing interfaces along the path to the root bridge. (Only outgoing interface costs are added; the receiving interface cost is not counted.)*

Spanning Tree Path Costs

What root cost is?

- Each interface has an associated spanning tree cost.

Interface Speed	Cost
10 Mbps (Ethernet)	100
100 Mbps (Fast Ethernet)	19
1 Gbps (Gigabit Ethernet)	4
10 Gbps (10 Gigabit Ethernet)	2

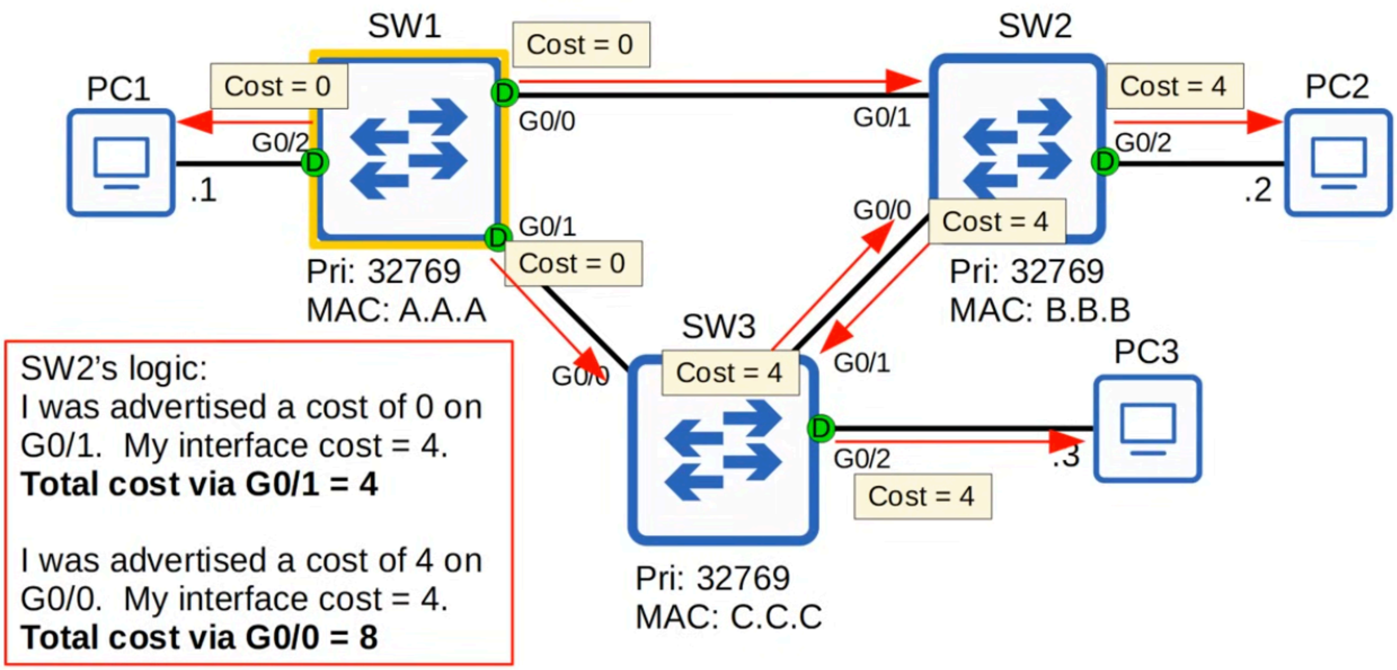
These costs are important for exam memory.

Port connected to a root port is always a designated port (forwarding).

What is Root Cost?

- The **root cost** is the total cost of the **outgoing interfaces** along the path to the root bridge.
- The root bridge itself has a root cost of **0** on all its interfaces.
- **Important:** You only count the cost of the **sending/outgoing** interface, not the receiving interface.

How Switches Calculate Root Cost



1. The root bridge (SW1) advertises a root cost of **0** in its BPDUs.
2. When a neighboring switch (e.g., SW2) receives the BPDU, it **adds the cost of its own outgoing interface** when flooding the BPDU further.
3. This accumulated cost is used to determine the best path to the root bridge.

Root Bridge

- **SW1** is the **Root Bridge**.
- The Root Bridge is the switch with the lowest Bridge ID.
- The Root Bridge always has a root path cost of **0**.
- **All ports on the Root Bridge are Designated Ports.**

Root Ports

A Root Port is the port on a non-root switch with the lowest cost path to the Root Bridge.

- **SW2 Root Port:** G0/1
- **SW3 Root Port:** G0/0

Each non-root switch has only **one Root Port**.

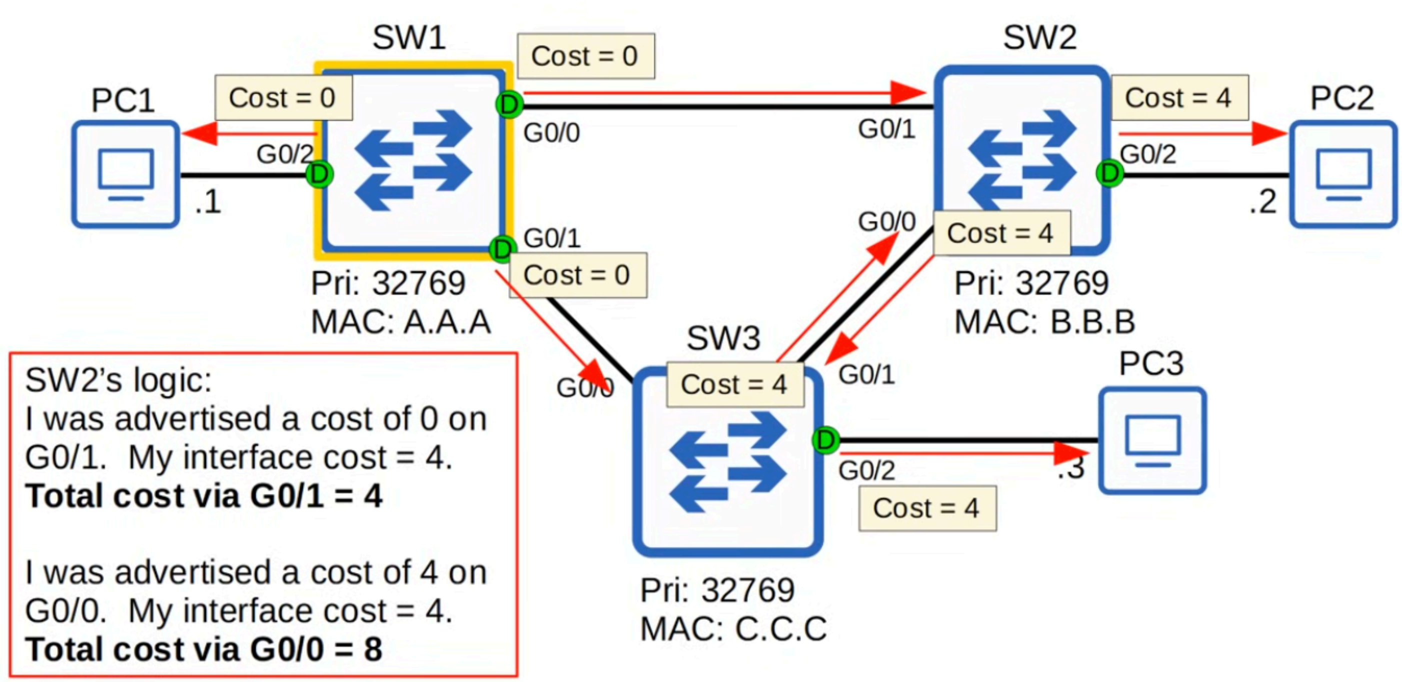
Designated Ports

A Designated Port is the port on a network segment that forwards traffic toward the Root Bridge.

- **SW1 G0/0** → Designated Port
- **SW1 G0/1** → Designated Port
- **SW2 G0/2** → Designated Port (toward PC2)
- **SW3 G0/2** → Designated Port (toward PC3)
- The port connected to another switch's Root Port must be a Designated Port.

SW2 Calculates the Cost Through SW3

The author is showing the total cost of each possible route a switch could take to reach the root bridge.



If SW2 wants to reach the root through SW3, it must send traffic out of its G0/0 interface:

The path toward the root is:

```
SW2 → SW3 → SW1
```

1. SW2 sends out G0/0 → cost 4
2. SW3 sends out G0/0 → cost 4
3. SW1 is the root → cost 0

Total: 4 + 4 = 8

If SW2 sends directly to SW1:

```
SW2 → SW1
```

1. SW2 sends out G0/1 → cost 4
2. SW1 is the root → cost 0

Total: 4

If SW3 wants to reach the root through SW2, it must send traffic out of its G0/1 interface:

```
SW3 → SW2 → SW1
```

SW3 sends out G0/1 → cost 4

1. SW2 sends out G0/1 → cost 4
2. SW1 is the root → cost 0

Total: 4 + 4 = 8

If SW3 sends directly to SW1:

```
SW3 → SW1
```

1. SW3 sends out G0/0 → cost 4
2. SW1 is the root → cost 0

Total: 4

Important: The ports connected to another switch's root port must be designated. Because the root is the switch's path to the root bridge, another switch must not block it.

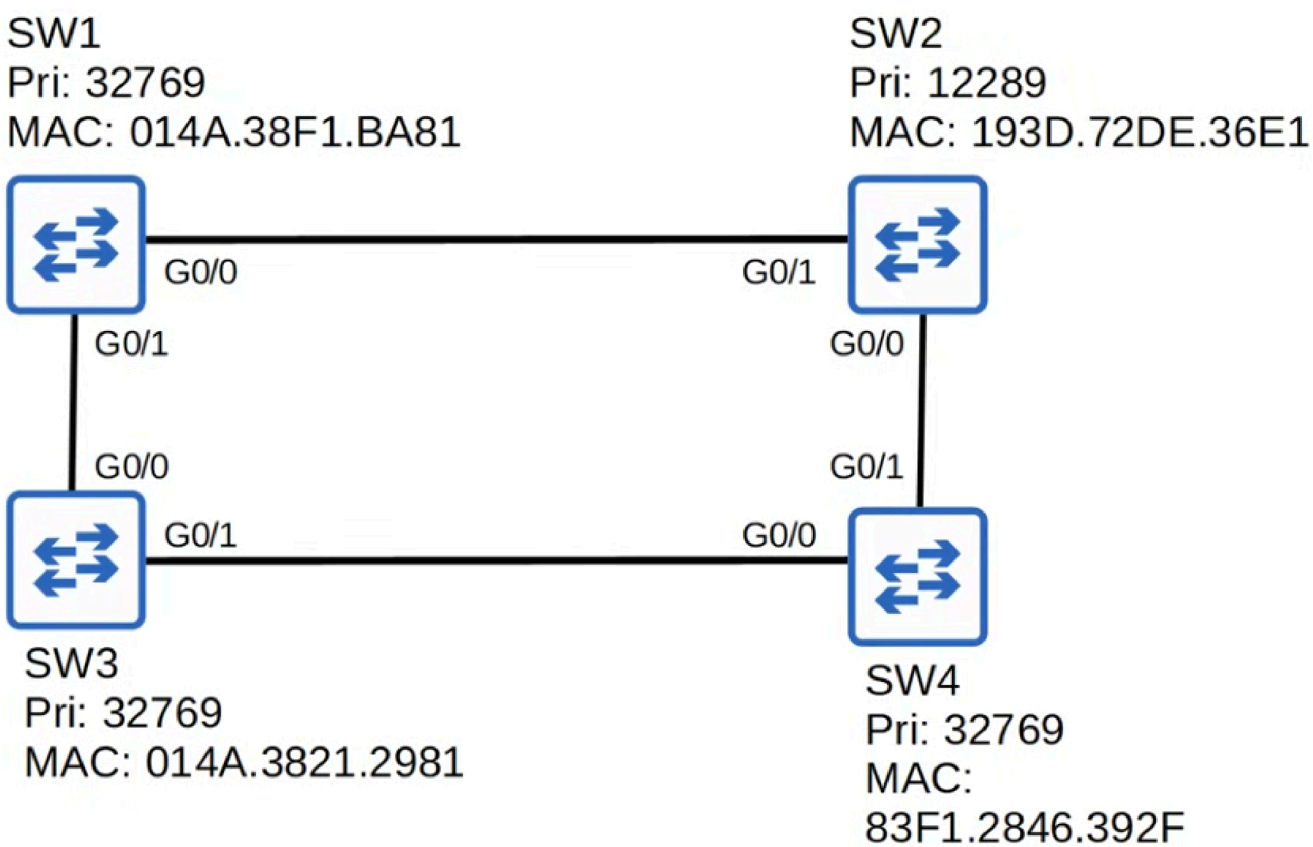
Summary 2 for STP:

1. One switch is elected as the root birdge. All ports on the root bridge are designayed ports (forwarding state). Root bridge selection:
    1. Lowest bridge ID
2. Each remaining switch will select one of its interfaces to be its root port (forwarding state). Ports acriss from the root port are always designated ports.

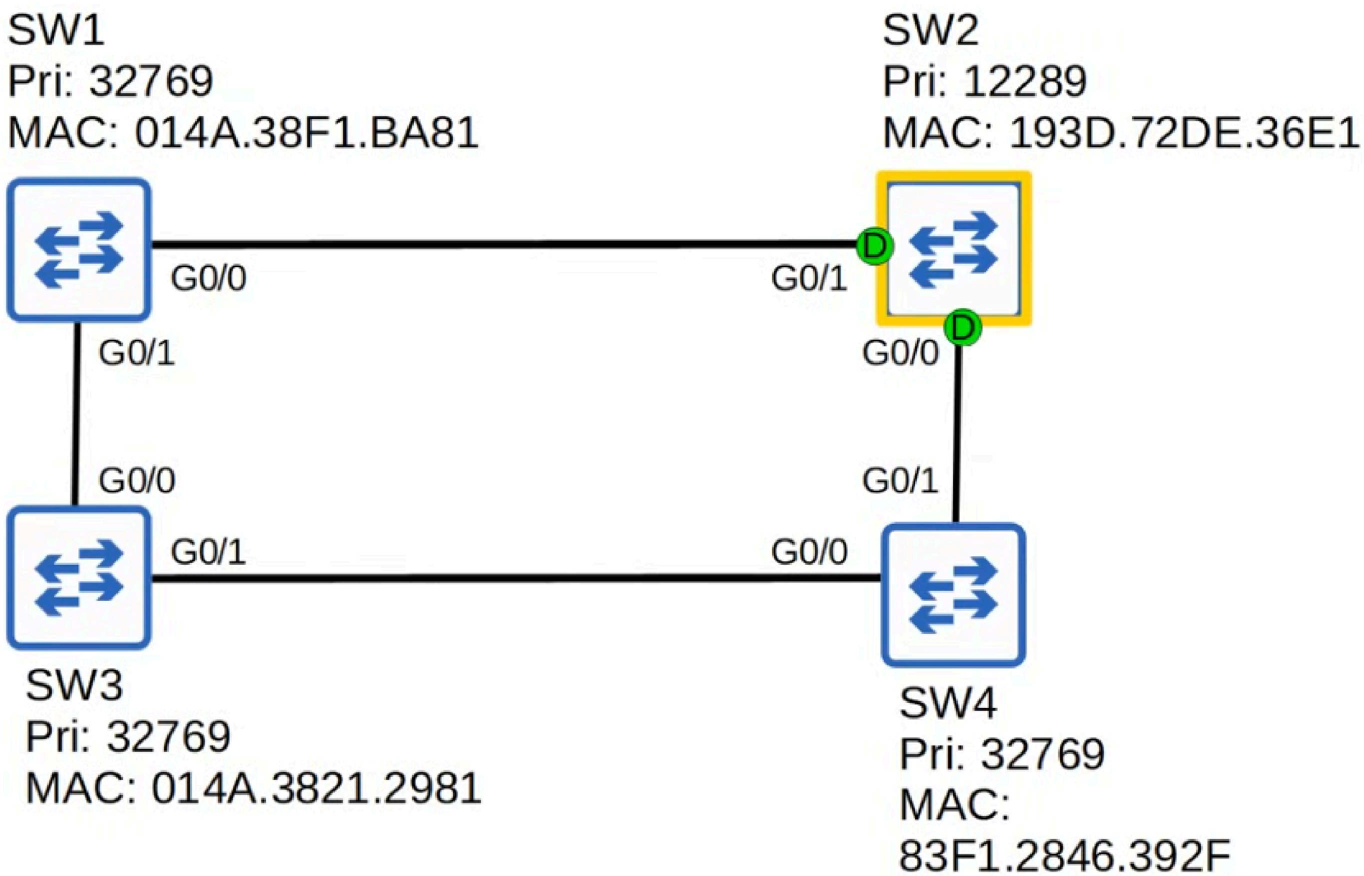
Root port selection:

1. Lowest root cost
2. Lowest neighbor bridge ID
3. Lowest neighbor port ID

Example for Lowest Neighbor Bridge ID



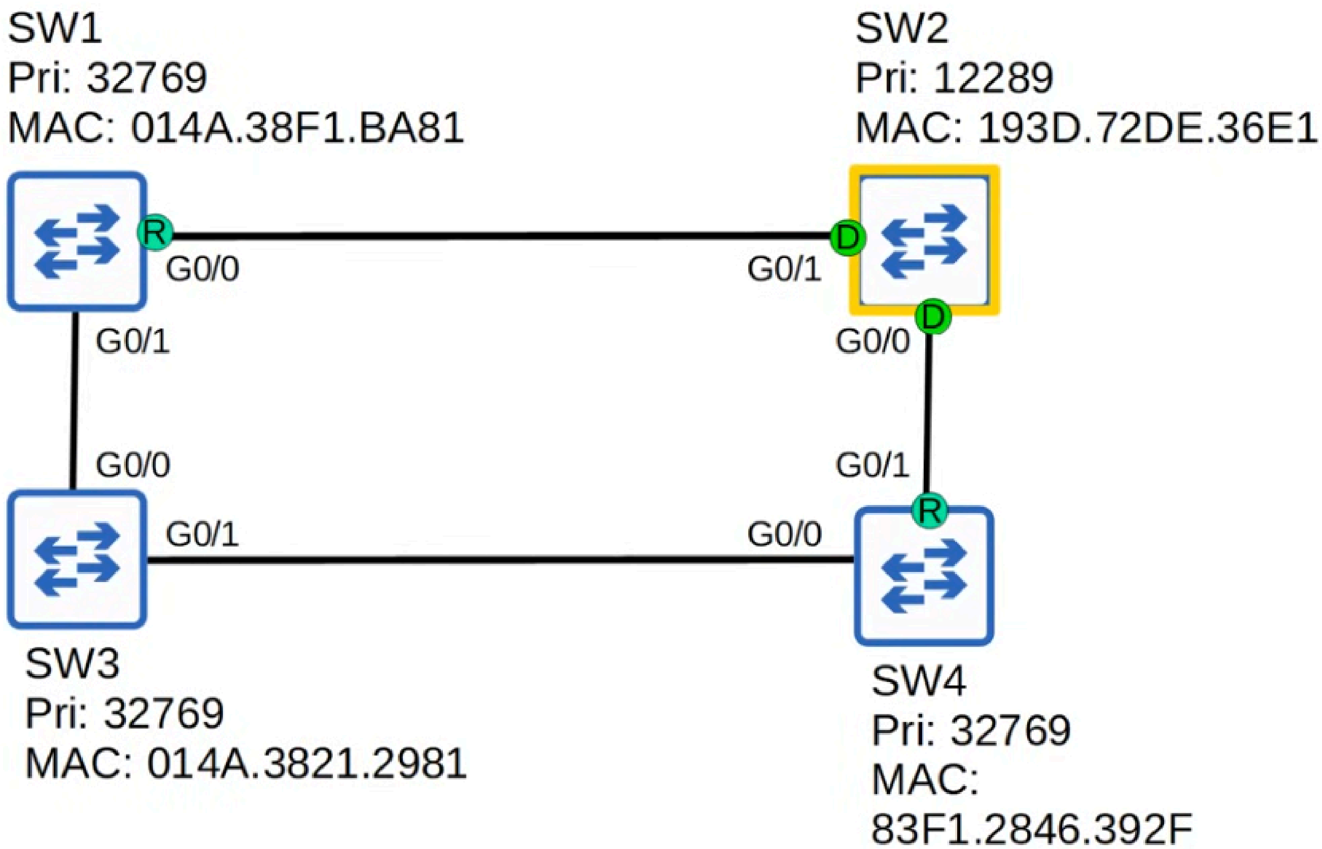
Step 1: Root Bridge Election



You're correct — **SW2** is the root bridge because it has the lowest priority. All of SW2's ports are **designated ports** (forwarding state).

Step 2: Root Port Selection

Each *non-root* switch (SW1, SW3, SW4) must choose **one** root port — the interface with the best path to SW2.



SW1's Choice

- **G0/0** → connects directly to SW2 → root cost = **4** ✅
- Any other port would have a higher cost.

Result: G0/0 is the root port (forwarding).

SW4's Choice

- **G0/1** → connects directly to SW2 → root cost = **4** ✅
- Any other port would have a higher cost.

Result: G0/1 is the root port (forwarding).

SW3's Choice (The Tricky One)

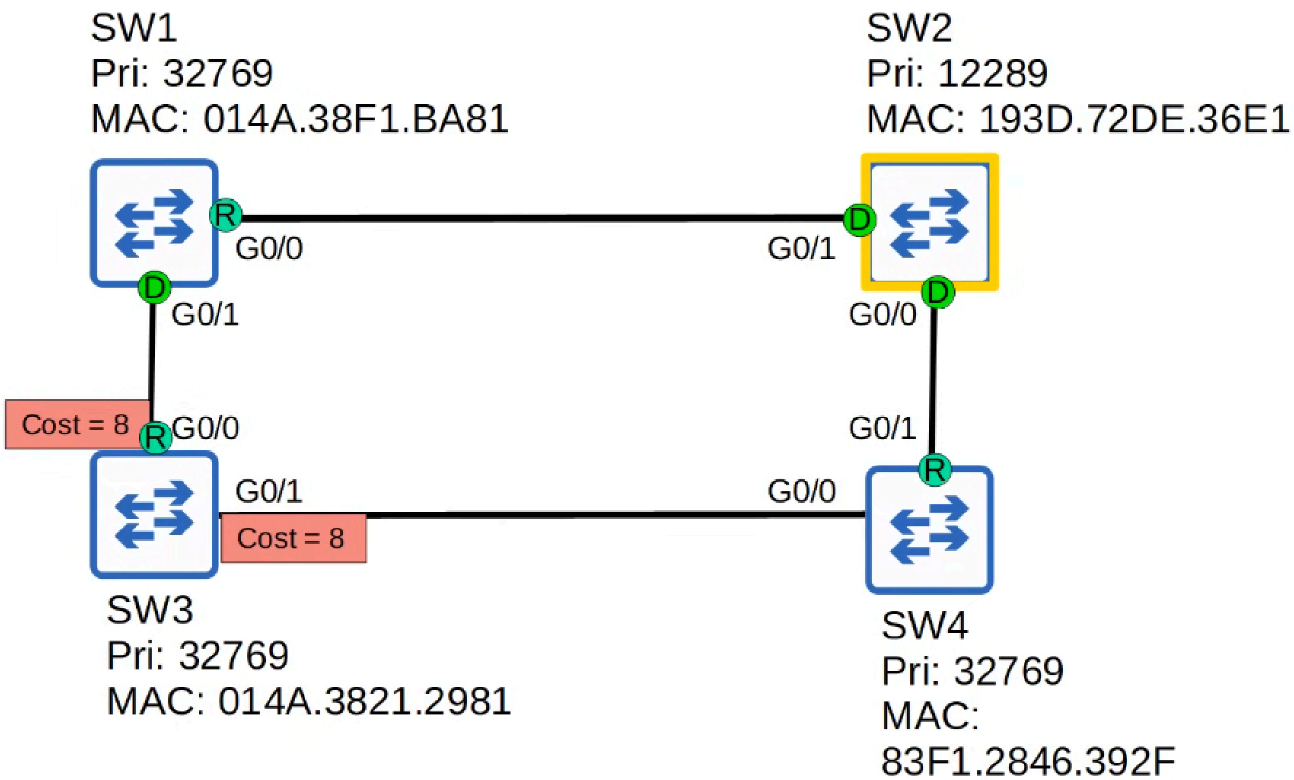
SW3 has two paths to SW2:

Path	Cost Calculation	Total Cost
Via G0/0 → SW1 → SW2	$4 + 4 4 + 4$	**8**
Via G0/1 → SW4 → SW2	$4 + 4 4 + 4$	**8**

Both paths have the same root cost of **8** — so we need a **tiebreaker**.

Tiebreaker: Lowest Neighbor Bridge ID

Path	Neighbor	Bridge ID Comparison
Via G0/0	**SW1**	Priority 32769, MAC `0000.001A`
Via G0/1	**SW4**	Priority 32769, MAC `0000.001D`



Same priority, but SW1's MAC address (`001A`) is lower than SW4's (`001D`).

Result: SW3 selects **G0/0** as its root port.

Why is SW1's G0/1 Designated?

SW3's root port is G0/0, which connects to **SW1's G0/1**. The rule says:

> ***A port connected to another switch's root port must be a designated port (forwarding).***

So SW1's G0/1 becomes **designated**, even though SW1 is not the root bridge.

The Key Point: It's the *Neighbor's* Port ID

The Port ID (Prio.Nbr)

Show spanning tree command, output:

```
Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Gi0/0            Desg FWD 4         128.1    P2p
Gi0/1            Desg FWD 4         128.2    P2p
```

The **Prio.Nbr** column shows the **STP Port ID**, which is made of two parts:

Part	Value	Meaning
Prio	128	Port priority (default, can be changed)
Nbr	1, 2, 3, etc.	Unique port number on the switch

So `128.1` = port priority `128` + port number `1`.

How to Read It

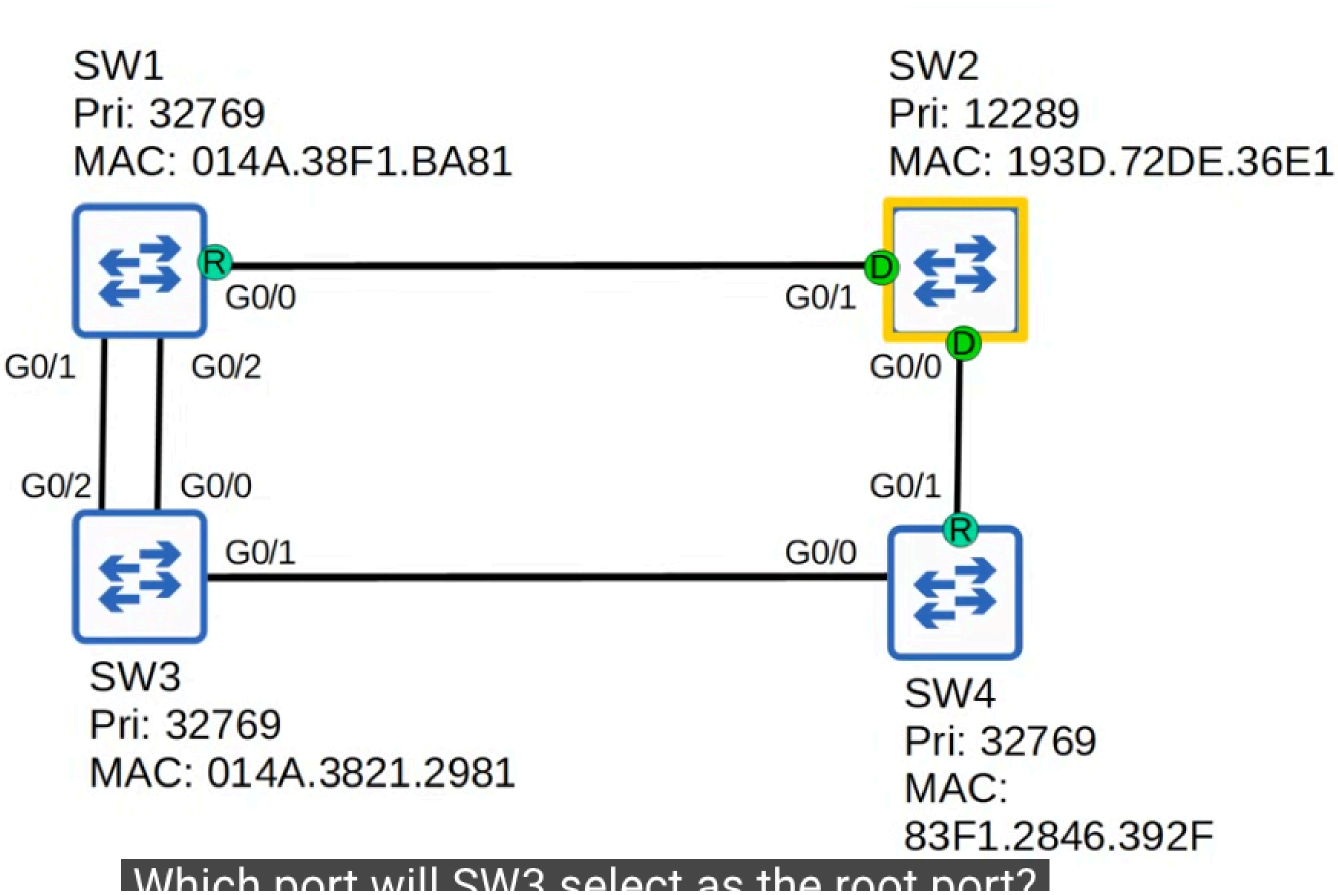
- **Gi0/0** → Prio.Nbr = **128.1** (port number 1)
- **Gi0/1** → Prio.Nbr = **128.2** (port number 2)
- **Gi0/2** → Prio.Nbr = **128.3** (port number 3)

This is the detail you must remember: **you look at the neighbor switch's port ID**, not your own switch's port ID.

Important: Similar to a bridge id, where the MAC address is used as a tie breaker if the priorities tie. In this case the port number is used as a tie breaker if the priorities tie.

Example: Two Links Between SW1 and SW3

Let's say SW1 is the root bridge, and both links are Gigabit Ethernet (cost 4 each):



SW3's Logic

- **Path via G0/0:** root cost = 4 (via SW1 G0/0)
- **Path via G0/2:** root cost = 4 (via SW1 G0/1)

Same root cost. And both go to the **same neighbor switch** (SW1), so the bridge ID is the same too.

The Deciding Factor: Neighbor's Port ID

SW3 compares the **ports on SW1** that each path connects to:

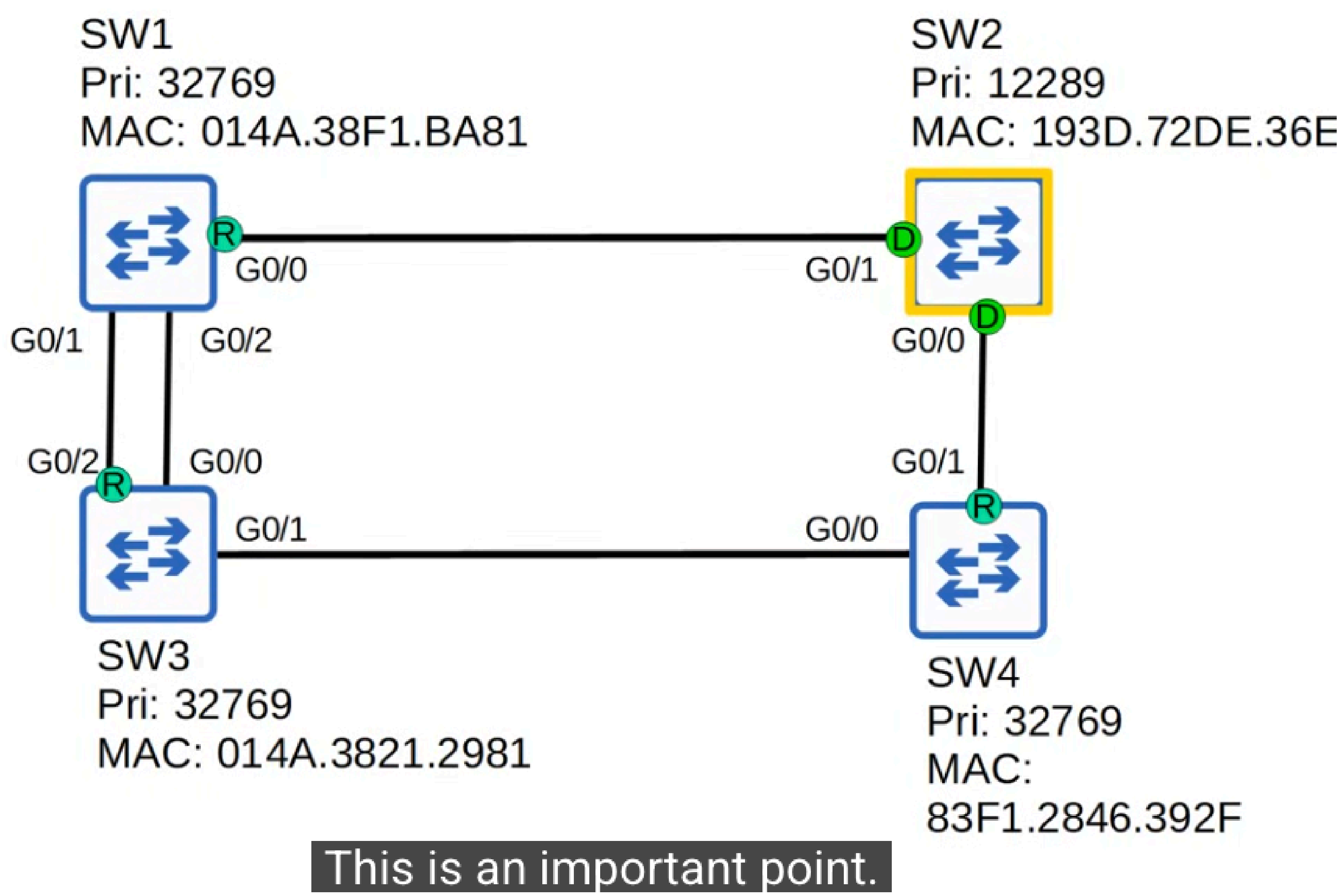
SW3 Port	Connected to SW1 Port	SW1 Port ID
G0/0	G0/2	128.3
G0/2	G0/1	128.2

SW1's G0/2 (port ID 128.3) is greater than G0/1 (128.2). But wait — the source says SW3 selects **G0/2**, which connects to SW1's **G0/1**...

Why G0/2 Wins?

Actually, let me re-read carefully. The source says:

> ***"The answer is G0/2, because it is connected to a lower port ID on the neighbor switch, SW1."***



Hmm, that would mean SW1's G0/1 has a **lower** port ID than G0/2. That's unusual since the default port priority is 128 for all ports. But the port number of G0/1 is 1 and G0/2 is 2, so G0/0 (128.1) should be lower.

This might be a case where:

- The port priorities were manually changed
- Or the port numbering is different (e.g., G0/2 might be port 128.3 and G0/1 might be 128.1)

Regardless, the **rule** is correct: **the neighbor's lower port ID wins**.

Step 3: Designated Port Selection (The Final Piece)

After the root bridge is elected and each non-root switch selects its root port, the remaining links need **one designated port** and **one non-designated (blocking) port** per link.

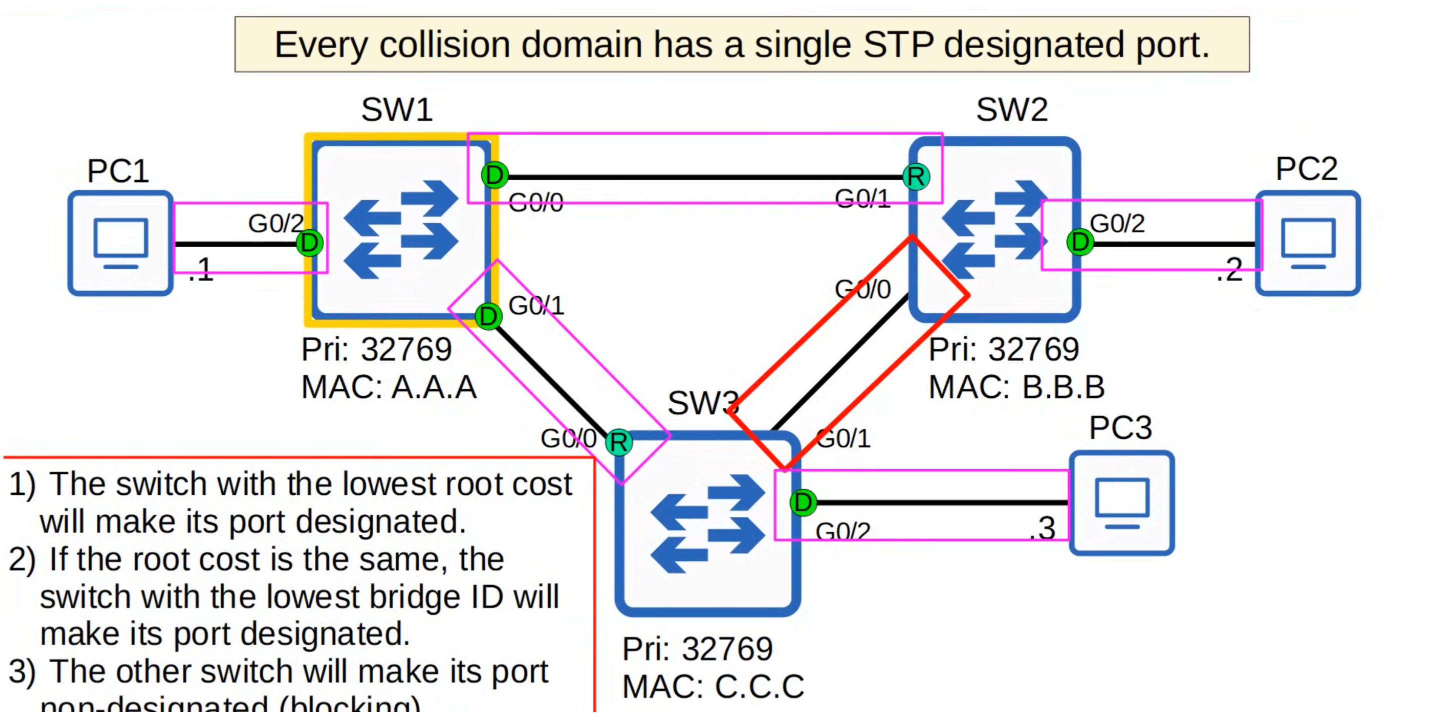
The Key Rule

> ***Every collision domain (link) has exactly one designated port.***

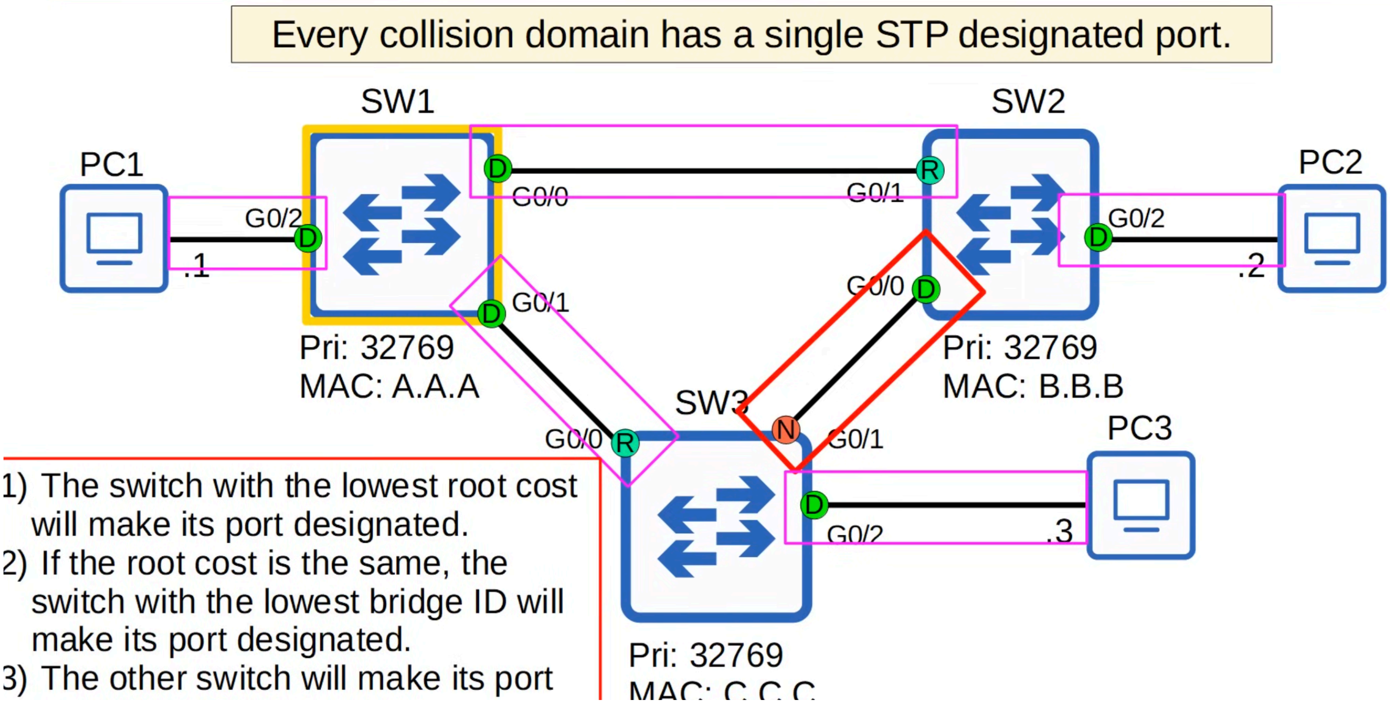
How to Choose the Designated Port

For each link between two switches, the port with the **lower root cost** becomes **designated** (forwarding). The other port becomes **non-designated** (blocking).

Your Example (SW2 ↔ SW3)



Switch	Root Cost	Bridge ID (Priority + MAC)
SW2	**4** (via G0/1)	Priority 32769, MAC `0000.001B`
SW3	**4** (via G0/0)	Priority 32769, MAC `0000.001C`



- **Root cost is tied** (both 4)
- **Tiebreaker: lowest bridge ID** → SW2 (001B < 001C) wins

Result	Port	Role	State
SW2's G0/0	Designated	Forwarding ✅	
SW3's G0/1	Non-Designated	**Blocking** 🛑	

Why Not Block Both?

You might ask: "Why not block both ports on that link?" Because that would completely isolate SW2 and SW3 with no backup path. STP blocks **only one** port per link — the one on the switch with the **higher root cost** (or higher bridge ID if costs tie). The other port stays forwarding, ready to take over if a primary link fails.

Complete STP Selection Process

Spanning Tree Protocol (STP) – Selection Process Summary

Step 1: Root Bridge Election

One switch is elected as the **Root Bridge**. All ports on the Root Bridge become **Designated Ports** (forwarding state).

Root Bridge Selection Criteria:

- Lowest Bridge ID

Step 2: Root Port Selection

Each remaining (non-root) switch selects **one** of its interfaces as its **Root Port** (forwarding state). Ports directly connected to a Root Port are always Designated Ports.

Root Port Selection Criteria (in order):

1. Lowest Root Cost
2. Lowest Neighbor Bridge ID
3. Lowest Neighbor Port ID

Step 3: Designated Port Selection

Each remaining collision domain (link between two switches) selects **one** interface as its **Designated Port** (forwarding state). The other interface on the same link becomes a **Non-Designated Port** (blocking state).

Designated Port Selection Criteria (in order):

1. Interface on the switch with the lowest Root Cost
2. Interface on the switch with the lowest Bridge ID

Final Port States Summary

Port Role	State	Notes
Root Bridge Ports	Forwarding	All ports on the Root Bridge
Root Port	Forwarding	One per non-root switch
Designated Port	Forwarding	One per collision domain
Non-Designated Port	Blocking	Remaining port on the link