

11W NET3011 - Detailed Review of STP (updated Jan 30/11)

Points to Highlight:

1. Everything is based on BPDUs, which themselves are based on link costs
 - a. if you customize link costs in any way, it MUST be done consistently throughout
2. Sequence of events for STP convergence:
 - a. elect **Root Bridge**
 - perspective: entire topology (or entire VLAN for PVST+)
 - a.i. who has the lowest BID, overall?
 - a.ii. results in a single Root Bridge in the topology (or in each VLAN for PVST+)
 - b. on each *non-root* device, elect a **Root port**
 - perspective: individually from each device
 - b.i. which port gives me the lowest root path cost (i.e. the cost to each the Root Bridge)?
 - b.ii. if multiple paths have the same least cost, use the one with the lowest Port ID
 - b.iii. results in a single Root Port per non-root device
 - b.iv. the “best” BPDU (the one associated with the root port) is stored
 - c. for each segment (i.e. each collision domain), elect a **Designated port**
 - perspective: individually from each segment
 - c.i. which port gives me the lowest cost to the Root Bridge?
 - c.ii. if multiple ports have the same least cost, break the tie using first the lowest BID, then the lowest PID
 - c.iii. clearly, the segment itself has no ability to “decide” ... it’s really the devices attached to the segment that collectively decide, based on the BPDUs received on each attached port, compared to the BPDU it sends out that same port
 - c.iv. results in a single Designated Port per segment (the device owning that port is called the Designated Bridge for that segment)
 - c.v. unless multiple Root Bridge ports connect to the same segment (“looped”), all Root Bridge ports will be Designated as they all have the lowest root path cost
3. Timers – defaults are derived from a *maximum* radius of 7 switches (root + 6 more):
 - a. **Hello time** (default=2 secs): interval at which Config BPDUs are sourced from Root and propagated along all Designated segments
 - b. **Max Age** (default=20 secs): timeout for flushing “best” BPDU, to make way for learning a new “best” BPDU
 - c. **Forwarding Delay** (default=15 secs): time spent in Listening and Learning states
 - d. Timers are set at the Root Bridge and propagated throughout the topology as fields in the Conf BPDU
 - e. **Bridge Table Aging** (default=300 secs): amount of time before learned MAC addresses are aged out (while Topology Change bit is set in Config BPDUs, this is shortened to Forwarding Delay to age out MAC addresses faster)
4. The “Best BPDU” stored at a device is discarded when:
 - a. the device receives a “superior” BPDU
 - b. the port on which it was received goes down
 - c. the device times out waiting for a Config BPDU (MaxAge) – remember, these are supposed to be sourced from the Root at Hello intervals, and propagated down all designated ports
5. Topology Change Notification
 - a. is a smaller “lightweight” BPDU (only 4 octets following LLC header)
 - b. sent to root whenever a device detects a topology change – a link goes up or down

- b.i. a port goes down
 - b.ii. a port transitions to Forwarding or Blocking state
- c. if root path is down, no TCN can be sent and device must wait MaxAge before re-converging to determine a new root port (once new root port starts Forwarding, TCN is sent)
 - c.i. if device can directly detect root port failure, it doesn't need to wait MaxAge