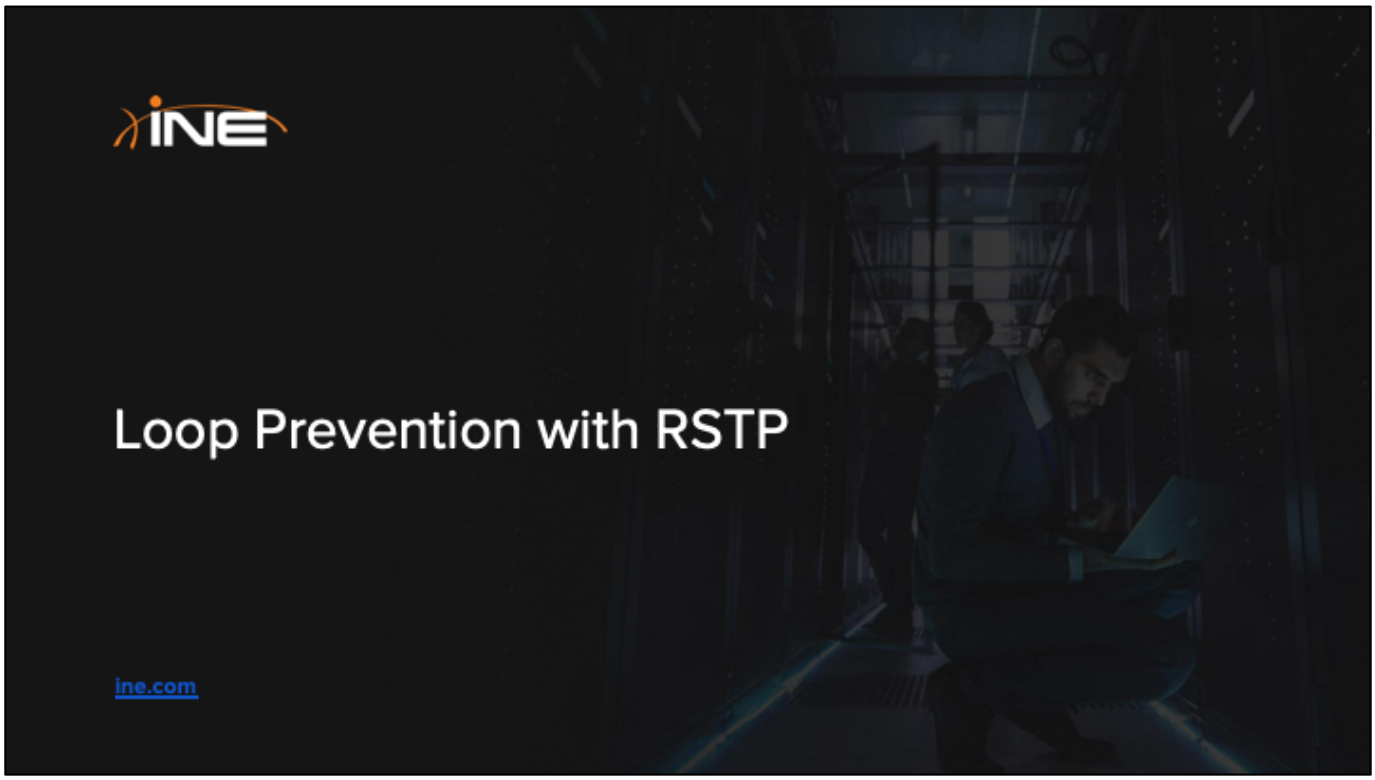

Loop Prevention with RSTP
ine.com

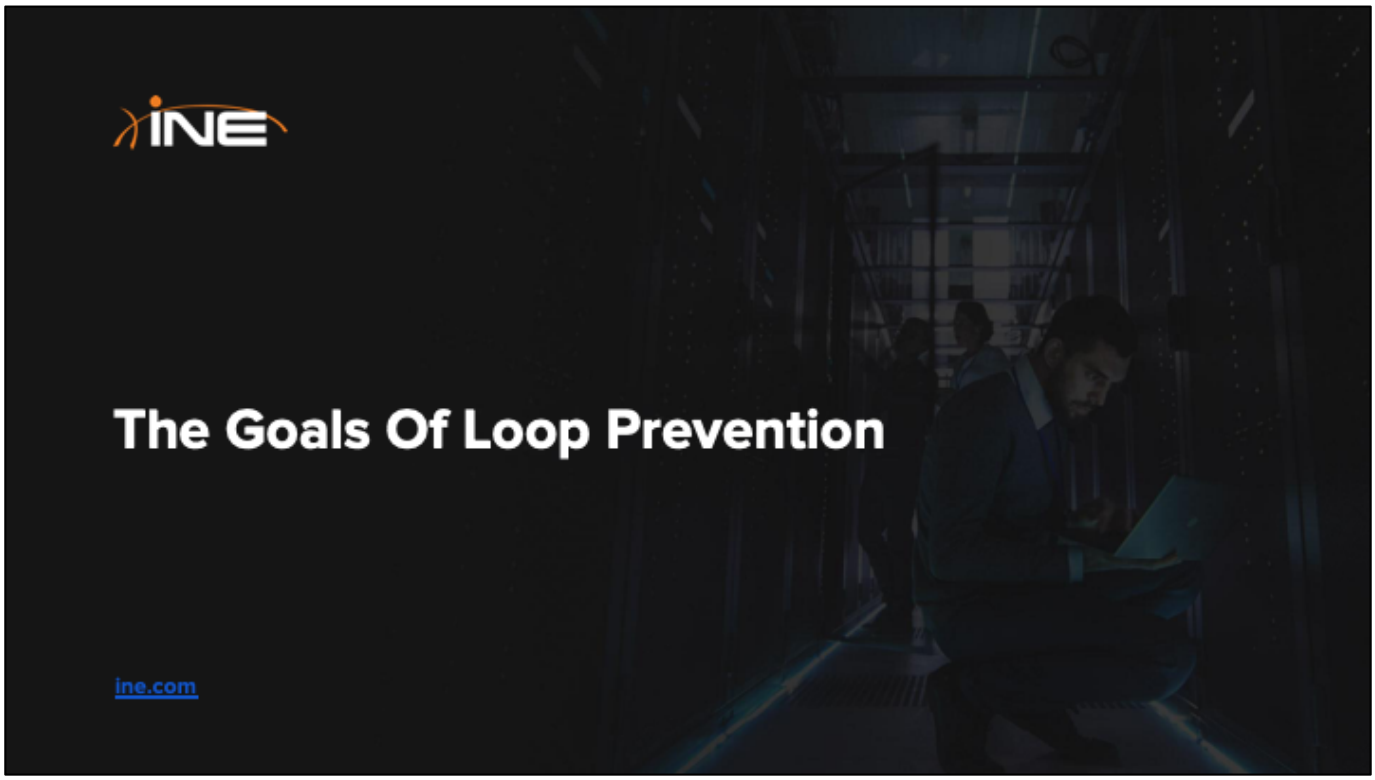

Keith Bogart
CCIE #4923
kbogart@ine.com
@keithbogart1
linkedin.com/in/keith-bogart-2a75042
CCIE Routing & Switching

Course Prerequisites

- Knowledge of the difference between unicast and broadcast traffic
- An understanding of the default forwarding behavior of switches
- Knowledge of VLANs, VLAN Trunks and their usage
- Ability to recognize Ethernet MAC addresses

Course Objectives

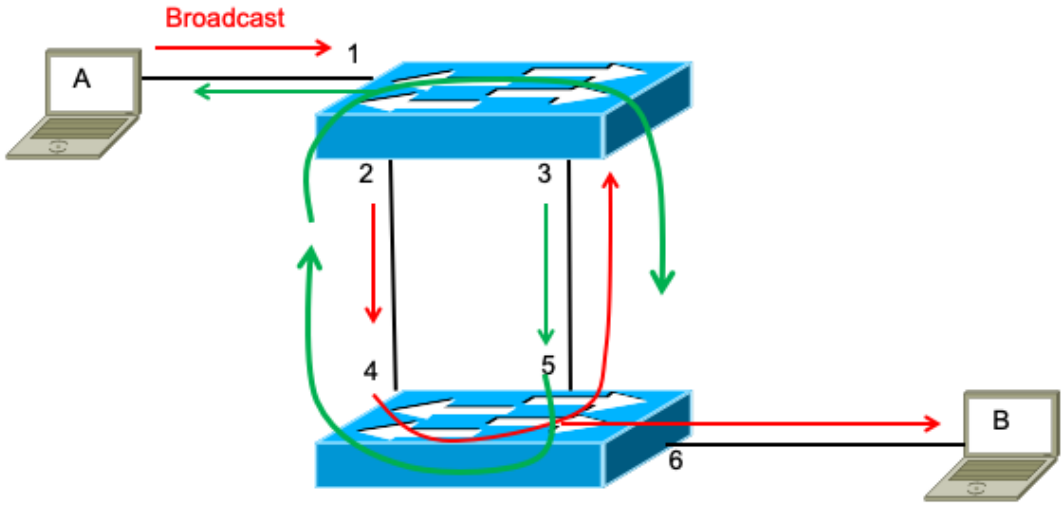
- Explain why bridging loops are harmful in a switched network
- Summarize the basic operation of RSTP Forwarding & Discarding ports
- Predict an RSTP topology if provided with the required minimum criteria
- Manipulate RSTP topologies for purposes of load-balancing
- Understand the RSTP Sync and Topology Change processes


The Goals Of Loop Prevention
ine.com

Topic Overview

+ What Is A Bridging Loop?
+ Introduction To Rapid Spanning-Tree
+ The Concept Of Trees As Related to RSTP

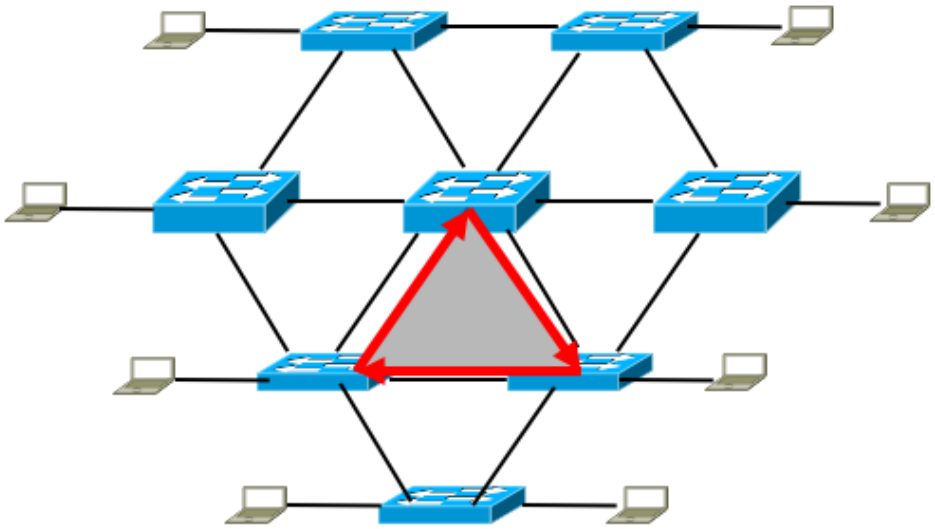
What Is A Bridging Loop?



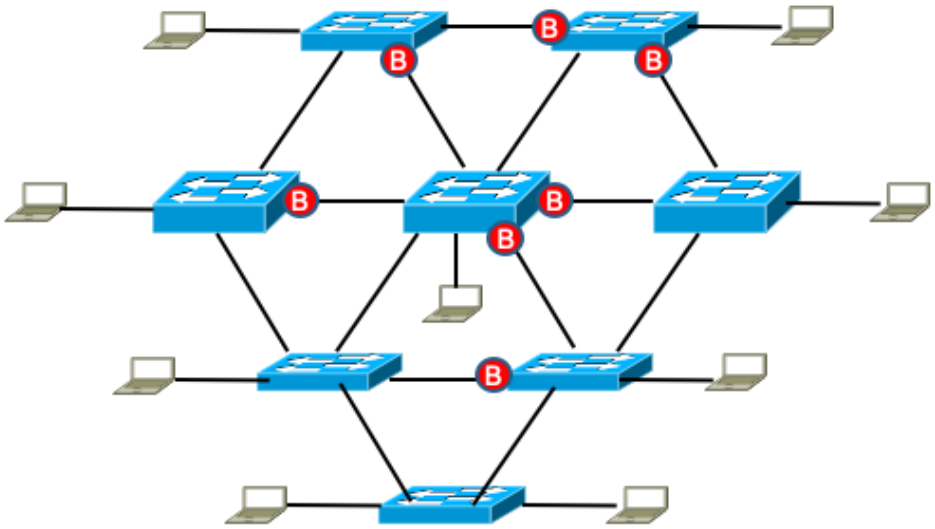
Introduction To Rapid Spanning-Tree

- Originally introduced in 2001 as IEEE 802.1w standard
- Incorporated into the IEEE 802.1d standard in 2004
- Usually called CST (Common Spanning Tree)
- No redundancy in traffic paths for frames

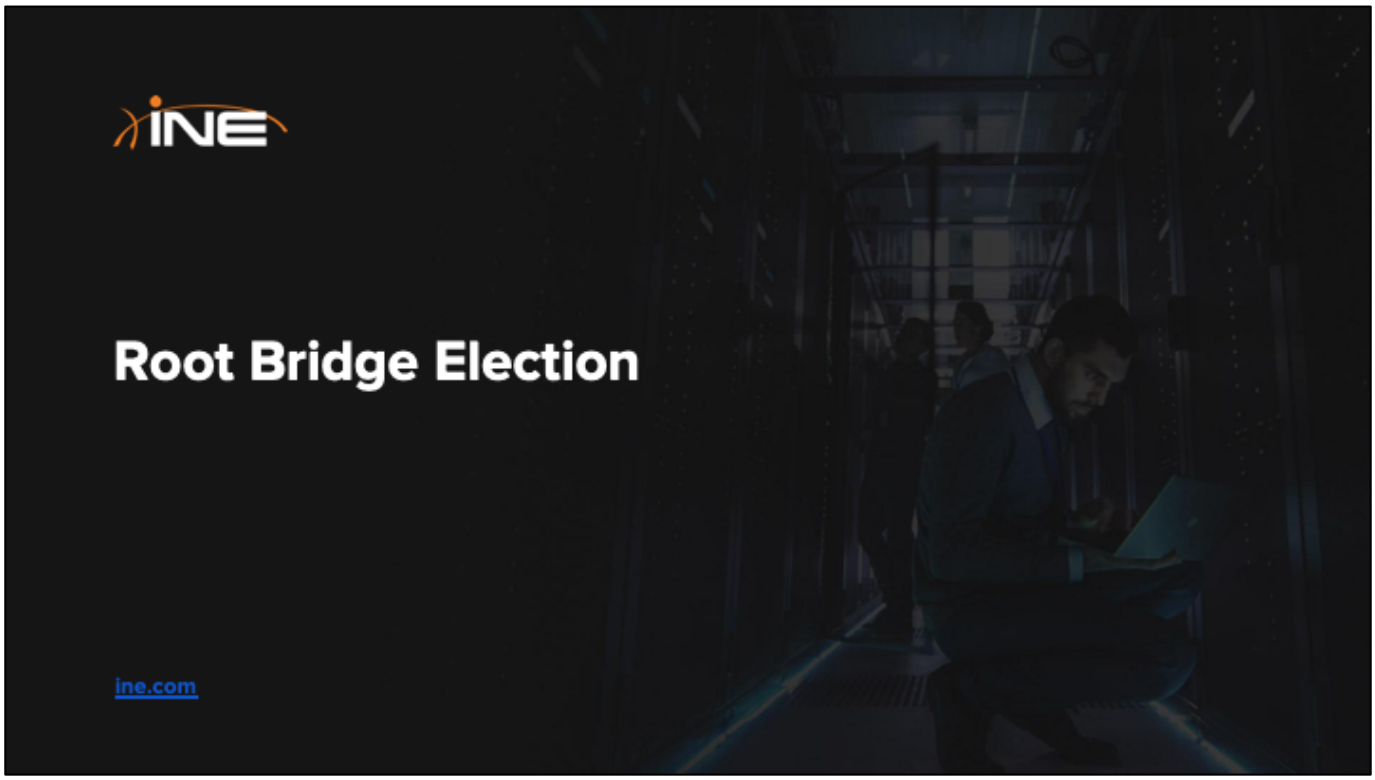
The Concept Of A Tree



The Concept Of A Tree



Thanks for Watching!


Root Bridge Election
ine.com

Topic Overview

+ How RSTP Works - Overview
+ Root Bridge Election
+ Understanding BPDU Structure
+ RSTP Port Roles & States

How RSTP Works

- Elect one Root Bridge
- Elect one Root Port per bridge
- Elect Designated Ports
- Block Non-Designated Ports

Root Bridge Election

- Switch with lowest Bridge ID in the network becomes Root Bridge
- Bridge ID contains...
  - Bridge Priority
    - 0 - 61440 in increments of 4096
  - System ID Extension
    - 0 - 4095
  - MAC Address

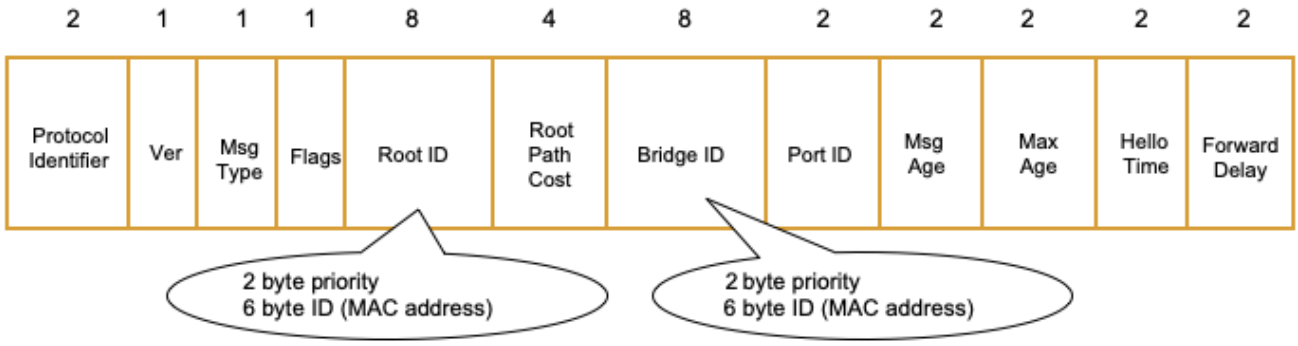
You cannot configure a Bridge-Priority unless it is an increment of 4096.

--

Root Bridge also in charge of dictating STP timers to be used within Broadcast Domain. We'll talk more about these timers in a moment.

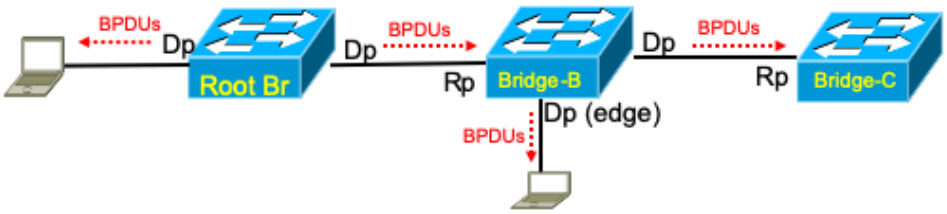
The Bridge Protocol Data Unit

- BPDU = Bridge Protocol Data Unit
- Required to determine, and maintain, STP topology



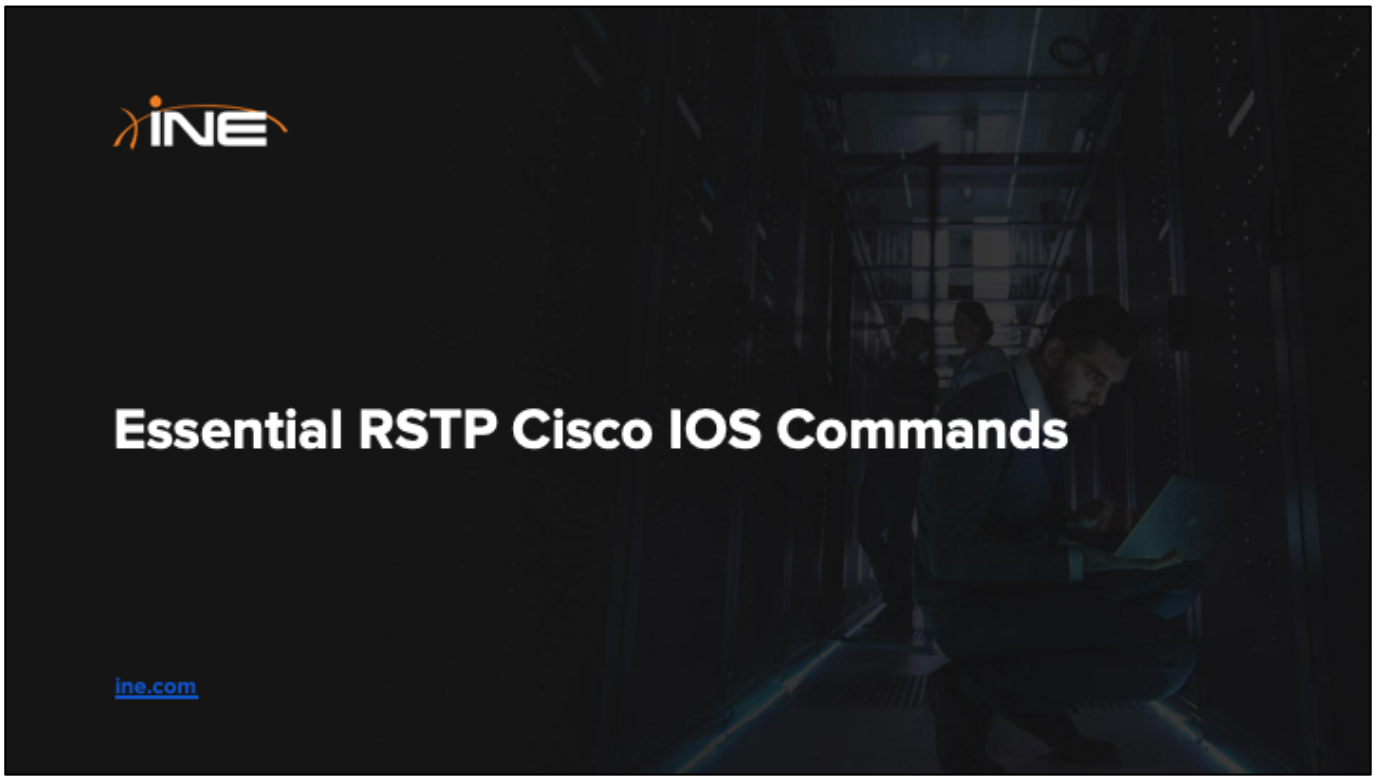
RSTP Port Roles & States

RSTP Port Role	RSTP Port State



We're concentrating right now on Designated and Root Ports. We'll talk about Alternate and Backup ports later on.
-
Main takeaways here:
---Designated Ports "deliver"
---Root Ports "receive"

Thanks for Watching!


Essential RSTP Cisco IOS Commands
ine.com

Topic Overview

+ RSTP Essential IOS Commands
+ Deterministic Placement Of The Root Bridge
+ Verifying The RSTP Root

RSTP Essential Commands

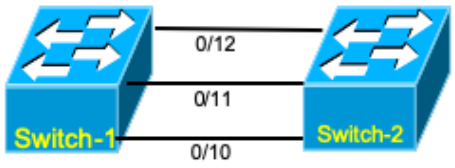
- Enable RSTP
  - (config)#spanning-tree mode rapid-pvst
  - Automatically backwards compatible with legacy STP
- Enable PortFast for RSTP Edge Port recognition
  - (config-if)#spanning-tree portfast [trunk]
  - ...or...
  - (config)#spanning-tree portfast default
- Deterministically set RSTP Root Bridge per VLAN
  - (config)#spanning-tree vlan <vlan-id> root primary
  - ...or...
  - (config)#spanning-tree vlan <vlan-id> priority <value>

If performing VLAN Trunking to something like a server or router-on-a-stick, you would want that trunk link to be recognized by RSTP as an Edge Port.

Spanning-tree vlan x root primary
--If bridge is already STP Root…lowers priority to 24576
--If current Root is already 24576, this command:
-----Lowers current priority to 24576 if this bridge has lower MAC than existing Root Br.
-----Selects next-lowest priority in increment of 4096 if this bridge has higher MAC than existing Root Br.
--If current Root is already 4096, this command fails with error message.
--
Spanning-tree vlan x root secondary
--Lowers priority to 28672
--Assumption is that NO other switches within broadcast domain are lower than this other than Root Bridge.

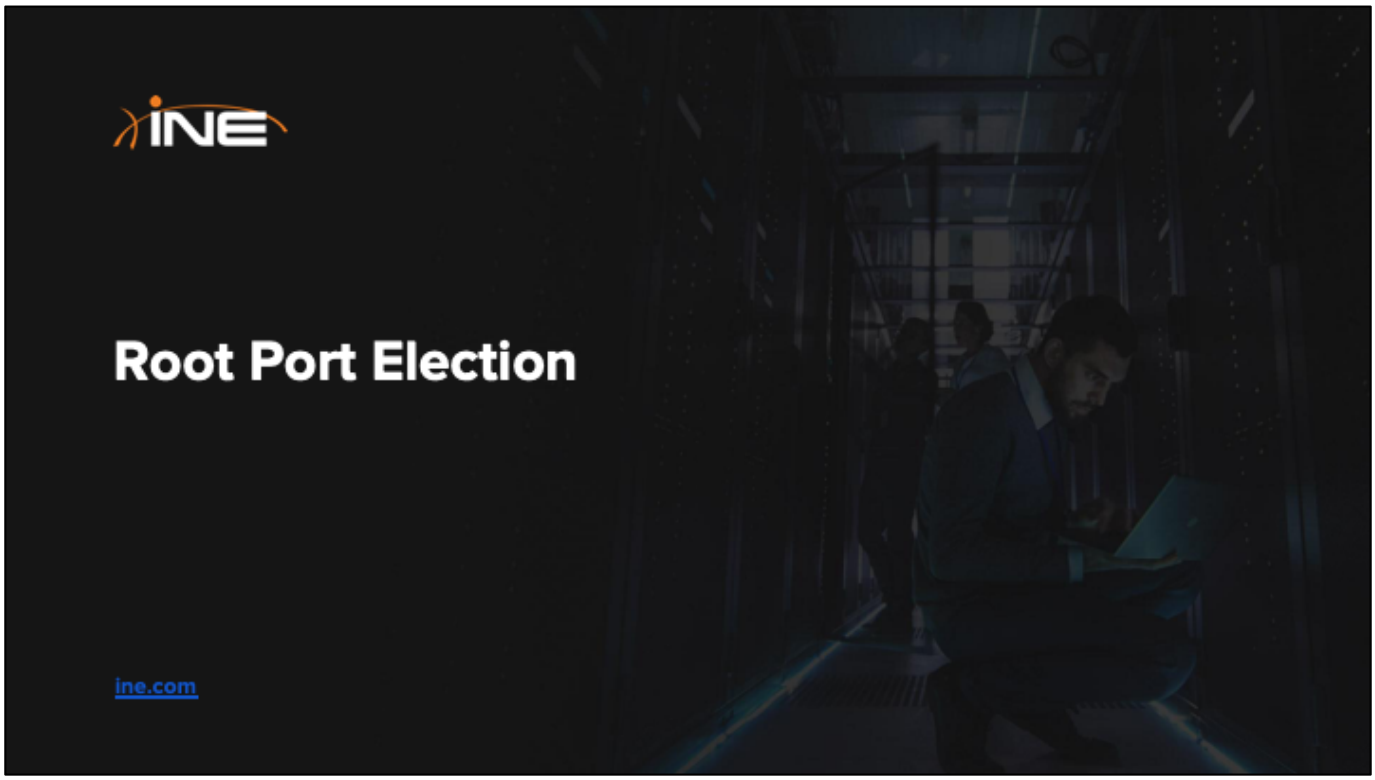
Verifying The RSTP Root Bridge

- Switch#show spanning-tree root
- Switch#show spanning-tree vlan <vlan-id>



Let’s do a quick demonstration to view the output of these commands.

Thanks for Watching!


Root Port Election
ine.com

Topic Overview

+ Root Port Election Process
+ RSTP Cost Values
+ Predicting The RSTP Root Ports

Root Port Election

- RP is upstream facing towards Root Bridge
- Elected based on lowest Root Path Cost
  - Cumulative cost of all links to get to the root
- Cost based on inverse bandwidth
  - i.e. higher bandwidth, lower cost
  - Not linear
- If tie in cost...
  - Choose lowest upstream BID
  - Choose lowest upstream Port ID

RSTP Cost Values

- Every switch that transmits/forwards a BPDU includes its own, local cost to reach STP Root
- 802.1d specified some default port costs but does NOT specify any formula for ports outside of these values

Table 8-5—Path Cost Parameter Values

Parameter	Link Speed	Recommended value	Recommended range	Range
Path Cost	4 Mb/s	250	100–1000	1–65 535
Path Cost	10 Mb/s	100	50–600	1–65 535
Path Cost	16 Mb/s	62	40–400	1–65 535
Path Cost	100 Mb/s	19	10–60	1–65 535
Path Cost	1 Gb/s	4	3–10	1–65 535
Path Cost	10 Gb/s	2	1–5	1–65 535

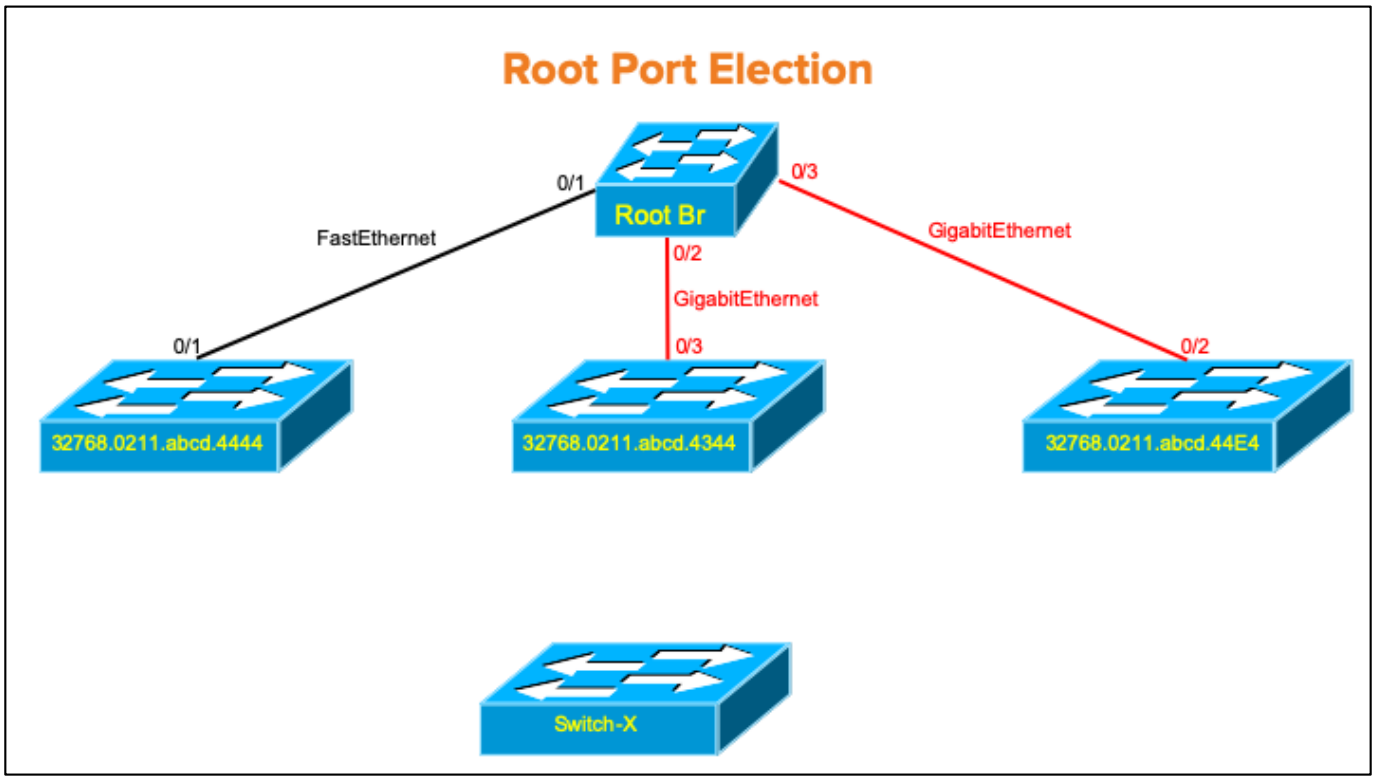
Image taken from page-109 of IEEE 802.1d (1998) specification.

-

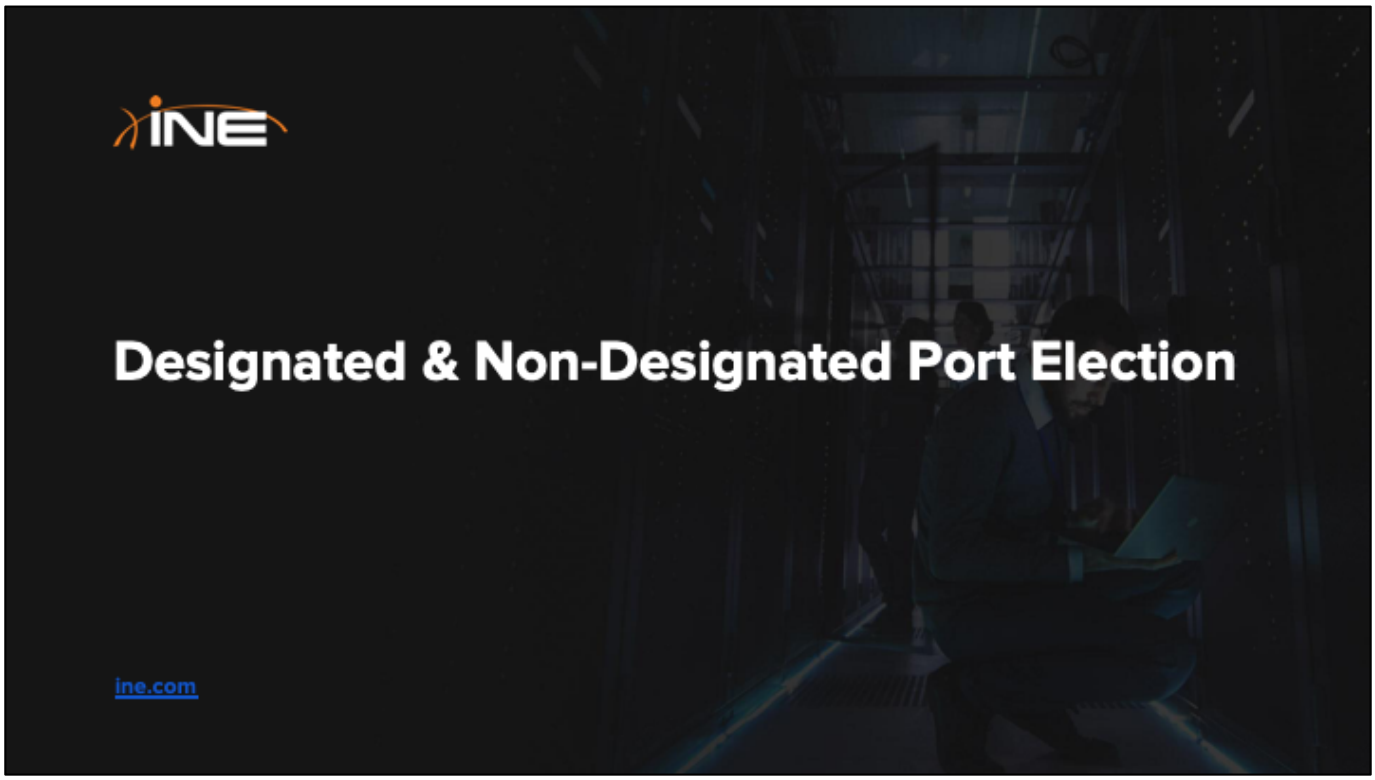
Interesting gotcha: On page 154 of the latest IEEE 802.1d revision (2004) the costs were upgraded to totally different values than what is shown here, however Cisco switches to this day (15-years later) are still using the costs shown in the graphic above.

-

On ports that are Blocking, you can see what value your upstream neighbor advertised in their BPDU by viewing the output of the “show spanning-tree detail” command.


Root Port Election
0/1
Root Br
0/3
FastEthernet
GigabitEthernet
0/2
GigabitEthernet
0/1
0/3
0/2
32768.0211.abcd.4444
32768.0211.abcd.4344
32768.0211.abcd.44E4
Switch-X

Thanks for Watching!

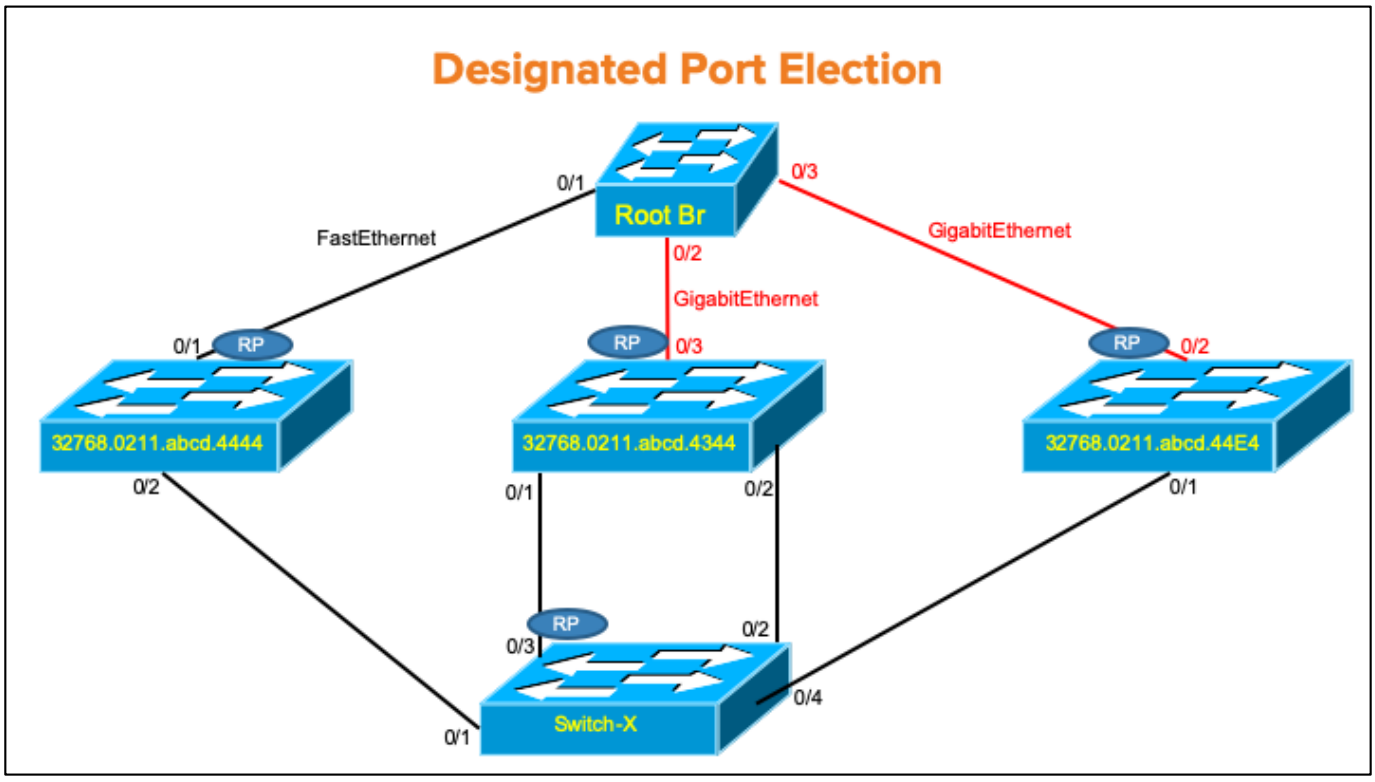

Designated & Non-Designated Port Election
ine.com

Topic Overview

+ Designated Port Elections
+ Predicting The Designated Port
+ BPDUs As Keepalives
+ Non-Designated Ports

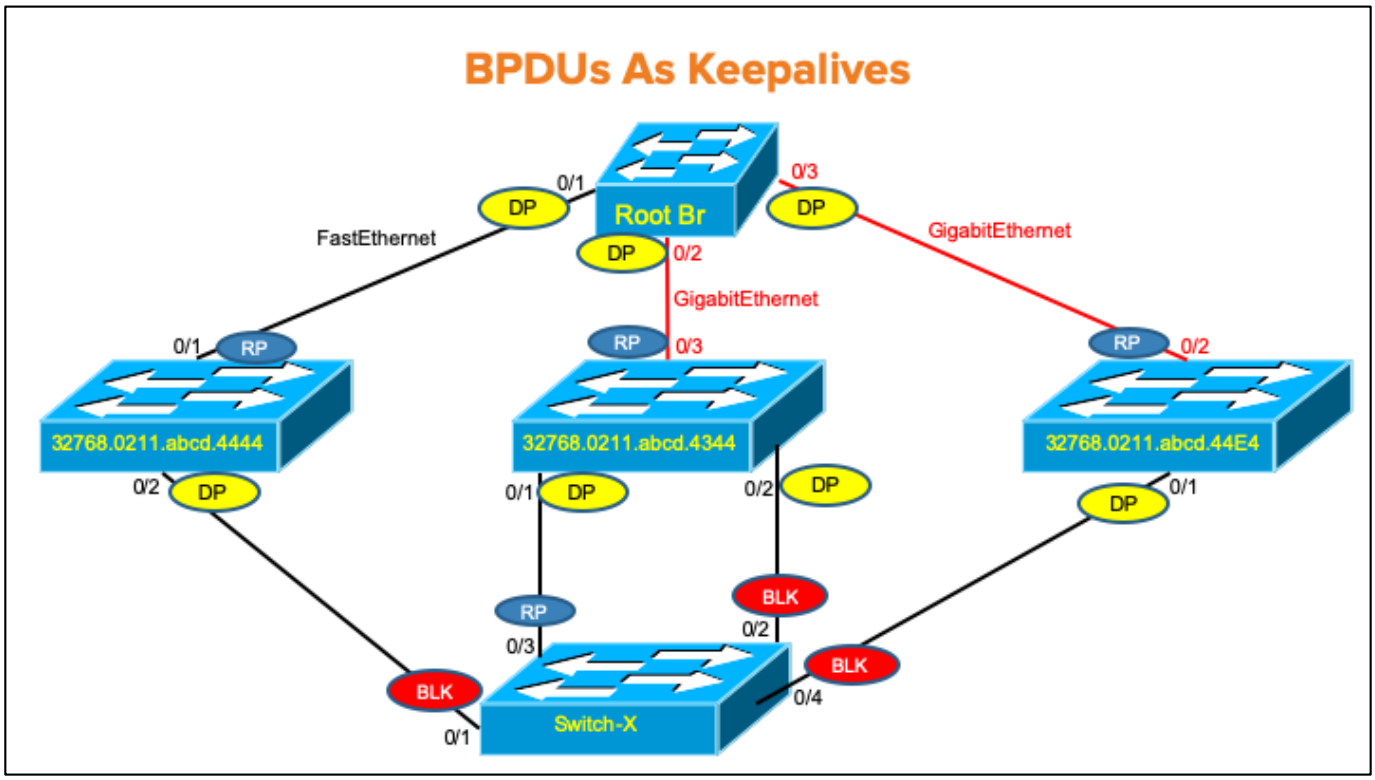
Designated Port Election

- DPs are downstream facing away from Root Bridge
- Like Root Port, elected based on...
  - Lowest Root Path Cost
  - Lowest BID
  - Lowest Port ID
- All other ports in the same collision domain go into either Root or Non_Designated (blocking) roles


Designated Port Election
Root Br
0/1
0/3
0/2
FastEthernet
GigabitEthernet
GigabitEthernet
RP
0/1
32768.0211.abcd.4444
0/2
RP
0/3
32768.0211.abcd.4344
0/1
0/2
RP
0/2
32768.0211.abcd.44E4
0/1
RP
0/3
0/2
Switch-X
0/1
0/4

BPDUs As Keepalives

- In classic (pre-2004) 802.1d, BPDUs:
  - Were only generated by the Root Bridge
  - Were forwarded by non-Root bridges
  - If a non-Root bridge were to stop receiving BPDUs it could NOT answer the question...
    - Did the Root Bridge fail?
    - Did the upstream (non-Root) neighbor fail?
- In RSTP:
  - BPDUs are used as keepalives
  - Designated Ports on non-Root bridges must always transmit BPDUs
  - Loss of 3-BPDUs triggers action on a downstream switch

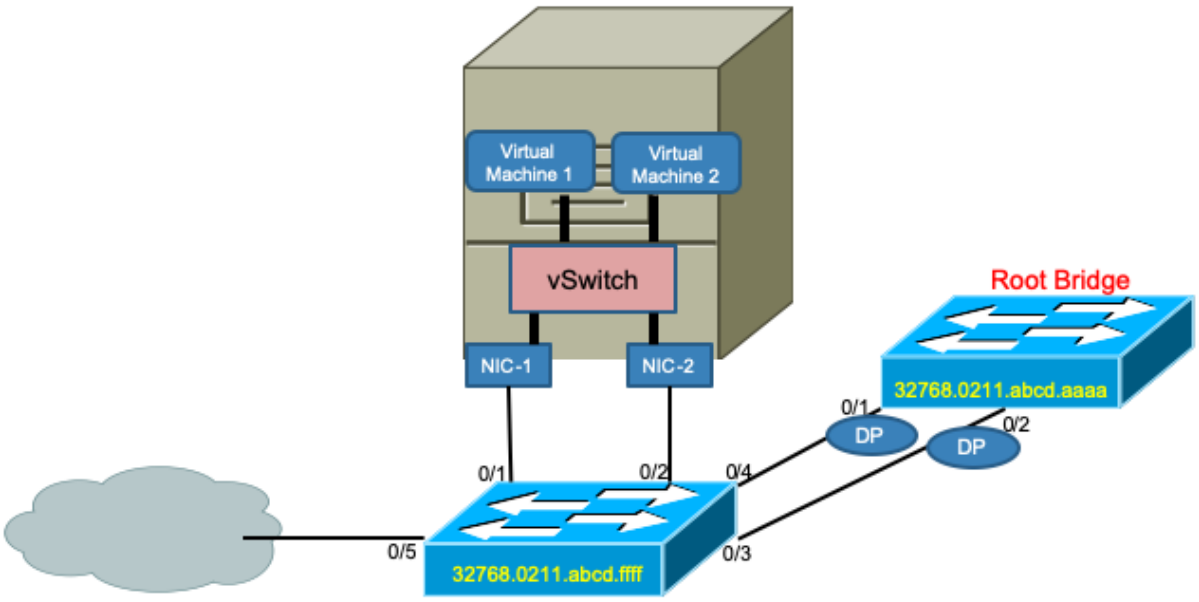

BPDUs As Keepalives
Root Br
0/1
0/3
0/2
DP
DP
DP
FastEthernet
GigabitEthernet
GigabitEthernet
0/1
RP
RP
0/3
RP
0/2
32768.0211.abcd.4444
32768.0211.abcd.4344
32768.0211.abcd.44E4
0/2
DP
0/1
DP
0/2
DP
DP
0/1
BLK
RP
0/2
0/3
BLK
BLK
0/4
Switch-X
0/1

Non-Designated Ports

- Ports that are neither Root Ports or Designated Ports become Non-Designated Ports
- This port role is always “Discarding”
- Two different types of Non-Designated Ports:
  - Alternate Port
    - Received superior BPDU on non-Root port
    - BPDU was received with a different sending Bridge-ID than local Bridge-ID
  - Backup Port
    - Received superior BPDU on non-Root port
    - BPDU was sent with a same sending Bridge-ID as local Bridge-ID

Discarding was formerly known in earlier versions of 802.1d as ”Blocking” and I use the terms interchangeably in this series.

Alternate & Backup



Thanks for Watching!

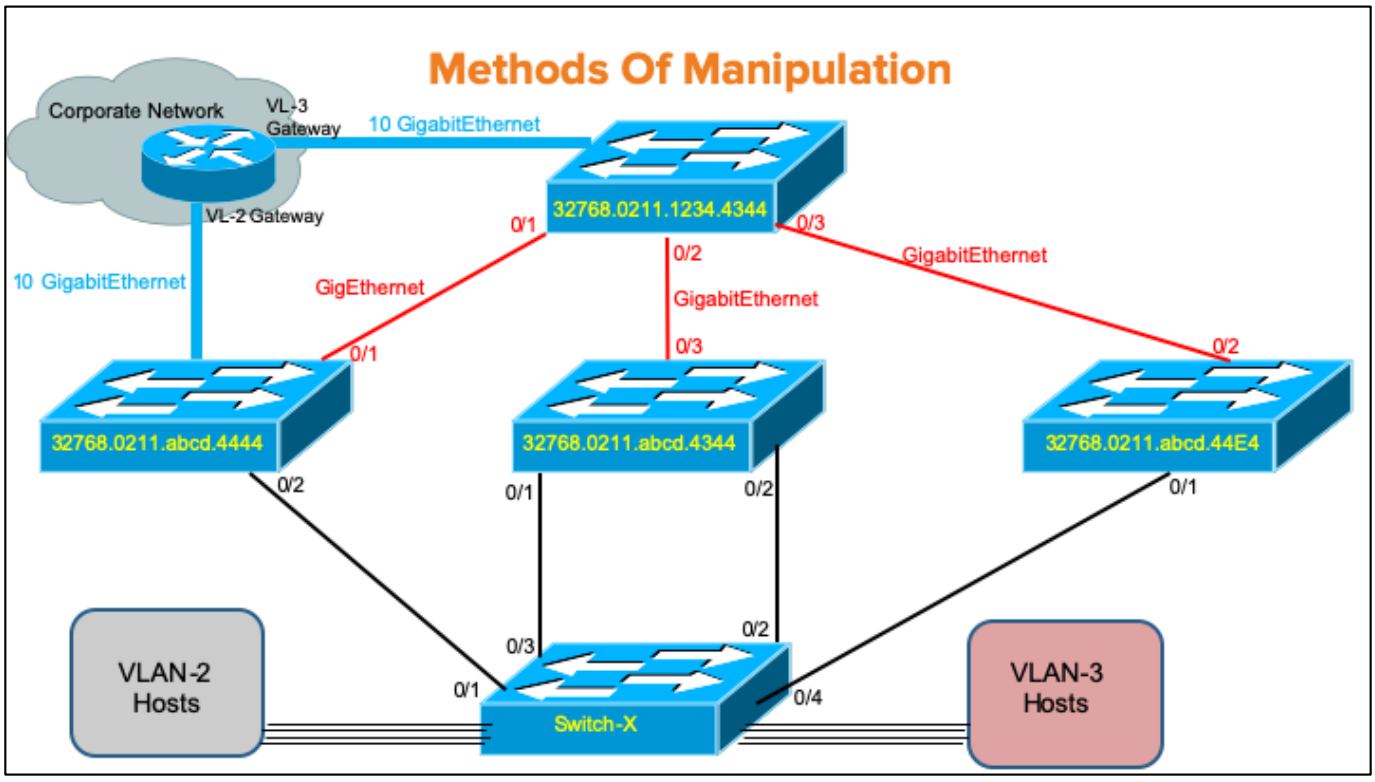

Rapid-PVST Topology Manipulation
ine.com

Topic Overview

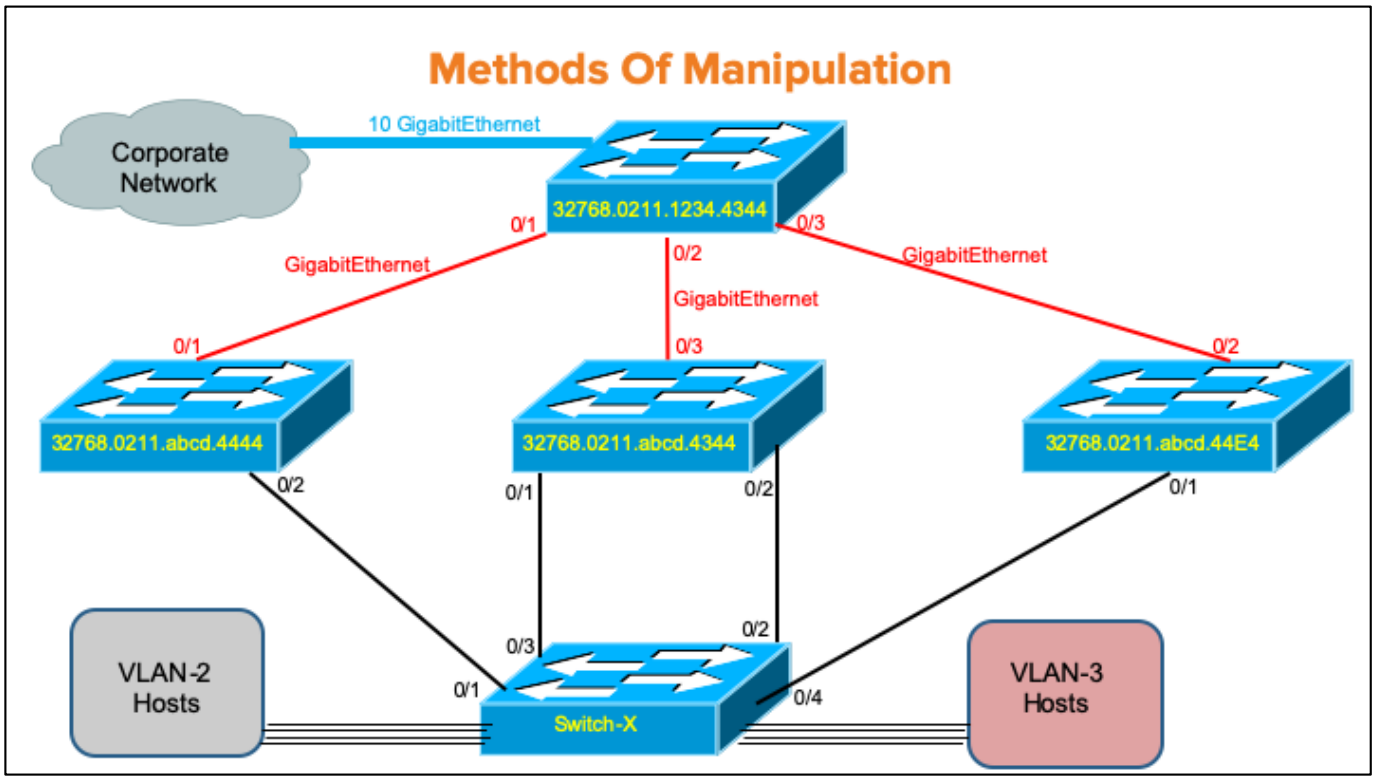
+ Manipulating RSTP Topologies

Manipulating Rapid-PVST Topologies

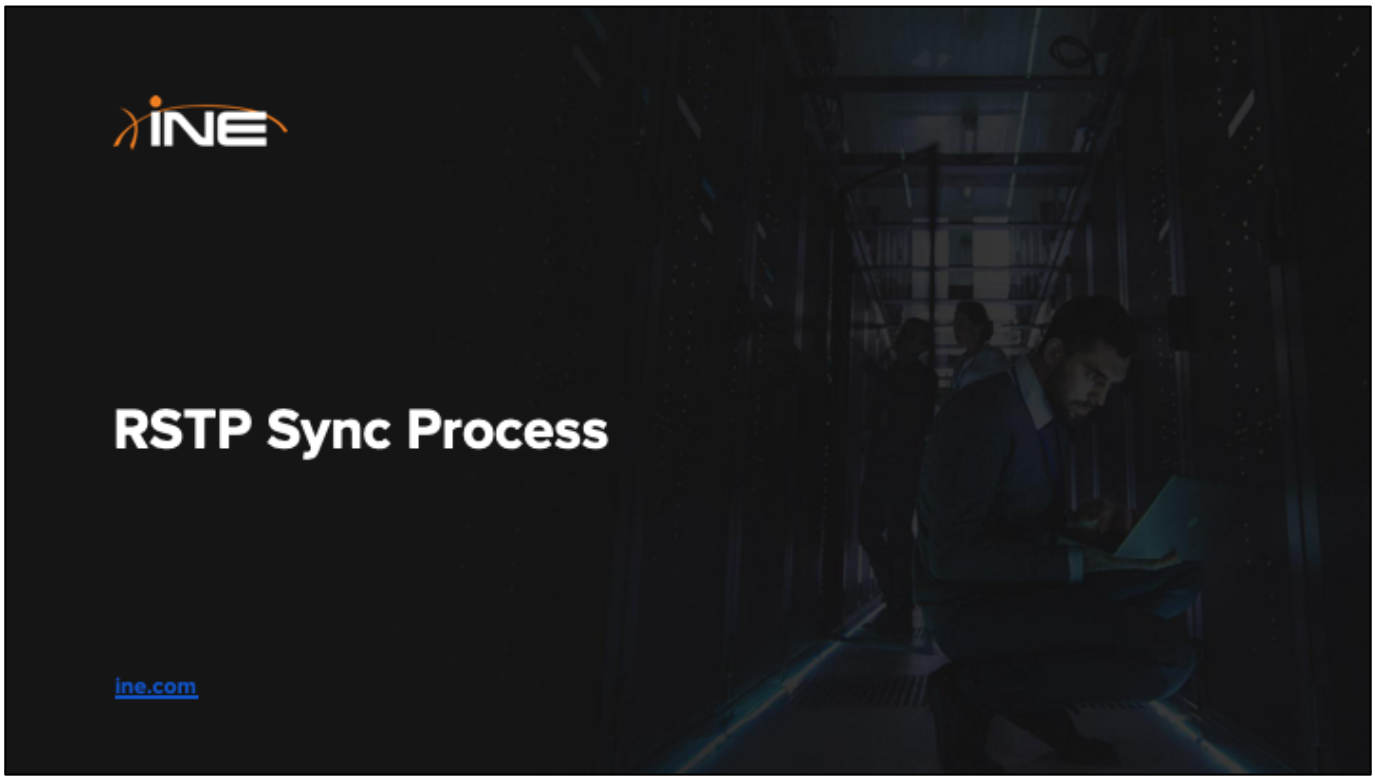
- Many switched topologies utilize several VLANs
- Typically all VLANs are allowed on VLAN trunks between switches
- Rapid-PVST involves creating a RSTP instance for each VLAN
- By default, all RSTP trees will follow the same path, potentially causing congested links
- VLAN traffic can be moved onto non-congested links by manipulating the RSTP topology



- In this topology we'll separate the Root Bridges for both VLANs to show how traffic going to the respective Default Gateway interfaces can avoid taking the exact same links.


Methods Of Manipulation
Corporate Network
10 GigabitEthernet
32768.0211.1234.4344
0/1
0/2
0/3
GigabitEthernet
GigabitEthernet
GigabitEthernet
0/1
0/3
0/2
32768.0211.abcd.4444
32768.0211.abcd.4344
32768.0211.abcd.44E4
0/2
0/1
0/2
0/1
0/3
0/2
0/1
0/4
Switch-X
VLAN-2 Hosts
VLAN-3 Hosts

Thanks for Watching!

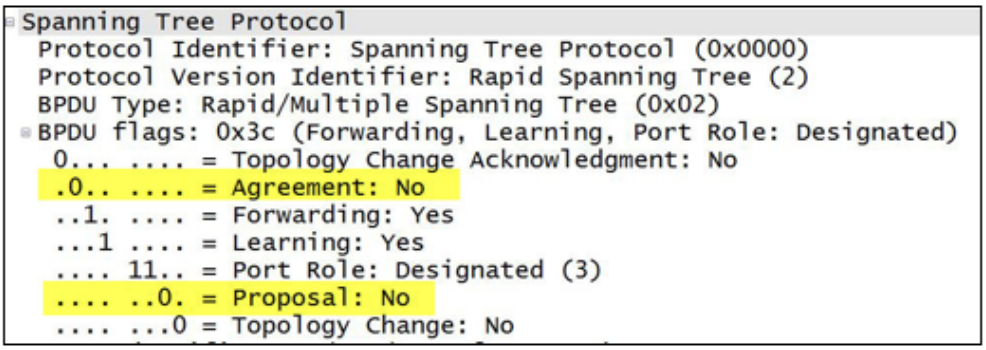

RSTP Sync Process
ine.com

Topic Overview

+ RSTP Sync Process Defined
+ Usage Of RSTP Flags
+ Propagating New Root Bridges Via The Sync Process

RSTP Sync Process

- RSTP goal is for a bridge to synchronize a new port with the rest of the topology as quickly as possible
- New flags of “Proposal” and “Agreement” within BPDU accomplish this

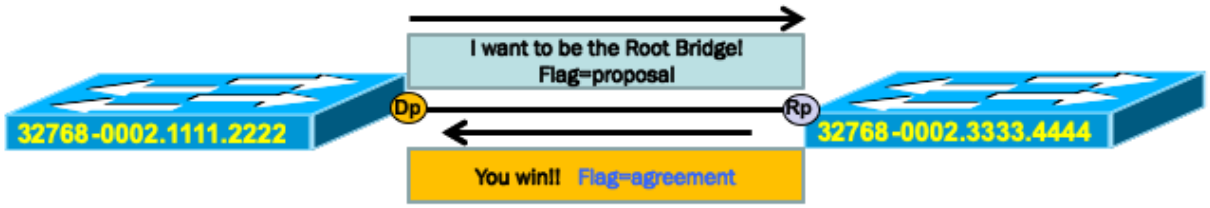


NOTE: These flags are only used on links identified by STP as “point-to-point”…normally full-duplex.

-

Links in half-duplex mode are considered “Shared” by RSTP and cannot use these bits/flags.

Electing The Root Bridge



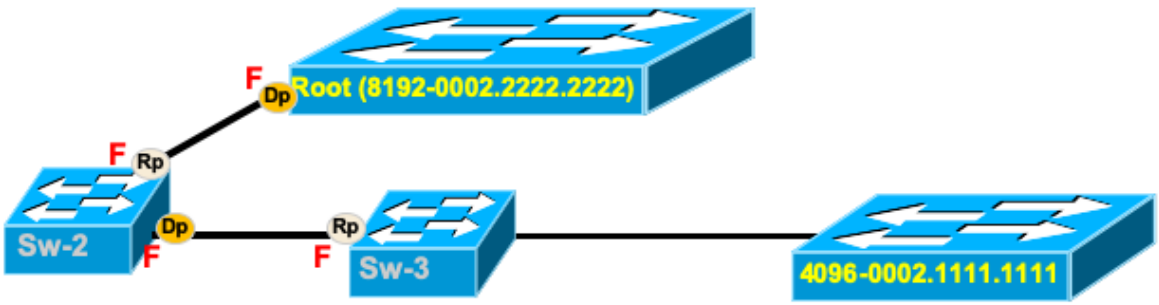
1. New link added connecting switches
2. Bi-directional BPDU transmission with “Proposal” bit set
3. BPDU transmission with “Agreement” bit set

RSTP: Root Bridge Changes

- A new Root Bridge introduces itself by way of BPDUs containing “proposal” bit
- On links connected to that new Root, process is same as previously described
- However, downstream switches take the following actions…

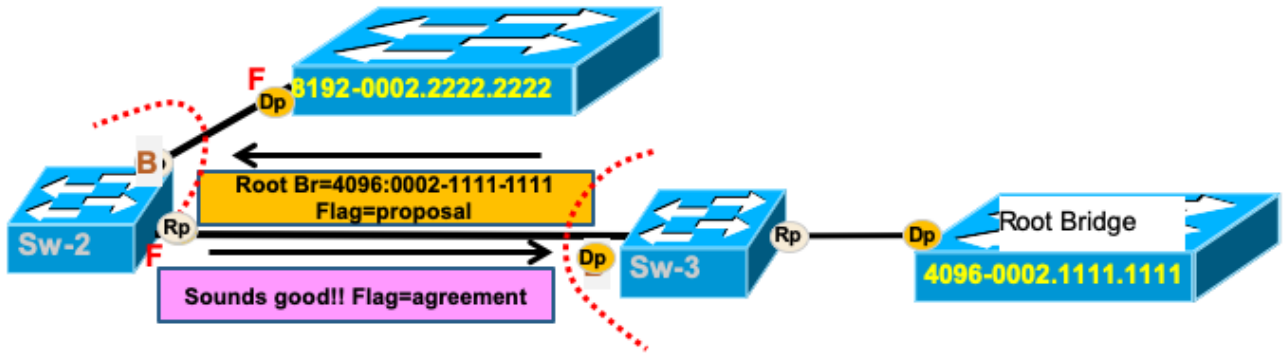
Synchronization Of Root Bridge Changes

- Introduction of new Switch/Bridge
- At this point, several things happen at once...

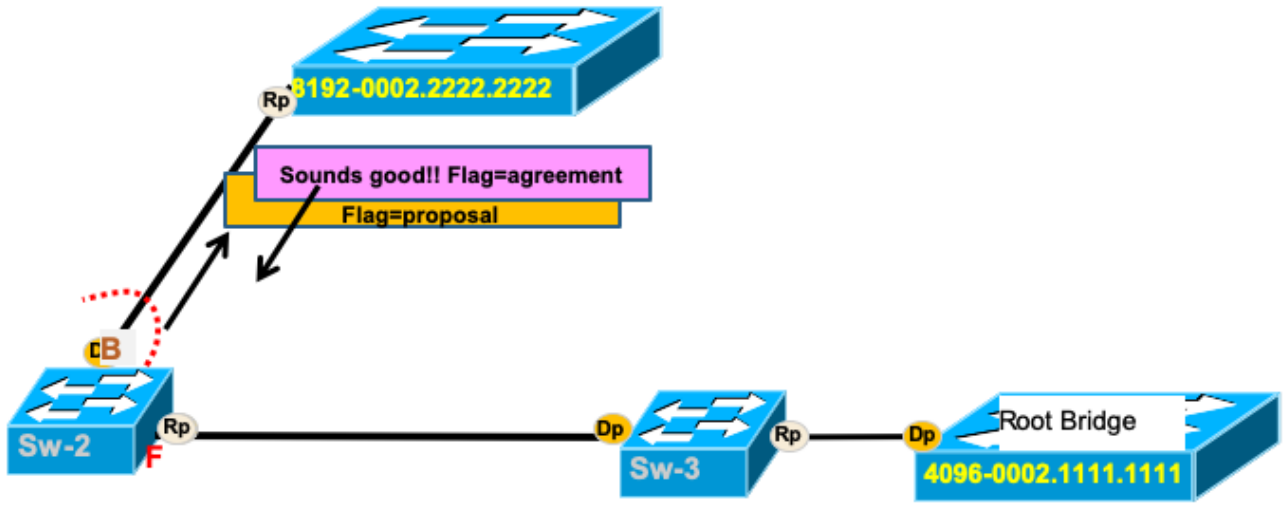


Synchronization Of Root Bridge Changes

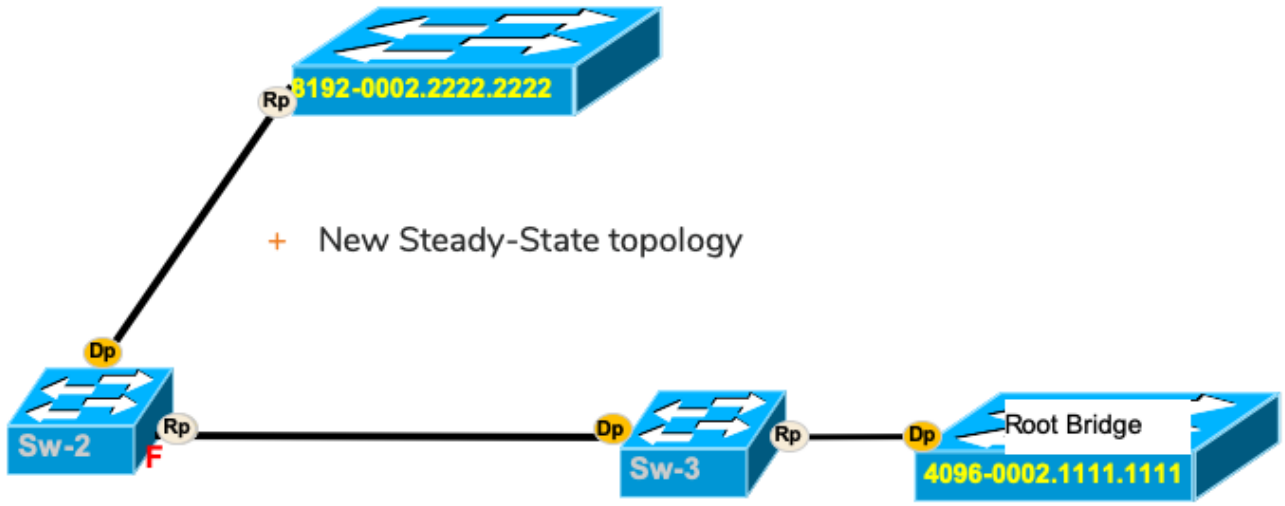
Synchronization Of Root Bridge Changes



Synchronization Of Root Bridge Changes



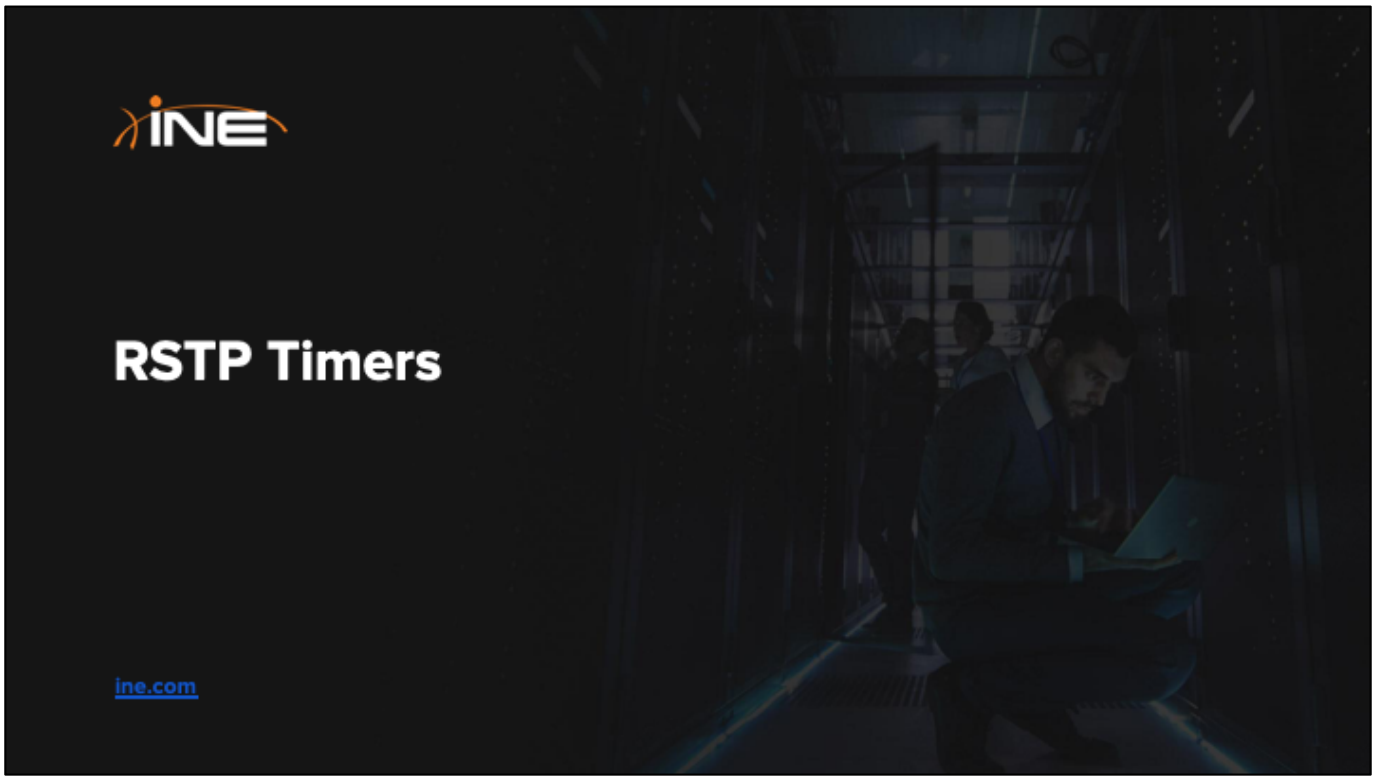
Synchronization Of Root Bridge Changes



RSTP & Root Ports

- RSTP allows for faster Root Port selection
- If Primary Root Port fails, switch can immediately promote an Alternate Port to become the new Root Port
- Similar to Cisco's UplinkFast feature
- Built-in to RSTP (no configuration commands)

Thanks for Watching!


RSTP Timers
ine.com

Topic Overview

+ Defining & Changing The Hello Timer
+ Defining & Changing The Forwarding Delay
+ Understanding The Message Age & Max_Age Timers

Hello Timer & Forwarding Delay

- Hello Timer
  - Dictates frequency of BPDU generation
  - 2-seconds (default)
  - Configurable
    - Switch(config)#spanning-tree vlan <vlan-id> hello-time <1-10 secs>
- Forwarding Delay
  - Used for non-edge host ports to control transition delays
  - Used between RSTP switches on half-duplex connections to control transition delays
  - Configurable
    - Switch(config)#spanning-tree vlan <vlan-id> forward-time <4-30 secs>

A non-edge host port would be any switchport that is NOT configured for Portfast that is connected to a non-STP device (router, laptop, server, etc).

Message Age & Max_Age

- Max_Age Timer
  - Not used for RSTP as a timer
  - Carryover from original 802.1d specification
  - 20-seconds by default
  - Switch(config)#spanning-tree vlan <vlan-id> max-age <6-40 secs>
- Message Age
  - Used as a hop count for RSTP
  - Starts at zero (0) from Root Bridge and incremented by one (1) for each subsequent bridge that BPDU passes through
  - If BPDU received with Message_Age = Max_Age then BPDU is discarded

In theory, if a BPDU is received by an RSTP switch in which the max-age equals the message-age that BPDU is supposed to be discarded. In practice I have seen the receiving switch WILL accept it however that switch will transmit back (to the sending switch) its own BPDU with the Proposal flag set. So the two switches will in a constant cycle of proposal/agreement.

Thanks for Watching!

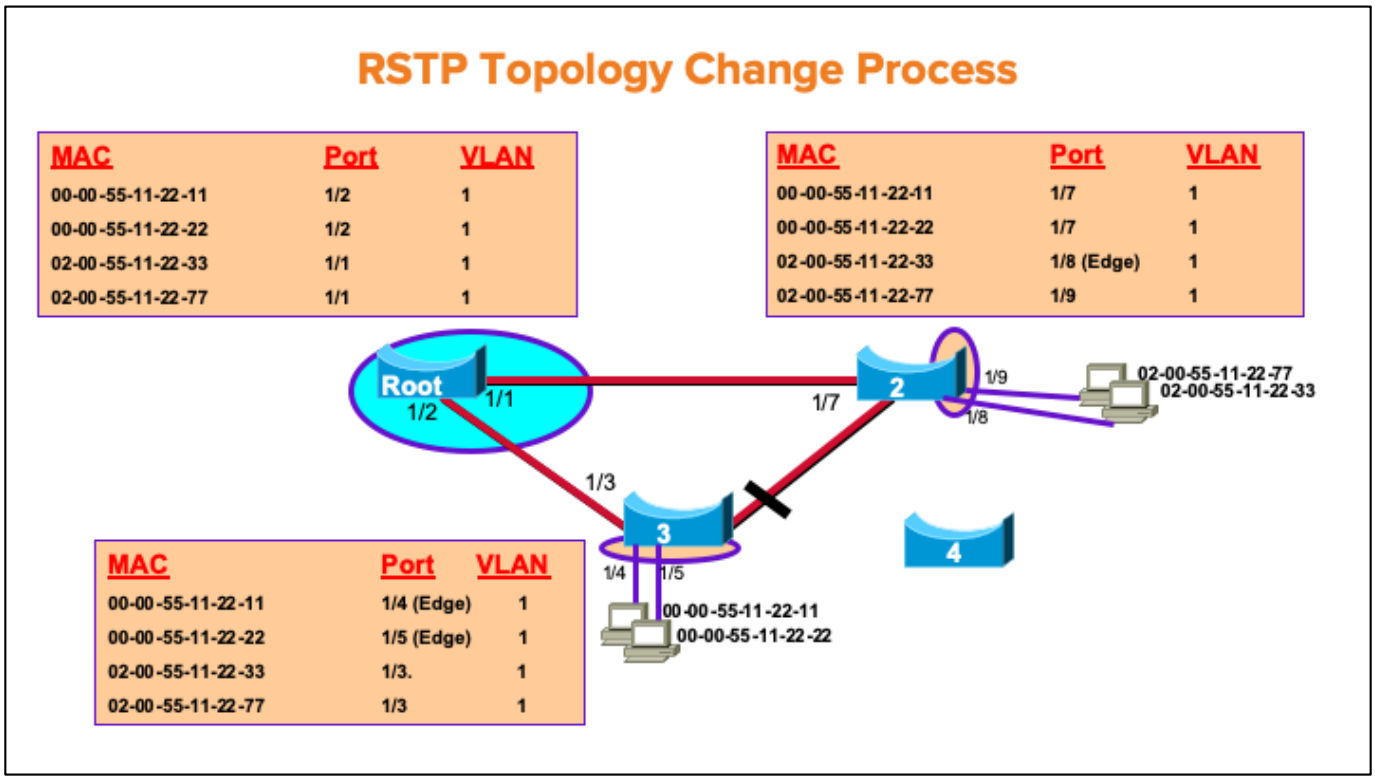

RSTP Topology Changes
ine.com

Topic Overview

+ RSTP Topology Change Process

RSTP Topology Change

- TC-BPDU when non-edge port goes forwarding only
- Transmitted on the network by the initiator (not by the Root any more)
  - Sent on all Root Ports
  - Sent on all Designated Ports that are Non-Edge
- Flush most of the mac-address entries in the network immediately
  - Edge ports are not flushed
  - MACs on ports receiving Topology Change BPDU are not flushed
  - MACs learned on ports from which TC-BPDU is transmitted are flushed



Notice that on Switch-2 we neglected to configure Portfast on 1/9 so this is NOT considered an Edge Port. This oversight WILL affect the RSTP Topology Change process.

Thanks for Watching!