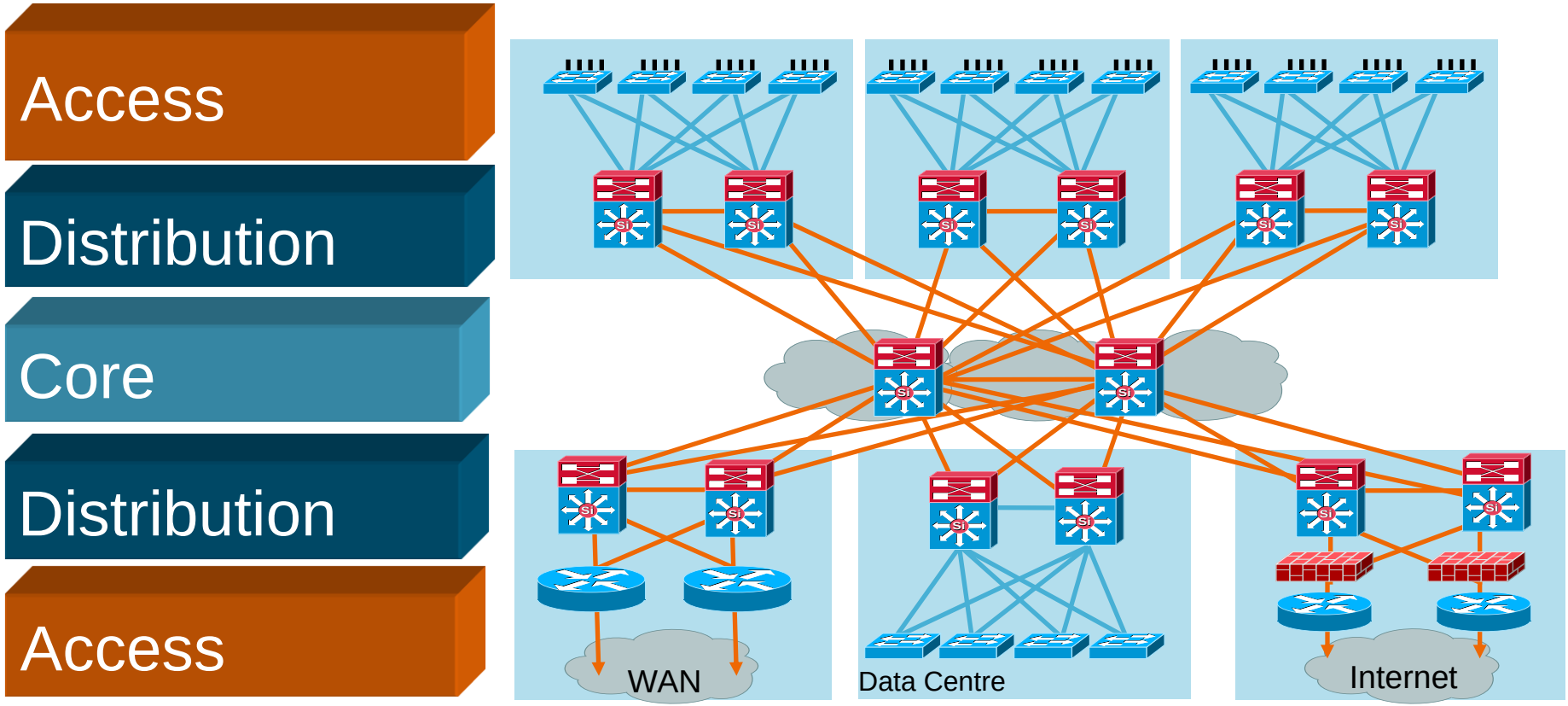
Troubleshooting LAN Switches and Protocols

BRKRST-2618

Session Agenda

- Trunking and Etherchannel
- VPC
- VSS
- Spanning Tree
- Unicast Flooding and MAC Flapping
- High CPU and Forwarding Issues
- QoS
- Q & A

Enterprise Composite Network Model



Troubleshooting Trunking and Etherchannel

Topics

- Dot1q Trunking
- VLAN Trunking Protocol
- Etherchannel

802.1Q Trunking

Operation Compatability Matrix

	Uses DTP	Forms Trunk with Off	Forms Trunk with Auto	Forms Trunk with Desirable	Forms Trunk with On	Forms Trunk with No Negotiate
off	No	No	No	No	No	No
auto	Yes	No	**No**	Yes	Yes	No
desirable	Yes	No	Yes	**Yes**	Yes	No
on	Yes	No	Yes	Yes	**Yes**	Yes
nonegotiate	No	No	No	No	Yes	Yes

802.1Q Trunking

Verify Configuration

- Verify your configuration
- Interface range command for consistent configuration!

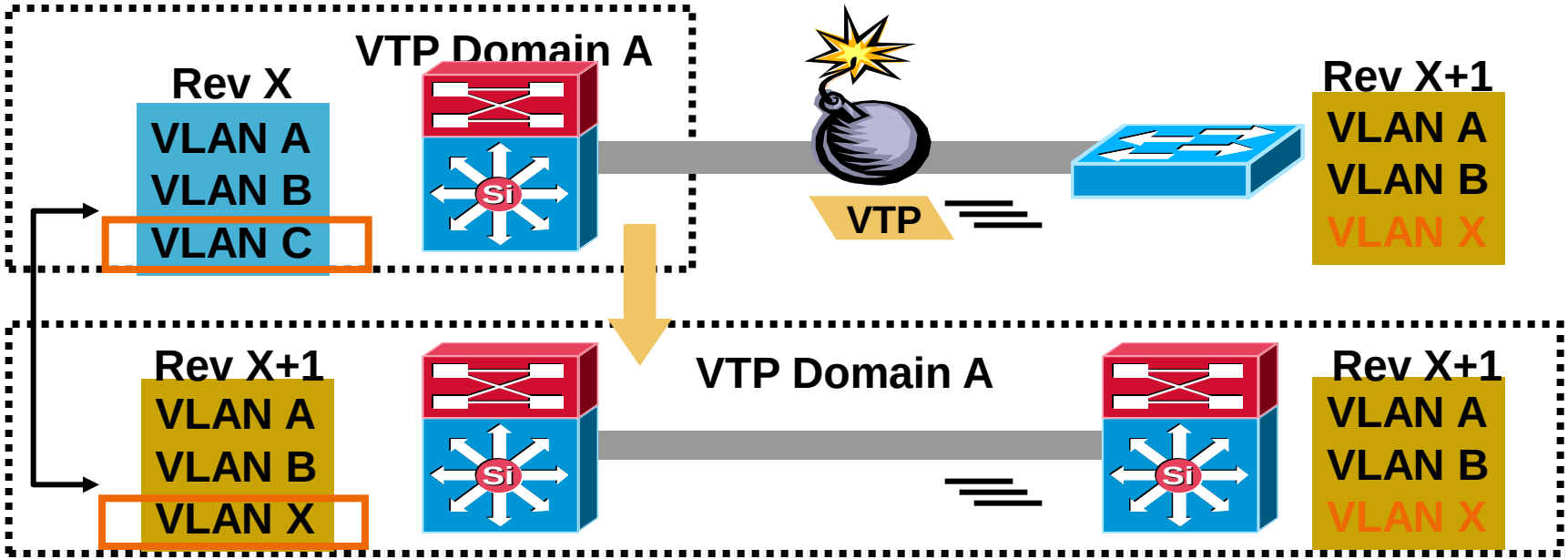
```
dist3750# show log | inc SPANTREE
%SPANTREE-7-RECV_1Q_NON_TRUNK: Received 802.1Q BPDU on non trunk
GigabitEthernet2/0/1 VLAN1.
%SPANTREE-7-BLOCK_PORT_TYPE: Blocking GigabitEthernet2/0/1 on VLAN0001.
Inconsistent port type.
```

```
vsscore# sh int trunk

Port        Mode     Encapsulation  Status    Native vlan
Gi1/8/3     on       802.1q         trunking  1
Gi1/8/4     on       802.1q         trunking  1
```

```
dist3750# sh int trunk

Port        Mode     Encapsulation  Status      Native vlan
? ? ?       ???
Gi2/0/2     on       802.1q         trunking    1
```

VTP – Wheres my Vlans!!

- VTP Server with high Configuration revision takes precedence.
- TIP: Always put a newly added switch into Transparent mode to erase the configuration revision!



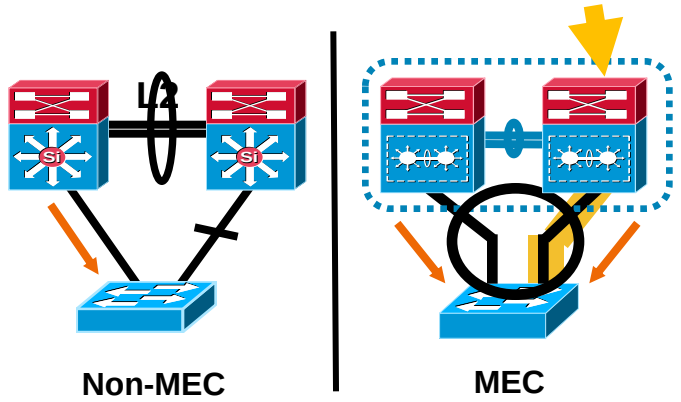
Etherchannel

Types of Etherchannel

- Etherchannel ports on the same module
- Distributed Etherchannel (DEC) ports using different modules on same switch, e.g. 1/1, 2/1 and 3/1
- Multichassis Etherchannel (MEC)

  Extending link aggregation to two separate physical switches

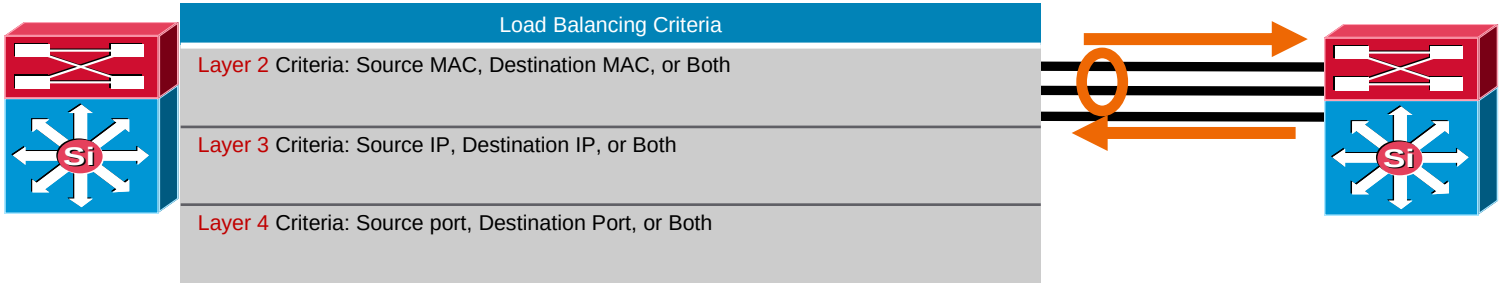
  VSS appears as single logical device
- Virtual Port-channel (vPC) – Two physical switches bonding one etherchannel.



Etherchannel

"channel interfaces are not load-balanced correctly"

```
6500# sh int gi1/37 | i rate
5 minute input rate 2760000 bits/sec, 556 packets/sec
  5 minute output rate 1265000 bits/sec, 1295 packets/sec
6500# sh int gi1/38 | i rate
  5 minute input rate 46000 bits/sec, 30 packets/sec
  5 minute output rate 641000 bits/sec, 408 packets/sec
6500# sh int gi1/40 | i rate
  5 minute input rate 148000 bits/sec, 40 packets/sec
  5 minute output rate 320000 bits/sec, 225 packets/sec
```

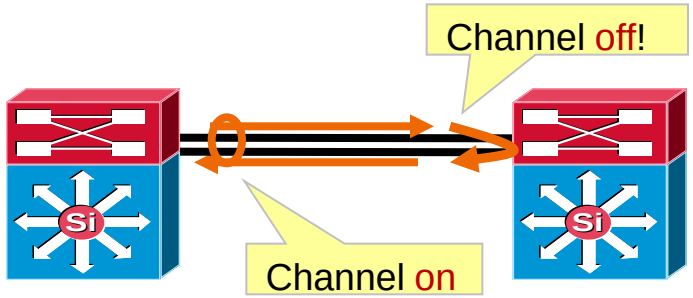


- Crucial to understand traffic profile
- L4 tends to achieve symmetry
- L3 Dst – not good when all clients are trying to access one link.

Etherchannel

Negotiation

Mode	Uses PAgP or LACP	Forms Channel with Off	Forms Channel with Auto	Forms Channel with Desirable	Forms Channel with On
On	No	No	No	No	Yes
Auto (Passive)	Yes	No	**No**	Yes	No
Desirable (Active)	Yes	No	Yes	**Yes**	No



- PAgP (Cisco) : Desirable – Desirable
- LACP (IEEE 802.3ad): Active – Active
- Prevent loop due to misconfig !

Etherchannel

Which link will be used?

- Hidden command in 12.2(18)SXF and 12.2(33)SXH

```
6500-SXF# remote command switch test EtherChannel load-balance
interface po 1 ip 1.1.1.1 2.2.2.2
Computed RBH: 0x5 => Would select Gi4/1 of Po1
```

- L3 - Look for any patterns – e.g. specific hosts...

show command in 12.2(33)SXI1

```
vsscore# show etherchannel load-balance hash-result interface port-channel 200 switch 1
   ip 192.168.1.1 192.168.1.36
Computed RBH: 0x4 => Would select Gi1/8/1 of Po200

vsscore# show etherchannel load-balance hash-result interface port-channel 200 switch 1
   ip 192.168.1.1 192.168.1.37
Computed RBH: 0x5 => Would select Gi1/8/2 of Po200
```

- L2 - MAC address load-balancing

```
3750# test etherchannel load-balance interface port-channel 1 mac 0012.4358.f080
001a.e281.2d06 => Would select Gi3/0/6 of Po1
```

Etherchannel - Continue

- Nexus 5000 / 7000 equivalent

```
nexus# show port-channel load-balance forwarding-path interface port-channel 25 src-ip 1.1.1.1 dst-ip 2.2.2.2 vlan 2

Module 2: Load-balance Algorithm: source-dest-ip-vlan

RBH: 0x6      Outgoing port id: Ethernet2/2
```

- Show command - load balancer

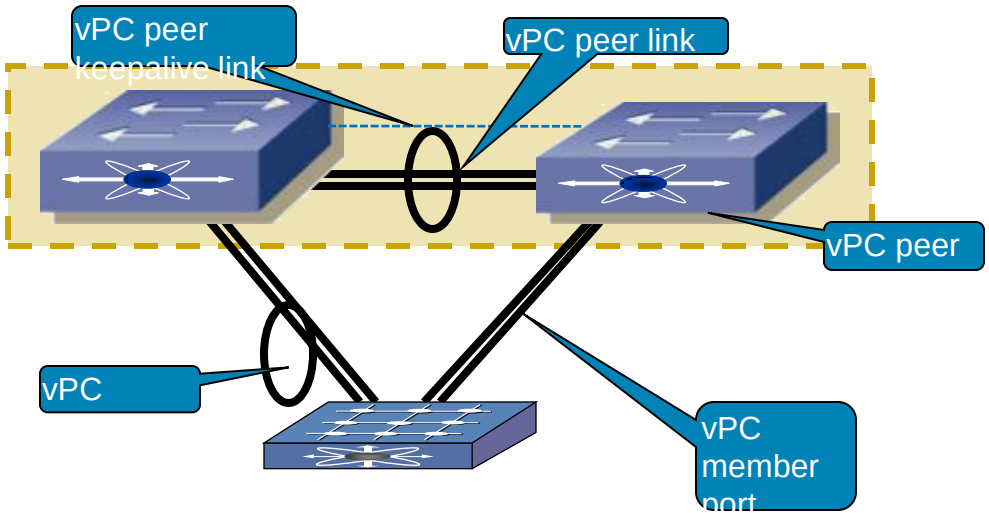
```
nexus# show port-channel load-balance
Port Channel Load-Balancing Addresses Used Per-Protocol:
Non-IP: source-dest-mac
IP: source-dest-ip-vlan
```

Troubleshooting VPC (Nexus)

Topics

- VPC Topologies
- Failure Symptom
- Show commands

Virtual Port-Channel
Terminology – Nexus 7000/5000

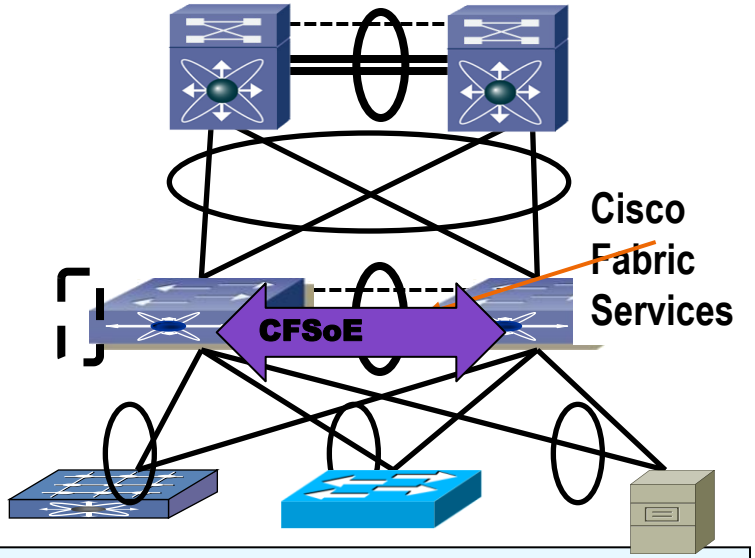


VPC Information

- VPC Peer
  - The remote Nexus switch
- vPC Member port
  - Channel member formed with its vPC peer.
- vPC Peer Link
  - Inter switch link which sends CFS messages.
- CFS
  - Stands for Cisco Fabric Services – Carries MAC DB for syncronisation.

Virtual Port-Channel
vPC Control Fabric – Cisco Fabric Services

- Cisco Fabric Services provides the control plane synchronisation between vPC peers
  - Configuration validation/comparison
  - MAC member port synchronisation
  - vPC member port status
  - IGMP snooping synchronisation
  - vPC status
- Highly Reliable - Inherited from MDS
- CFS messages are encapsulated in standard Ethernet frames (with CoS 6)

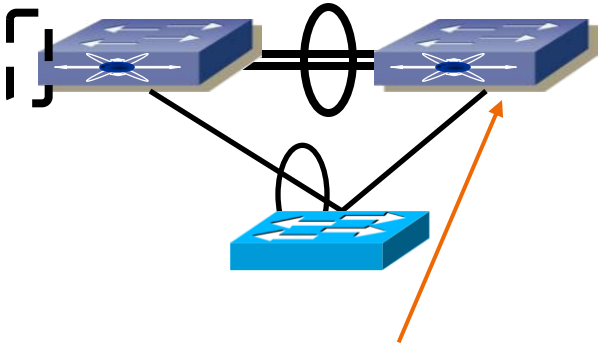



```
dc11-5020-2# show cfs status
Distribution : Enabled
Distribution over IP : Disabled
IPv4 multicast address : 239.255.70.83
IPv6 multicast address : ff15::efff:4653
Distribution over Ethernet : Enabled
```

Virtual Port-Channel

vPC Control Plane – Type 1 Consistency Check

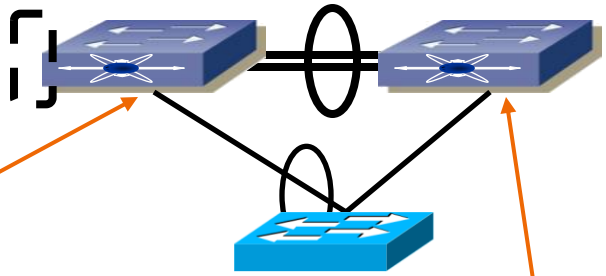
- Type 1 Consistency Checks are intended to prevent network failures
  - Incorrectly forwarding of traffic
  - Physical network incompatibilities
- vPC will be suspended



```
dc11-5020-2# show vpc brief
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
<snip>
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
201    Po201       up     failed      vPC type-1 configuration   -
                                      incompatible - STP
                                      interface port guard -
                                      Root or loop guard
                                      inconsistent
```

Virtual Port-Channel
vPC Control Plane – Type 2 Consistency Check

- Type 2 Consistency Checks are intended to prevent undesired forwarding
- vPC will be modified in certain cases (e.g. VLAN mismatch)



```
dc11-5020-1# sh run int po 201

interface port-channel201
switchport trunk allowed vlan 100-105
  vpc 201
  spanning-tree port type network
```

```
dc11-5020-2# sh run int po 201

interface port-channel201
switchport trunk allowed vlan 100-104
  vpc 201
  spanning-tree port type network
```

```
dc11-5020-1# show log

2009 May 17 21:56:28 dc11-5020-1 %ETHPORT-5-IF_ERROR_VLANS_SUSPENDED: VLANs 105 on
Interface port-channel201 are being suspended. (Reason: Vlan is not configured on remote
vPC interface)
```

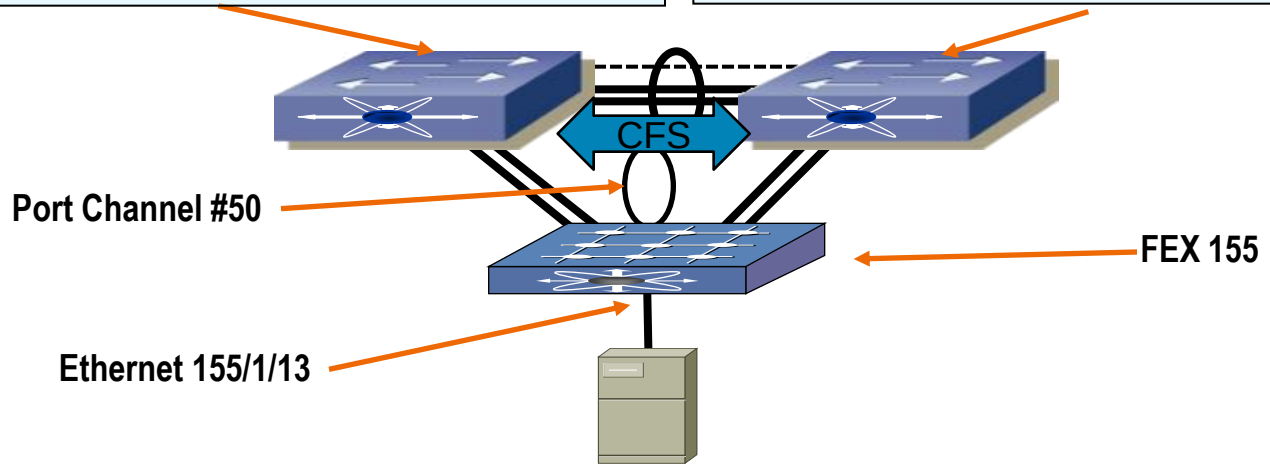
Nexus Virtualised Access Switch

Nexus 2000 vPC Host Ports

- A port on a dual homed Nexus 2000 is known as a vPC Host Port

```
dc11-5020-3# sh vpc
<snip>
vPC status
----------------------------------------------------------------
id     Port        Status Consistency Reason    Active vlans
------ ----------- ------ ----------- --------- -----------
<snip>
157708 Eth155/1/13 up     success     success      105
```

```
dc11-5020-4# sh vpc
<snip>
vPC status
----------------------------------------------------------------
id     Port        Status Consistency Reason    Active vlans
------ ----------- ------ ----------- --------- -----------
<snip>
157708 Eth155/1/13 up     success     success      105
```

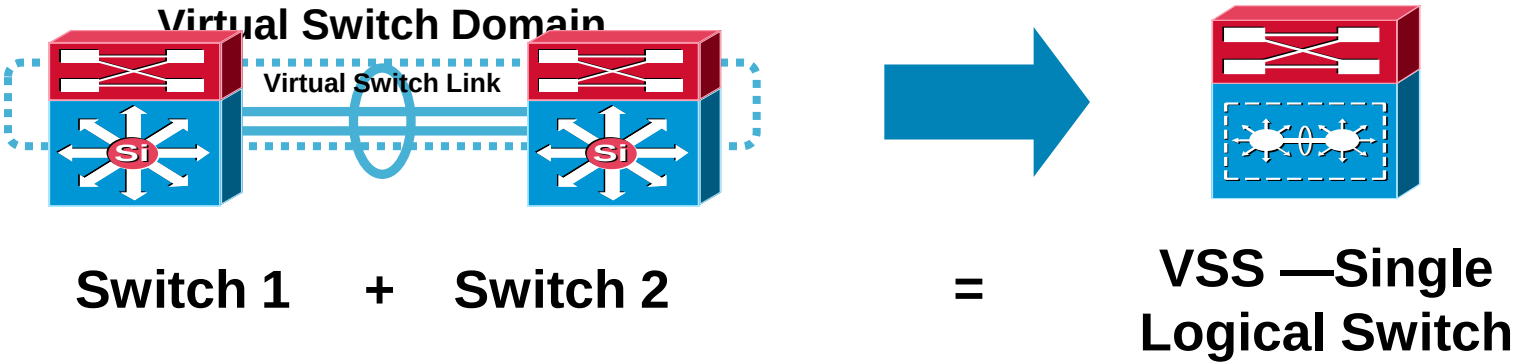


Troubleshooting VSS (6500)

VSS (Virtual Switching System)

Introduction

- Virtual Switching System consists of two Cisco Catalyst 6500 Series defined as members of the same virtual switch domain
- Single control plane with dual active forwarding planes
- Design to increase forwarding capacity while increasing availability by eliminating STP loops
- Reduced operational complexity by simplifying configuration



VSS

Verifying Redundancy Status

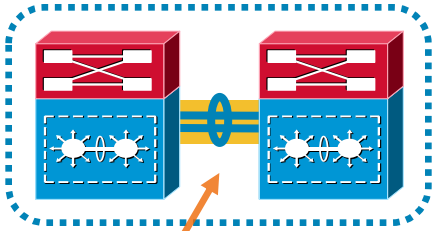
```
vsscore# sh switch virtual
Switch mode                    : Virtual Switch
Local switch operational role: Virtual Switch Active
Peer switch operational role : Virtual Switch Standby
...
vsscore# sh red states
       my state = 13 -ACTIVE
     peer state = 8  -STANDBY HOT
Redundancy Mode (Operational) = sso

vsscore# sh switch virtual link
VSL Status : UP
VSL Uptime : 1 day, 21 hours, 25 minutes
VSL SCP Ping : Pass
VSL ICC Ping : Pass
VSL Control Link : Te2/6/5
```

2

3

1

switch/slot/port



VSL Initialization

- Link Bringup
  - which ports form the VSL
- Link Management Protocol (LMP)
  - track and reject uni-dir links
  - exchange info such as chassis ID
- Role Resolution Protocol (RRP)
  - determine compatible hw/sw versions
  - Active/Standby role

VSS

Verifying VSL Status

```
*20:37:09: %VSLP-SW2_SP-3-VSLP_LMP_FAIL_REASON: Te2/6/5: Link down
*20:37:09: %VSLP-SW1_SPSTBY-3-VSLP_LMP_FAIL_REASON: Te1/6/5: Link down
```

1

```
vsscore# sh switch virtual link port
Peer Peer           Peer    Peer       Timer(s)running
Interface Flag State         Flag MAC          Switch Interface (Time remaining)
-------------------------------------------------------------------------------
Te2/6/4   vfsp operational   vfsp 0003.6c56.4800 1    Te1/6/4   T4(308ms)
                                                                 T5(59.98s)
Te2/6/5   v    link_down     -    -                    -              -
```

2

```
vsscore# ping vslp output interface ten2/6/5 count 10
% TenGigabitEthernet2/6/5 is down

vsscore# ping vslp output interface ten2/6/4 count 10
Sending 10, 100-byte VSLP ping to peer-sup via output port 2/6/4,
timeout is 2 seconds:
!!!!!!!!!!
```

3

Verify connectivity across VSL link

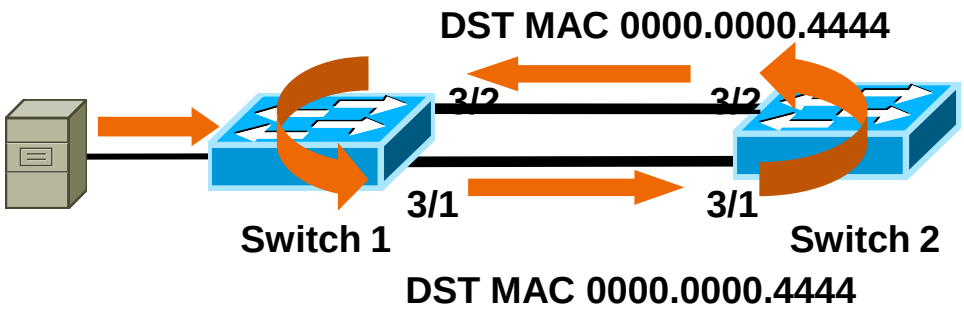
Spanning-tree

Topics

- Layer 2 Loops
- STP Standards and Features
- Troubleshooting STP
- Debugging STP
- MST Regions
- STP Problem!

Characteristics of Layer 2 Loops

- L2 has no native mechanism to recover
  - –IP has TTL
  - –Layer 2 has nothing!



- Symptoms include
  - –High link utilisation
  - –High CPU utilisation
  - –MAC flapping/Duplicate HSRP log messages

Spanning-Tree Instability Methodology

1. Topology: Know the spanning-tree topology of the network and the location of Root Switch – root ports and blocked ports

2. Syslog: Rely on syslog (spanning-tree, loopguard, dispute ,…) to find a starting point of investigation

3. Expected Behaviour: Understand BPDU flow
   - Normal BPDU flow is Designated port → (correct ports sending bpdu's)

4. show spanning-tree … [det]: look for
   - TCN
   - BPDU flowing upstream (TX by supposed Root or Blocked ports)
   - Port role flapping

5. Debug: use debug when you have isolated to the device in question.

Troubleshooting STP

Logging what is going on in your network

```
interface GigabitEthernet1/8/3
...
 logging event spanning-tree status
 logging event link-status
 logging event trunk-status
...
```

- Getting better visibility in your network.
- 1 link state change can cause sequential set system messages.

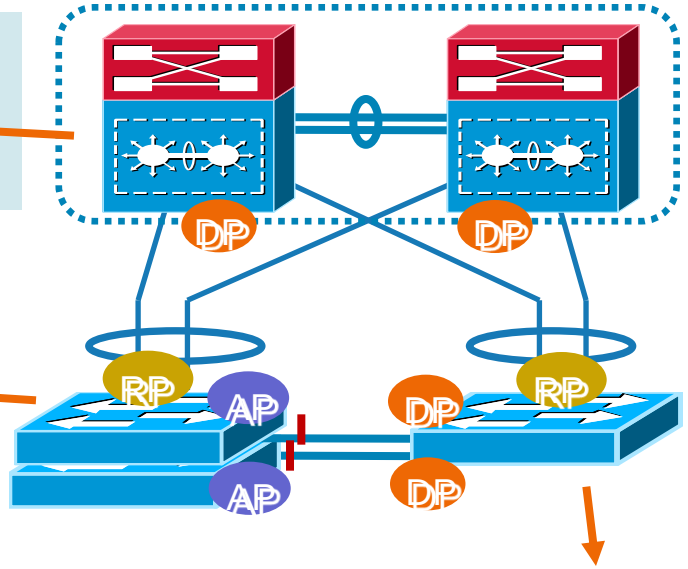
```
18:26:52: %SPANTREE-SW1_SP-6-PORT_STATE: Port Po1 instance 0 moving from forwarding to
disabled
...
18:26:52: %LINK-SW1_SP-3-UPDOWN: Interface Port-channel1, changed state to down
18:26:52: %LINK-SW1_SP-3-UPDOWN: Interface GigabitEthernet1/8/3, changed state to down
18:26:52: %DTP-SW1_SP-5-NONTRUNKPORTON: Port Gi1/8/3 has become non-trunk
18:26:52: %DTP-SW2_SPSTBY-5-NONTRUNKPORTON: Port Gi1/8/3 has become non-trunk
18:26:53: %STANDBY-3-DUPADDR: Duplicate address 10.25.33.3 on Vlan9, sourced by 0019.a95d.9c00
18:26:53: %STANDBY-3-DUPADDR: Duplicate address 10.25.33.3 on Vlan9, sourced by 0019.a95d.9c00
```

Spanning Tree

Know your port states in a stable environment…

```
vsscore# sh spanning-tree vlan 99
Interface  Role Sts Cost      Prio.Nbr Type
---------- ---- --- --------- -------- -----
Po2        Desg FWD 10000     128.5763 P2p
Po1        Desg FWD 10000     128.5764 P2p
```

```
dist3750# sh spanning-tree vlan 99
Interface  Role Sts Cost      Prio.Nbr Type
---------- -------- -------- ------------
Fa2/0/1    Altn BLK 200000    128.57   P2p
Po1        Root FWD 20000     128.488  P2p
Fa3/0/1    Altn BLK 200000    128.111  P2p
```



- Root forwarding port for ST topology
- Designated forwarding port for LAN segment
- Alternate blocking alternate path to root bridge
- Backup blocking redundant path to a bridge segment

```
dist4500# sh spanning-tree vlan 99
Interface  Role Sts Cost      Prio.Nbr Type
---------- ---- --- --------- -------- ----
Fa4/1      Desg FWD 200000    128.193  P2p
Fa4/2      Desg FWD 200000    128.194  P2p
Po2        Root FWD 6660      128.642  P2p
```

Understanding the STP Process Output

What does it tell me?

```
dist3750# sh spanning-tree vlan 99 de

 MST0 is executing the mstp compatible Spanning Tree protocol
...
Current root has priority 32768, address 0003.6c56.4800
  Root port is 488 (Port-channel1), cost of root path is 0
  Topology change flag not set, detected flag not set
  Number of topology changes 429 last change occurred 00:00:39 ago
          from Port-channel1
...

 Port 57 (FastEthernet2/0/1) of MST0 is alternate blocking
...
   Number of transitions to forwarding state: 8
   BPDU: sent 290, received 27469

 Port 488 (Port-channel1) of MST0 is root forwarding
...
   Number of transitions to forwarding state: 1
   BPDU: sent 498, received 110725
```

MST in use

TCN notification

port state

BPDU count

Understanding the STP Process Ouput

Very useful shortcut command

```
dist3750# show spanning-tree vlan 99 detail | inc (MST|BPDU)

 MST0 is executing the mstp compatible Spanning Tree protocol
 Port 57 (FastEthernet2/0/1) of MST0 is alternate blocking
   BPDU: sent 0, received 27
 Port 59 (FastEthernet2/0/3) of MST0 is designated forwarding
   BPDU: sent 26, received 0
 Port 488 (Port-channel1) of MST0 is root forwarding
   BPDU: sent 0, received 25
```

- Taken approx 10 seconds later

▪RSTP/MST when stable sends BPDUs on designated ports
▪All switches send BPDUs (as opposed to 802.1d)

```
dist3750# show spanning-tree vlan 99 detail | inc (MST|BPDU)
...
 Port 59 (FastEthernet2/0/3) of MST0 is designated forwarding
   BPDU: sent 30, received 0
```

```
dist3750# clear spanning-tree counters
```

Very useful!

Spanning Tree – Event History

Nexus 5000 / 7000

- Being able view the history of every port role change

```
Nexus# show spanning-tree internal event-history tree 25 interface port-channel 1
VDC01 VLAN0025 <port-channel1>

0) Transition at 477482 usecs after Mon Feb 21 11:53:27 2011
     State: BLK  Role: Root  Age:  0 Inc: no  [STP_PORT_STATE_CHANGE]

1) Transition at 478062 usecs after Mon Feb 21 11:53:27 2011
     State: BLK  Role: Desg  Age:  0 Inc: no  [STP_PORT_ROLE_CHANGE]

2) Transition at 445194 usecs after Mon Feb 21 11:53:28 2011
     State: BLK  Role: Root  Age:  1 Inc: no  [STP_PORT_ROLE_CHANGE]

3) Transition at 445543 usecs after Mon Feb 21 11:53:28 2011
     State: FWD  Role: Root  Age:  1 Inc: no  [STP_PORT_STATE_CHANGE]
```

Troubleshooting Topology Changes (TC)

TC Principle

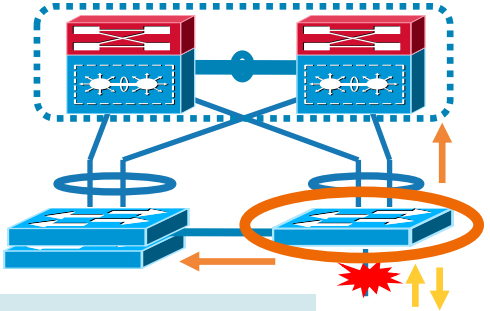
- TC on link moving to forwarding only
- Sent out by initiator (not by root)
- Propagated along active topology
- Uses TC bit in BPDU set for 2 x hello_time
- Flushes CAM immediately

```
dist3750# show spanning-tree vlan 99 de

 MST0 is executing the mstp compatible Spanning Tree protocol
. . .
Root port is 488 (Port-channel1), cost of root path is 0
  Topology change flag not set, detected flag not set
  Number of topology changes 96 last change occurred 00:11:19 ago
          from Port-channel1
  Times:  hold 1, topology change 35, notification 2
. . .
```

Troubleshooting TC

TC Example

- Example flow of TC through the network



```
dist4500# sh spanning-tree vlan 99 de | inc (Port 642|BPDU)
Port 642 (Port-channel2) of MST0 is root forwarding
   BPDU: sent 0, received 79

dist4500# sh spanning-tree vlan 99 de | inc (Port 642|BPDU)
Port 642 (Port-channel2) of MST0 is root forwarding
   BPDU: sent 2, received 83
```

```
vsscore-sp# debug spanning-tree mstp tc
MSTP Topology Change Notifications debugging is on
Aug 27 10:21:00: SW2_SP: MST[0]: port Po2 received internal tc
Aug 27 10:21:02: SW2_SP: MST[0]: port Po2 received internal tc
Aug 27 10:21:35: SW2_SP: MST[0]: tc timer expired
```

```
dist3750# debug spanning-tree mstp tc
Aug 27 10:20:59: MST[0]: port Fa2/0/1 received internal tc
Aug 27 10:21:01: MST[0]: port Fa2/0/1 received internal tc
Aug 27 10:21:34: MST[0]: tc timer expired
```

Troubleshooting Topology Changes (TC)

TC Troubleshooting Steps

- RTSP Topology Change Detection
  - starts “TC while timer”
  - initiator floods TC information
  - clears mac address table (potential for flooding)
- Remember “portfast” (edge port) on host ports

```
dist4500#show spanning-tree vlan 99
. . .
Fa5/1            Desg FWD 200000    128.257  P2p Edge
```

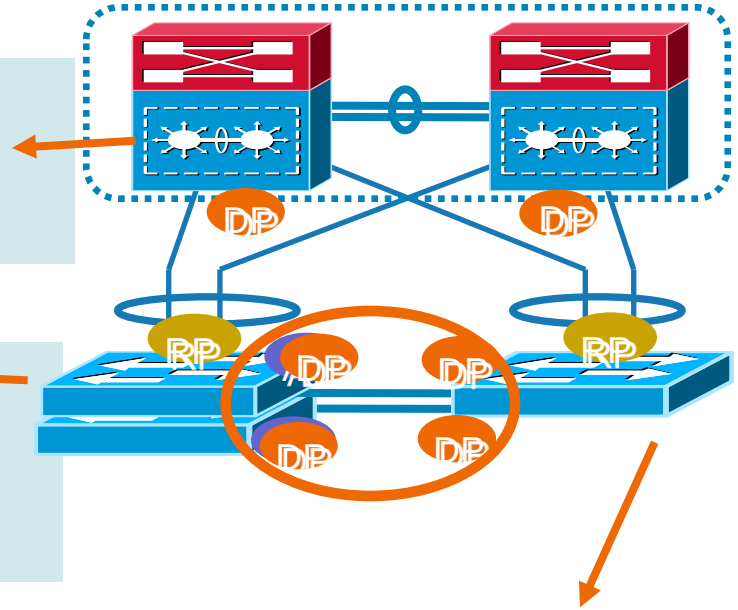
- Track the source of the TC
  - start from the root
  - “show spanning-tree vlan” -> Topology Change
  - work downstream towards the “initiator”
  - use “sh cdp neighbors” to help you

Spanning Tree

Do those ports states look correct?

```
vsscore# sh spanning-tree vlan 99
Interface  Role Sts Cost      Prio.Nbr Type
---------- ---- --- --------- -------- ------
Po2        Desg FWD 10000     128.5763 P2p
Po1        Desg FWD 10000     128.5764 P2p
```

```
dist3750# sh spanning-tree vlan 99
Interface  Role Sts Cost     Prio.Nbr Type
---------- ---------- -------- ---------------
Fa2/0/1    Desg BLK 200000  128.57   P2p
Po1        Root FWD 20000   128.488  P2p
Fa3/0/1    Desg BLK 200000  128.111  P2p
```



- What changed?
- What should the port states be?
- What are the port states now?

```
dist4500# sh spanning-tree vlan 99
Interface  Role Sts Cost      Prio.Nbr Type
---------- ---- --- --------- -------- ----
Fa4/1      Desg FWD 200000    128.193  P2p
Fa4/2      Desg LRN 200000    128.194  P2p
Po2        Desg FWD 6660      128.642  P2p
```

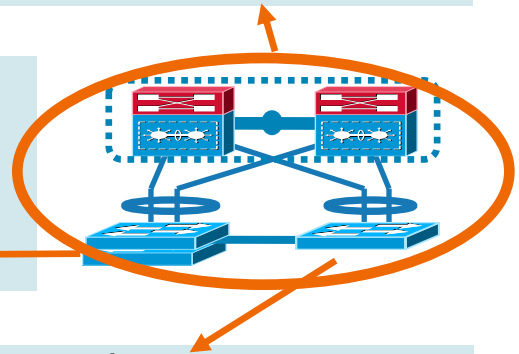
Debugging spanning-tree

How to see a STP snapshot of your network

```
vsscore-sp# debug spanning-tree events snapshot
Spanning Tree snapshot debugging is on

Aug 10 13:32:49: SW1_SP: MST[0]: snapshot: Po2->Desg.FWD Po1->Desg.FWD
Aug 10 13:32:58: SW1_SP: MST[0]: snapshot: Po2->Desg.BLK Po1->Desg.FWD
Aug 10 13:32:59: SW1_SP: MST[0]: snapshot: Po2->Desg.FWD Po1->Desg.FWD
```

Clue….Po2 port role is changing state frequently..

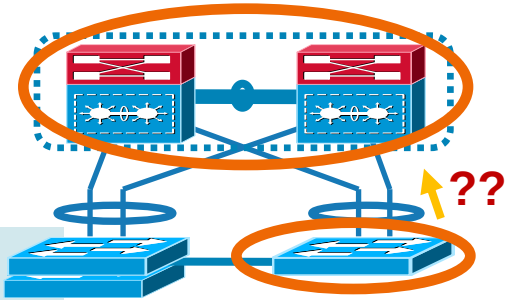
```
dist3750# debug spanning-tree events snapshot
Spanning Tree snapshot debugging is on
Aug 10 01:32:54: MST[0]: snapshot: Fa2/0/1->Desg.BLK
Aug 10 01:32:56: MST[0]: snapshot: Fa2/0/1->Desg.FWD
Aug 10 01:32:56: MST[0]: snapshot: Fa2/0/1->Altn.BLK
Aug 10 01:33:09: MST[0]: snapshot: Fa2/0/1->Desg.BLK
```



```
dist4500#debug spanning-tree events snapshot
Spanning Tree snapshot debugging is on

Aug 10 13:23:56: MST[0]: snapshot:
Aug 10 13:23:56: MST[0]: snapshot:
```

Debugging Spanning Tree

Confirming the BPDU flow




```
vsscore-sp# debug spanning-tree bpdu receive
Spanning Tree BPDU Received debugging is on
vsscore-sp# debug condition interface port-channel 2
Condition 1 set
```

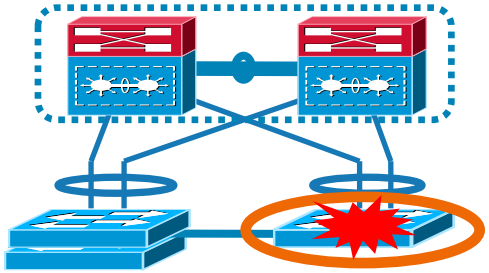
- Why are seeing BPDUs on what should be the designated port?
- What could be the reasons?

```
Aug 30 11:03:06: SW2_SP: STP: MST0 rx BPDU: config protocol = mstp, packet from Port-channel2
, linktype IEEE_SPANNING , enctype 2,

Aug 30 11:03:08: SW2_SP: STP: MST0 rx BPDU: config protocol = mstp, packet from Port-channel2
, linktype IEEE_SPANNING , enctype 2,
```

Debugging Spanning Tree

Reasons for the disruption of BPDU flow

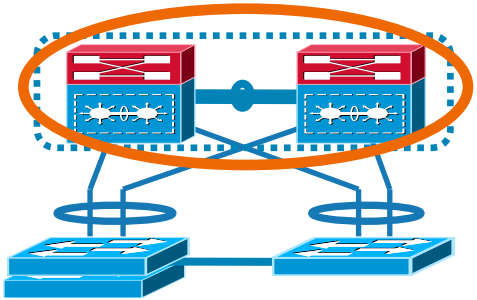
- Check the adjacent neighbour
- What disrupt the flow of BPDUs?
  - High link utilisation (input/output drops)
  - Interface Errors
  - Uni-directional link (faulty cabling/SFP issue)
  - High CPU



```
dist4500# sh proc cpu sorted
CPU utilization for five seconds: 99%/1%; one minute: 99%; five minutes: 98%
 PID Runtime(ms)   Invoked      uSecs   5Sec   1Min   5Min TTY Process
  43    19627192  15249915       1287 84.91% 85.10% 84.08%   0 Cat4k Mgmt LoPri
  42    20219100  59800788        338  8.45%  8.81%  8.88%   0 Cat4k Mgmt HiPri
  84      483372    460016       1050  3.59%  3.42%  3.37%   0 IP Input
```

MST Regions

MST Principles



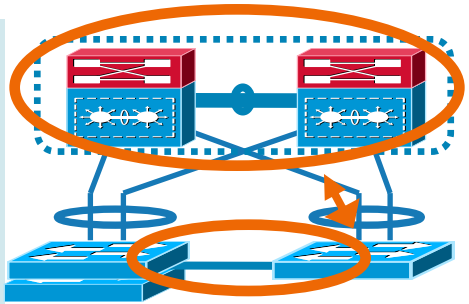
- Maps x # of vlans to a single STP instance
- Why use Regions?
  - Different Administrator control
  - Not all switches may run MSTP
- Configuration of Consistency
  - Switches must have the SAME configuration
  - Uses digest sent in BPDU.

```
vsscore# sh spanning-tree mst configuration digest
Name      [NETWORKERS]
Revision  0     Instances configured 2
Digest          0xF585D2E4EE371D9AC35F9DB6D3BAD9A8
Pre-std Digest  0x492BEF4B417C2B862888BDF23DDF4ADA
```

MST Regions

Mismatched Configuration Example

```
vsscore# sh spanning-tree mst configuration
Name      [NETWORKERS]
Revision  0     Instances configured 2
Instance  Vlans mapped
--------  ------------------------------------
0         1-89,100-4094
9         90-99
-------------------------------------------------
vsscore# sh spanning-tree mst configuration digest
Name      [NETWORKERS]
Digest            0xF585D2E4EE371D9AC35F9DB6D3BAD9A8
```



Non root switch will see Bound(RSTP)

- Check your config!
- Any pruned vlans?
- Boundary port means legacy switch or different region
- boundary port flap will trigger a TC in every instance as CIST flaps!

```
dist4500# sh spanning-tree mst configuration
Name      []
Revision  0     Instances configured 1
Instance  Vlans mapped
--------  ------------------------------------
0         1-4094
-------------------------------------------------
dist4500# sh spanning-tree mst configuration digest
Name      []
Digest            0xAC36177F50283CD4B83821D8AB26DE62
```

Spanning Tree Protocol

Troubleshooting summary

- Logging – Getting a sequential set of system messages.
- Understanding your Spanning-tree topology (bpdu flow)
- Network Diagram (that also reflects STP)
- CDP – Is your friend when chasing the source of TC
- Avoid Boundary ports as much as possible
- Loop – Always start off at Root bridge and work your way down from Core to distribution.

Spanning Tree Hardening and Why!

Topics

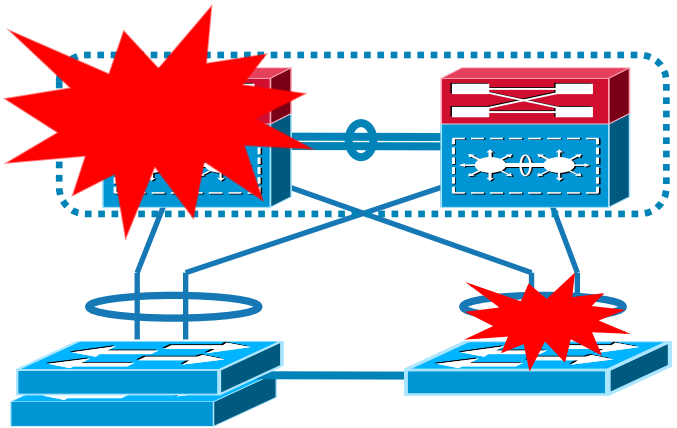
- Protecting Spanning Tree
- UDLD
- Loop Guard
- Dispute
- Dead Brain Switch
- Bridge Assurance

Protecting Spanning Tree

New and Established

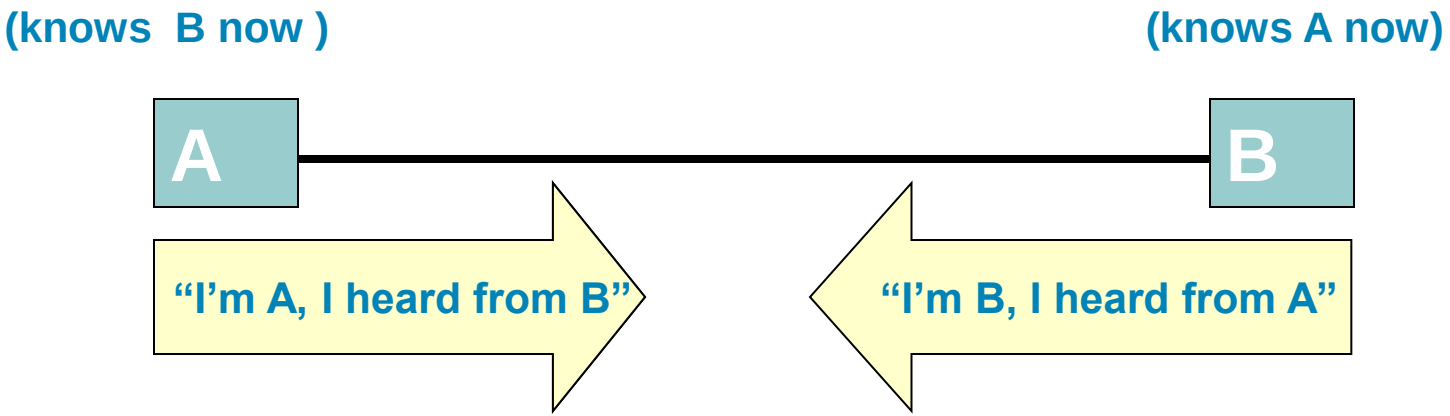
- Exisiting STP 802.1D stability methods are still recommended such as PortFast, RootGuard and BPDU Guard

- UDLD echo based protocol to detect link problems
- LoopGuard prevents alternate or root port from becoming designated in absence of BPDUs
- Dispute similar to LoopGuard but now implemented into MST and RSTP IEEE standard
- Bridge Assurance is a Cisco enhancement to STP similar to combined UDLD and loop guard functionality



How UDLD Builds Bidirectional Link Status

Bidirectional status is achieved

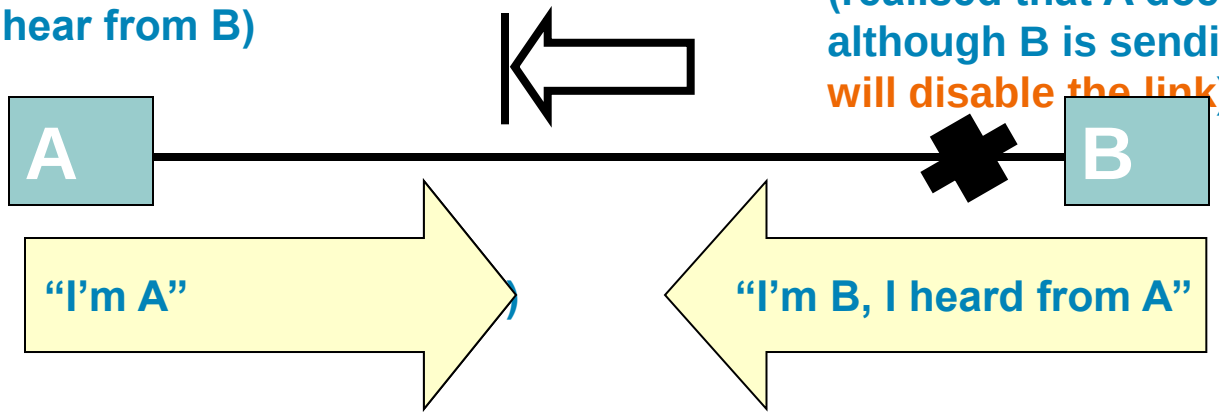


Regular UDLD

Unidirectional link –
Errdisabled by regular UDLD
When Transition from Any state (unknown or Bidir)
to UniDir

(didn't hear from B)

(realised that A doesn't know B although B is sending hellos will disable the link)

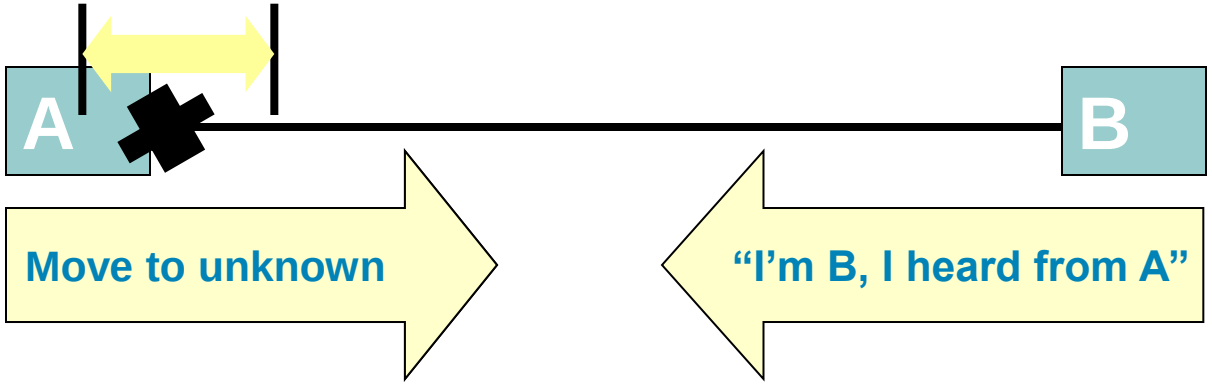


Aggressive UDLD

Bidir link moving to unknown
Errdisabled by Aggressive UDLD

(didn't hear anything for 3 hello)

Port stuck: no traffic comes in/out



UDLD

Debugging UDLD Issues

```
vsscore-sp# debug udld events
UDLD events debugging is on
vsscore-sp# debug udld packets
UDLD packets debugging is on
```

```
Aug 11 08:11:04: SW1_SP: Checking if link is bidirectional (Gi1/8/1)
Aug 11 08:11:04: SW1_SP: Found my own ID pair in 2way conn list (Gi1/8/1)
Aug 11 08:11:04: SW1_SP: Checking if multiple neighbors (Gi1/8/1)
Aug 11 08:11:04: SW1_SP: Single neighbor detected (Gi1/8/1)
Aug 11 08:11:04: SW1_SP: Checking if link is bidirectional (Gi1/8/1)
Aug 11 08:11:04: SW1_SP: Found my own ID pair in 2way conn list (Gi1/8/1)
```

```
Aug 11 08:14:26: SW1_SP: Zero IDs in 2way conn list (Gi1/8/1)
...
Aug 11 08:14:26: SW1_SP: Udld receive packet *END*.  (Gi1/8/1)
...
Aug 11 08:14:27: %UDLD-SW1_SP-4-UDLD_PORT_DISABLED: UDLD disabled interface Gi1/8/1,
unidirectional link detected
Aug 11 08:14:27: %PM-SW1_SP-4-ERR_DISABLE: udld error detected on Gi1/8/1, putting Gi1/8/1
in err-disable state
```

```
vsscore# sh int status err-disabled
Port      Name               Status       Reason
Gi1/8/1                      err-disabled udld
```

UDLD

Key points

- Ensure mode matches both ends
- Check status with “sh udld neighbour”
- Always recommended to use on Inter-Switch Links (Although IPS uses it)
- With aggressive mode, use an errdisable recovery timer
- 21s-42s to detect failure (15s default = 42s, reduce to 7s)

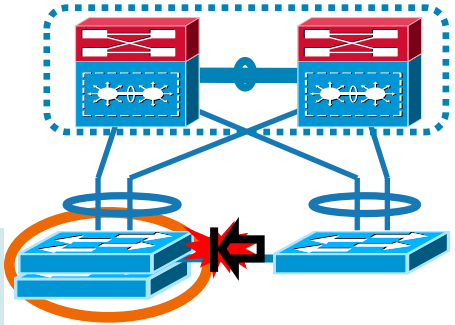
- VSS Specific
  - Link management protocol (LMP) checks for unidirectional Links

LoopGuard

- Protects alternate (blocked) or root (forwarding) ports from moving to forwarding upon no receipt of BPDU's
- P2P links losing BPDU's can indicate uni-directional issue
- Enable LoopGuard and UDLD

```
dist3750(config-if)# spanning-tree guard loop
Aug 31 16:21:52: %SPANTREE-2-LOOPGUARD_BLOCK:
Loop guard blocking port FastEthernet2/0/1 on MST0.
```

```
dist3750# show spanning-tree vlan 99
. . .
Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------
Gi2/0/1             Desg FWD 20000     128.55   P2p
Fa2/0/1             Desg BKN*200000    128.57   P2p *LOOP Inc
. . .
```



Dispute

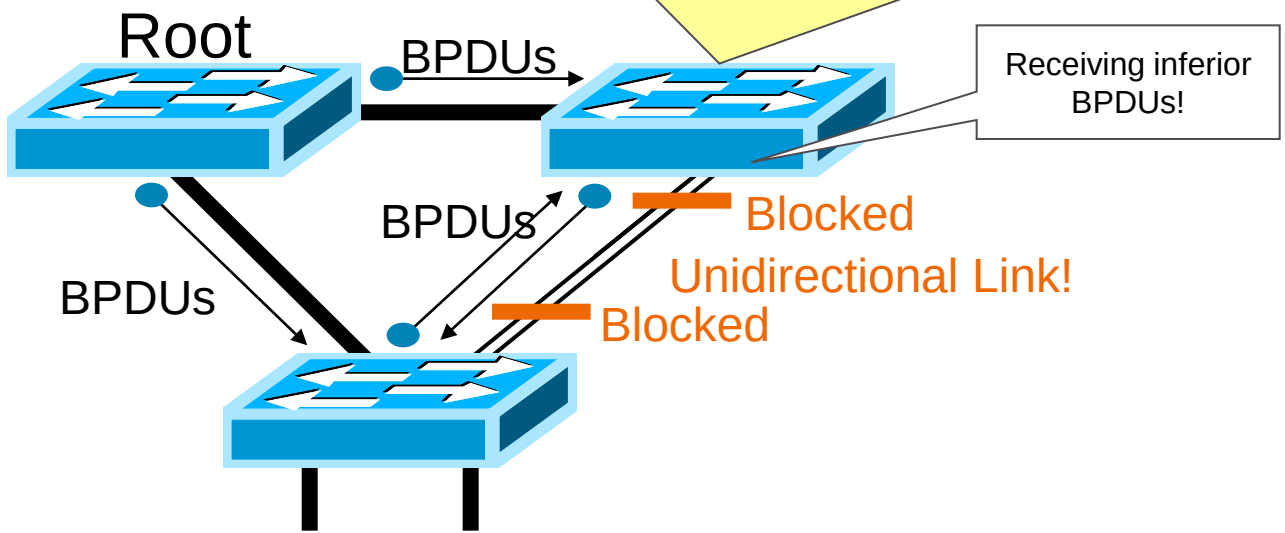
The Mechanism

- New mechanism in RSTP 801.d 2004 and MST 802.1Q
- Implemented in standard MST code 12.2(18)SXF
- Checks consistency of the port role and state
- Not configurable!
- Very efficient protection against unidirectional link failures
- Quote from 802.1d 2004 specification

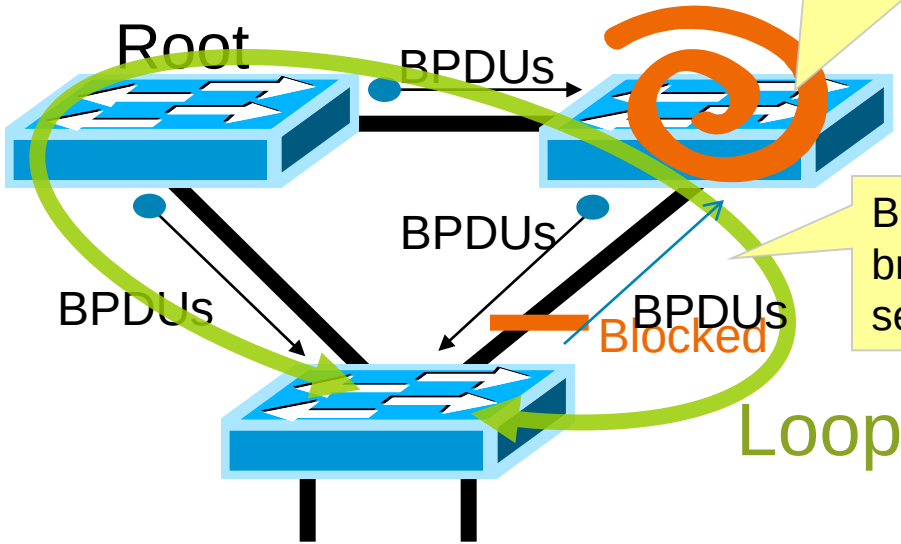
If a Port Receive state machine receives an inferior RST BPDU from a Port that believes itself to be a Designated Port AND is Learning or Forwarding it will set disputed, causing this state machine to transition a Designated Port to Discarding

Dispute Mechanism

```
%STP-2-DISPUTE_DETECTED: Dispute detected on port Ethernet1/2 on VLAN0700

6500# sh spanning vlan 700 | in BLK
Eth1/2              Desg BLK 2000      128.130  Network P2p
```



The Brain Dead Switch Problem

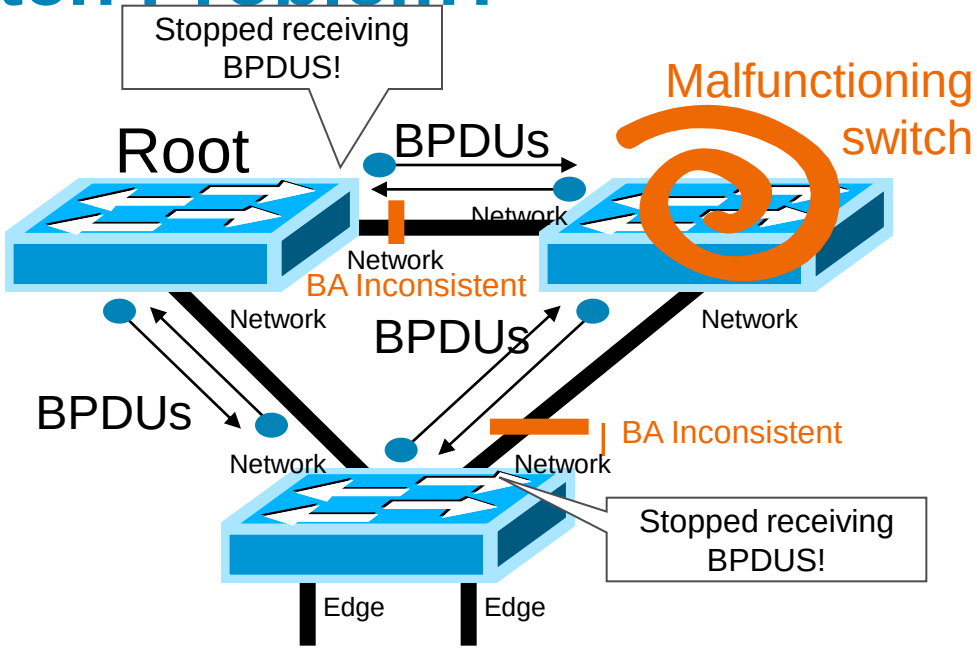


Not capable to send bpdu, process received bpdu or disable port with dispute

Brain dead switch

BPDU not processed by brain Dead switch – can't set dispute

How Does Bridge Assurance Solve Brain Dead Switch Problem?




```
%STP-2-BRIDGE_ASSURANCE_BLOCK:Bridge Assurance blocking port Ethernet2/48 VLAN700

6500# sh spanning vlan 700 | inc BA
Eth2/48          Altn BKN*4         128.304  Network P2p *BA_Inc
```

Bridge Assurance

Cisco Enhancement - Example

```
vsscore(config)# int po2
vsscore(config-if)# spanning-tree portfast network

*Aug 12 15:27:46: %SPANTREE-SW2_SP-2-BRIDGE_ASSURANCE_BLOCK:
Bridge Assurance blocking port Port-channel2.
```

```
vsscore# sh spanning-tree vlan 99
...
Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------
Po1                 Desg FWD 10000     128.5761 P2p
Po2                 Desg BKN*5000      128.5762 P2p Network *BA_Inc
```

note network type

- Configurable globally for “network” ports
- Must be enabled both ends of p2p link

Analysis - Complete

- Symptoms on dist4500 displayed root port flapping.
- Who is suppose to be root? Isolation!
- Next Steps – Once isolated to the problematic switch who was also claiming to be root, check the following:
  - Uni directional link
  - High Cpu
  - Interface drops

Troubleshooting Flooding and MAC Flapping

Topics

- Flooding
  - –Unicast Flooding
- MAC Flapping
  - – Vmware ESX Vswitch

Validate Flooding at the Core

6500/7600 – Supervisor 720

```
cat6500#remote loging switch
```

- Before flooding

```
Cat6500-sp#show earl statistics | inc Dst Mac misses
Dst Mac misses                          = 0x00000000005A9DF0 (5938672)
```

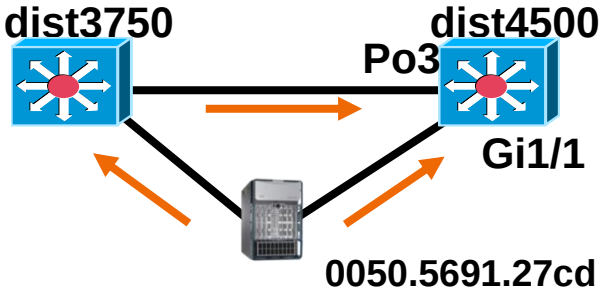
- After flooding – look for high increment in misses

```
Cat6500-sp#show earl statistics | inc Dst Mac misses
Dst Mac misses                          = 0x00000000005A9DF0 (5939542)
```

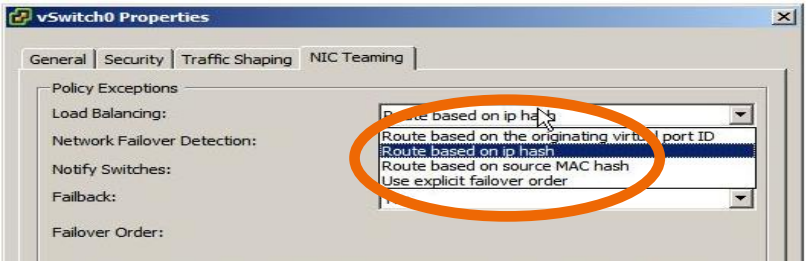
MAC Flapping

NIC Teaming

- Teaming of Network Interface Cards
- Server virtual address (SVA)
  - Use same MAC address
  - Fault Tolerant
  - Must use Active/Standby mode
- Separate NIC MACs
  - Use separate MAC addresses
  - Load Balance
- Check server configuration




```
Nov 11 15:39:27 DST: %MAC_MOVE-SP-4-NOTIF: Host 0050.5691.27cd in vlan 99 is flapping between
port Po3 and port Gi1/1
Nov 11 15:39:27 DST: %MAC_MOVE-SP-4-NOTIF: Host 0050.5691.27cd in vlan 99 is flapping between
```



00-50-56 (hex) VMware, Inc.
(from http://standards.ieee.org/cgi-bin/ouisearch)

MAC Flapping - Continue

Nexus

- Nexus 5000 requires globally enabling mac-move

```
Nexus5000(config)# mac address-table notification mac-move
```

- Nexus 7000 is more unique ☺ - Worked this out at 3am!

```
Nexus7000# show system internal l2fm l2dbg macdb vlan 25

VLAN 25 MAC 0002.3d40.0a02:
    Time                       If         Db Op Src Slot
    Tue Feb 15 16:28:32 2011 0x1600001a 0  1  0   0
    Tue Feb 15 17:53:04 2011 0x16000063 0  0  3   1
    Tue Feb 15 17:54:29 2011 0x1600001a 0  3  0   0
```

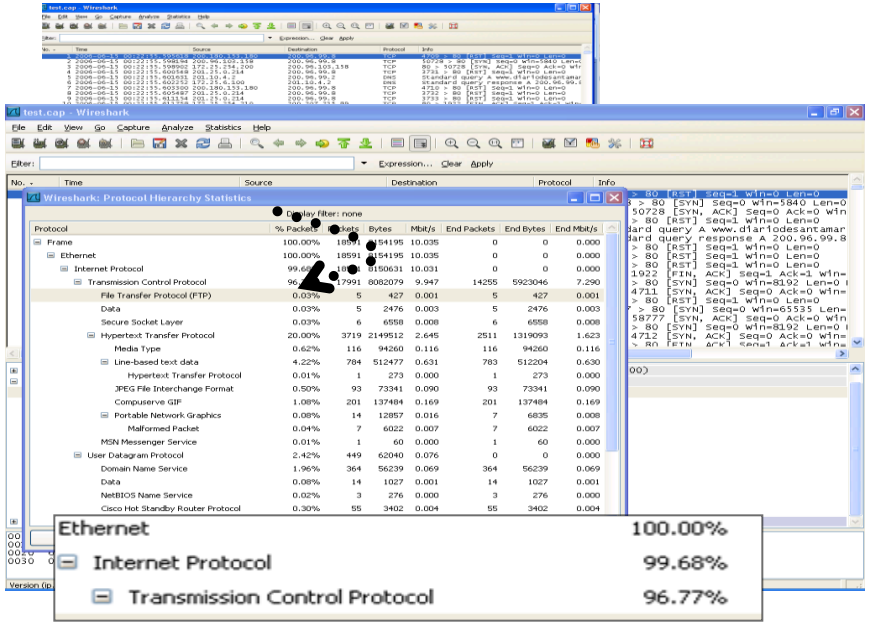
```
Nexus7000# show system internal pktmgr interface cache | inc 01a
Port-channel27, ordinal 64, if-index: 1600001a, up/up
```

Topics

- Traffic Captures
- Troubleshooting Tools
- High CPU
  - –Causes of High CPU
  - –IP Input
  - –Platform Specific Commands
- Forwarding Issues
  - –ELAM
  - –SUP720 TCAM
  - –3750 SDM Template

Wireshark Captures

What does it mean to me?



- Protocol Hierarchy
- TCP Pattern i.e. Retransmission
- Application replay
- Specific Bit set
- RTP stream

Troubleshooting Tools

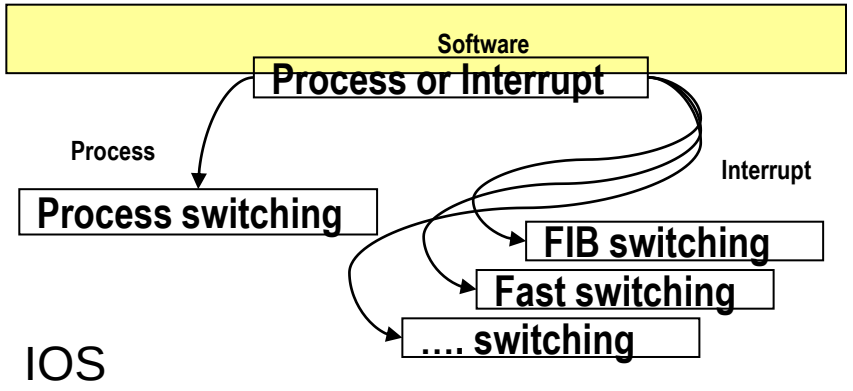
Many different tools…

- Which tools for which platform?

Tool\Platform	cat3750 family	cat4500 family	cat6500 family
SPAN	Yes	Yes	Yes
SPAN of inband	No	Yes	Yes
Mini Protocol Analyzer	No	No	Yes
VACL capture	No	No	Yes
CPU traffic capture	No	Yes (cpu buffer)	Yes (netdriver)
CPU queue dump	Yes	No	No

High CPU

Causes of High CPU



- Traffic that should be punted
  - Fragmentation
  - TTL
  - Redirects or Unreachables
  - ACL logging etc
  - Mitigate using rate-limiters or control plane policing where supported)
- Traffic that should NOT be punted
  - Forwarding issue – misprogramming between hw and sw
  - Resource issue
  - Feature conflict, e.g. NAT

High CPU

IP Input (process) driven

```
vsscore#show proc cpu sorted | exc 0.00
CPU utilization for five seconds: 99%/58%; one minute: 86%; five minutes: 75%
 PID Runtime(ms)   Invoked      uSecs   5Sec   1Min   5Min TTY Process
 315      789424   6020868        131 40.71% 27.42% 12.22%   0 IP Input
 528      708796   4488558        157  0.15%  0.11%  0.10%   0 Port manager per ...
```

1

99% Total,
58% Interrupt
41% Process

```
vsscore# show interfaces switching | inc (Vlan|Drops|Flushes)
...
Vlan99
        Drops          RP        37636          SP           0
  SPD Flushes        Fast        37636         SSE           0
```

2

```
vsscore# sh buffers input-interface vlan 99 header
. . .
 if_input 0x4ABF3278 (Vlan99), if_output 0x0 (None)
. . .
  source: 10.1.99.50, destination: 10.1.99.1, id: 0x0000, ttl: 63,
  TOS: 0 prot: 6, source port 0, destination port 0
```

3

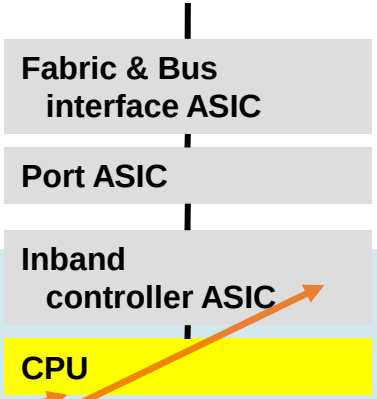
```
vsscore#remote command switch sh proc cpu sorted | exc 0.00
CPU utilization for five seconds: 13%/0%; one minute: 8%; five minutes: 8%
```

Remember to check SP CPU

High CPU

6500/VSS/7600 IBC

- Inband Channel
- Carries “process switched” traffic to RP
  - Check for high traffic levels
  - Any inband throttling?




```
vsscore#show ibc
Interface information:
   Interface IBC0/0(idb 0x515983C0)
   Hardware is Mistral IBC (revision 5)
   5 minute rx rate 65749000 bits/sec, 10387 packets/sec
   5 minute tx rate 7000 bits/sec, 15 packets/sec
   17952578 packets input, 14745548780 bytes
   ...
   Potential/Actual paks copied to process level 17014701/16996627
   (18074 dropped, 18074 spd drops)
 ...
MISTRAL ERROR COUNTERS
LBIC RXQ Drop pkt count = 15586          LBIC drop pkt count  = 0
```

Mini Protocol Analyzer

6500/VSS/7600 - VLAN capture example

```
vsscore(config)# monitor session 1 type capture
vsscore(config-mon-capture)# source vlan 99 rx
vsscore# do sh monitor
Session 1
---------
Type                    : Capture Session
Source VLANs            :
    RX Only             : 99
...
vsscore# monitor capture start for 30 seconds
Aug 11 11:14:20: %SPAN-5-PKTCAP_START: Packet capture session 1 started
```

1

- Inbuilt "Mini protocol analyzer"
- 6500 - 12.2(33)SXI / 7600 - 12.2(33)SRD
- Export files in PCAP format for external use
- Filter vlan/mac/ethertype/packet size

```
vsscore# sh monitor capture status
capture state       : BUFFER_FULL
capture mode        : Linear
Number  of packets
          captured : 21845
...

vsscore# sh monitor capture buffer
 1     IP: s=10.1.99.50 , d=224.0.0.2, len 982
 2     IP: s=10.1.99.50 , d=224.0.0.2, len 982
```

2

That same host again!! ☹

Netdriver Capture

6500/VSS/7600 - CPU bound example

- CPU-bound/originated traffic
- Non-Intrusive (safe to use during high CPU)
- All supervisors as of 12.2(18)SXF

```
vsscore# debug netdr capture rx
vsscore# sh netdr captured-packets
A total of 4096 packets have been captured
The capture buffer wrapped 0 times
Total capture capacity: 4096 packets

------- dump of incoming inband packet -------
interface Vl99, routine mistral_process_rx_packet_inlin, timestamp 00:00:12
dbus info: src_vlan 0x63(99), src_indx 0xB43(2883), len 0x3E8(1000)
...
destmac 01.00.5E.00.00.02, srcmac 00.00.05.00.09.00, protocol 0800
protocol ip: version 0x04, hlen 0x05, tos 0x00, totlen 982, identifier 0
  df 1, mf 0, fo 0, ttl 1, src 10.1.99.50, dst 224.0.0.2
    udp src 63, dst 63 len 962 checksum 0xDC0C
```

```
vsscore-sp#test mcast ltl-info index B43
index 0xB43 contain ports 24/3, 38/R,5
```

CPU Buffer Capture

4500 – CPU bound example

- Cat4500/4500E 12.2(18)EW1 onwards
- Captures CPU-bound traffic
- Similar to netdriver in 6500
- Implemented in sw

```
dist4500# sh platform cpu packet buffered
Total Received Packets Buffered: 1024
-------------------------------------
Index 0:
6 days 1:41:10:992654 - RxVlan: 99, RxPort: Gi3/2
Priority: Normal, Tag: Dot1Q Tag, Event: Input Acl, Flags: 0x40, Size: 1000
Eth: Src 00:00:05:00:09:00 Dst 01:00:5E:00:00:02 Type/Len 0x0800
Ip: ver:4 len:20 tos:0 totLen:982 id:0 fragOffset:0 ttl:1 proto:udp
  src: 10.1.99.50 dst: 224.0.0.2 firstFragment lastFragment
...
dist4500#undebug all
```

1

```
dist4500# traceroute mac 0021.5589.9f90  0000.0500.0900
Source 0021.5589.9f90 found on dist4500
1 dist4500 (10.66.91.188) : Fa5/1 => Po2
2 vsscore (10.1.99.1) : Po2 => Po1
3 dist3750 (10.66.91.237) : Po1 => Fa2/0/3
Destination 0000.0500.0900 found on dist3750
Layer 2 trace completed
```

Host in vlan 99 connected to this switch

2

CPU Queues on cat3750/E

3750 – CPU bound example

```
dist3750# debug platform cpu-queues ?
  broadcast-q
  cbt-to-spt-q
  cpuhub-q
...
```

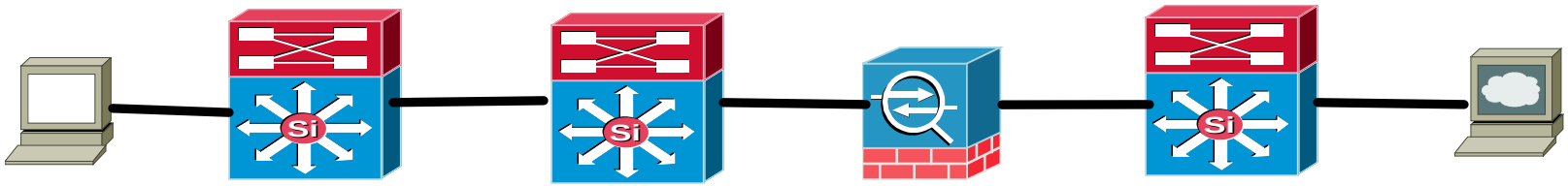
- Cat3750/E
- Captures CPU-bound traffic
- Various cpu-queues to debug
- Per-packet debugs!, do not log console
- Implemented in sw

```
dist3750#debug platform cpu-queues igmp-snooping-q

debug platform cpu-queue igmp-snooping-q debugging is on

Aug 11 00:24:28: Pak recvd on IGMP-SNOOP-Q: Local Port Fwding L3If:Vlan99
L2If:FastEthernet2/0/3 DI:0x1304, LT:7, Vlan:99   SrcGPN:59, SrcGID:59, ACLLogIdx:0x0,
MacDA:0100.5e00.0002, MacSA: 0000.0500.0900   IP_SA:10.1.99.50 IP_DA:224.0.0.2 IP_Proto:17
...

dist4500#undebug all
```

Forwarding – Independent Approach

Where did my packets go?

- The core problem needs to be narrowed down to device level where possible
  - –Trace path through network .
  - –Find the easiest way. Take a holistic view.
  - – Once isolated see if mac is learning i.e. show mac-address-table



Web Client

ASA / FWSM

Web Server

Firewall Packet Capture

- Create a capture

```
ASA(config)# access-list web permit ip host 10.1.1.2 host 198.133.219.25
ASA(config)# access-list web permit ip host 198.133.219.25 host 10.1.1.2
```

```
ASA# capture inside access-list web interface inside
```

- Find the packet in the capture you want traced

```
ASA# show capture inside

68 packets captured
   1: 15:22:47.581116 10.1.1.2.31746 > 198.133.219.25.80: S
   2: 15:22:47.583465 198.133.219.25.80 > 10.1.1.2.31746: S ack
   3: 15:22:47.585052 10.1.1.2.31746 > 198.133.219.25.80: . ack
   4: 15:22:49.223728 10.1.1.2.31746 > 198.133.219.25.80: P ack
   5: 15:22:49.223758 198.133.219.25.80 > 10.1.1.2.31746: . Ack
      ...
```

Forwarding

Using ping with IP options

- Normally host traffic is hardware switched
- Traffic with IP options (record route, DF….) is punted via cpu (software) path
- Verifies software and hardware programming
- Look for any patterns !.!.!.!.!

```
vsscore# ping ip
Target IP address: 10.1.99.2
...
Extended commands [n]: y
...
Loose, Strict, Record, Timestamp, Verbose[none]: T
...
Time= 00:00:00.000 UTC (00000000)
...
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/4 ms
```

1000Base Link Negotiation

Verifying GBIC/SFP Operation

- Displaying GBIC/SFP details

```
vsscore# show interfaces gig1/8/12
GigabitEthernet1/8/12 is up, line protocol is up (connected)
...
Full-duplex, 1000Mb/s, media type is SX
  input flow-control is off, output flow-control
...
```

1

```
vsscore# show interface gig1/8/12 transceiver properties
Name : Gi1/8/12    Admin Speed Nego : Disable,
Admin Duplex Nego : Disable
```

2

```
vsscore# show idprom interface gigabitEthernet 1/8/12 detail
...
 Vendor Name          : CISCO-FINISAR
 Vendor PN            : FTRJ8519P1BNL-C2
 Vendor rev           : A
...
```

3

Vendor name and part #.
Useful if working with TAC

Interface Errors

Link Clocking Errors

- Typically fibre or hardware issue
  - –Test the fibre
  - –Try another GBIC/SFP
  - –Move to another Port/Card
  - –Test using “known” working fibre

```
dist4500#sh int gi6/22 counters errors

Port        Align-Err    FCS-Err   Xmit-Err    Rcv-Err UnderSize OutDiscards
Gi6/22              0         55          0        133         0            0

Port      Single-Col Multi-Col  Late-Col Excess-Col Carri-Sen     Runts    Giants
Gi6/22             0         0         0          0         0         6         0

Port       SQETest-Err Deferred-Tx IntMacTx-Err IntMacRx-Err Symbol-Err
Gi6/22               0           0            0            0         61
```

ELAM (Embedded Logic Analyser Module)

6500/VSS/7600

TAC Tool

```
vsscore(config)# service internal
vsscore# remote login switch
vsscore-sp# sh platform capture elam asic superman slot 6
vsscore-sp# sh platform capture elam trigger dbus ipv4 if IP_DA=10.1.99.3
```

1

```
vsscore-sp# sh platform capture elam start
vsscore-sp# sh plat cap elam status
Active ELAM info:
Slot Cpu   Asic    Inst Ver PB Elam
---- --- -------- ---- --- -- ----
6    0   ST_SMAN  0    3.2    Y
DBUS trigger: FORMAT=IP L3_PROTOCOL=IPV4 IP_DA=10.1.99.3
```

Trigger based on IP_SA, IP_DA, vlan etc.... (combinations allowed)

2

```
vsscore-sp# sh platform capture elam data
DBUS data:
VLAN                    [12] = 99
SRC_FLOOD               [1] = 0
SRC_INDEX               [19] = 0xB43
DMAC                    = 000f.232c.e93f
SMAC                    = 0013.c4e4.0342
IP_SA                   = 10.1.99.2
IP_DA                   = 10.1.99.3

RBUS data:
DEST_INDEX              [19] = 0xB42
VLAN                    [12] = 99
```

3

Use "remote command sw test mcast ltl-info index" to find specific ingress/egress interface

- Check Packet Lookup/Forwarding
- Enable on Ingress (PFC or DFC)
- Be as specific as possible with trigger
- One Packet only

3750 SDM Templates

Verifying TCAM

- Ethernet controller ASIC uses single TCAM
- Subdivided into different areas

- Check required template
- Based on requirements
- “sdm prefer” and reload

1

```
dist3750#sh ip mroute count
1322 routes using 573520 bytes of memory
...
```

2

```
dist3750# sh sdm prefer
 The current template is "desktop default" template.
...
  number of unicast mac addresses:                  6K
  number of IPv4 IGMP groups + multicast routes:    1K
...
```

3

```
dist3750# sh platform tcam utilization
CAM Utilization for ASIC# 0                  Max                  Used
                                          Masks/Values         Masks/values

 Unicast mac addresses:                    784/6272             73/510
 IPv4 IGMP groups + multicast routes:      144/1152             132/1025
...
```

QoS

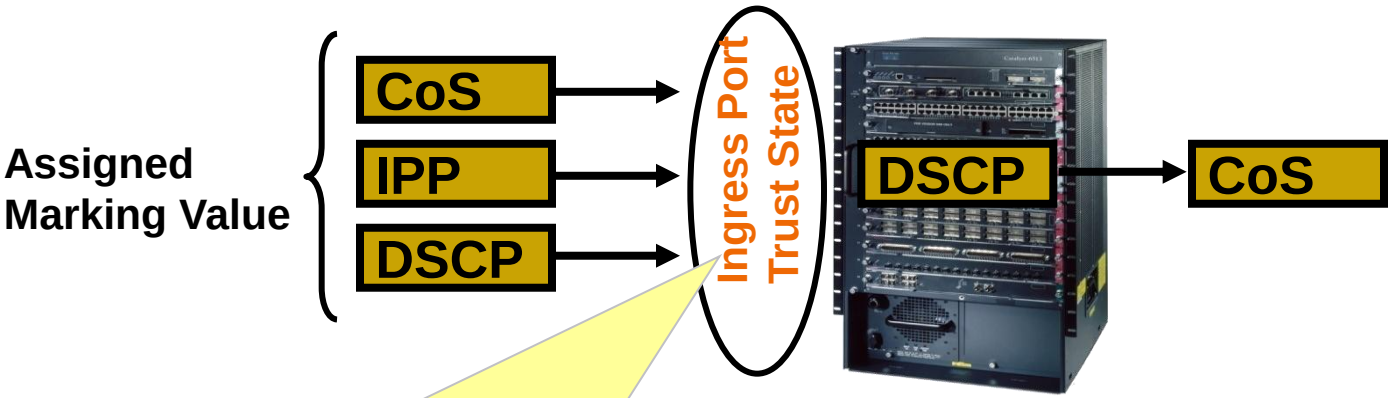
Topics

- Campus QOS Considerations
- Verify Voice Signalling QOS Markings
- Verifying QOS
  - –3750 – Trusting DSCP
  - –3750 – Output Drops
  - –6500 – VLAN Based Policing
- Building Natural Fault Boundaries

Campus QoS Considerations

Internal Mapping Tables

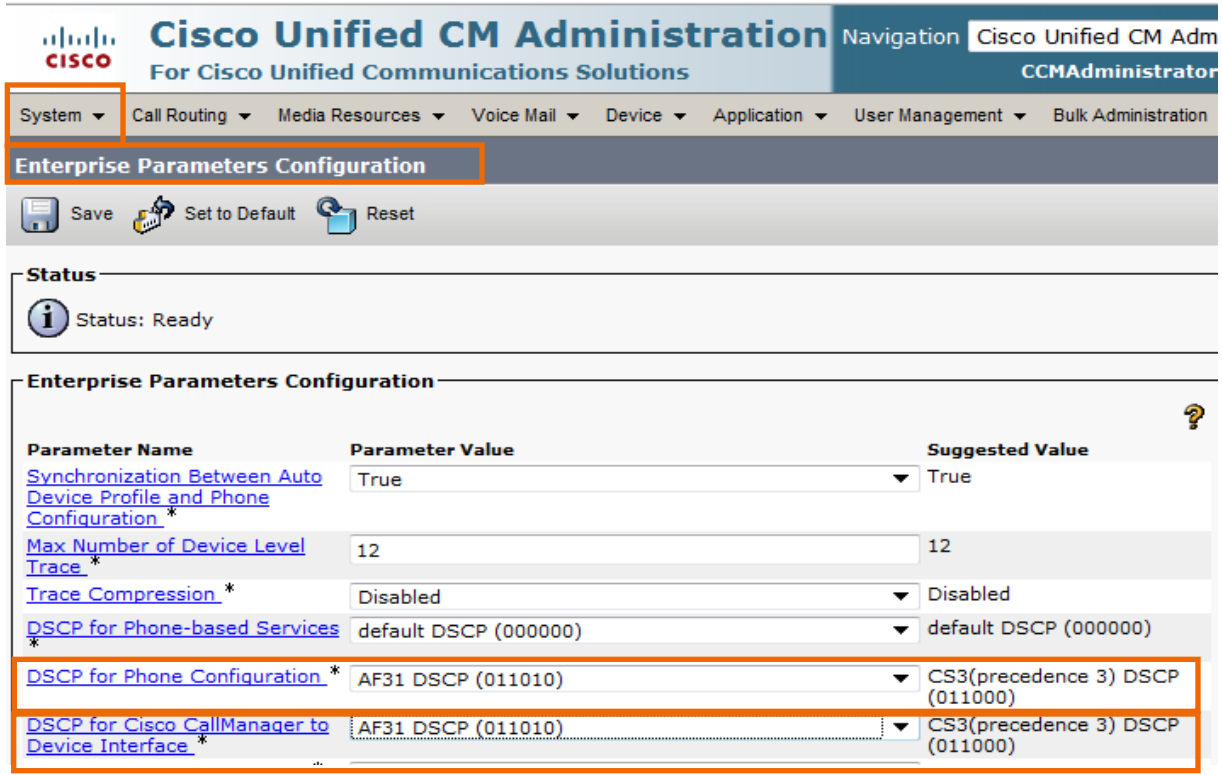
Ingress mapping tables are used to take an existing layer 2 or layer 3 marking and map it to an internal DSCP value used by the switch to assign service levels to the frame as it is in transit.



- Check path of packet
- Ensure to "trust" on **all** interfaces
- Check wire capture to confirm 100%

Egress mapping tables are used to rewrite CoS for applicable frames from the internal DSCP on egress from the switch.

Voice Signalling Verification - CUCM



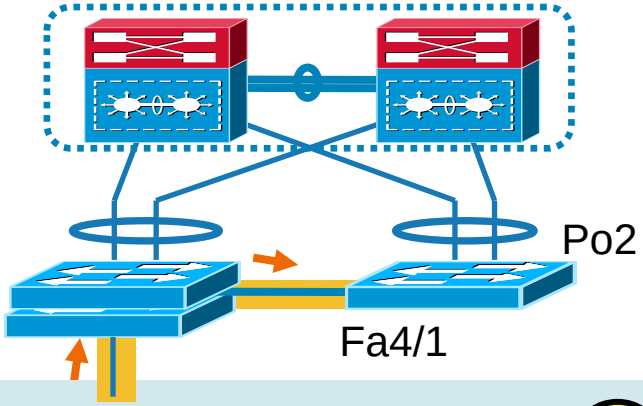
- Shortcuts are good!
- One less Wireshark Capture!
- Auto QoS mismatch.

Verifying QoS

Cat3750 – Trusting DSCP

```
dist3750# sh mls qos interface fa2/0/3
trust state: trust dscp
trust mode: trust dscp
trust enabled flag: ena
.  .  .
```

1




```
dist3750# sh mls qos maps dscp-output-q
   Dscp-outputq-threshold map:
     d1 :d2    0     1     2     3     4     5     6     7     8     9
     ------------------------------------------------------------
      0 :     02-01 02-01 02-01 02-01 02-01 02-01 02-01 02-01 02-01 02-01
     ...
      4 :     01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 04-01 04-01
```

2

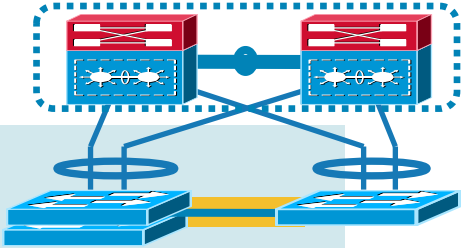
```
dist3750# sh mls qos interface fa2/0/1 statistics
...
dscp: outgoing
-------------------------------
 0 -  4 :        0         0         0         0         0
...
45 - 49 :        0       100         0         0         0
```

3

- 100 packets sent with DSCP 46
- DSCP 46 maps to Q1/T1 on egress
- Matches shown in egress Q

Verifying QoS

Cat3750 – Output Drops



```
dist3750# show mls qos queue-set
Queueset: 1
Queue     :       1       2       3       4
----------------------------------------------
buffers   :      25      25      25      25
threshold1:     100     200     100     100
threshold2:     100     200     100     100
reserved  :      50      50      50      50
maximum   :     400     400     400     400
```

hw refers queues 0-3 (0-based)
sw refers queues 1-4

1

```
dist3750# sh interfaces fa2/0/1
FastEthernet2/0/1 is up, line protocol is up (connected)
Last clearing of "show interface" counters 00:09:16
  Input queue: 0/75/0/0 (size/max/drops/flushes);
Total output drops: 55169029
```

2

```
dist3750# show platform port-asic stats drop fa2/0/1

  Interface Fa2/0/1 TxQueue Drop Statistics
    Queue 0
      Weight 0 Frames 172573975
. . .
```

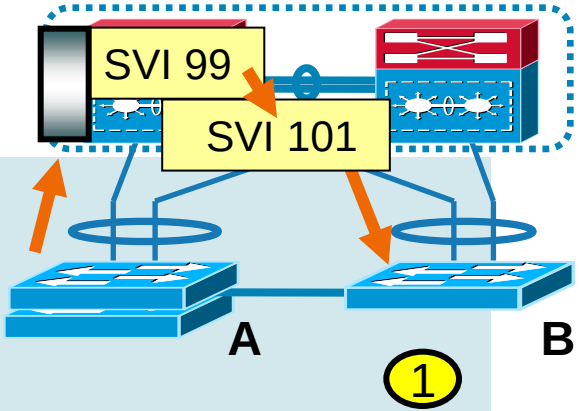
3

Queue 1, threshold 1 seeing drops

- Know your traffic profile
- Buffer tuning may help

Verifying QoS

6500/VSS/7600 – VLAN based Example




```
vsscore# show policy-map interface vlan 99
 Vlan99
. . .
      Match: access-group 101
      police :
        10000000 bps 312500 limit 312500 extended limit
. . .
Earl in switch 2, slot 6 :
 aggregate-forwarded 18445800 bytes action: transmit
        exceeded 11828300 bytes action: drop
```

```
vsscore# show mls qos ip vlan 99
. . .
Int   Sid Mod Dir  Class-map ... AgForward-By   AgPoliced-By
-------------------------------------------------------------
Vl99    1   6  In NETWORKERS     0               0            0
Vl99    2   6  In NETWORKERS     0        18445800     11828300
```

2

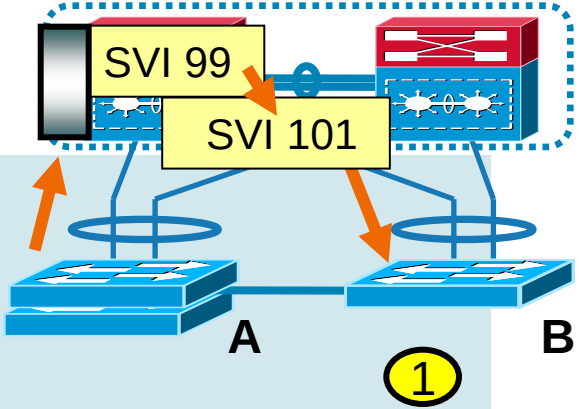
```
vsscore# show tcam interface vlan 99 qos type1 ip
. . .
    AT      udp host 10.1.99.50 any
    MU      ip any any
```

3

Remember
"mls qos vlan-based"
if using SVI policy

Verifying QoS

6500/VSS/7600 – VLAN based Example




```
vsscore# show policy-map interface vlan 99
 Vlan99
. . .
      Match: access-group 101
      police :
        10000000 bps 312500 limit 312500 extended limit
. . .
Earl in switch 2, slot 6 :
 aggregate-forwarded 18445800 bytes action: transmit
       exceeded 11828300 bytes action: drop
```

```
vsscore# show mls qos ip vlan 99                                    2
. . .
Int   Sid Mod Dir  Class-map ... AgForward-By   AgPoliced-By
-------------------------------------------------------------
Vl99    1   6  In NETWORKERS     0                 0          0
```

```
vsscore# show tcam interface vlan 99 qos type1 ip
. . .                                                              3
    AT      udp host 10.1.99.50 any
    MU      ip any any
```

Remember
"mls qos vlan-based"
if using SVI policy

Fault Boundaries and Storm Control

- Manage unexpected flooding.
- Mitigate network latency.

```
Interface GigabitEthernet0/1
  storm-control broadcast level 50.00
```

Storm-control activated as traffic exceeded over 50%

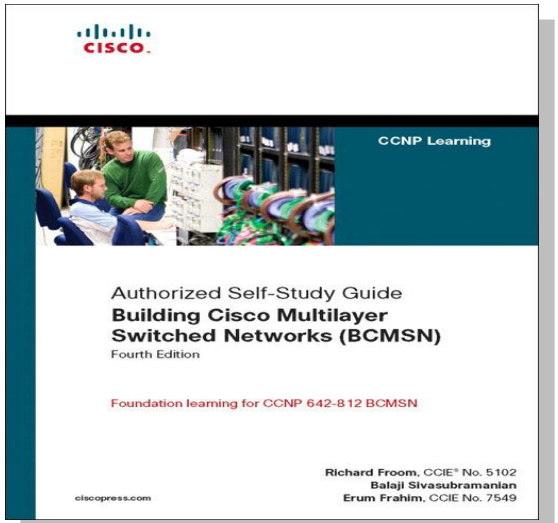
```
switch#show storm-control broadcast
  Interface  Filter State   Trap State     Upper    Lower    Current  Traps Sent
  ---------  -------------  -------------  -------  -------  -------  ----------
  Fa0/1      Forwarding     inactive       50.00%   50.00%    30.00%           0
```

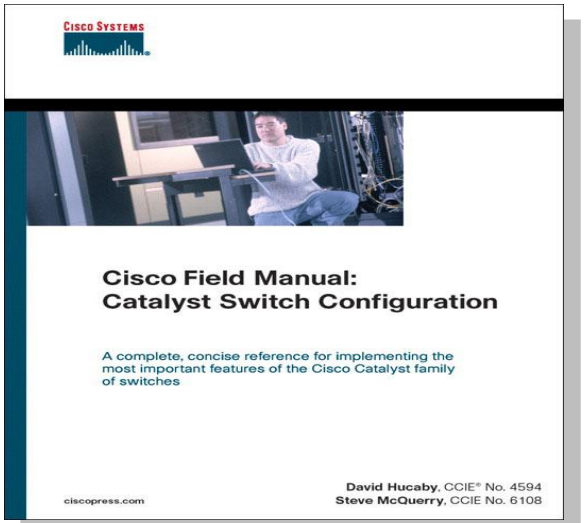
```
switch#show storm-control broadcast
  Interface  Filter State   Trap State     Upper    Lower    Current  Traps Sent
  ---------  -------------  -------------  -------  -------  -------  ----------
  Fa0/1      Blocking       inactive       50.00%   50.00%    53.00%           0
```

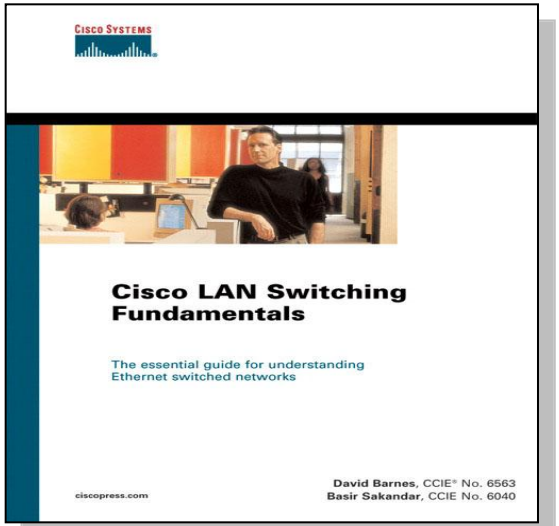
Key Take Away

- Identifying the network path
- Operational view of Dot1q Trunk and Etherchannel
- Techniques to isolate Spanning tree loops
- Why you need to harden spanning-tree
- How network flooding can occur.
- Identify common issues with VPC / VSS
- Understanding the reasons of High CPU and next steps
- QoS – How to identify packet classification

Recommended Reading







Source: Cisco Press

Some Useful Links

- NIC Teaming & Unicast Flooding

http://www.cisco.com/en/US/tech/tk648/tk362/technologies_tech_note09186a0080094afd.shtml#t8
http://www.cisco.com/en/US/products/hw/switches/ps700/products_tech_note09186a00801d0808.shtml#cause1

- STP

http://www.cisco.com/en/US/tech/tk389/tk621/technologies_tech_note09186a0080136673.shtml

- High CPU

http://www.cisco.com/en/US/products/hw/switches/ps708/products_tech_note09186a00804916e0.shtml
http://www.cisco.com/en/US/products/hw/switches/ps5023/products_tech_note09186a00807213f5.shtml
http://www.cisco.com/en/US/products/hw/switches/ps663/products_tech_note09186a00804cef15.shtml

- QOS

http://www.cisco.com/warp/public/cc/pd/si/casi/ca6000/prodlit/buffe_wp.pdf
http://www.cisco.com/en/US/products/hw/switches/ps708/products_tech_note09186a008074d6b1.shtml

Q & A

Complete Your Online Session Evaluation

Complete your session evaluation:

- Directly from your mobile device by visiting www.ciscoliveaustralia.com/mobile and login by entering your badge ID (located on the front of your badge)
- Visit one of the Cisco Live internet stations located throughout the venue
- Open a browser on your own computer to access the Cisco Live onsite portal


CISCO

Appendix

Identify and Avoid Unicast Flooding

```
SUP720# sh int | inc is up | rate | Input queue:
SUP720# sh mac-address-table | i flood
SUP720# sh mac-address-table add 000d.56b9.ecdb all detail
SUP720# remote command sw show mac-add add 000d.56b9.ecdb all detail
SUP720# remote command mod x show log
SUP720# show mac-address-table synchronize statistics
```

- Enabled if WS-X6708-10GE Present , otherwise disabled by default
- SXF onwards

- Recommended to be 3 times sync activity (160 sec default)
- Should be greater than ARP timeout!

```
Example :-

SUP720# conf t
SUP720(config)# mac-address synchronize
SUP720(config)# mac-address aging-time 0 routed-mac
SUP720(config)# mac-address aging-time 480
SUP720(config)# interface Vlan360
SUP720(config-if)# arp timeout 300
```

Reduce ARP time to less than mac aging time (default is 4 hours)

Frames routed by Sup720 and have SA re-written to MSFC

Troubleshooting catalyst 3750 QoS

Cheatsheet

- **General QoS command :**

```
Sh running-config
Sh mls qos
Sh platform tcam utilization
```

- **Aggregate Policer – Marking in policy-map**
  - –Check Configuration
  - –Sh mls qos int gig x/y statistics
  - –!!! NOT SUPPORTED :
    - •sh policy-map interface

- Queueing and scheduling :

```
–show platform port-asic stats drop  gig x/y
–show platform port-asic stats enqueue  gig x/y
```

Troubleshooting catalyst 4500 QoS

Cheatsheet

- **General QoS command :**
  - `Sh running-config`
  - `Sh qos`
  - `Sh platform hardware acl statis util`
- **Queueing and scheduling :**
  - `–Sh interface slot/port capability`
  - `–Sh interface slot/port counter detail`
  - `–Sh qos interface slot/port`
  - `–Sh plaform hardware int all`
  - `–Sh platform software int all`
- **Aggregate Policer – Marking in policy-map**
  - `–Check Configuration`
  - `–sh policy-map interface (software view)`
  - `–show platform hardware qos policers utilization`

Catalyst 6500 QoS

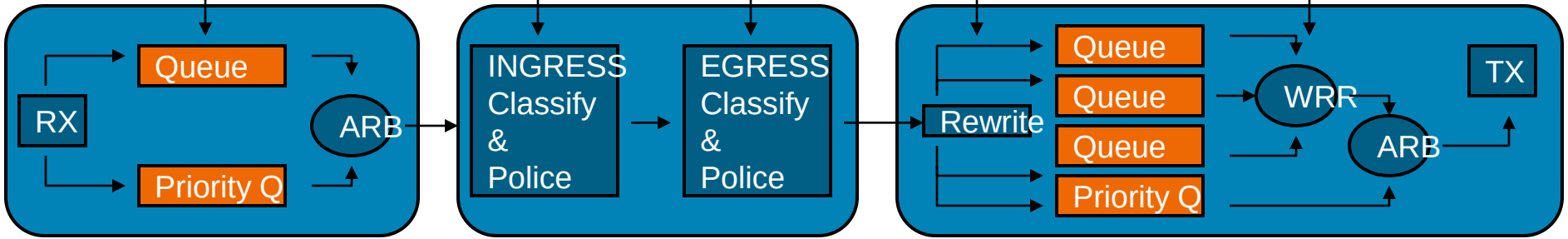
QoS Model

Scheduling: Queue and Threshold - select based on received CoS through configurable MAP I/F - CoS can be overwritten if port untrusted

Police via ACLs - Police actions include Forward, Mark and Drop. Based on Burst (Token Bucket) and Byte Rate

Scheduling: Queue and Threshold selected based on CoS through a Map

De-queue uses WRR or SRR between the round robin queues



Incoming encap can be ISL, 802.1Q or None

DSCP based classification based on “trusted port” and layer 2 info with ACL, layer 3 info with ACL and layer 4 info with ACL

Rewrite TOS field in IP Header and 802.1p/ISL CoS field

Each queue has configurable thresholds - some have WRED (except PQ)

Outgoing encap can be ISL, 802.1Q or None

Troubleshooting catalyst 6500 QoS

Cheatsheet

- **General QoS command** :

```
Sh running-config
Sh mls qos
Sh tcam count
```

- **Aggregate Policer – Marking in policy-map**
  - –Check Configuration – verify
  - –sh policy-map interface (software view)
  - –sh mls qos ip (hardware)
  - –sh tcam int (hardware)

- **Queueing and scheduling :**
  - –Sh interface *slot/port* capability
  - –Sh queuing interface *slot/port*
  - –Remote com sw sh qm-sp port-data *slot port*

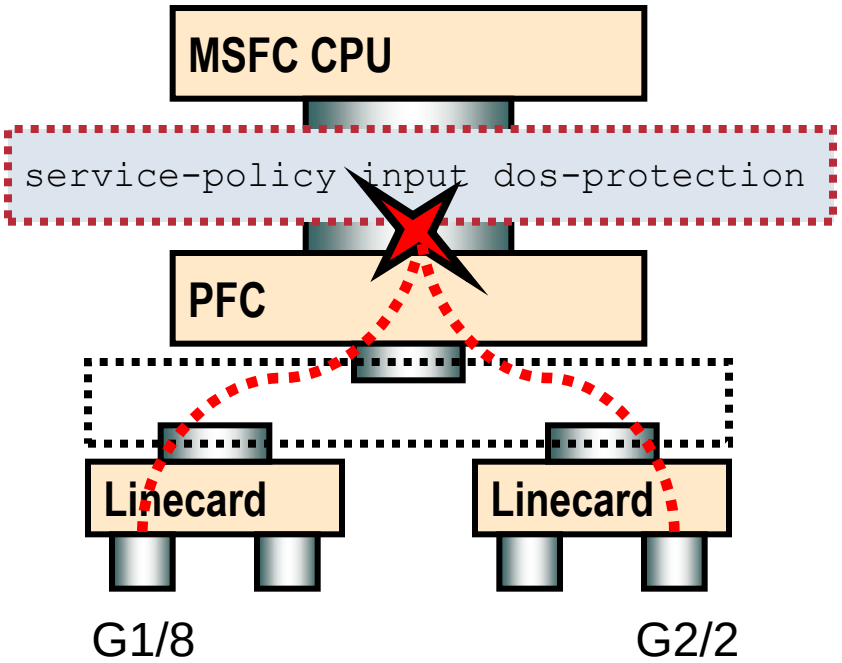
- **Microflow policer specifics** :

```
Check Configuration – verify
sh policy-map interface (software view)
sh mls qos ip (hardware)
sh tcam int (hardware)
Sh mls netflow ip …
```

Catalyst 6500 Features –

Control Plane Policing

- **Platforms Sup-720 and Sup-6E**
  - control-plane interface
  - Provides QoS control for Control Plane packets
  - Uses MQC CLI
  - Preserves existing interface configuration
  - Is easy to configure
  - Hardware implementation provides scalable protection against DoS



Control Plane Policing

Control Plane Policing Deployment

Control Plane Policing Deployment (3 steps)

- **Define a packet classification criteria**

```
Core-01-NW08(config)# class-map <traffic_class_name>
Core-01-NW08(config-cmap)# match <access-group>
```

- **Define a service policy**

```
Core-01-NW08(config-pmap)# policy-map <service_policy_name>
Core-01-NW08(config-pmap)# class <traffic_class_name>
Core-01-NW08(config-pmap)# police <rate> conform-action transmit exceed-action drop
```

- **Apply QoS Policy**

```
Core-01-NW08(config)# control-plane
Core-01-NW08(config)# service-policy input <service_policy_name>
```


CISCO