Post 2

Thursday, January 24, 2019 11:18 PM

LAN REDUNDANCY PT 1

Multiple paths: Manage so L2 loops don't happen: Best Paths chosen → Alternates available in case of failure
STP: Spanning Tree Protocol: Manages L2 redundancy
First Hop Redundancy Protocols: Manages how clients are assigned to default GW **||** Uses alternate GW if primary fails
Redundancy at L1/L2: 3tier design (core/access/distribution) attempts to eliminate single point of failure

- Improves reliability/availability **|** Allows alternate physical paths for data: Despite disruption

L2 Logical Loops: Natural operation of switches: Learning/fwding process: Multiple paths exist/no STP implemented on switches
Considerations w/Implementing Redundancy

MAC Database Instability	**Copies of same frame are received on different ports of switch:** • Data fwding impaired when switch consumes resources coping w/instability in MAC table **Broadcast Ethernet frames don't have a TTL attribute:** • If no mechanism to block continued propagation of frames • They continue to propagate between switches endlessly • Or until link is disrupted/breaks loop • Occurs due to broadcast frames fwding
Broadcast Storms	**W/out a loop-avoidance process:** • Each switch may flood broadcasts endlessly **Broadcast frames fwded out all switch ports except ingress:** • Ensures all devices in broadcast domain able to receive frame • If more than 1 path to fwd: Endless loop can result • When loop occurs: Possible for MAC table on switch to constantly update from frames **A host caught in a loop is not accessible to other hosts:** • Due to the constant changes in MAC table: ○ Switch doesn't know out which port to fwd unicast frames **Broadcast Storm:** When so many frames caught in L2 loop that all BW is consumed: • No BW left for traffic: Network becomes unavailable for data comm • Effective DoS/Inevitable on looped network • As more devices send broadcasts: More traffic gets caught in loop: Causes failure **Broadcast traffic fwded out every port on switch:** • All connected devices must process all traffic being flooded endlessly • *Can cause end devices to malfunction*: *High traffic load on NIC* • Can happen in seconds b/c devices regularly send broadcasts (ARP reqs)
Multiple frame transmission	**Multiple copies of unicast frames may be delivered to destination stations:** • Many protocols expect to receive only 1 copy of each transmission • Multiple copies of same frame cause errors **Duplicate Unicast Frames:** • Unicasts sent onto looped network can result in duplicate frames arriving at dest device • Most upper-layer protocols: Not designed to recognize duplicates **Protocols that use sequence-numbering:** • Assume transmission failed/sequence # recycled for another comm session Other protocols: • Attempt to hand duplicate transmission to upper-layer protocol to be processed/discarded

	L2 LAN protocols (Ethernet): • Lacks mechanism to recognize/eliminate endlessly looping frames Some L3 protocols: Implement TTL that limits # of times L3 device can retransmit packet • L2 devices: Don't have this ability: They retransmit loops indefinitely • Spanning tree enabled by default on Cisco switches to prevent L2 loops

Spanning Tree Algorithm: Spanning Tree Protocol (STP): Developed to address these issues

STP: Based on an alg invented by Radia Perlman: While working for Digital Equipment Corp

- Published: 1985: "*An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN*"

Ensures only 1 logical path bet all destinations on network:

- Intentionally blocks redundant paths that cause loops
- Doesn't include **BPDU: Bridge Protocol Data Unit** frames used by STP to prevent loops
- Blocking redundant paths: Critical to preventing loops

Physical paths: Still exist to provide redundancy: Disabled to prevent loops

If path needed (cable/switch failure): STP recalculates paths/unblocks necessary ports

STP recalculation: Prevents loops from occurring by config loop-free path using strategically placed "**blocking-state**" ports

- Compensates for failures by dynamically unblocking previously blocked ports
- Permits traffic to alternate paths

RSTP: Rapid Spanning Tree Protocol

MSTP: Multiple Spanning Tree Protocol

Latest IEEE doc	**STP superseded by RSTP** **• IEEE: Original STP: 802.1D | RSTP: 802.1D-2004**

Port Roles: STP/RSTP use STA

STA: Spanning Tree Algorithm: Determines which switch ports must be put in blocking state to prevent loops

- Designates single switch as **root bridge**: Uses it as ref point for all path calcs
- **All switches in STP: Exchange BPDU frames:** Determines which switch has lowest **BID: Bridge ID** on network
- Switch w/lowest BID auto becomes root bridge for STA calcs

BPDU	Msging frame exchanged by switches for STP • Each BPDU contains BID that ID's switch that sent BPDU **BID contains:** **• Priority value** **• MAC address of sending switch** **• Optional extended system ID** Lowest BID value: Determined by combo of 3 fields

After root bridge determined: STA calcs shortest path to root bridge

- Each switch uses STA to determine which ports to block
- While STA determines best paths to root bridge for all ports in broadcast domain: Traffic not fwded
- STA considers BOTH path/port costs when determining which ports to block

Path costs	**Calc:** Uses port cost values associated w/port speeds for each switch port along given path **Sum of port cost values:** Determine overall cost to root bridge **If more than 1 path to choose:** STA chooses path w/lowest cost When STA determines which paths most desirable w/each switch: • Assigns port roles to participating switch ports

Port roles describe relation in network to root bridge/whether they are allowed to fwd traffic:

Root ports	**Switch ports closest to root bridge:** **• Selected on per-switch basis**
Designated ports	**All non-root ports that still permitted to fwd traffic:** **• Selected on per-trunk basis** • If 1 end of trunk root port: Other end designated port • All ports on root bridge are designated ports
Alternate/backup ports	**Config'd to be in blocking state to prevent loops:** **• Selected only on trunk links where neither end is root port**

	• Only 1 end of trunk blocked • Allows faster transition to fwding state when necessary • Blocking ports only come into play when 2 ports on same switch provide redundant links
Disabled ports	**Switch port that is shut down**

STA: Root Bridge

Root bridge: Serves as ref point for all spanning tree calcs to determine which redundant paths to block

Election process determines which switch becomes root bridge:

BID fields: Made up of: Priority value, Ext sys ID, MAC of switch

- All switches in broadcast domain participate in election process
- After switch boots: Begins to send out BPDU frames every 2 seconds

BPDUs contain:

Switch BID	Root ID

As switches fwd BPDU frames:

- Adjacent switches in broadcast domain read root ID info from BPDU frames
- If root ID from BPDU received is lower than on receiving switch: Receiving switch updates root ID
- ID adjacent switch as root bridge
- It may not be adjacent switch, but any other switch in broadcast domain
- Switch fwds new BPDU frames w/lower root ID to other adjacent switches
- Eventually: Switch w/lowest BID ends up being ID'd as root bridge for spanning tree instance

A root bridge is elected for each spanning tree instance

- Possible to have multiple distinct root bridges
- If all ports on all switches members of VLAN 1: Only 1 spanning tree instance
- Extended sys ID: Plays role in how spanning tree instances determined

STA: Path Cost

Link Speed	Cost (Revised IEEE Spec)	Cost (Previous IEEE Spec)
10 Gb/s	2	1
1 Gb/s	4	1
100 Mb/s	19	10
10 Mb/s	100	100

When root bridge elected for spanning tree instance:

- STA starts process of determining best paths to root bridge from all destinations in broadcast domain
- Path info determined by summing up individual port costs along path from destination to root bridge
- Each "destination" a switch port
- Default port costs defined by speed port operates at

Although switch ports have default port cost associated: Port cost is config

- Ability to config individual port cost: Gives flexibility to control spanning tree paths to root bridge

To config:

```
S2 (config-if)# spanning-tree cost <value>
S2 (config-if)# end
```

- Can be bet 1-200,000,000

To restore to default:

```
S2 (config-if)# no spanning-tree cost
S2 (config-if)# end
```

Path cost = Sum of all port costs along path to root bridge

- Paths w/lowest cost preferred: All other redundant paths blocked

To verify port/path cost to root bridge:

```
S2# show spanning-tree
```

- Cost field near top of output: Total path cost to root bridge
- Value changes depending on how many switch ports must be traversed to get to root bridge
- In output: Each int also ID'd w/individual port cost of 19

Port Role Decisions for RSTP

After switch determines which ports config in root port: Needs to know which ports for designated/alternate roles

- Root bridge auto configs all its ports in designated role
- Other switches config non-root ports as designated/alternates

Designated ports	**Config'd for ALL LAN segments** • When 2 switches connected to same LAN segment **AND** root ports already defined • 2 switches decide which port to config as designated/alternate **Switches on LAN segment exchange BPDU frames:** • Contain switch BID • Switch w/lower BID config designated **||** Higher BID config alternate 1st priority is lowest path cost to root bridge: **Sender's BID used only if port costs equal** • Each switch determines which roles assigned to each of its ports to create loop-free spanning tree

Designated/Alternate Ports

When determining root port on switch: Switch compares path costs on all switch pots in the spanning tree

- Switch port w/lowest overall path cost to root bridge auto assigned root port role b/c its closest to root bridge
- All non-root bridge switches have single root port chosen: Port provides lowest cost path back to root bridge

802.1D BPDU Frame Fmt

STA depends on exchange of BPDUs to determine root bridge

BPDU frame contains 12 distinct fields that convey the path/priority info used to determine root bridge/paths to it

1st 4 fields	**ID:** **• Protocol** **• Version** **• Msg type** **• Status flags**
Next 4 fields	**ID:** **• Root bridge** **• Cost of path to root bridge**
Last 4 fields	**All timer fields:** **• Determine how freq BPDU msgs sent** **• How long info received through BPDU process retained**

802.1D BPDU Propagation/Process

Each switch in broadcast domain initially assumes it's root bridge for spanning tree instance

- BPDU frames sent contain BID of local switch as root ID
- Default: Every 2 seconds: Value of Hello timer specified in frame
- Each switch maintains local info about its BID/root ID/path cost to root

When adjacent switches receive BPDU frame:

- Compare root ID from BPDU frame w/local root ID
- If root ID in BPDU is lower than local root ID: Switch updates local root ID/ID in its BPDU msgs
  - ○ Msgs indicate new root bridge on network
  - ○ Distance to root bridge is also indicated by path cost update

Example: If BPDU received on Fast Ethernet port: Path cost would increment by 19

- If local root ID lower than root ID received in BPDU frame: Frame is discarded
- After r/ID updated to ID new root bridge: All subsequent BPDU's sent from switch contain new root ID/path cost

All adjacent switches: Able to see lowest root ID all times

- As BPDU frames pass bet other adjacent switches: Path cost updated to indicate total path cost to root bridge
- Each switch in spanning tree uses its path costs to ID best possible path to root bridge

Extended System ID BID used to determine root bridge on network:

BID field of BPDU contains 3 separate fields:

- **Bridge priority**

- Extended system ID
- MAC address

Bridge Priority Customizable value that can be used to influence which switch becomes root bridge

- Switch w/lowest priority: Lowest BID: Becomes root bridge b/c precedence
- 0 takes precedence over all other bridge priorities

Extended System ID

When 2 switches are config w/same priority/have same extended system ID:

- Switch having MAC w/lowest hex value will have the lower BID
- Initially, all switches config'd w/same default priority value
- MAC is then deciding factor