Chapter 4
Rapid Spanning Tree
Multiple Spanning Tree
NET3011 – 17W

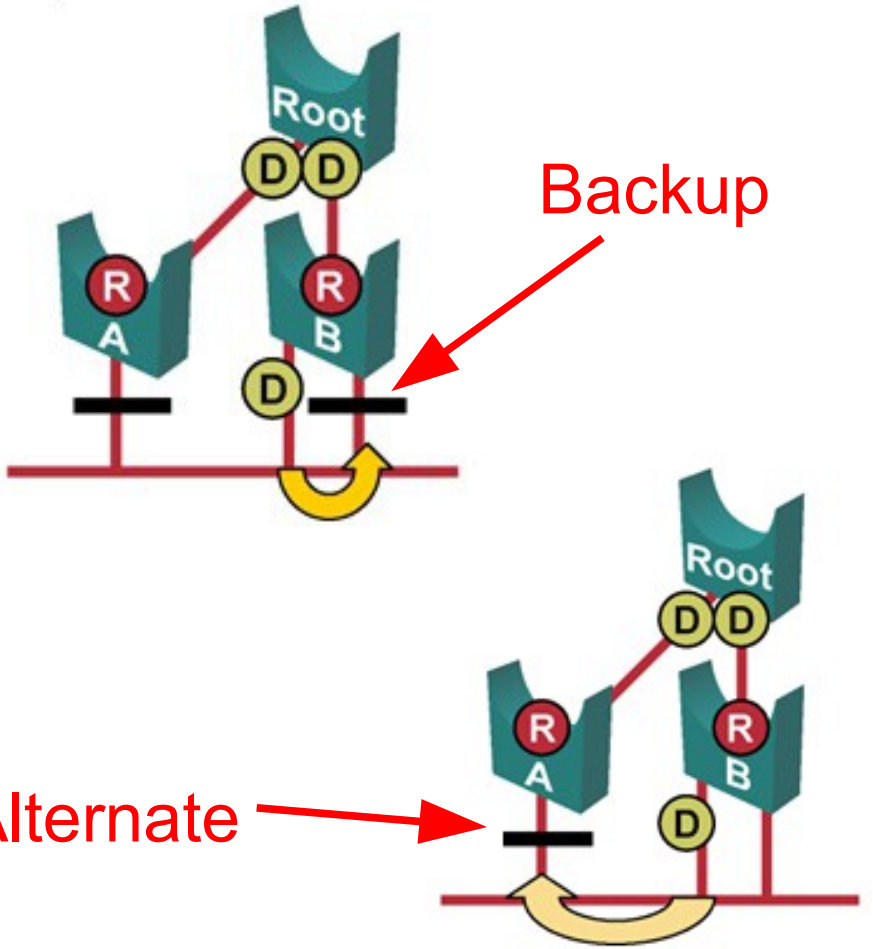
Rapid Spanning Tree – Basics

- *It's all about the speed!* The main goal of Rapid Spanning Tree Protocol (IEEE 802.1w-2001) is to improve convergence time from STP's 30-50 secs
- RSTP improves convergence by:
  – being more proactive in transmitting BPDUs, and
  – adding another negotiation process,
  – thereby (mostly) eliminating the need for STP delay timers
  … **but** at the expense of more bandwidth & CPU resources
- RSTP is backwards compatible since it uses the same frame format as 802.1D BPDUs with the following changes:
  – Version field is set to 2 to indicate RSTP
  – BPDU (message) type also set to 2, for RSTP (or MST)
  – new "Version 1 Length" field is appended
- RSTP elections, ie. the spanning tree algorithm (STA), still use exactly the same metrics as STP; MST incorporates RSTP

RSTP Port Roles

- Root and Designated port roles are identical to STP
- RSTP eliminates the Blocked role, replacing it with either **Backup** or **Alternate**

- Backup is a port on Designated bridge that receives superior BPDUs **from the same bridge**
- Alternate is a port on the non-designated bridge that receives superior BPDUs **from another bridge**



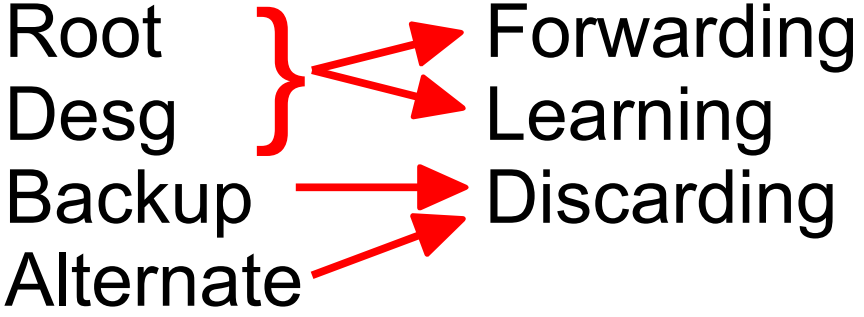
- Both Backup and Alternate are in the *discarding* state

RSTP Port States

- In RSTP, the STP Listening & Blocking states have been replaced with a single new *Discarding* state

STP	RSTP
(Disabled)	(Disabled)
Blocking	**Discarding**
Listening	**Discarding**
Learning	Learning
Forwarding	Forwarding

Be sure to distinguish:
port ROLES vs port STATES



RSTP Neighbour Detection

- When **three BPDUs** in a row (3x Hello intervals) are missed:
  - the neighbor switch is presumed down
  - All MAC address entries pointing to that switch (out that port) are immediately aged out (flushed)
  - See later slides for additional action taken
- RSTP can detect any neighbor down (brain-dead or multi-access segment) in ~6 secs instead of MaxAge (~20 secs)

RSTP Convergence – 4 cases

- All Spanning Tree convergence is a two step process:
  - Elect a Root Bridge
  - Process BPDUs on all switch ports (which default to Designated role) and progress to the appropriate state
- STP requires the expiration of several timers before switch ports can be moved to Forwarding state.
- RSTP takes a different approach: when a switch joins the topology (powered-up) or reacts to a failure in the existing topology, decisions are based on the type of port:
  - **Edge Port**: configured with PortFast
  - **Root Port**: created from Designated port during convergence
  - **Point-to-Point Port**: default for ports that are *full-duplex*
  - **Shared Medium Port**: default for ports that are *half-duplex*

http://www.cisco.com/en/US/docs/switches/lan/catalyst2950/software/release/12.1_9_ea1/configuration/guide/swmstp.html

Case 1: Edge Ports

- Edge port should never have a switch connected, so cannot form bridging loops
- Edge ports are similar to STP PortFast ports:
  – they immediately transition to forwarding state;
  – never flag a topology change when the port transitions between disabled & enabled
- If an Edge port receives a BPDU:
  – the Edge port status is lost, and
  – the port reverts to normal RSTP behaviour
- The command is the same as before:

```
ALS1(config-if)# spanning-tree portfast
```

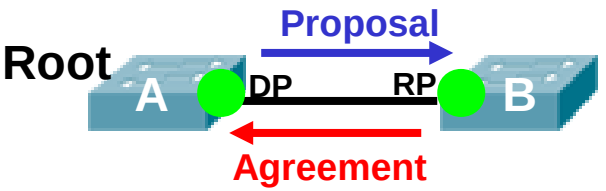
Case 2: Root Port

- Exactly as with STP, a Root port is created from a Designated port immediately upon receipt of a superior BPDU
- See next slide for how that happens

(Wow, that was easy!)

Case 3: P-to-P with Quick Handshake

- Two switches connected P-to-P
  - both ports start in Designated state
  - one switch (A) has a lower BID



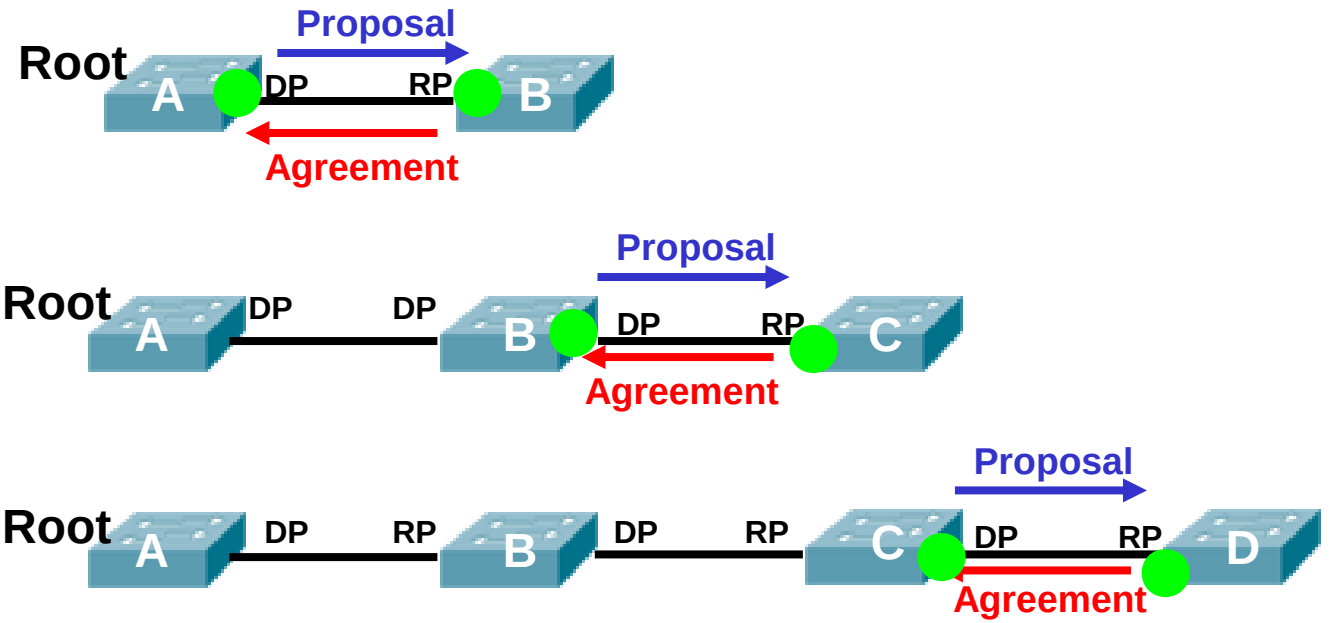
- Switch A sends a **proposal message** (Config BPDU+flag), proposing itself as Root Path/Designated switch for the link
- Switch B:

Can you find these flags in Wireshark??

  - Immediately puts all non-edge ports in Discarding state
  - Sends an **agreement message** (Config BPDU+flag)
  - Creates its new Root port where the **proposal message** was received and immediately goes into Forwarding State
- Switch A:
  - upon receiving the **agreement message**, immediately transitions its Designated port to the forwarding state
  - if ***no*** *agreement received*, follow the normal STP process

Case 3: Proposal-Agreement waves

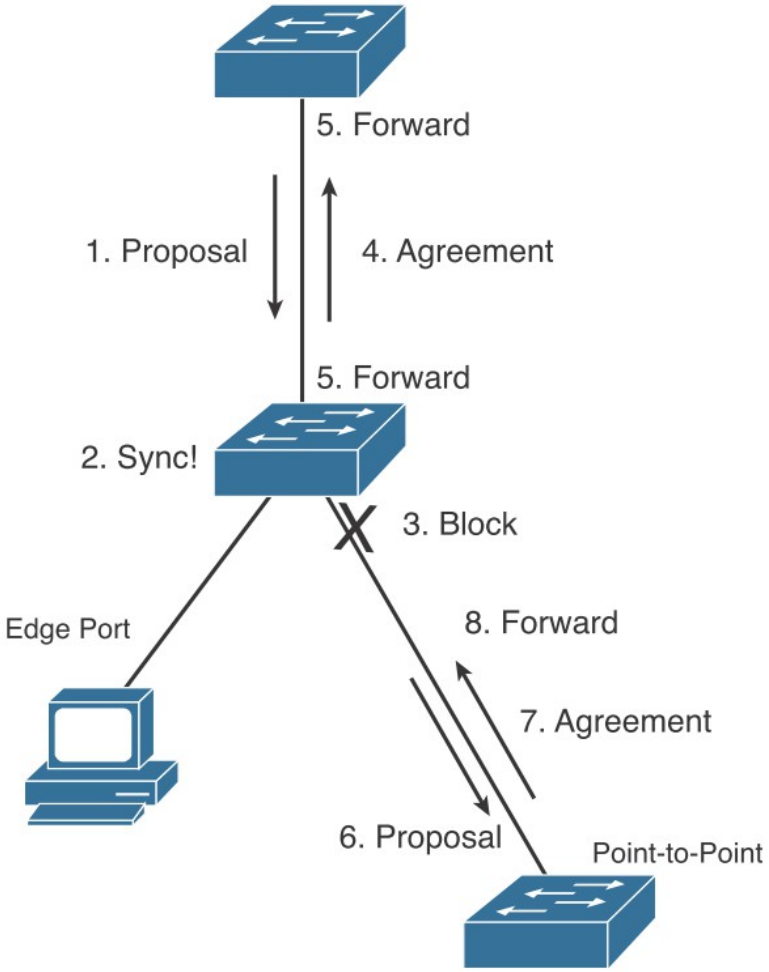
- The Proposal-Agreement propagates rapidly like a wave throughout the network (or VLAN if RPVST!)
- No loops are ever formed: each switch blocks all of its non-edge ports during transition(s) and there are no other switches on links (ie. point-to-point!)



Proposal-Agreement "Synchronization"

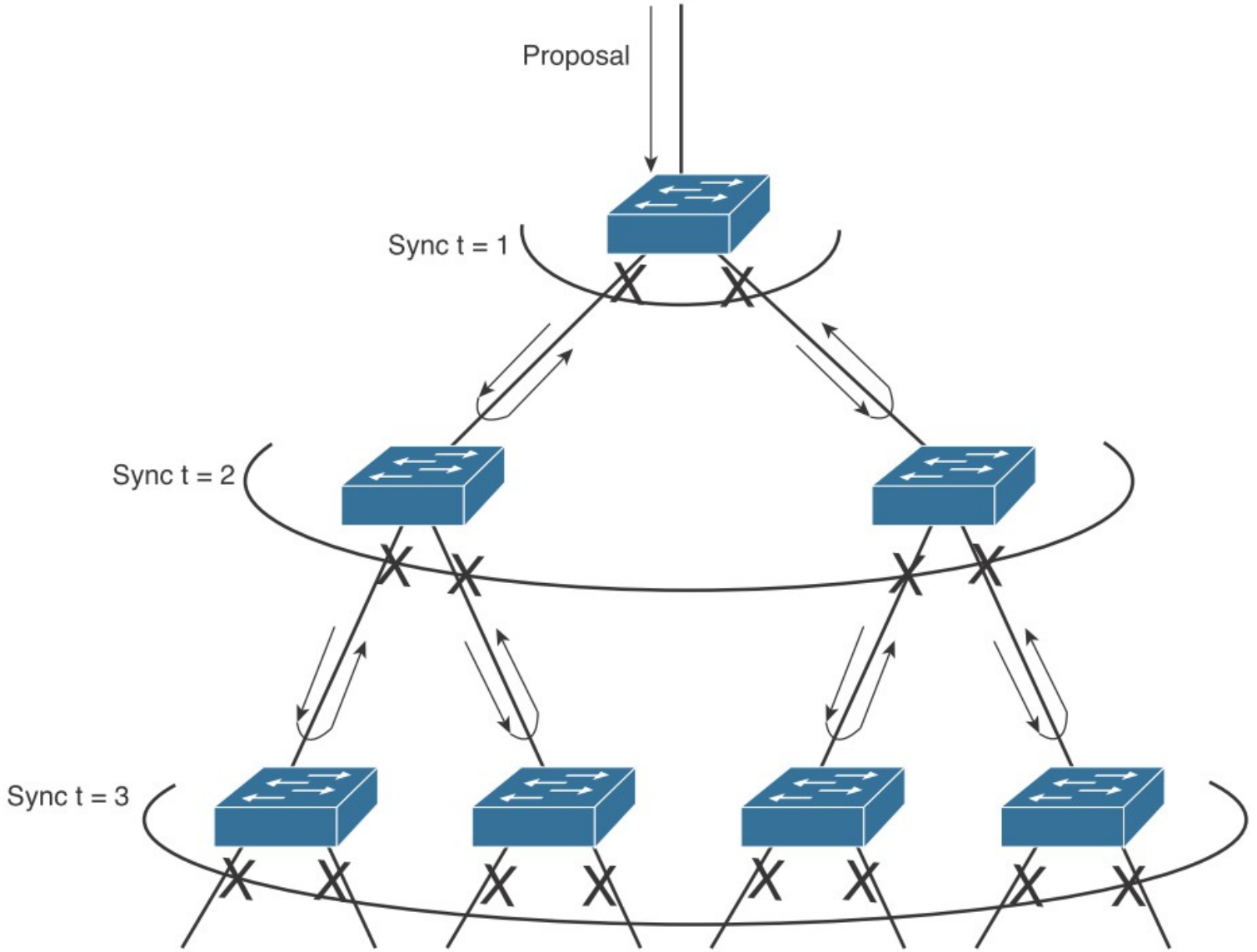
Upon receiving a proposal message on a port, the following sequence of events occurs.

1. If the BPDU is superior, the receiver switch determines that the sender should be Designated and that it's receiving port is it's new root port
2. Before the switch agrees to anything, it must synchronize itself with the topology.
3. All non-edge ports immediately are moved into the Discarding state to guarantee that no bridging loops form
4. An agreement message (configuration BPDU) is sent back to the sender, indicating that the switch is in agreement with the new designated port choice. This also tells the sender that the switch is in the process of synchronizing itself.
5. The root port immediately is moved to the Forwarding state. The sender's port also immediately can begin forwarding.
6. For each non-edge port that is currently in the Discarding state, a proposal message is sent to the respective neighbor.
7. An agreement message is expected and received from a neighbor on a non-edge port.
8. The non-edge port is immediately moved to Forwarding state.



From: CCNP Routing & Switching - Switch 300-115 Official Cert Guide, p. 226-228

Synchronization Propagation



From: CCNP Routing & Switching - Switch 300-115 Official Cert Guide, p. 229

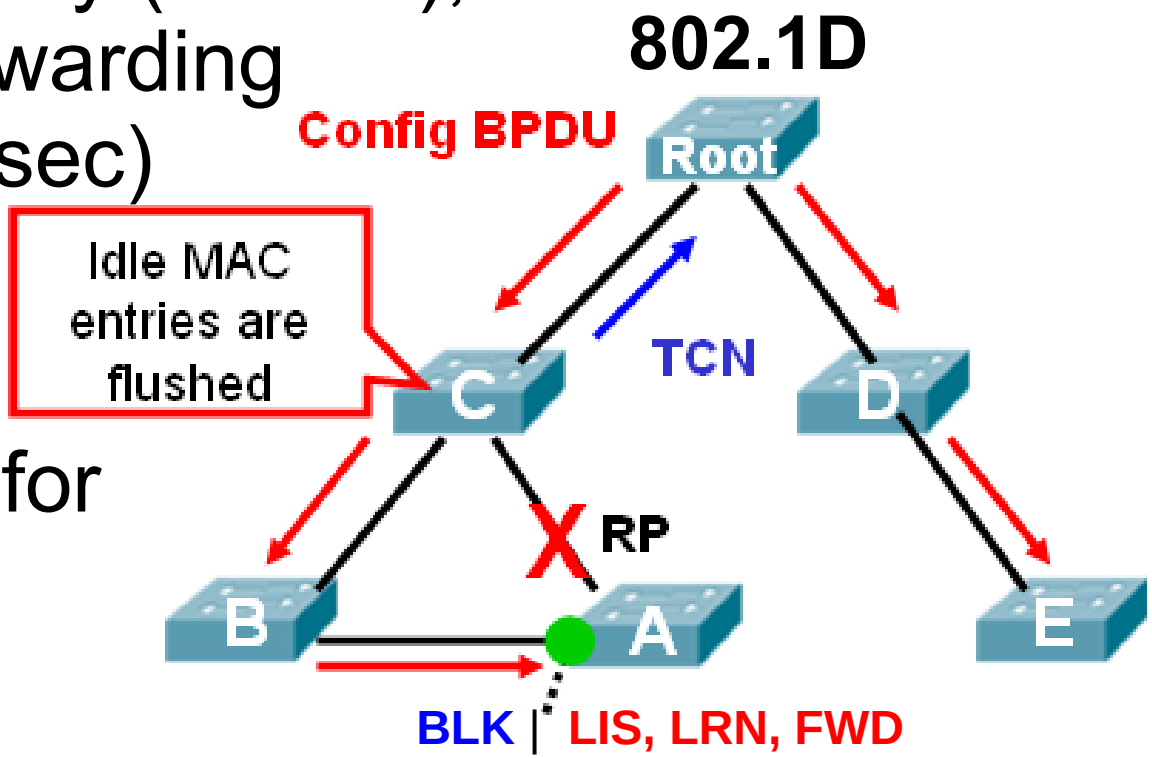
Case 4: Multi-access/Shared Segment

- Multi-access or shared segments is the default for all ports operating in half-duplex mode
- The Spanning-tree operation is based on the assumption that multiple switches might exist
- Port role and states converge using the same process as conventional STP

… So multi-access RSTP segments (and P-to-P when no agreement is received) are *no faster* than STP.

(STP TCN reminder)

- Switch detects state change (port **down** *or* moved to Fwd), sends TCN which is propagated to Root
- Root Bridge sends Config BPDU with TC bit set
  - entire network/VLAN functions with extra flooding for Max Age + Fwd Dly (35 sec);
  - failed areas non-forwarding for 2 x Fwd Dly (30 sec)
- PortFast ports (non-Edge) do *not* generate TCN BPDU for up/down transitions



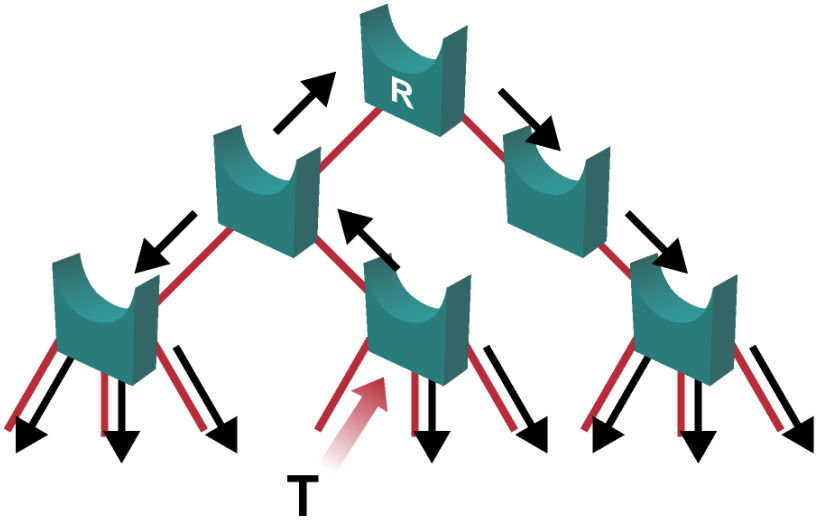
RSTP Generation of TC BPDUs

- Unlike STP, a loss of connectivity does not trigger a topology change i.e. a port down or moved to blocking no longer generates a TCN BPDU
- Instead **only** when, and not until, a non-edge port transitions to **Forwarding** is a config BPDU generated with TC flag (ie. when all ports are ready to go)
- **All** bridges, not just root, may generate the TC flag
- RSTP uses its regular mechanisms to detect topology changes, prevent bridging loops while re-converging, and trigger a TC flag
- The only purpose for detecting and sending the TC flag is to update MAC address tables

RSTP Response to TC BPDU

Exactly two responses:

1. Flush all MAC table entries for non-Edge ports (except for port on which TC flag is received*)
2. Propagate the TC flag out all non-Edge ports in Fwd state (for 2x Hello interval*)



- BPDUs with TC propagate rapidly in an ever-expanding wave, similar to Proposal-Agreement synchronization
  – TCN BPDUs are not used except for "Peer(STP)" links
- In a few secs, most CAM table entries in the entire VLAN are flushed. It may result in *more temporary flooding*, but quickly clears any stale information and provides rapid convergence

* Not verified; from: CCNP Routing & Switching - Switch 300-115 Official Cert Guide, p. 229

RSTP – Config & Verification

- Switching Spanning-Tree operating mode:

```
Sw(config)# spanning-tree mode rapid-pvst !RSTP
Sw(config)# spanning-tree mode pvst       !STP
```

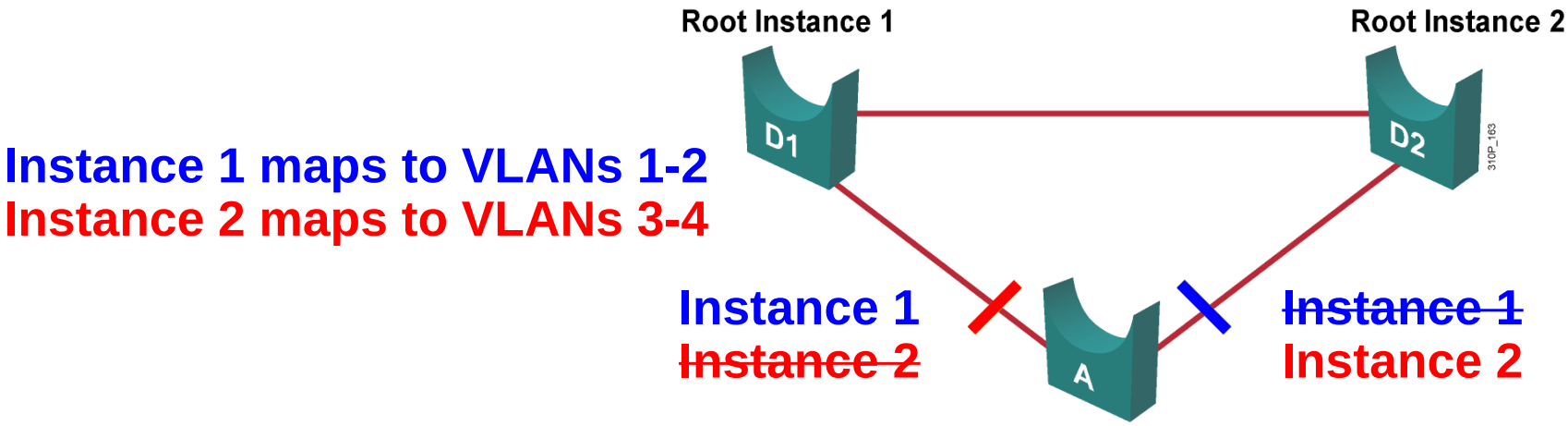
- P-to-P link (regardless of duplex setting):

```
Sw(config-if)# spanning-tree link-type point-to-point
```

```
Sw# show spanning-tree
VLAN0001
  Spanning tree enabled protocol rstp
  Root ID     Priority    24577
              Address     0001.C945.A573
              Cost        4
[... output omitted ...]
```

Multiple Spanning Tree – Basics

- Multiple Spanning Tree (MST – IEEE 802.1s) extends 802.1w RSTP to multiple VLANs
- The purpose of MST is to:
  - administratively define, and thus *reduce*, the total *number of spanning-tree instances* to match the physical topology of the network
  - and thus reduce the CPU loading on the switch



MST - Characteristics

- MST operates in **regions**:
  - a region is a group of switches under a single administration (think AS)
  - a region is defined by:
    - **region name** (32 chars max)
    - **revision number** (0 – 65535)
    - **instance-to-VLAN** mapping (up to 4094 VLANs per instance)

    … and all switches in the region have identical values configured
  - in most networks, a single MST **region** is sufficient
- A region typically has multiple **instances**:
  - instances may be numbered 0-4094 (**but** instance 0 always exists!)
  - 12-bit *Extended System ID* field of the Bridge ID is the instance #
  - total number of instances may be limited by the switch make & model
  - each instance is responsible for running STA for a range of VLANs
- MST can inter-operate with STP (802.1D-199x) & RSTP (802.1w)
  - **Instance 0** [(C)IST] is special: it exists to ensure a loop-free topology between (a) devices inside the MST Region and (b) those outside it.
  - CIST presents the entire MST region as a single virtual switch (bridge) to all external devices (other MST regions, RSTP, or STP)

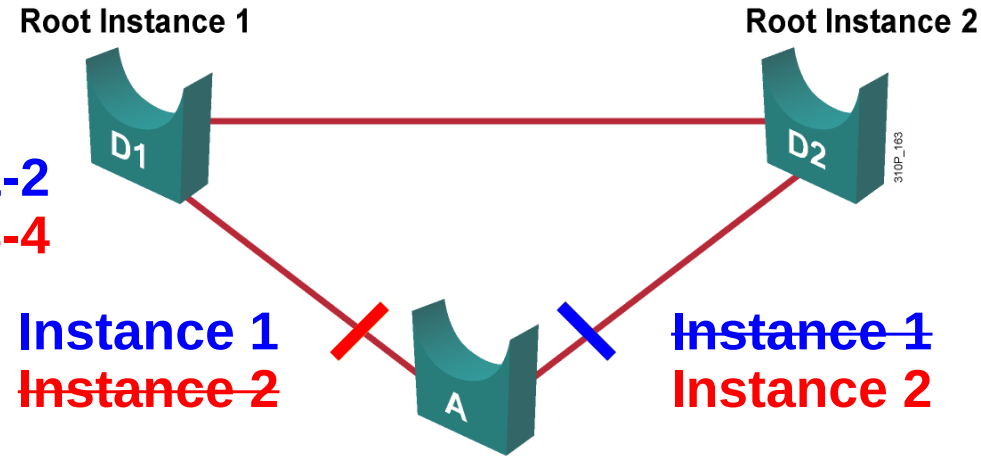
MST – Configuration (1)

- MST configuration on Cisco switches uses a special *editing* mode:
  - start editing: `spanning-tree mst configuration`
  - see active: `show current`
  - see edited: `show pending`
  - commit: `exit` or `end`
  - quit: `abort`
- VTP v3 can propagate the full MST config between switches as long as they are set correctly:
  - activate VTP on **both**: trunk, name, ver 3
  - activate MST on **both**: `spanning-tree mode mst`
  - set VTP for MST on **both**: `vtp mode server mst`
  - set **one** primary server: `vtp primary mst`

MST – Configuration (2)

```
DLS2(config)# spanning-tree mode mst                !Set MST
DLS2(config)# spanning-tree mst configuration       !Edit mode
DLS2(config-mst)# name region1                      !Define
DLS2(config-mst)# revision 10                       !  region
DLS2(config-mst)# instance 1 vlan 1-2               !  values
DLS2(config-mst)# instance 2 vlan 3,4
DLS2(config-mst)# exit                              !Commit
DLS2(config)# spanning-tree mst 2 root primary
DLS2(config)# spanning-tree mst 0-1 root secondary
```

Instance 1 maps to VLANs 1-2
Instance 2 maps to VLANs 3-4



MST – Verification (1)

- At least 4 different signs that MST is running!

```
DLS1# show spanning-tree
MST0
  Spanning tree enabled protocol mstp
  Root ID     Priority   32768
              Address    1c17.d3d2.df00
              Cost       0
              Port       9 (FastEthernet0/7)
              Hello Time 2sec  Max Age 20sec  Fwd Dly 15sec

  Bridge ID   Priority   32768 (priority 32768 sys-id-ext 0)
              Address    7010.5c16.0b00
              Hello Time 2sec  Max Age 20sec  Fwd Dly 15sec

Interface    Role Sts Cost      Prio.Nbr Type
------------ ---- --- --------- -------- --------------------
Fa0/7        Root FWD 200000    128.9    P2p BOUND(STP)
```

MST – Verification (2)

```
DLS2# show spanning-tree mst
##### MST0    vlans mapped:   5-4094
Bridge        address 7010.5c16.0b00  priority   32768 (32768 sysid 0)
Root          address 1c17.d3d2.df00  priority   32768 (32768 sysid 0)
              port    Fa0/7           path cost      0
Regional Root address 1c17.d3d2.df00 priority 32768 (32768 sysid 0)
                                     internal cost 200000 rem hops 19
Operational hello time 2, forward delay 15, max age 20, txholdcount 6
Configured  hello time 2, forward delay 15, max age 20, max hops   20

Interface    Role Sts Cost      Prio.Nbr Type
---------    ---- --- --------- -------- --------------------------
Fa0/7        Root FWD 200000    128.9    P2p

##### MST1    vlans mapped:   1-2
Bridge        address 7010.5c16.0b00  priority   32769 (32768 sysid 1)
Root          address 1c17.d3d2.df00  priority   32769 (32768 sysid 1)
              port    Fa0/7           cost      200000     rem hops 19

Interface    Role Sts Cost      Prio.Nbr Type
---------    ---- --- --------- -------- --------------------------
Fa0/7        Root FWD 200000    128.9    P2p
```

MST – Verification (3)

```
DLS2# show spanning-tree mst interface fa0/7

FastEthernet0/7 of MST0 is root forwarding

Edge port: no     (default)  port guard : none      (default)
Link type: p-to-p (auto)     bpdu filter: disable   (default)
Boundary : internal          bpdu guard : disable   (default)

Bpdus sent 6, received 762

Instance Role Sts Cost      Prio.Nbr Vlans mapped
-------- ---- --- --------- -------- --------------------
0        Root FWD 200000    128.9    5-4094
1        Root FWD 200000    128.9    1-2
```

MST – Verification (4)

```
DLS2# show spanning-tree mst 1 detail

##### MST1    vlans mapped:   1-2
Bridge address 7010.5c16.0b00  priority  32769 (32768 sysid 1)
Root   address 1c17.d3d2.df00  priority  32769 (32768 sysid 1)
       port    Fa0/7           cost     200000    rem hops 19

FastEthernet0/7 of MST1 is root forwarding
Port info    port id       128.9  priority   128 cost   200000
Desgn root   addr 1c17.d3d2.df00  priority 32769 cost        0
Desgn bridge addr 1c17.d3d2.df00  priority 32769 port id 128.7
Timers: message expires in 4 sec, fwd dly 0, fwd transitions 1
Bpdus (MRecords) sent 6, received 873
```

Spanning-Tree Best Practices (1)

- Unlike DTP and VTP ("turn Off"), the recommended best practice for STP is to turn it On even if the designed topology doesn't require it:

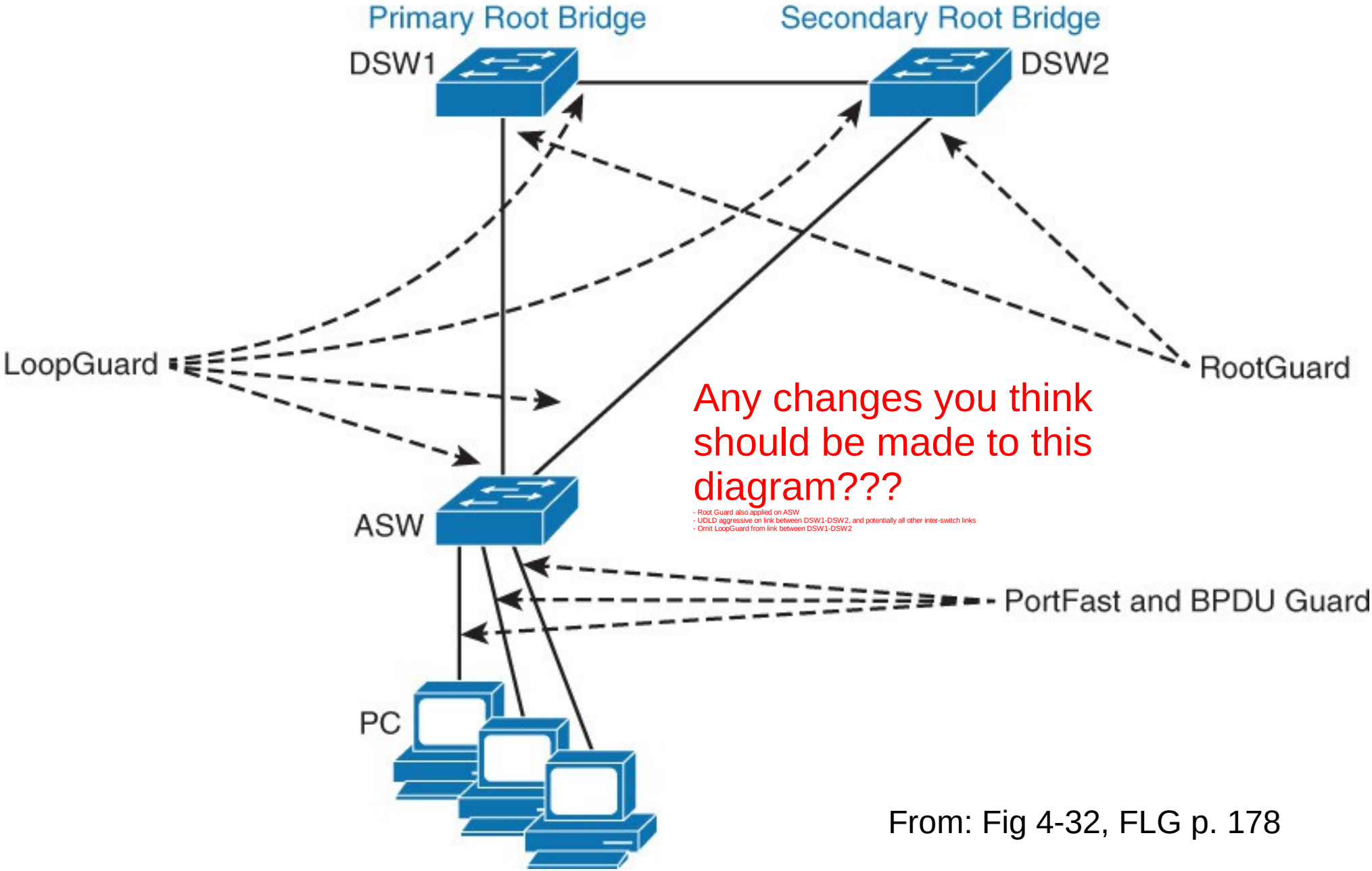
  "Some security personnel have recommended disabling STP at the network edge. This practice is not recommended because the risk of lost connectivity without STP is far greater than any STP information that might be revealed." FLG p.177

- Use RSTP in preference to STP (only *1 line* of config!)
- Use STP enhancements generously:
  - use Portfast on user ports; combine with BPDU Guard
  - use Root Guard widely, especially at the access layer
  - use Loop Guard, but *never* in combination with root guard
  - use UDLD, especially on inter-switch trunks / links
  - avoid BPDU Filter, unless specifically needed
  - *never* configure BPDU Guard + BPDU Filter on same port

Spanning-Tree Best Practices (2)

- For UDLD:
  - generally, use UDLD for all fiber-optic links
  - use UDLD aggressive mode
  - configure globally to avoid ID10T errors or misses
  - ... but it may not be of any use/benefit for 10GE
- If a network includes switches from mixed vendors, isolate the different STP domains with L3 routing to avoid STP compatibility issues.
- Keep L2, and thus STP, at the access layer
- Review the pros/cons of End-to-end vs Local VLANs to decide whether you should limit L2 to individual switches and thus eliminate STP entirely!

Spanning-Tree Best Practices (3)



From: Fig 4-32, FLG p. 178

Reminder

- *LOTS* of details do not appear in these slides
- You are responsible for *reading* the textbook to gain the knowledge (memorization) and understanding (apply the knowledge)