НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

МИЭМ НИУ ВШЭ

ИНФОКОММУНИКАЦИОННЫЕ СИСТЕМЫ И СЕТИ

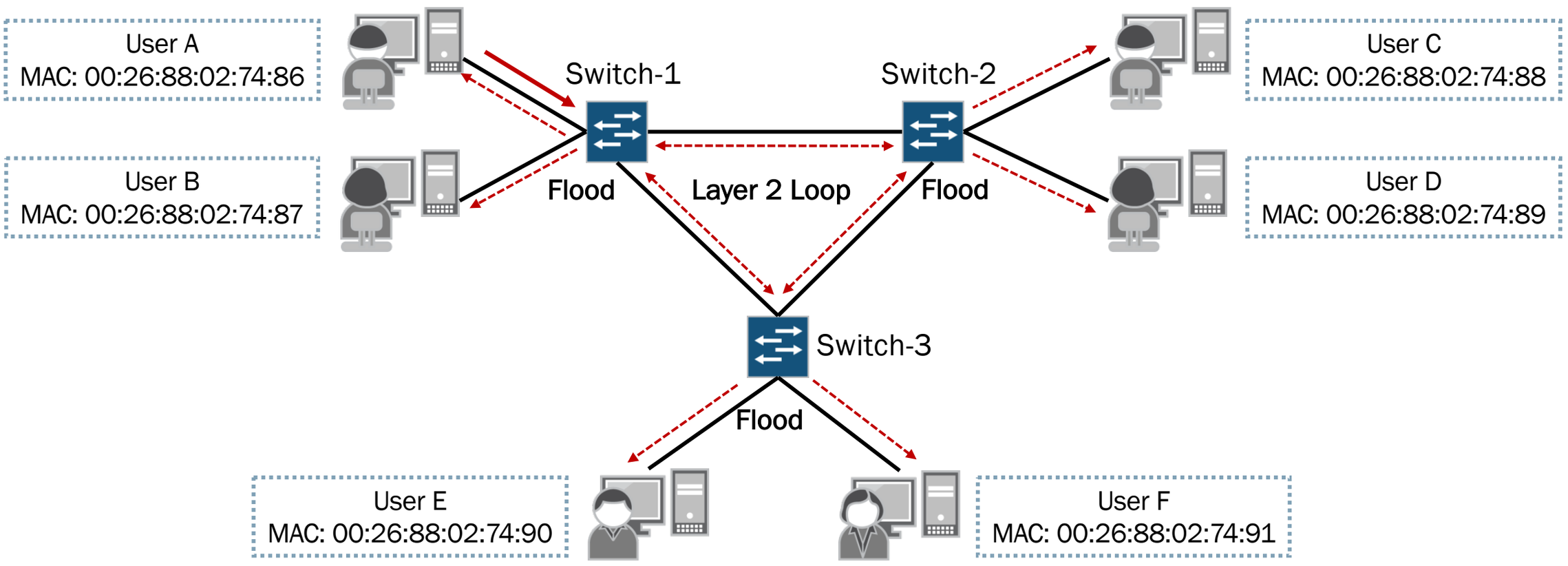
STP

Москва, 2018

What If ...?

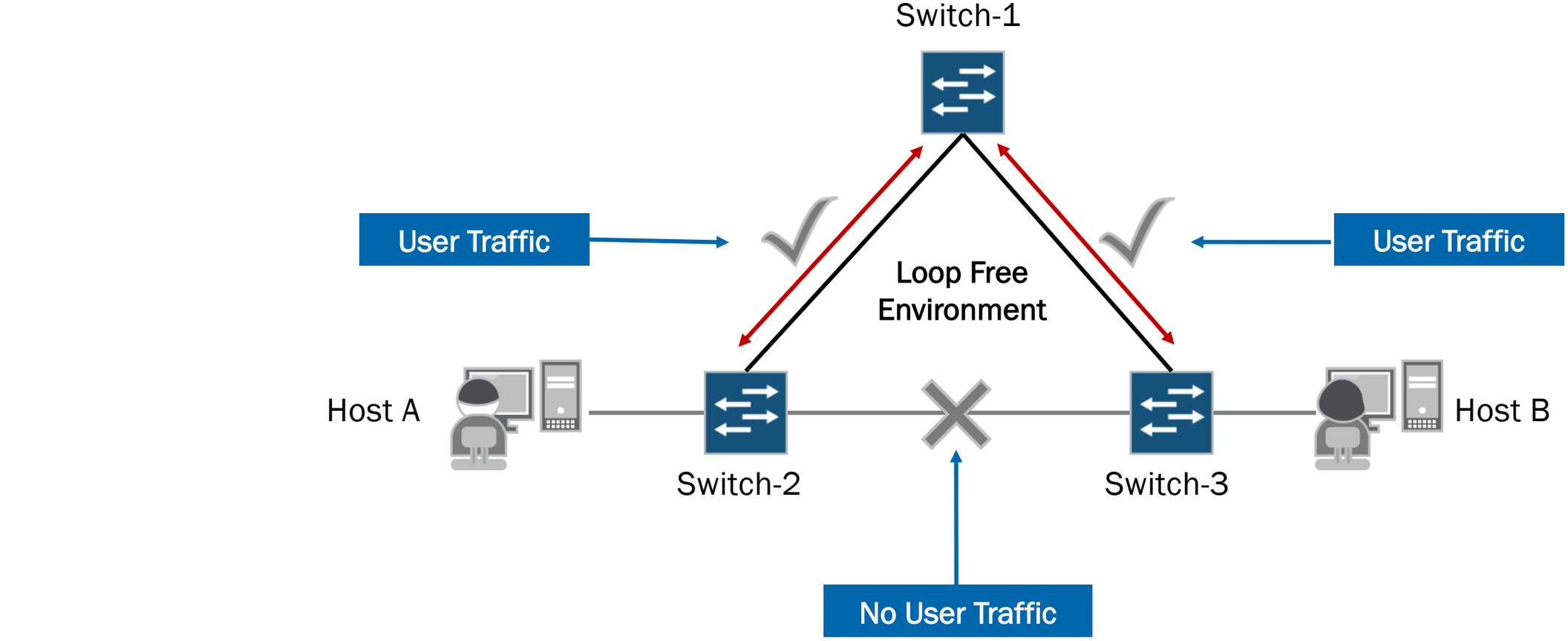
- What if a broadcast frame or a frame with an unknown destination MAC address were sent into a Layer 2 network with redundant paths?

Example: Source MAC: 00:26:88:02:74:86 / Destination MAC: 00:26:88:02:74:95



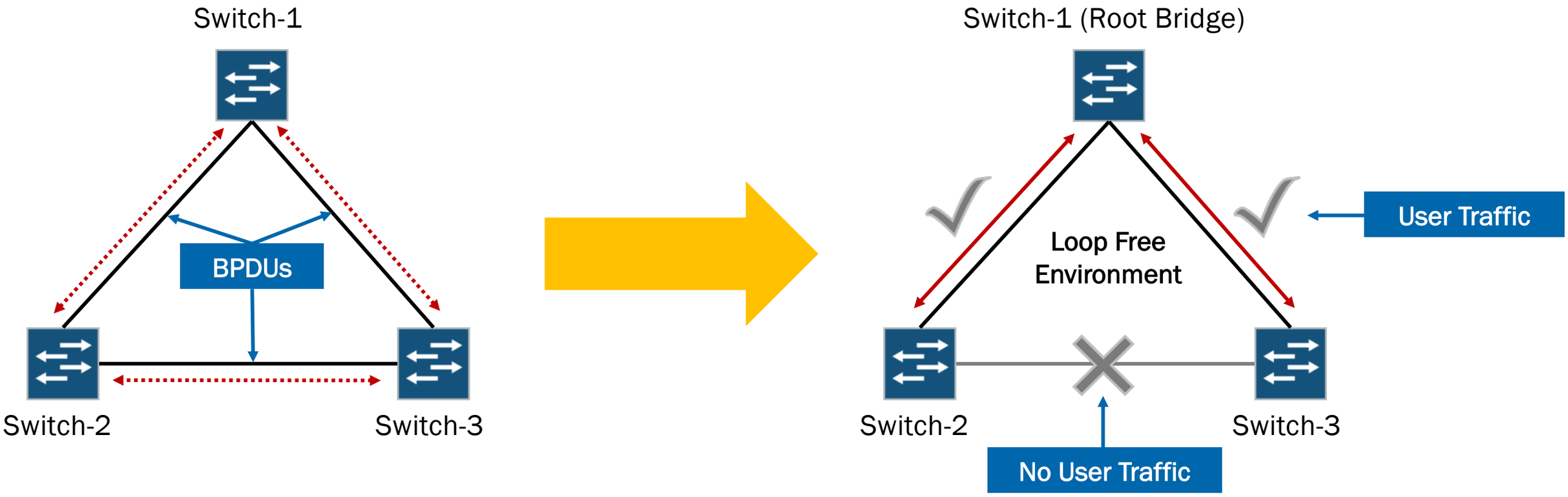
Spanning Tree Protocol

- STP
  - Defined in the IEEE 802.1D-1998 specification
  - Builds loop-free paths in redundant Layer 2 networks
  - Automatically rebuilds tree when topology changes



How Does it Work?

- Steps for creating a spanning tree include:
  1. Switches exchange bridge protocol data units (BPDUs)
  2. Root bridge is elected
  3. Port role and state are determined
  4. Tree is fully converged



Terms and Concepts (1 of 2)

- Key terms and concepts of STP:
  - *Bridge ID*: Unique identifier for each switch
  - *Root bridge*: Switch with the lowest bridge ID
  - *Root port*: The port on each bridge closest to the root bridge
  - *Root path cost*: A bridge's calculated cost to get from itself to the root bridge
    - Equal to the received root path cost from configuration BPDUs plus the port cost of the root port on the bridge
  - *Port cost*: Every interface on a bridge has an assigned port cost value
    - Used in the calculation of the root path cost for the local bridge
    - Configurable value (1–200000000)
    - The default value is 20000 for 1 Gigabit Ethernet

Terms and Concepts (2 of 2)

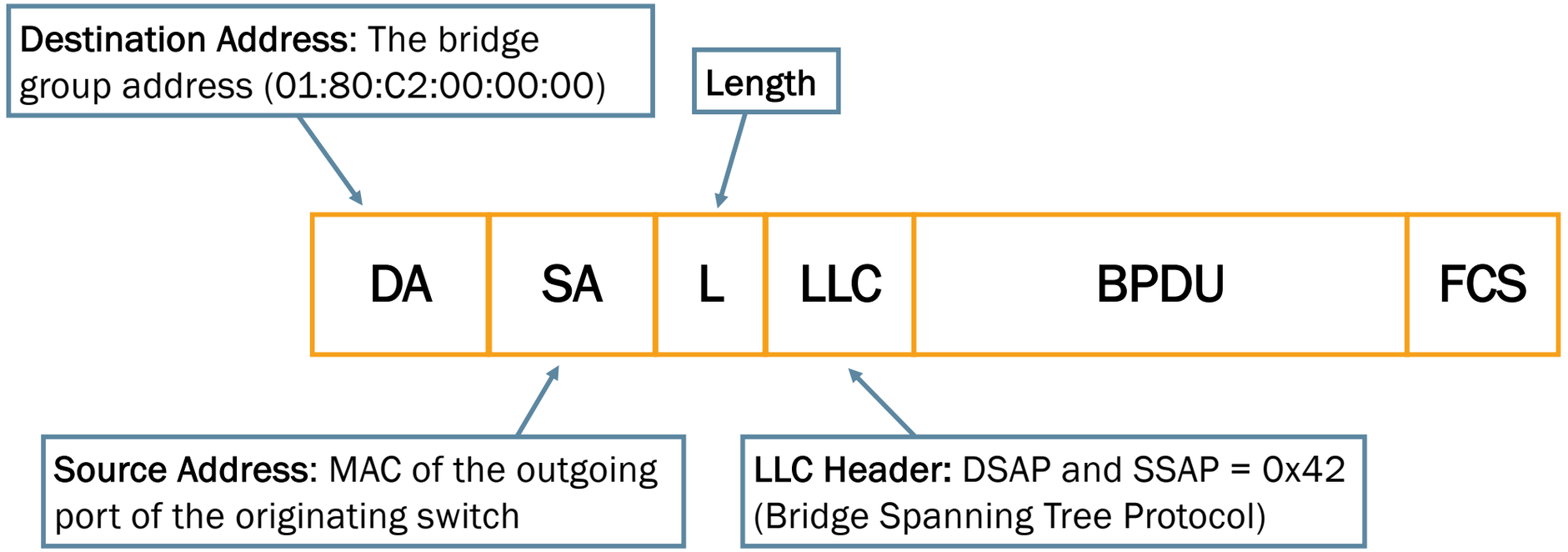
- Key terms and concepts of STP (contd.):
  - *Designated bridge*: A switch representing the LAN segment
  - *Port ID*: A unique identifier for each port on each switch
  - *Designated port*: The designated bridge's forwarding port on a LAN segment
    - The port used by a designated bridge to send traffic from the direction of the root to the LAN or from the LAN toward the root
  - *Bridge protocol data unit*: Packets used to exchange information between switches
    - Configuration BPDU
    - Topology change notification BPDU

Port States

- Each individual port of each bridge can be in one of four states:
  - Blocking
    - The port drops all data packets and listens to BPDUs
    - The port is not used in active topology
  - Listening
    - The port drops all data packets and listens to BPDUs
    - The port is transitioning and will be used in active topology
  - Learning
    - The port drops all data packets and listens to BPDUs
    - The port is transitioning and the switch is learning MAC addresses
  - Forwarding
    - The port receives and forwards data packets and sends and receives BPDUs
    - The port has transitioned and the switch continues to learn MAC addresses

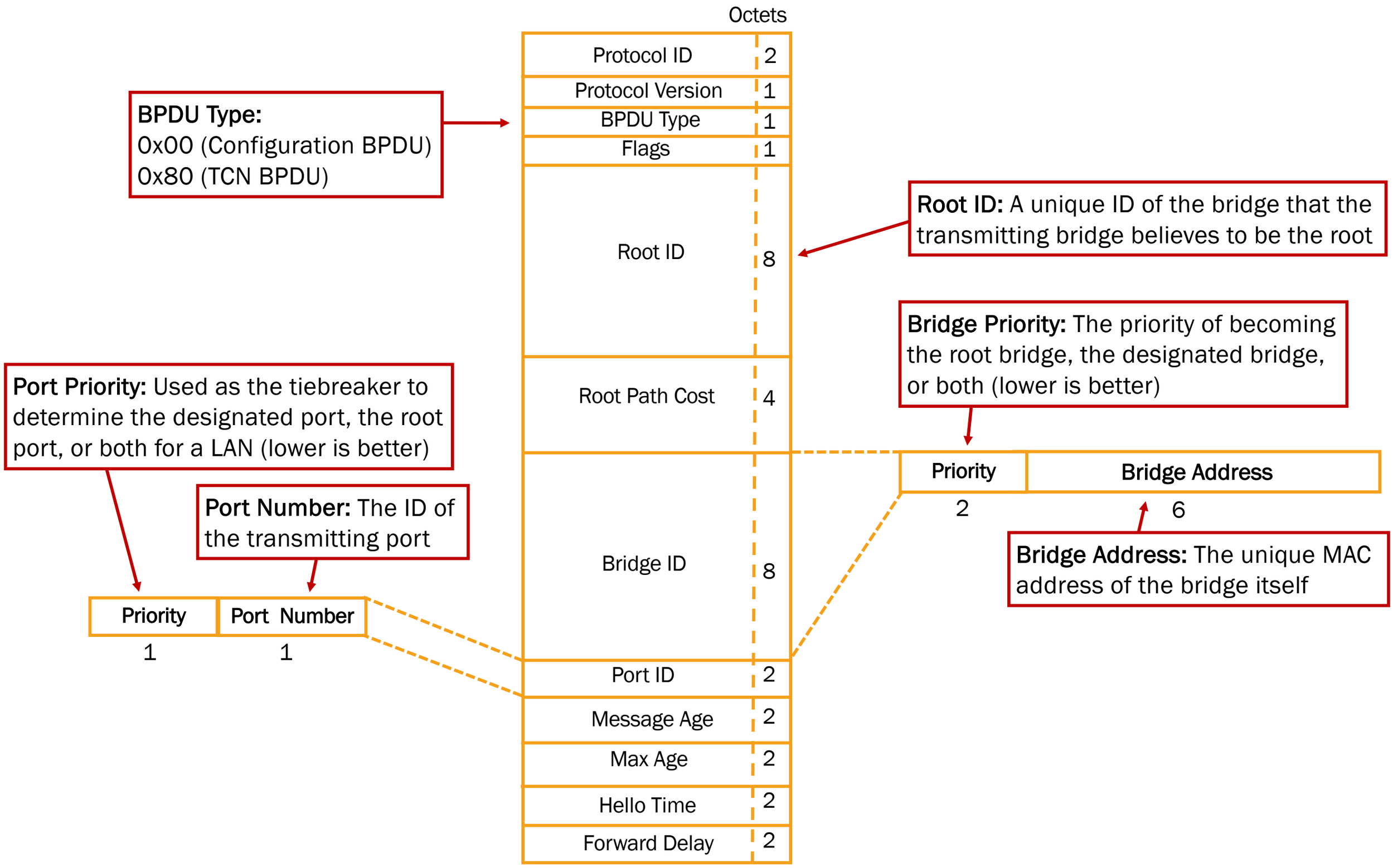
BPDU—Ethernet Frame Format

- Ethernet Frame



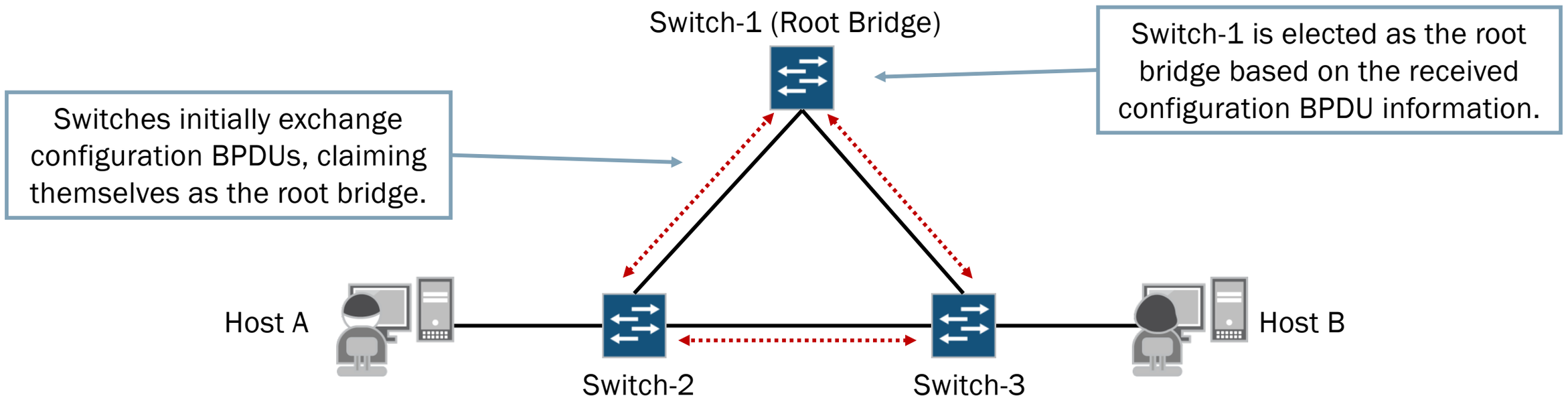
- BPDU types:
  - Configuration BPDUs
    - Used to build the spanning-tree topology
  - Topology change notification (TCN) BPDUs
    - Reports topology changes

BPDU Format


Octets
Protocol ID 2
Protocol Version 1
BPDU Type 1
Flags 1
Root ID 8
Root Path Cost 4
Bridge ID 8
Port ID 2
Message Age 2
Max Age 2
Hello Time 2
Forward Delay 2
BPDU Type:
0x00 (Configuration BPDU)
0x80 (TCN BPDU)
Root ID: A unique ID of the bridge that the transmitting bridge believes to be the root
Bridge Priority: The priority of becoming the root bridge, the designated bridge, or both (lower is better)
Priority 2
Bridge Address 6
Bridge Address: The unique MAC address of the bridge itself
Port Priority: Used as the tiebreaker to determine the designated port, the root port, or both for a LAN (lower is better)
Port Number: The ID of the transmitting port
Priority 1
Port Number 1

Building a Spanning Tree (1 of 3)

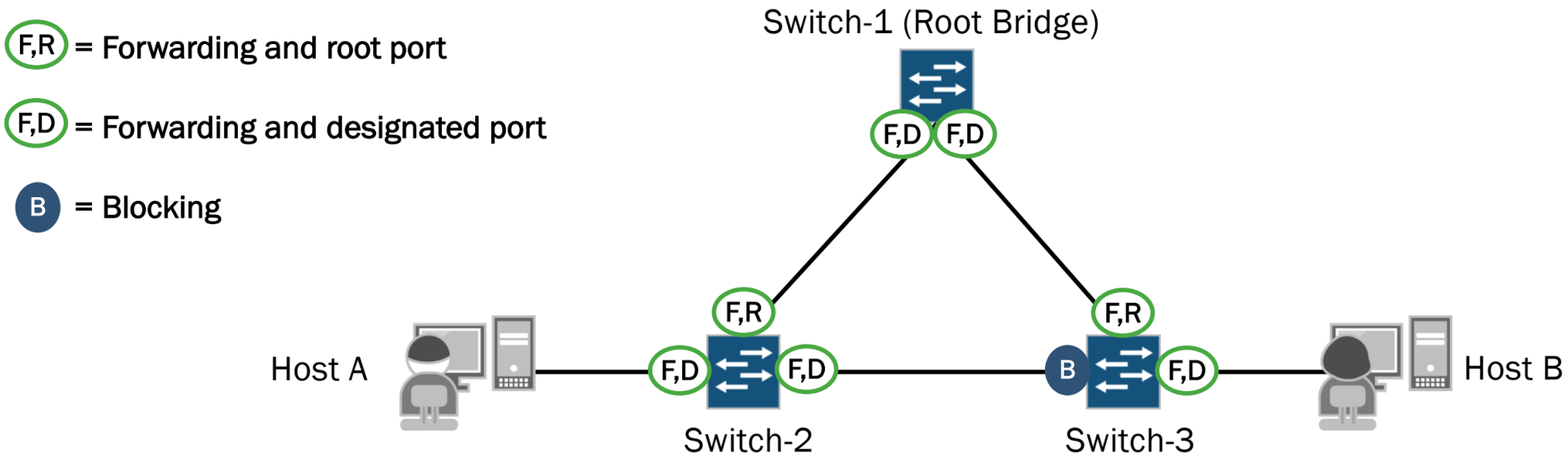
- Switches exchange configuration BPDUs:
  - They do not flood—instead each bridge uses information in the received BPDUs to generate its own
- Root bridge is elected based on BPDU information:
  - Criterion for election is the bridge ID
    - The election process reviews priority first—lowest priority wins
    - If the priority values are the same, bridge addresses (MAC) are compared—the lowest identifier wins



Building a Spanning Tree (2 of 3)

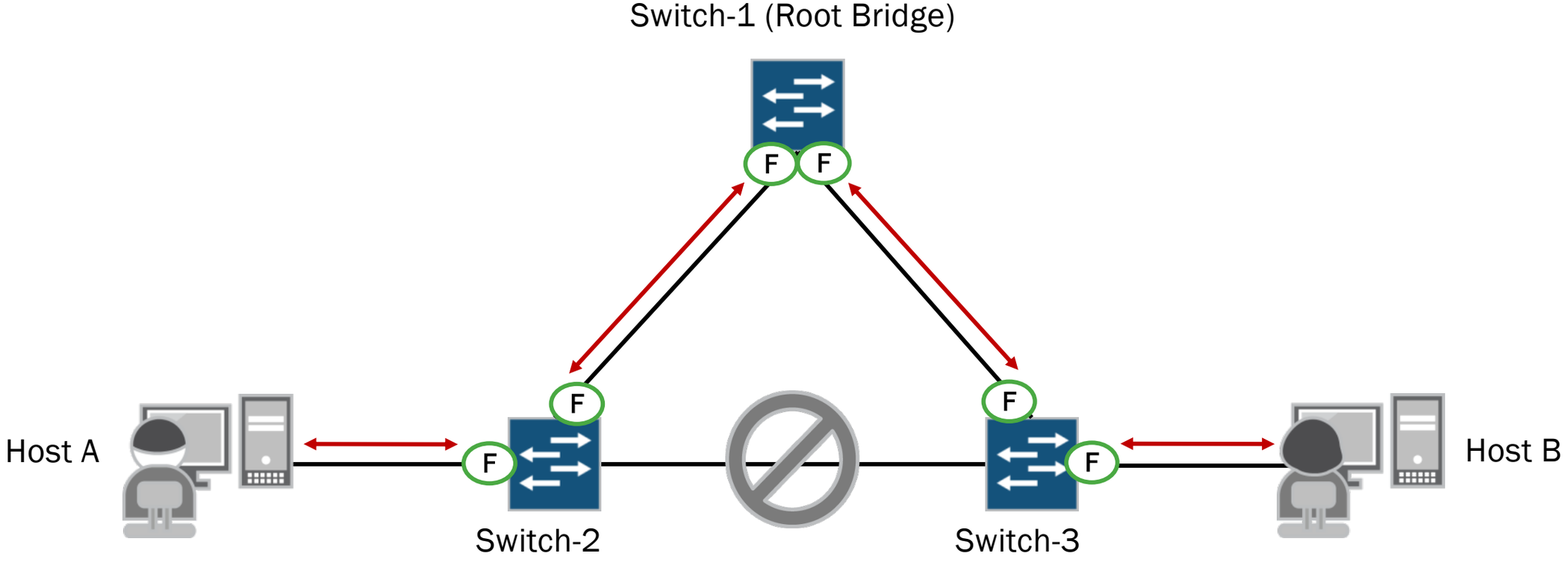
- Least-cost path calculation to root bridge determines port role; port role determines port state:

Port Role and State Designations
All ports on root bridge assume designated port role and forwarding state
Root ports on switches are placed in the forwarding state; root bridge has no root ports
Designated ports on designated bridges are placed in the forwarding state
All other ports are placed in the blocking state



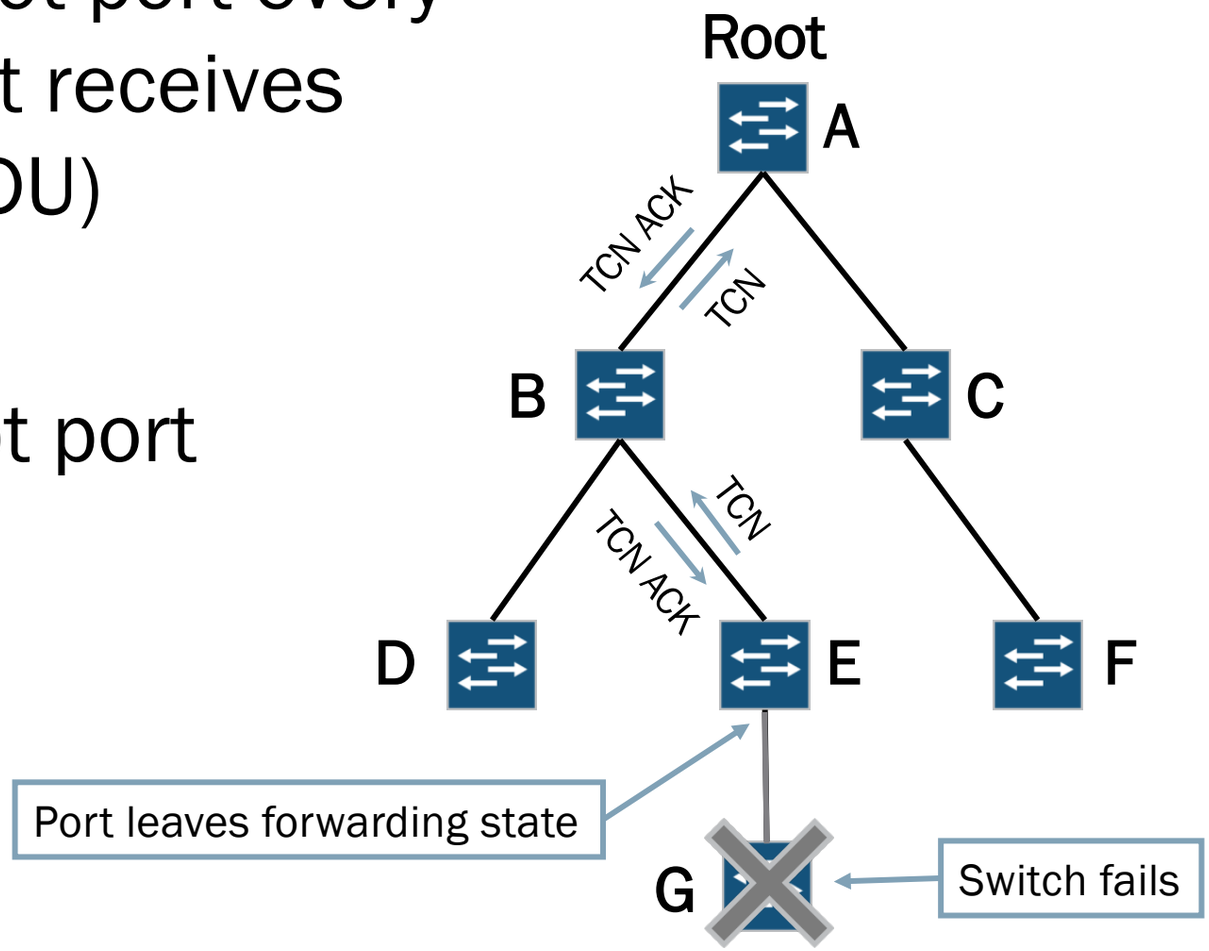
Building a Spanning Tree (3 of 3)

- The tree is fully converged
  - All traffic between Host A to Host B flows through the root bridge (Switch-1)



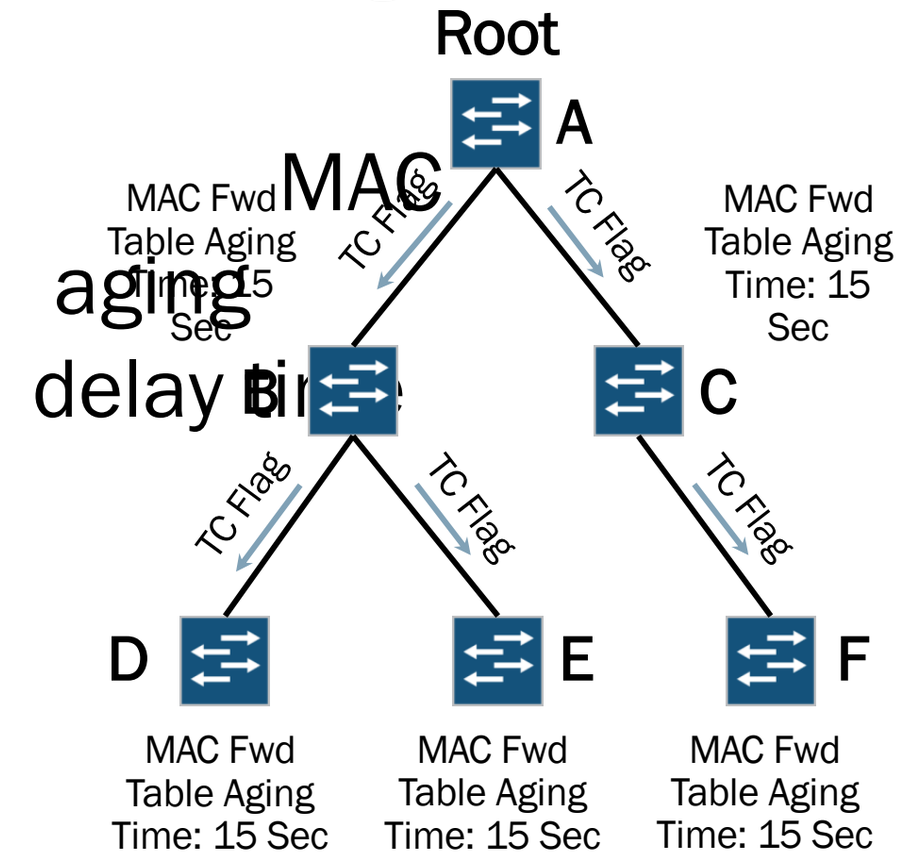
Reconvergence Example (1 of 2)

- Steps:
  1. Switch G fails
  2. Switch E's port leaves forwarding state
  3. Switch E sends TCNs out root port every 2 seconds until E's root port receives TCN ACK (configuration BPDU)
  4. Switch B sends TCN ACK
  5. Switch B sends TCN out root port
  6. Switch A sends TCN ACK



Reconvergence Example (2 of 2)

- Steps (contd.):
  7. The root bridge sets the topology change flag and sends an updated configuration BPDU
  8. Switches B and C relay the topology change flag to downstream switches
  9. All nonroot bridges change the address forwarding table timer to equal the forwarding delay time (default: 15 seconds)



STP Drawbacks

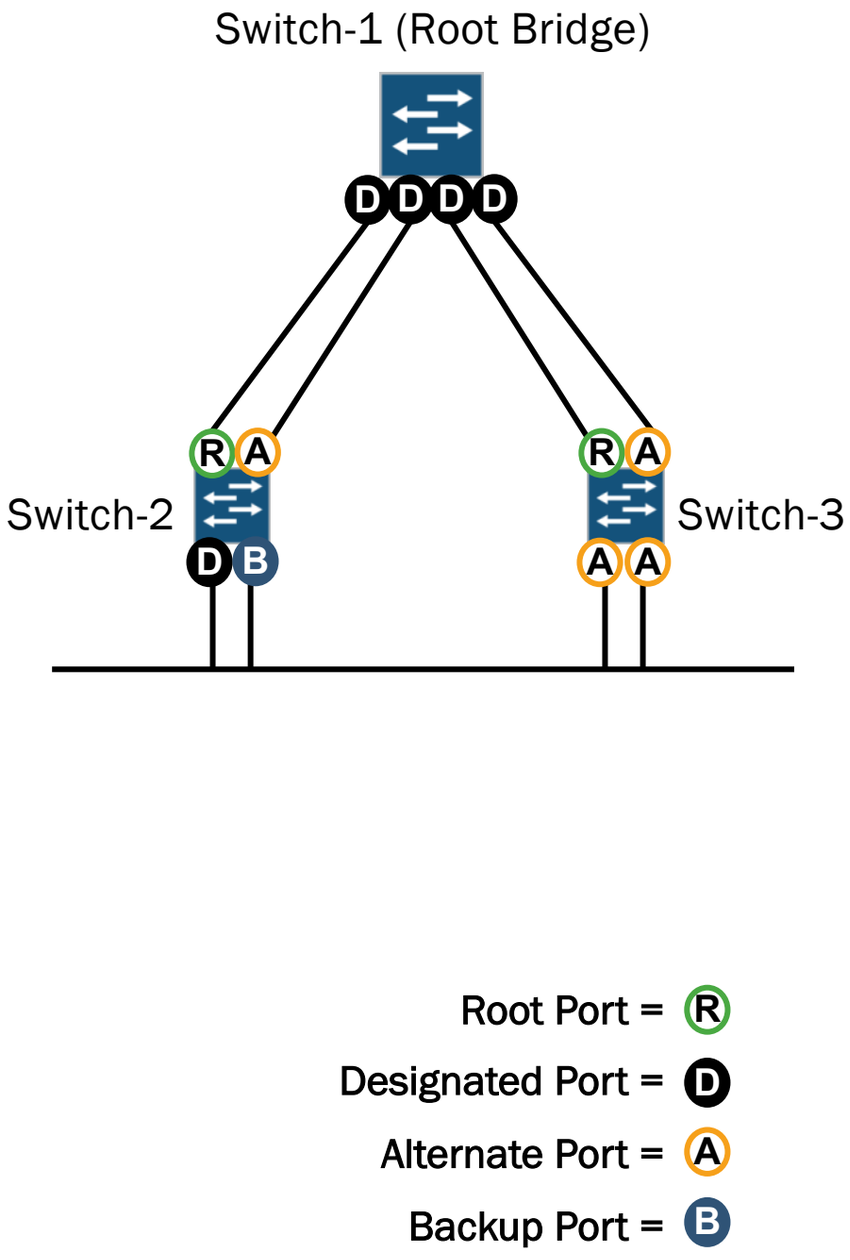
- Slow convergence time
  - STP uses timers to transition between port states
    - STP can take 30 to 50 seconds to respond to a topology change (20 seconds for a BPDU to age out, 15 seconds for the listening state, and 15 seconds for the learning state)
  - Root bridge is responsible for communicating the current tree topology

Rapid Spanning Tree Protocol

- RSTP was first defined in IEEE 802.1w and later incorporated into IEEE 802.1D-2004
- Convergence improvements:
  - Point-to-point link designation
  - Edge port designation
    - A port that connects to a LAN with no other bridges attached
    - It is always in the forwarding state
- Allows for rapid recovery from failures
    - A new root port or designated port can transition to forwarding without waiting for the protocol timers to expire
- Direct and indirect link failure and recovery

RSTP Port Roles

- RSTP introduces new port roles:
  - Alternate port:
    - Provides an alternate path to the root bridge (essentially a backup root port)
    - Blocks traffic while receiving superior BPDUs from a neighboring switch
  - Backup port:
    - Provides a redundant path to a segment (on designated switches only)
    - Blocks traffic while a more preferred port functions as the designated port
- RSTP continues to use the root and designated port roles



STP and RSTP Port States

- RSTP uses fewer states than STP but has the same functionality

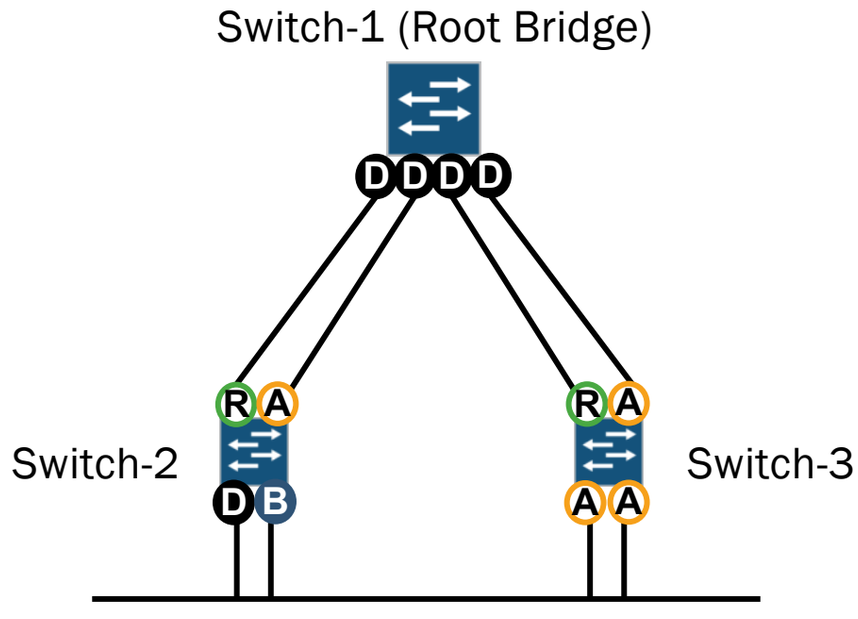
802.1D-1998 STP	802.1D-2004 RSTP
Blocking	Discarding
Listening	
Learning	Learning
Forwarding	Forwarding

Alternate Backup, and Disabled Ports

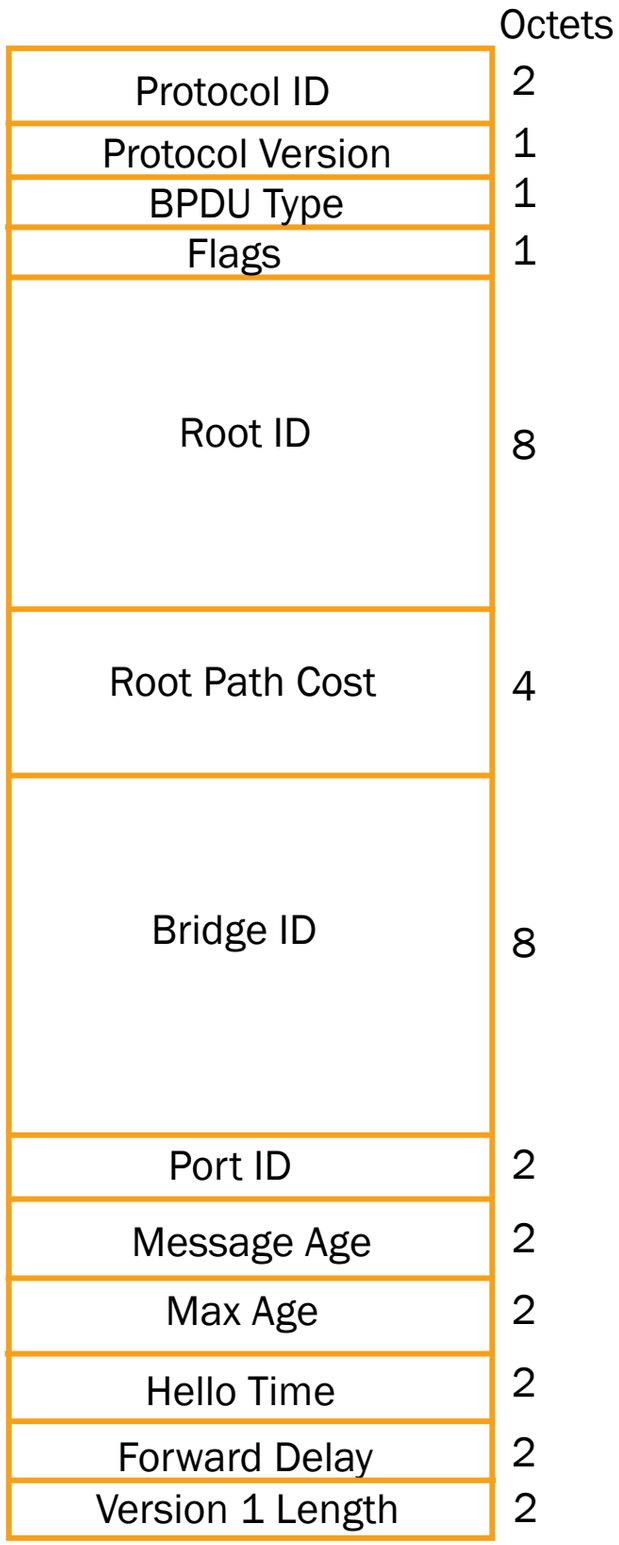
Root and Designated Ports

Rapid Spanning Tree BPDUs

- Rapid Spanning Tree BPDUs:
  - Act as keepalives
    - RSTP-designated ports send Configuration BPDUs every hello time (default of 2 seconds)
  - Provide faster failure detection
    - If a neighboring bridge receives no BPDU within 3 times the hello interval (3 x 2 = 6 seconds), connectivity to the neighbor is faulty



RST BPDU Format



- RST BPDU fields that differ from STP:
  - Protocol Version—0x02 (IEEE 802.1D-2004)
  - BPDU Type—0x02 (RST BPDU)
  - Flags
    - Topology Change Acknowledgement Flag (Bit 8)
    - Agreement Flag (Bit 7)
    - Forwarding Flag (Bit 6)
    - Learning Flag (Bit 5)
    - Port Role (Bits 3 and 4)
    - Proposal Flag (Bit 2)
    - Topology Change Flag (Bit 1)
  - Version 1 Length—0x0000

Transitioning to the Forwarding State

- STP:
  - Takes 30 seconds before the ports start forwarding traffic after port enablement
    - 2x forwarding delay (listening + learning)
- RSTP:
  - Uses a proposal-and-agreement handshake on point-to-point links instead of timers
    - Exceptions are alternate ports that immediately transition to root, and edge ports that immediately transition to the forwarding state
    - Nonedge-designated ports transition to the forwarding state once they receive explicit agreement

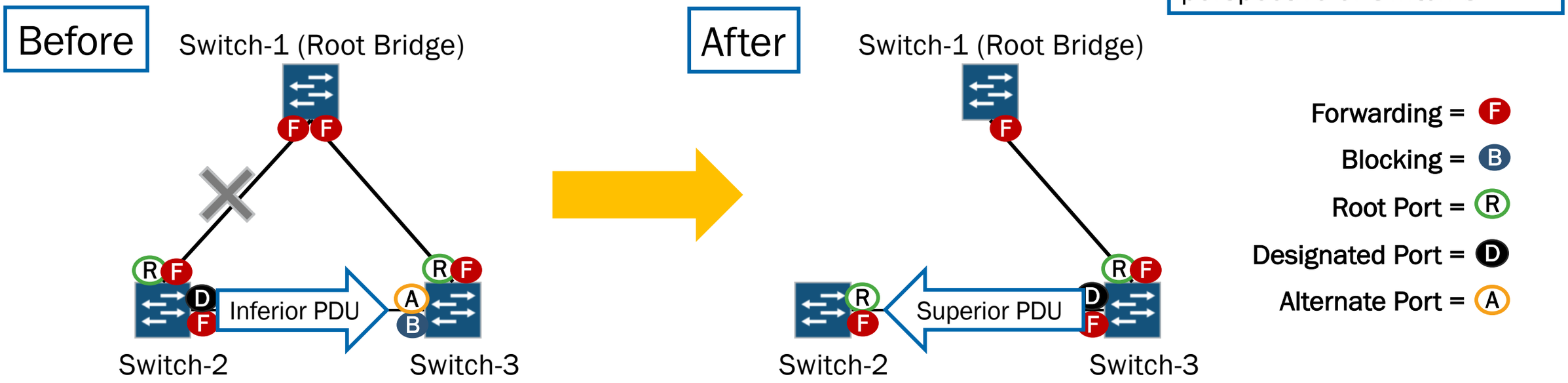
Topology Change Reconvergence

- Topology changes occur only when nonedge ports transition to the forwarding state:
  - Port transitions to the discarding state no longer trigger the STP TCN/TCN Acknowledgment sequence
  - The initiator sends RSTP TCNs (RST BPDU with TCN flag set) out of all designated ports as well as out of the root port
  - Because of the received RSTP TCN, switches flush the majority of MAC addresses in the bridge table
    - Switches do not flush MAC addresses learned from edge ports
    - Switches do not flush MAC addresses learned on port receiving TCN

Indirect Link Failure

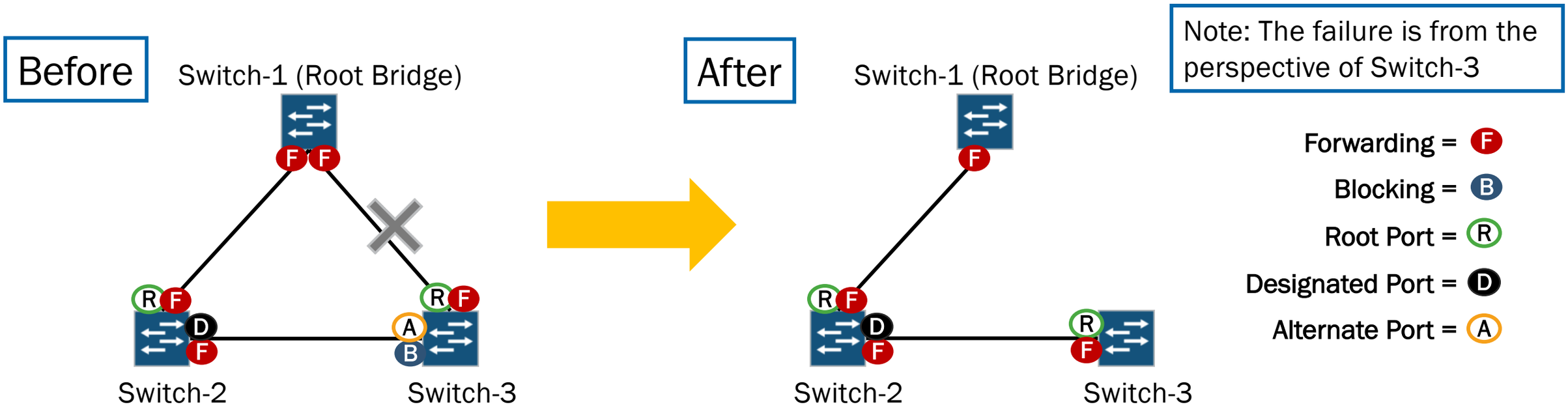
- When an indirect link failure occurs:
  - Switch-2's root port fails—it assumes it is the new root
  - Switch-3 receives inferior BPDUs from Switch-2—it moves the alternate port to the designated port role
  - Switch-2 receives superior BPDUs, knows it is not the root, and designates the port connecting to Switch-3 as the root port

Note: The failure is from the perspective of Switch-3



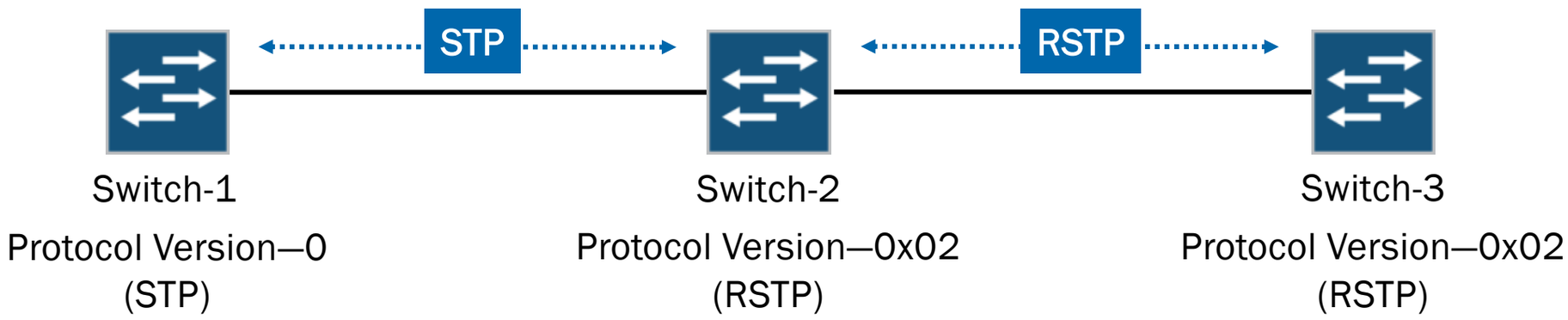
Direct Link Failure

- When a direct link failure occurs:
  - Alternate port transitions to forwarding state and assumes root port role following the failure of the old root port
  - Switch-3 signals upstream switches to flush their MAC tables by sending RSTP TCNs out new root port
  - Upstream switches only flush MAC entries that they learned on active ports that did not receive the RSTP TCNs (except edge ports)



RSTP Interoperability with STP

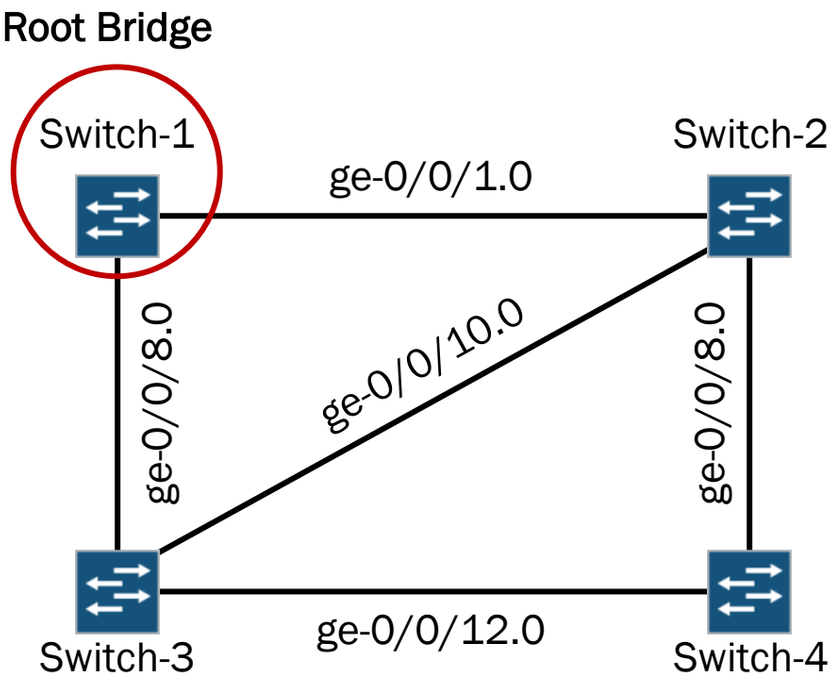
- STP and RSTP interoperability considerations:
  - If a switch supports only the STP protocol, it discards any RSTP BPDUs it receives
  - If an RSTP-capable switch receives BPDUs, it reverts to STP mode on the receiving interface only and sends STP BPDUs



Test Your Knowledge (1 of 4)

- Which switch will be elected the root bridge?

```
{master:0}[edit protocols rstp]
user@Switch-1# show
bridge-priority 4k;
interface ge-0/0/8.0 {
    cost 1;
}
interface all {
    priority 128;
    cost 200000;
}
```




```
{master:0}[edit protocols rstp]
user@Switch-2# show
bridge-priority 8k;
interface ge-0/0/10.0 {
    cost 1;
}
interface all {
    priority 16;
    cost 20000;
}
```

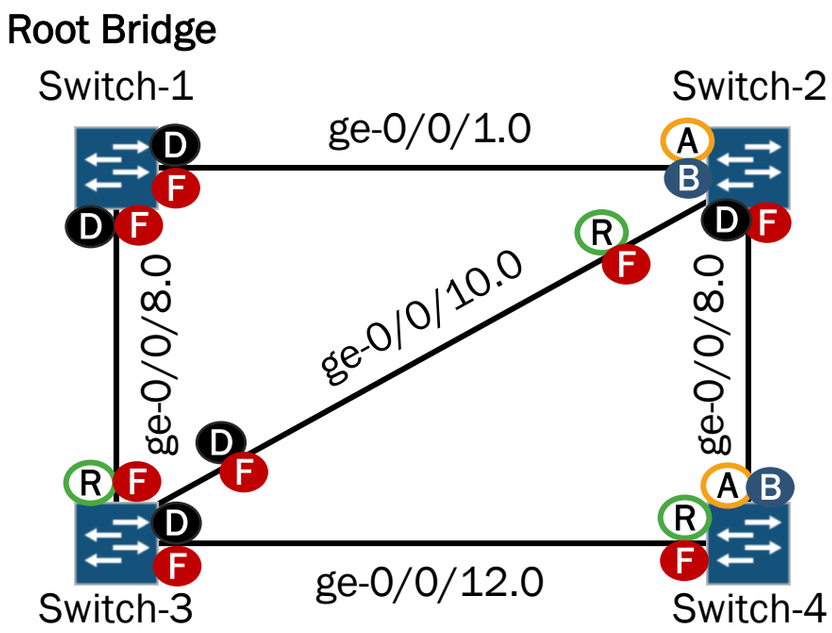
```
{master:0}[edit protocols rstp]
user@Switch-3# show
bridge-priority 32k;
interface all {
    priority 16;
    cost 2000;
}
```

```
{master:0}[edit protocols rstp]
user@Switch-4# show
bridge-priority 36k;
interface all {
    priority 128;
    cost 20000;
}
```

Test Your Knowledge (2 of 4)

- What role and state will be assigned to the various switch ports?

```
{master:0}[edit protocols rstp]
user@Switch-1# show
bridge-priority 4k;
interface ge-0/0/8.0 {
    cost 1;
}
interface all {
    priority 128;
    cost 200000;
}
```




```
{master:0}[edit protocols rstp]
user@Switch-2# show
bridge-priority 8k;
interface ge-0/0/10.0 {
    cost 1;
}
interface all {
    priority 16;
    cost 20000;
}
```

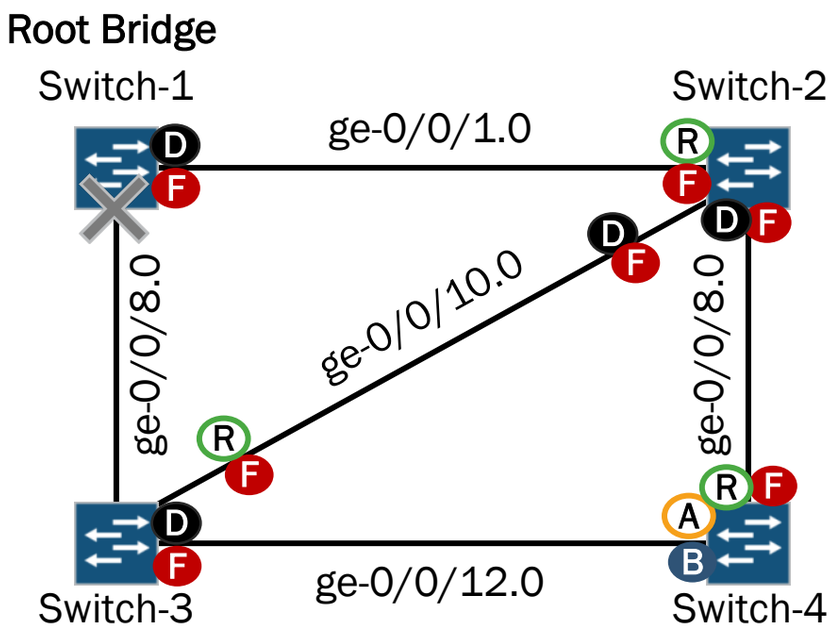
```
{master:0}[edit protocols rstp]
user@Switch-3# show
bridge-priority 32k;
interface all {
    priority 16;
    cost 2000;
}
```

```
{master:0}[edit protocols rstp]
user@Switch-4# show
bridge-priority 36k;
interface all {
    priority 128;
    cost 20000;
}
```

Test Your Knowledge (3 of 4)

- Assume ge-0/0/8 on Switch-1 has failed, what role and state will be assigned to the remaining ports?

```
{master:0}[edit protocols rstp]
user@Switch-1# show
bridge-priority 4k;
interface ge-0/0/8.0 {
    cost 1;
}
interface all {
    priority 128;
    cost 200000;
}
```




```
{master:0}[edit protocols rstp]
user@Switch-2# show
bridge-priority 8k;
interface ge-0/0/10.0 {
    cost 1;
}
interface all {
    priority 16;
    cost 20000;
}
```

```
{master:0}[edit protocols rstp]
user@Switch-3# show
bridge-priority 32k;
interface all {
    priority 16;
    cost 2000;
}
```

Forwarding = F

Blocking = B

Root Port = R

Designated Port = D

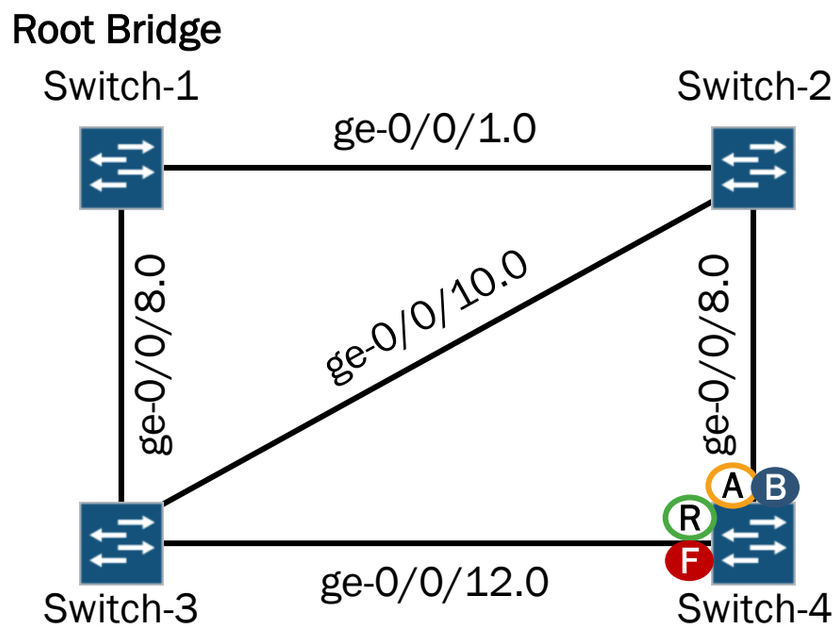
Alternate Port = A

```
{master:0}[edit protocols rstp]
user@Switch-4# show
bridge-priority 36k;
interface all {
    priority 128;
    cost 20000;
}
```

Test Your Knowledge (4 of 4)

- Based on the modified configurations, what role and state will be assigned to Switch-4's ports?

```
{master:0}[edit protocols rstp]
user@Switch-1# show
bridge-priority 4k;
interface all {
    priority 128;
    cost 20000;
}
```



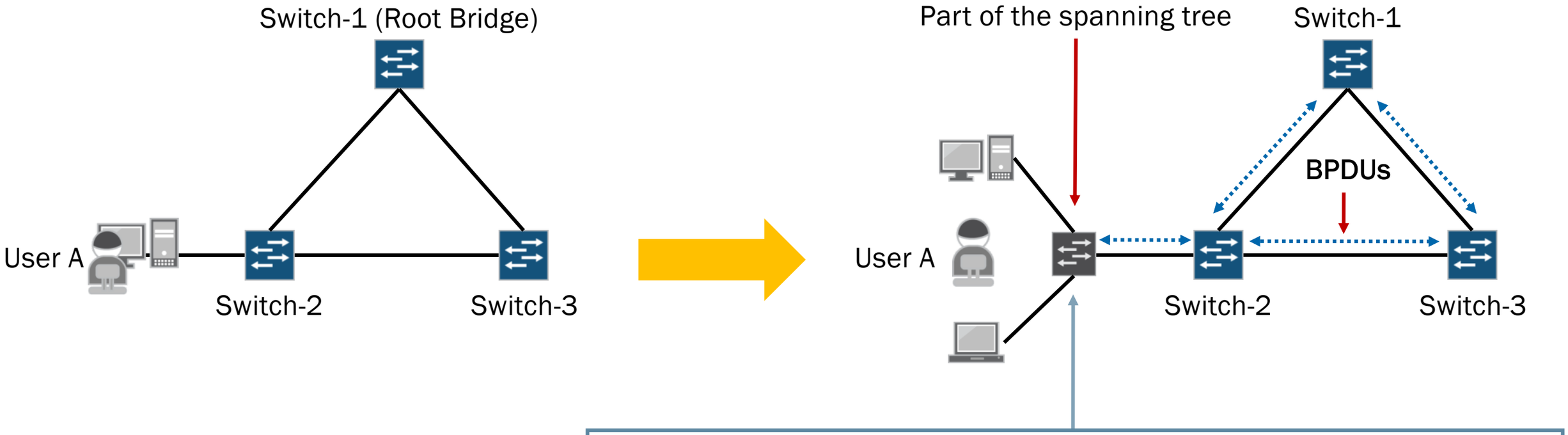

```
{master:0}[edit protocols rstp]
user@Switch-2# show
bridge-priority 32k;
interface all {
    priority 16;
    cost 20000;
}
```

```
{master:0}[edit protocols rstp]
user@Switch-3# show
bridge-priority 32k;
interface all {
    priority 16;
    cost 20000;
}
```

```
{master:0}[edit protocols rstp]
user@Switch-4# show
bridge-priority 36k;
interface ge-0/0/8.0 {
    priority 32;
}
interface ge-0/0/12.0 {
    priority 16;
}
```

What If…?

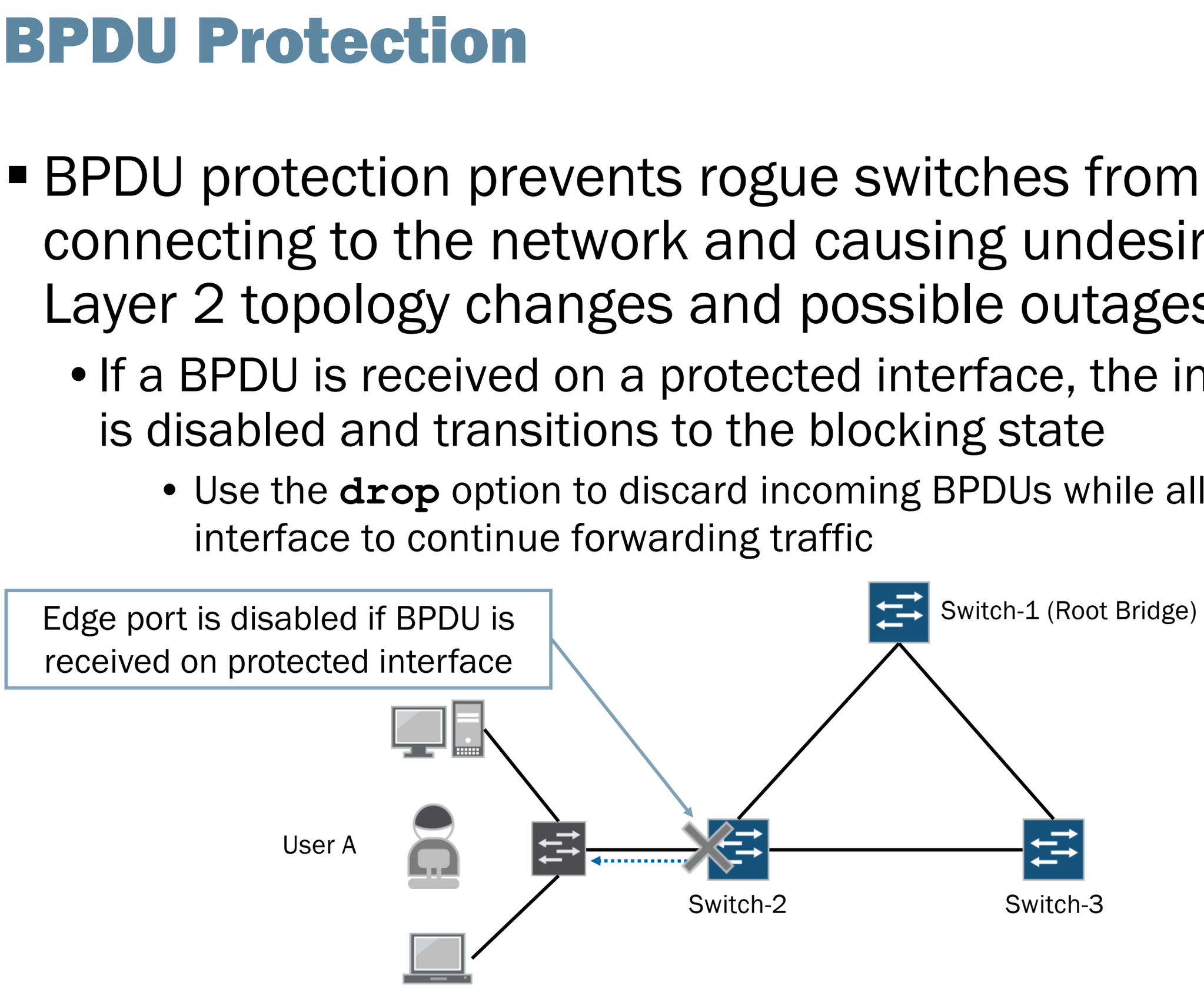
- Given the topology below, what if User A connects a personal (unauthorized) switch running the spanning tree protocol to Switch-2?



BPDUs would be exchanged, a new STP calculation would occur, and the rogue switch would become part of the spanning tree, potentially leading to a network outage

BPDU Protection

- BPDU protection prevents rogue switches from connecting to the network and causing undesired Layer 2 topology changes and possible outages
  - If a BPDU is received on a protected interface, the interface is disabled and transitions to the blocking state
    - Use the **`drop`** option to discard incoming BPDUs while allowing the interface to continue forwarding traffic



Configuring BPDU Protection

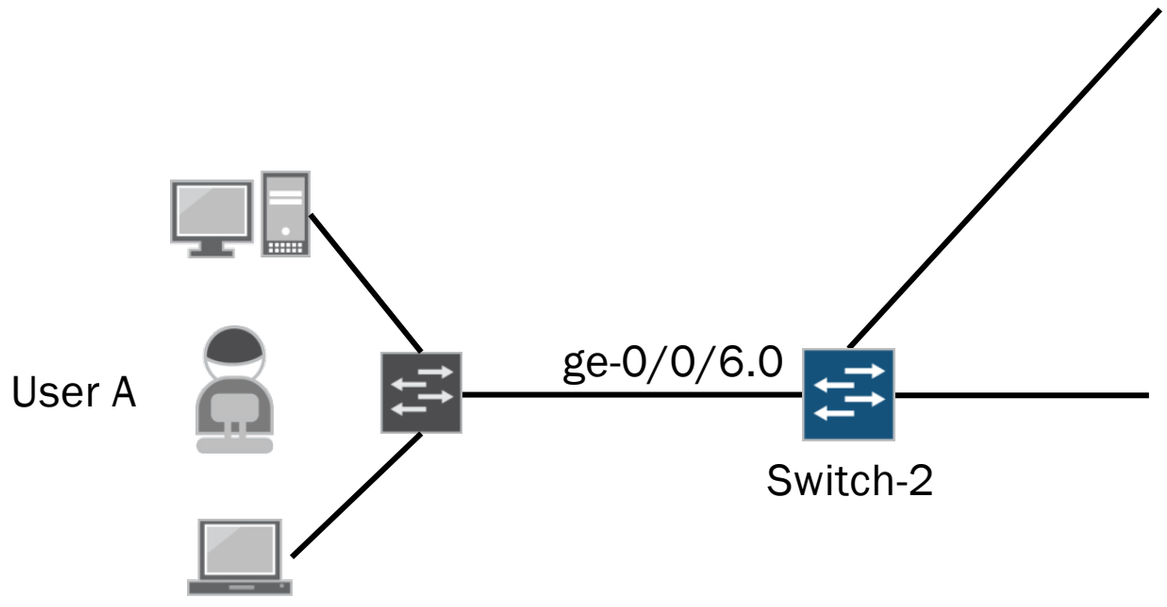
- BPDU protection can be enabled on switches whether or not the spanning tree protocol enabled:

```
{master:0}[edit protocols rstp]
user@Switch-2# show
interface ge-0/0/6.0 {
    edge;
}
bpdu-block-on-edge;
```

Use **bpdu-block-on-edge** option when spanning tree protocol is enabled

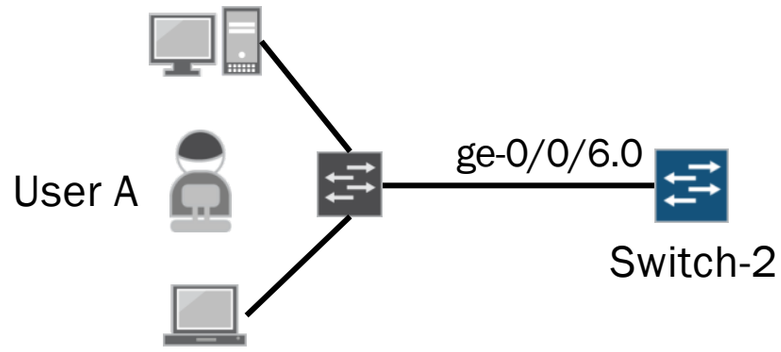
```
{master:0}[edit ethernet-switching-options]
user@Switch-2# show
bpdu-block {
    interface ge-0/0/6.0;
}
```

Use **bpdu-block** option when spanning tree protocol is not enabled



Monitoring BPDU Protection

Before BPDU is received on protected interface




```
{master:0}
user@Switch-2> show spanning-tree interface ge-0/0/6.0

Spanning tree interface parameters for instance 0

Interface    Port ID    Designated      Designated          Port    State  Role
                         port ID         bridge ID           Cost
ge-0/0/6.0     128:519      128:519  32768.0019e2516580      20000  FWD    DESG

{master:0}
user@Switch-2> show ethernet-switching interfaces ge-0/0/6.0
Interface    State  VLAN members         Tag    Tagging  Blocking
ge-0/0/6.0   up     default                     untagged unblocked
```

Before BPDU violation

After BPDU is received on protected interface

```
{master:0}
user@Switch-2> show spanning-tree interface ge-0/0/6.0
```

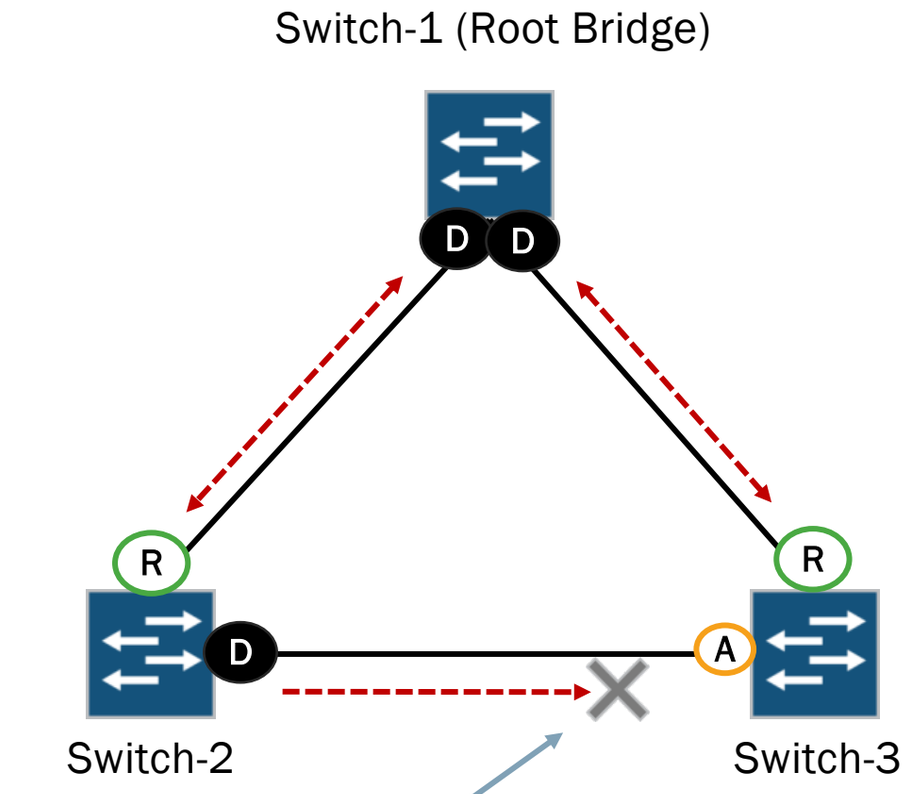
After BPDU violation

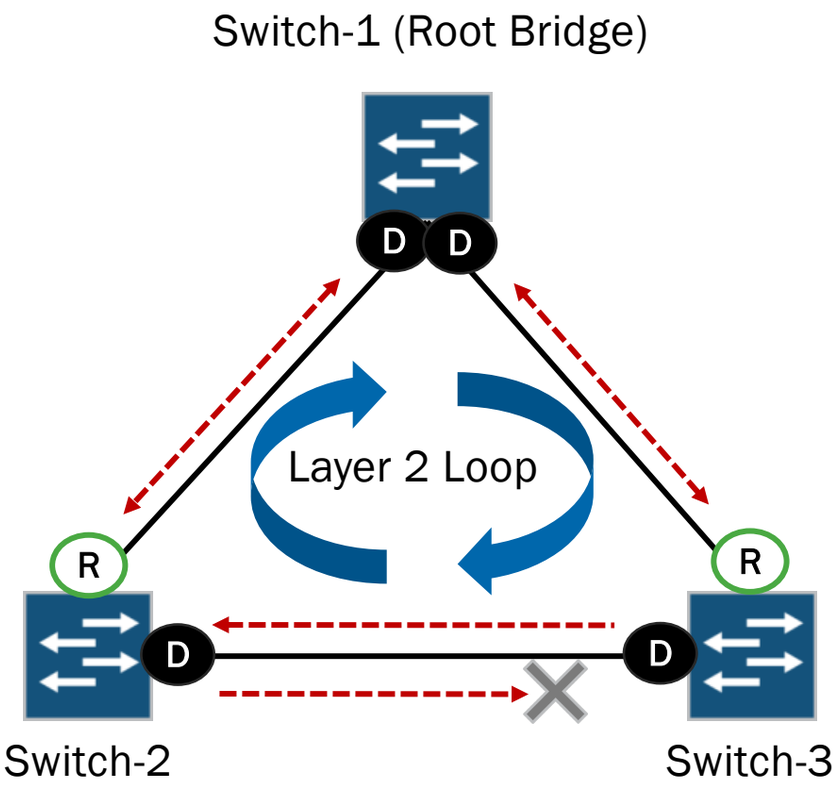
```
{master:0}
user@Switch-2> show ethernet-switching interfaces ge-0/0/6.0
Interface    State  VLAN members         Tag    Tagging  Blocking
ge-0/0/6.0   down   default                     untagged Disabled by bpdu-control


{master:0}
user@Switch-2> clear ethernet-switching bpdu-error interface ge-0/0/6.0
```

Re-enables interface

What If…?

- Given the topology below, what if BPDUs sent by Switch-2 were not received by Switch-3?

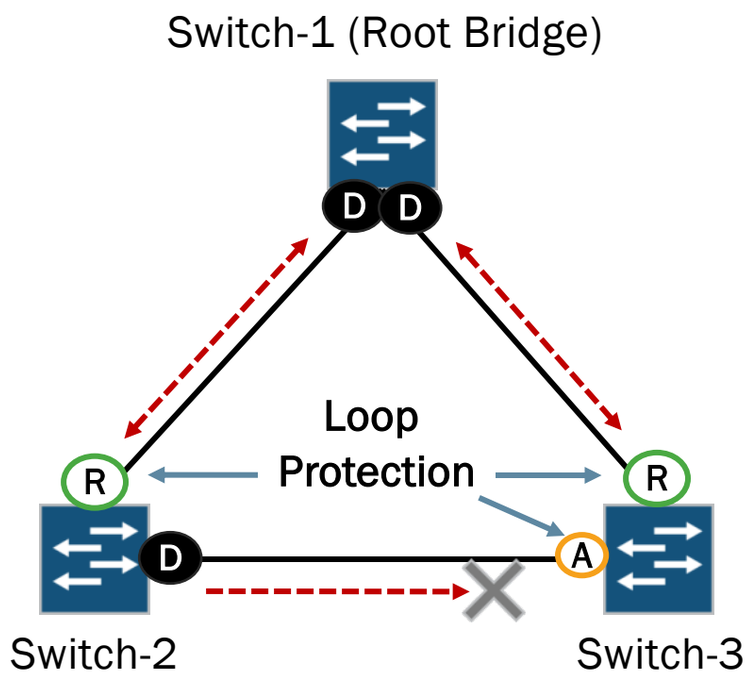




Switch-3 waits until the max-age timer expires then transitions its alternate port to the designated port role and the forwarding state thus removing the blocked port and causing a Layer 2 loop

Loop Protection

- The loop protection feature provides additional protection against Layer 2 loops by preventing non-designated ports from becoming designated ports
  - Enable loop protection on all non-designated ports
    - Ports that detect the loss of BPDUs transition to the “loop inconsistent” role which maintains the blocking state
    - Port automatically transitions back to previous or new role when it receives a BPDU

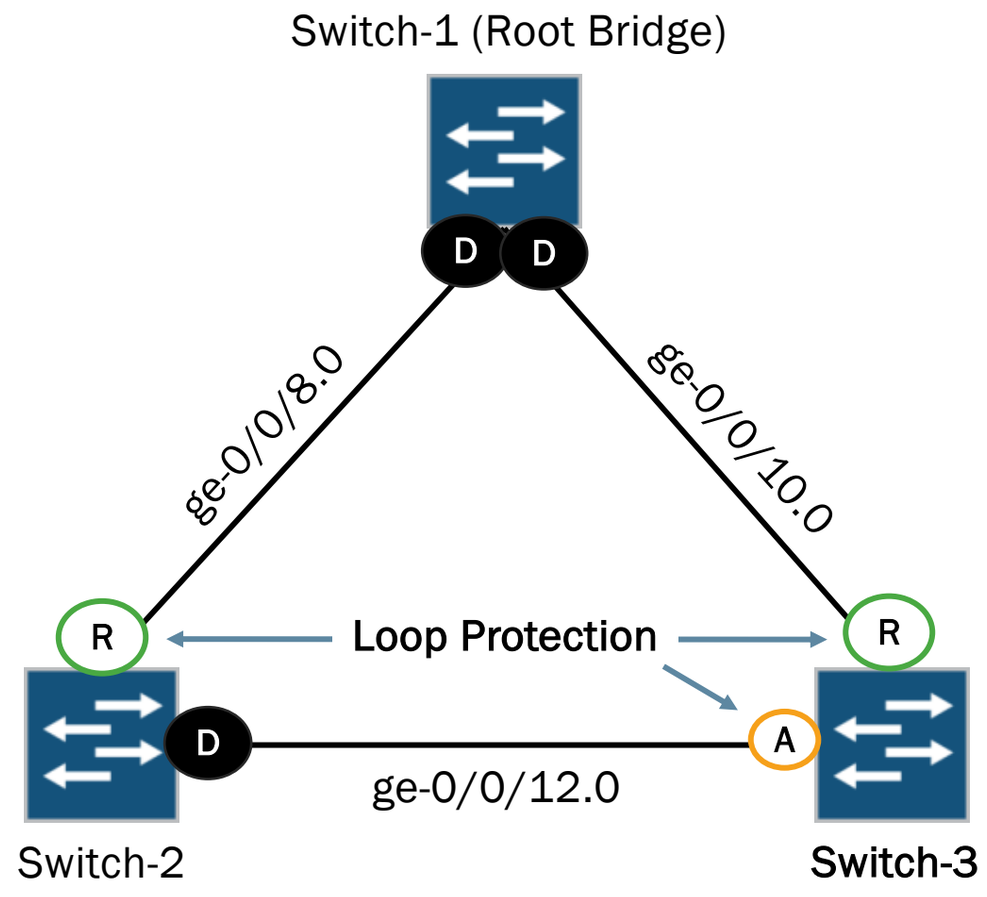


Configuring Loop Protection

- Configure loop protection on non-designated ports (root and alternate ports):

```
{master:0}[edit protocols rstp]
user@Switch-3# show
interface ge-0/0/10.0 {
    bpdu-timeout-action {
        block;
    }
}
interface ge-0/0/12.0 {
    bpdu-timeout-action {
        block;
    }
}
```

Use the `block` or `alarm` action in conjunction with the loop protection feature



Monitoring Loop Protection

When BPDUs are received on protected interface:

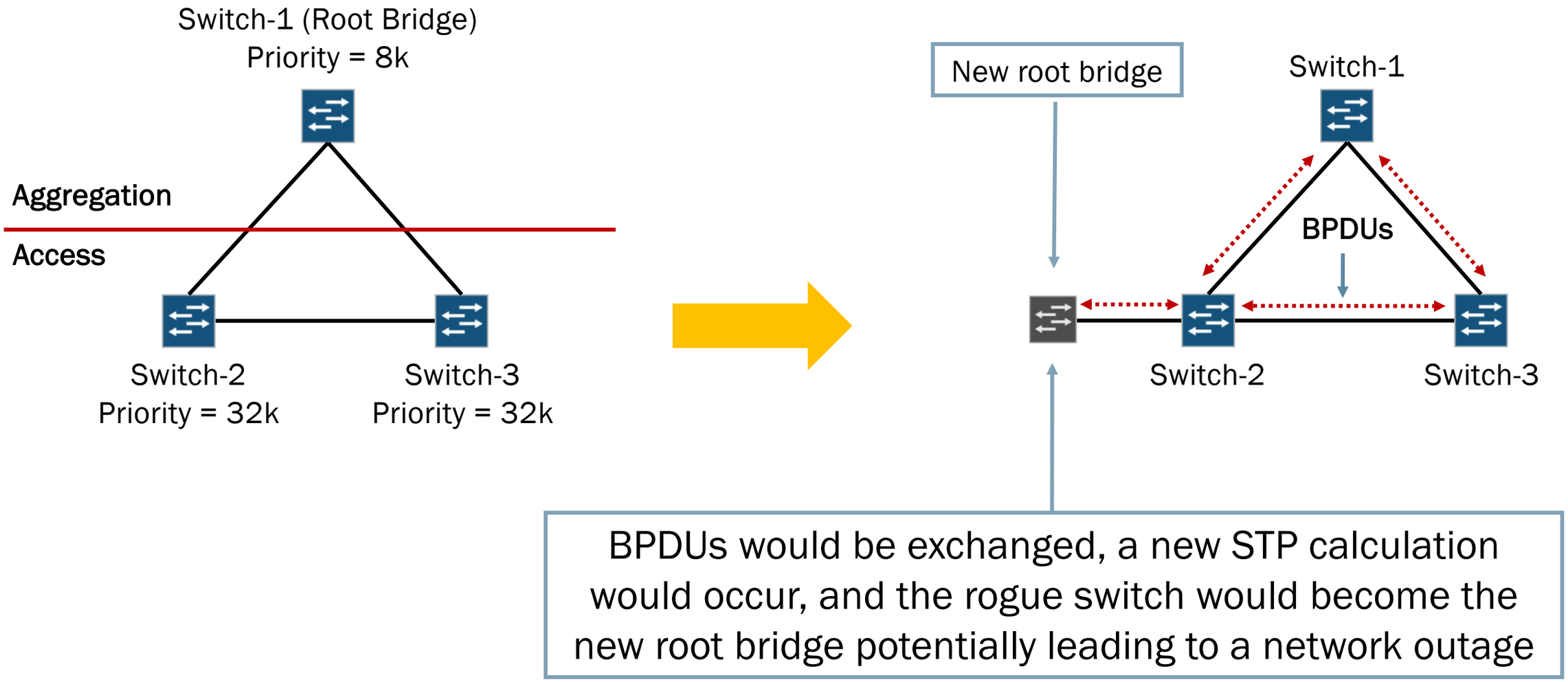
```
{master:0}
user@Switch-3> show spanning-tree interface

Spanning tree interface parameters for instance 0

Interface    Port ID    Designated      Designated          Port    State  Role
                         port ID         bridge ID           Cost
ge-0/0/10.0    128:523      128:523   4096.002688027490      20000  FWD    ROOT
ge-0/0/12.0    128:525      128:525  16384.0019e2516580      20000  BLK    ALT
```

When BPDUs are not received on protected interface:

```
{master:0}
user@Switch-3> show spanning-tree interface

Spanning tree interface parameters for instance 0

Interface    Port ID    Designated      Designated          Port    State  Role
                         port ID         bridge ID           Cost
ge-0/0/10.0    128:523      128:523   4096.002688027490      20000  FWD    ROOT
ge-0/0/12.0    128:525      128:525  32768.0019e2553600      20000  BLK    DIS (Loop-Incon)
```

What If…?

- Given the topology and details below, what if a rogue switch with a bridge priority of 4K was connected to the Layer 2 network?

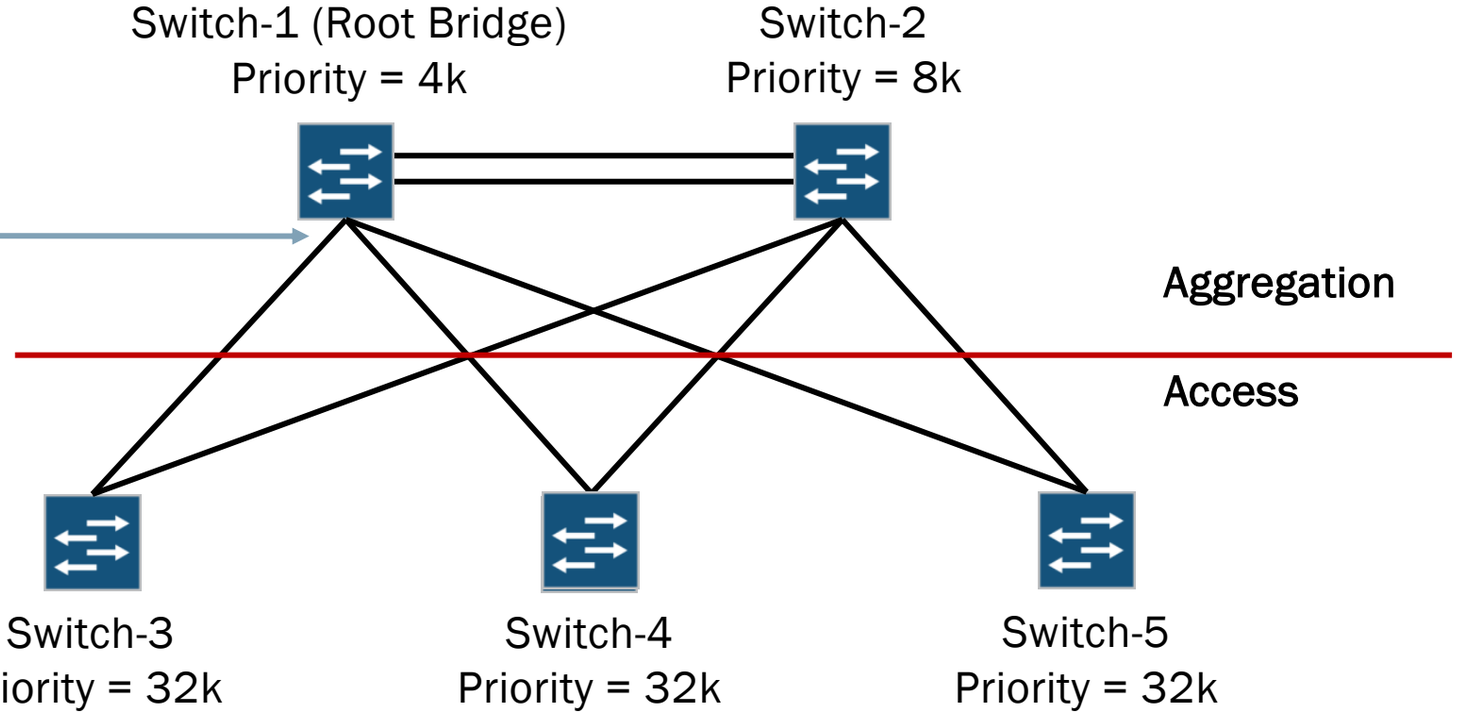


BPDUs would be exchanged, a new STP calculation would occur, and the rogue switch would become the new root bridge potentially leading to a network outage

Root Protection

- Enable root protection to avoid unwanted STP topology changes and root bridge placement
  - If a superior BPDU is received on a protected interface, the interface is disabled and transitions to the blocking state

Root protection is typically configured on the ports of aggregation switches that connect to access switches

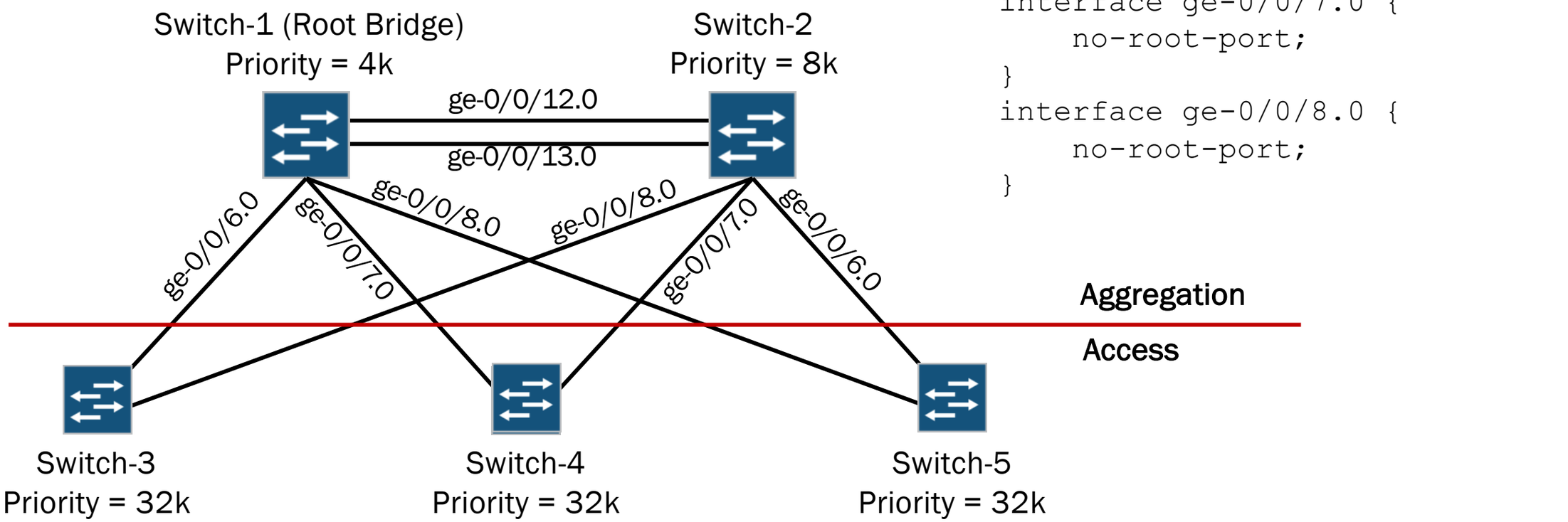


Configuring Root Protection

- Enable root protection on ports that should not receive superior BPDUs from the root bridge and should not be elected as the root port:

```
{master:0}[edit protocols rstp]
user@Switch-1# show
bridge-priority 4k;
interface all {
    no-root-port;
}
```

```
{master:0}[edit protocols rstp]
user@Switch-2# show
bridge-priority 8k;
interface ge-0/0/6.0 {
    no-root-port;
}
interface ge-0/0/7.0 {
    no-root-port;
}
interface ge-0/0/8.0 {
    no-root-port;
}
```



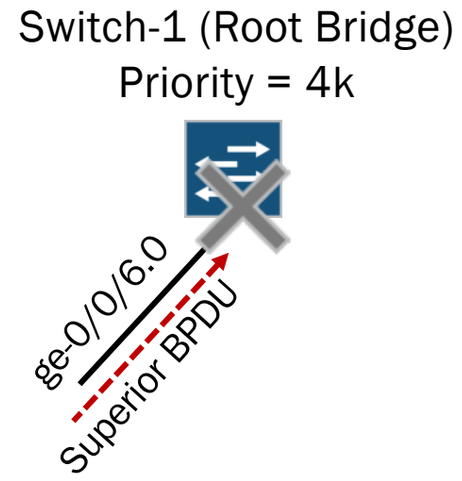
Monitoring Root Protection

Before superior BPDU is received on protected interface

```
{master:0}
user@Switch-1> show spanning-tree interface

Spanning tree interface parameters for instance 0

Interface    Port ID    Designated      Designated          Port    State  Role
                         port ID         bridge ID           Cost
ge-0/0/6.0     128:519      128:519   4096.0019e2516580      20000  FWD    DESG
ge-0/0/7.0     128:520      128:520   4096.0019e2516580      20000  FWD    DESG
ge-0/0/8.0     128:521      128:521   4096.0019e2516580      20000  FWD    DESG
ge-0/0/12.0    128:525      128:525   4096.0019e2516580      20000  FWD    DESG
ge-0/0/13.0    128:526      128:526   4096.0019e2516580      20000  FWD    DESG
```

After superior BPDU is received on protected interface

```
{master:0}
user@Switch-1> show spanning-tree interface

Spanning tree interface parameters for instance 0

Interface    Port ID    Designated      Designated          Port    State  Role
                         port ID         bridge ID           Cost
ge-0/0/6.0     128:519      128:519     0.002688027490       20000  BLK    ALT (Root-Incon)
ge-0/0/7.0     128:520      128:520   4096.0019e2516580      20000  FWD    DESG
ge-0/0/8.0     128:521      128:521   4096.0019e2516580      20000  FWD    DESG
ge-0/0/12.0    128:525      128:525   4096.0019e2516580      20000  FWD    DESG
ge-0/0/13.0    128:526      128:526   4096.0019e2516580      20000  FWD    DESG
```



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

miem.hse.ru Телефон.: +7 (495) 771-32-32 Адрес: г. Москва, ул. Таллинская, д.34