ACI Fabric Protection

John Weston
Technical Marketing Engineer, Data Center Networking
July 2019

ACI Hardening

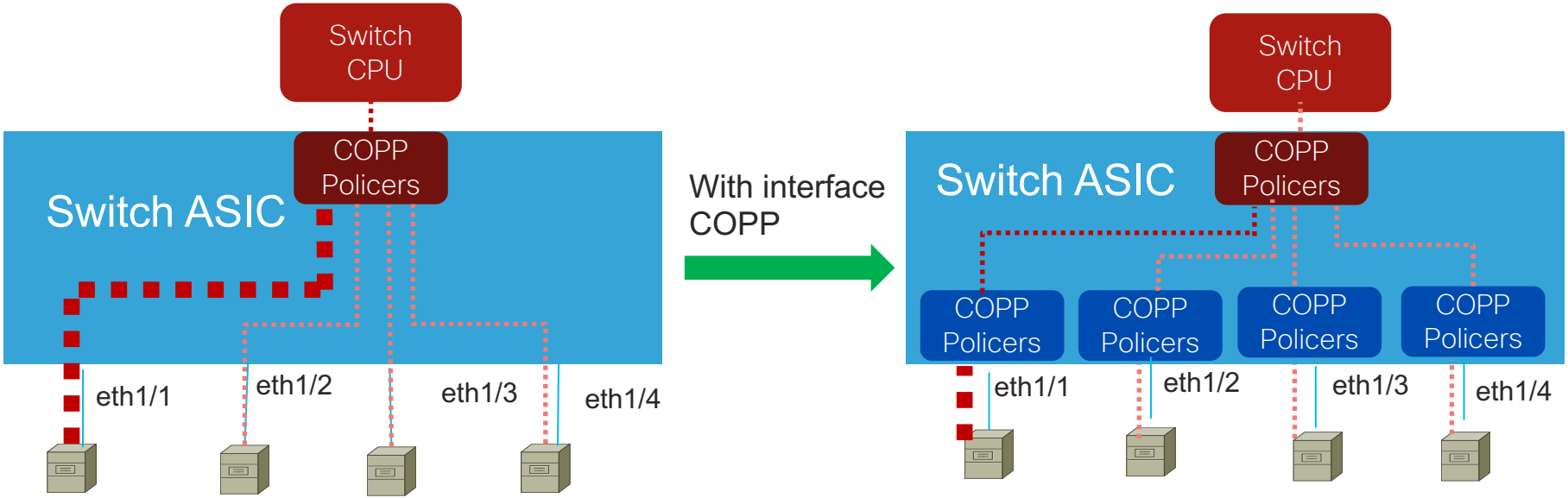
ACI Hardening Every Major and Minor SW Release

Flooding Attacks
SYN-FLOOD: Remain stable during SYN flooding attack
EST-FLOOD: Remain stable during ESTABLISHED flooding attack
LASTACK-FLOOD: Remain stable during LASTACK flooding attack
FINWAIT-FLOOD: Remain stable during FINWAIT flooding attack
CLOSING-FLOOD: Remain stable during CLOSING flooding attack
Port and Service Scans
DEF-CRED: No default authentication credentials
RECON-PORT-TCP: Remain stable during TCP port scan
RECON-PORT-UDP: Remain stable during UDP port scan
RECON-OSID: Remain stable during OS Fingerprinting
RECON-IP-PROT: Remain stable during IP protocol scan
NESSUS-SCAN: Known vulnerability scanner- Nessus
WEB-DEFECT: Known webserver and application defects
WEB-ID: Remain stable during web fingerprinting
Fuzzing
ESIC: UUT must endure malformed Ethernet packets
ICMPSIC: UUT must endure malformed ICMP packets
ISIC: UUT must endure malformed IPv4 packets
TCPSIC: UUT must endure malformed TCP packets
UDPSIC: UUT must endure malformed UDP packets
ICMPSIC6: UUT must endure malformed ICMPv6 packets
ISIC6: UUT must endure malformed IPv6 packets
TCPSIC6: UUT must endure malformed TCP over IPv6 packets
UDPSIC6: UUT must endure malformed UDP over IPv6 packets
Web Scan
Nexpose
IBM AppScan
OpenVas

Interface level COPP Policer

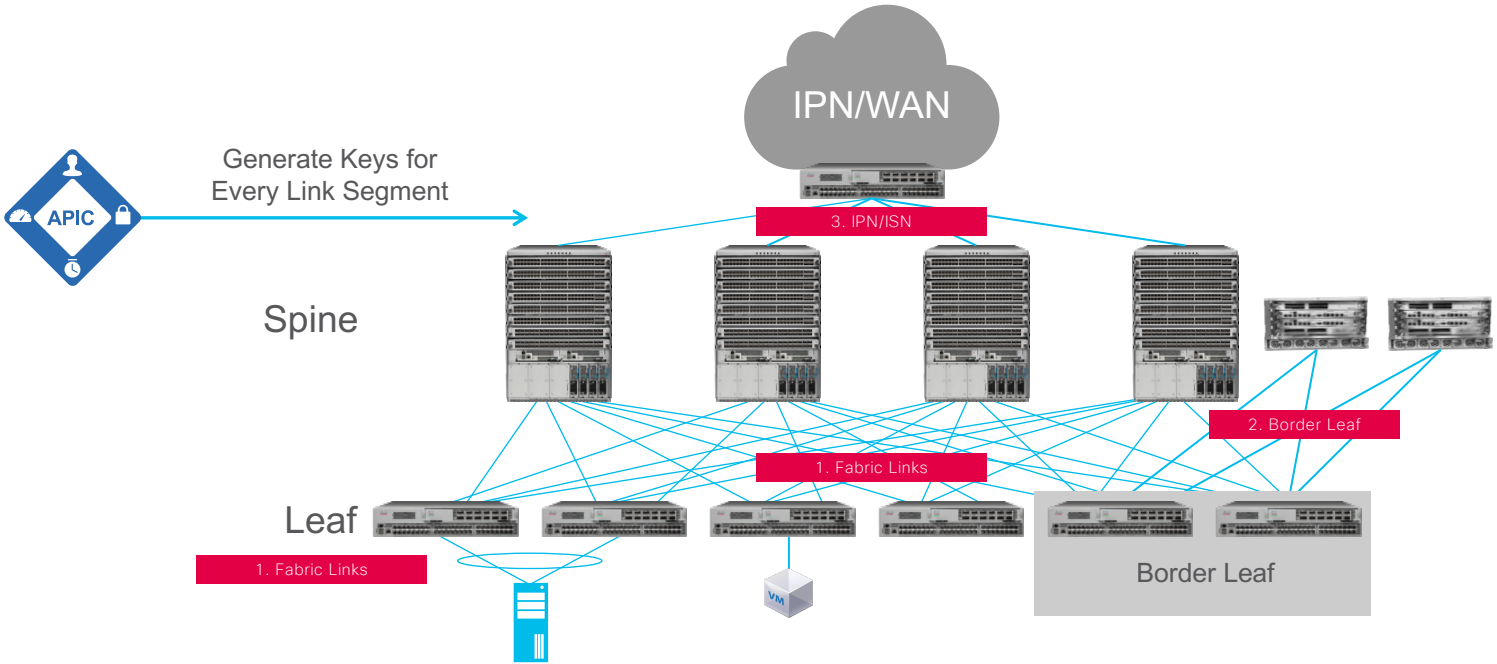
- Control traffic is rate limited first by interface level policer before it hits the aggregated COPP policer
  - Prevent traffic from one interface from “flooding” aggregate COPP policer. Ensure control traffic from other interfaces can reach CPU in case of DDOS from one or more interfaces
  - Per-interface-per-protocol policer supports following protocols
    - ARP, ICMP, CDP, LLDP, LACP, BGP, STP, BFD, OSPF
  - Support on 2nd generation leaf switches
  - Up to 256 policers per leaf switch for this feature
- Pre-Filtering Control Plane traffic before CoPP policies
  - Whitelist policy to explicitly list the type of traffic allowed

CoPP Per Interface Policers



ACI Encryption

MAC Sec

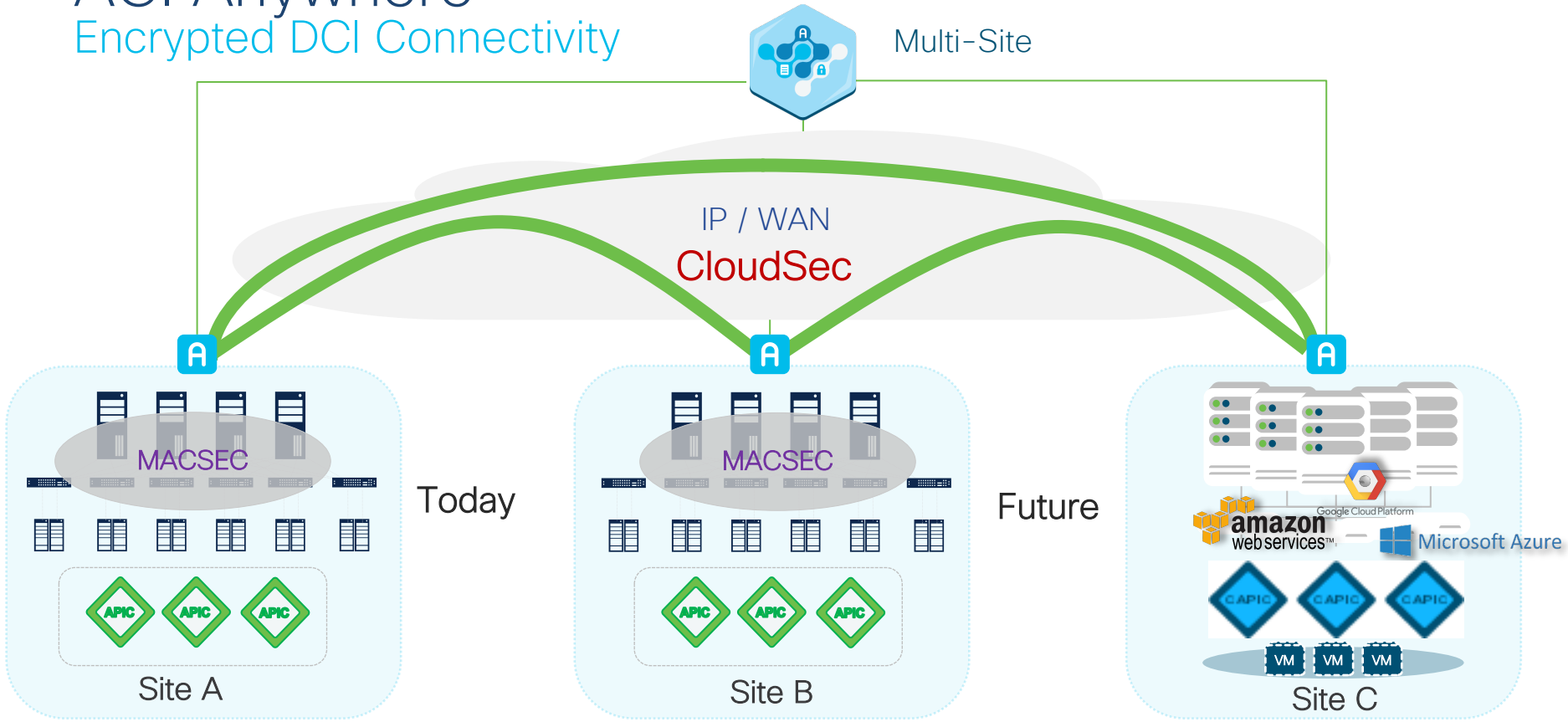


MACSEC Link Encryption
MKA Key Exchange

APIC Centralized Key
Management

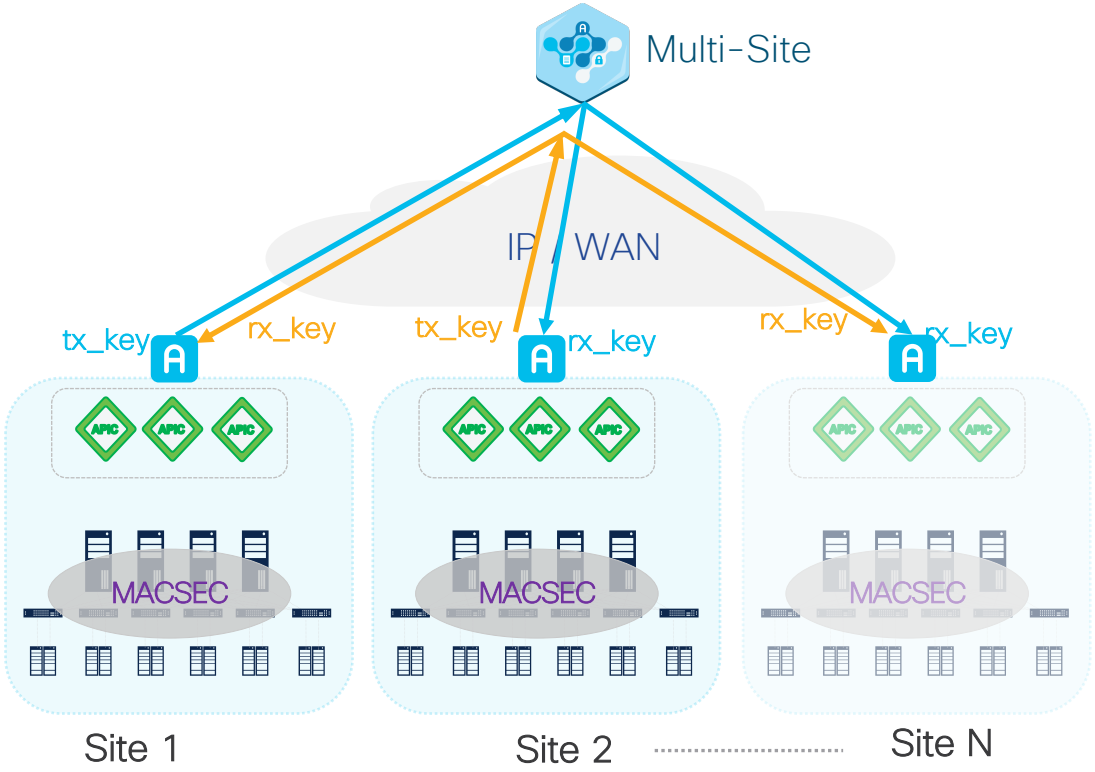
ACI Anywhere

Encrypted DCI Connectivity



Cloud Sec

Automated Key Distribution & Re-Key

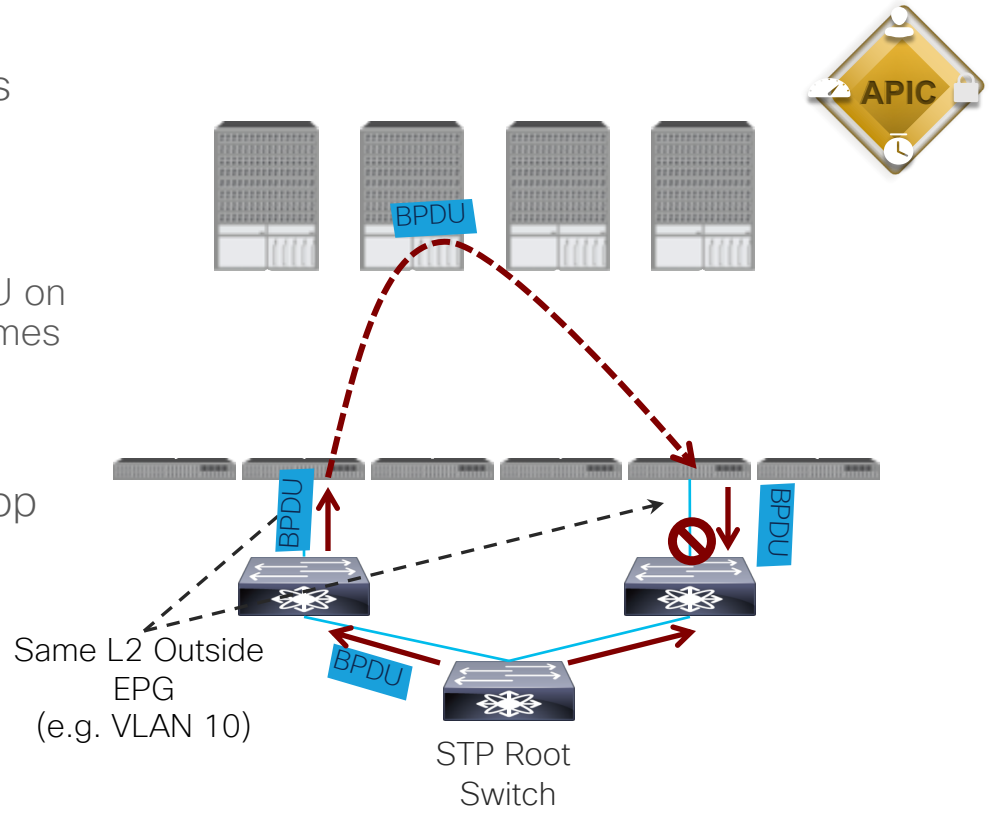


- Multi-site controller driven
  - No protocol dependency
- Reliable and secure key transport
  - rx_key installed before tx_key
- Non disruptive re-key
  - Hitless: make before break
- Always encrypted
  - In case of programming errors, stay encrypted with previous key

ACI Fabric Policies

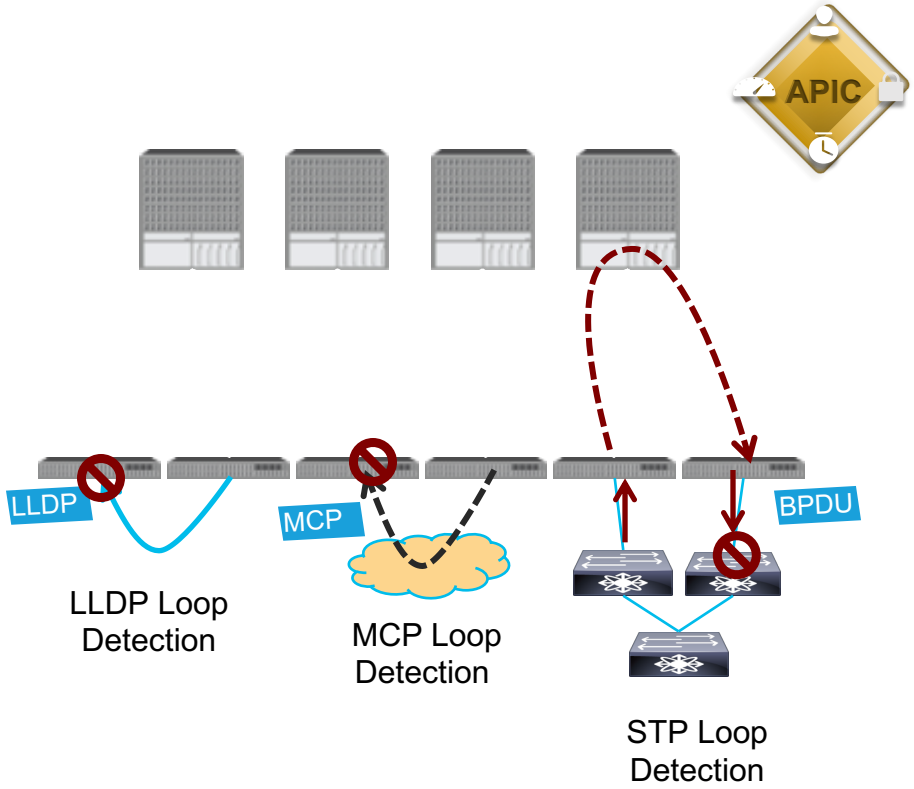
ACI Interaction with STP

- No STP running within ACI fabric
- BPDU frames are flooded between ports configured to be members of the same external L2 Outside (EPG)
  - No Explicit Configuration required
  - Hardware forwarding, no interaction with CPU on leaf or spine switches for standard BPDU frames
  - Protects CPU against any L2 flood that is occurring externally
- External switches break any potential loop upon receiving the flooded BPDU frame fabric
- BPDU filter and BPDU guard can be enabled with interface policy



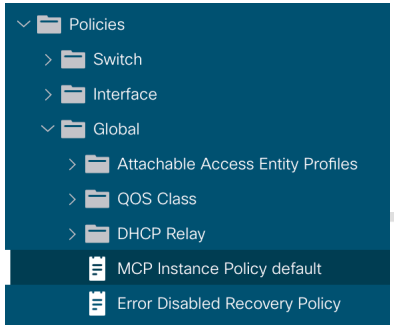
ACI Fabric Loopback Protection

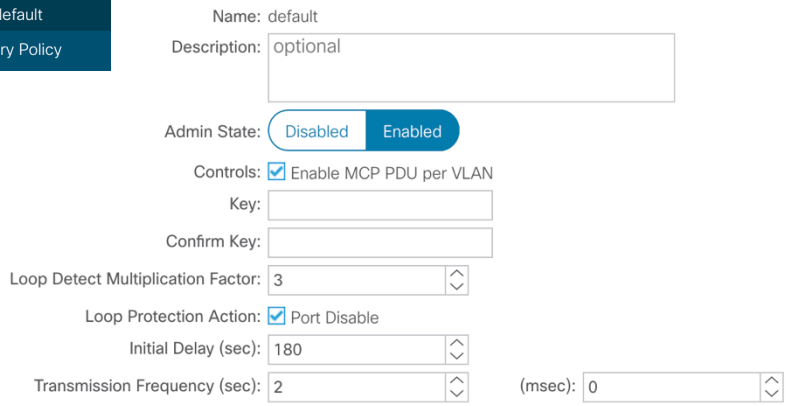
- Multiple Protection Mechanisms against external loops
- LLDP detects direct loopback cables between any two switches in the same fabric
- Mis-Cabling Protocol (MCP) is a link level loopback packet that detects an external L2 forwarding loop
  - MCP frame sent on all VLAN's on all Ports
  - If any switch detects MCP packet arriving on a port that originated from the same fabric the port is err-disabled
- External devices can leverage STP/BPDU
- MAC/IP move detection and learning throttling and err-disable



Miscabling Protocol (MCP)

- After ports go up, MCP waits before sending MCP PDUs
- This is so that Spanning-Tree IF present can converge
- *If during that time there is a temporary loop*, Rogue EP will quarantine the IP/MAC of the hosts on the BD that are experiencing the loop.
- This protects the fabric from the effects of the temporary loop

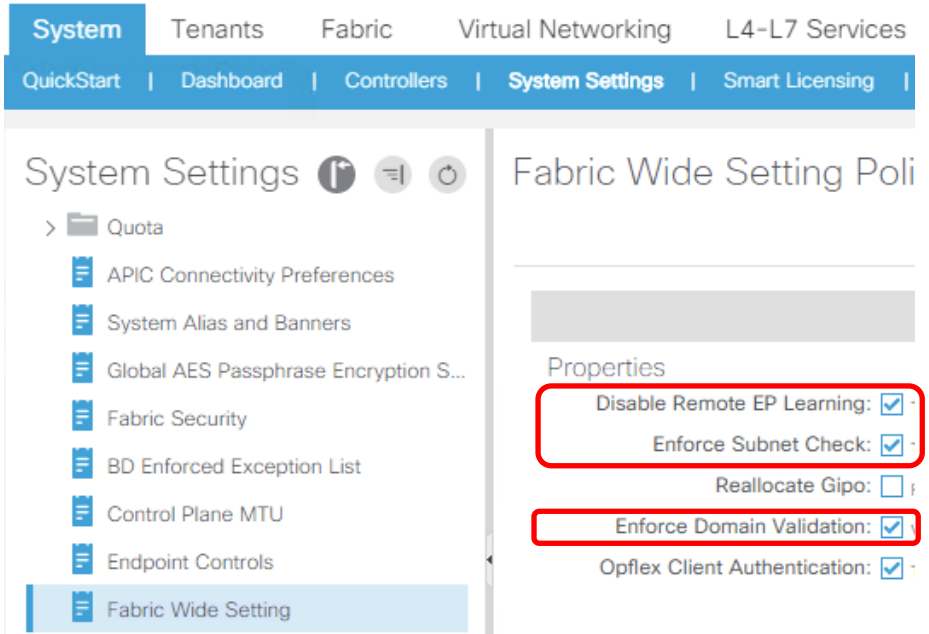




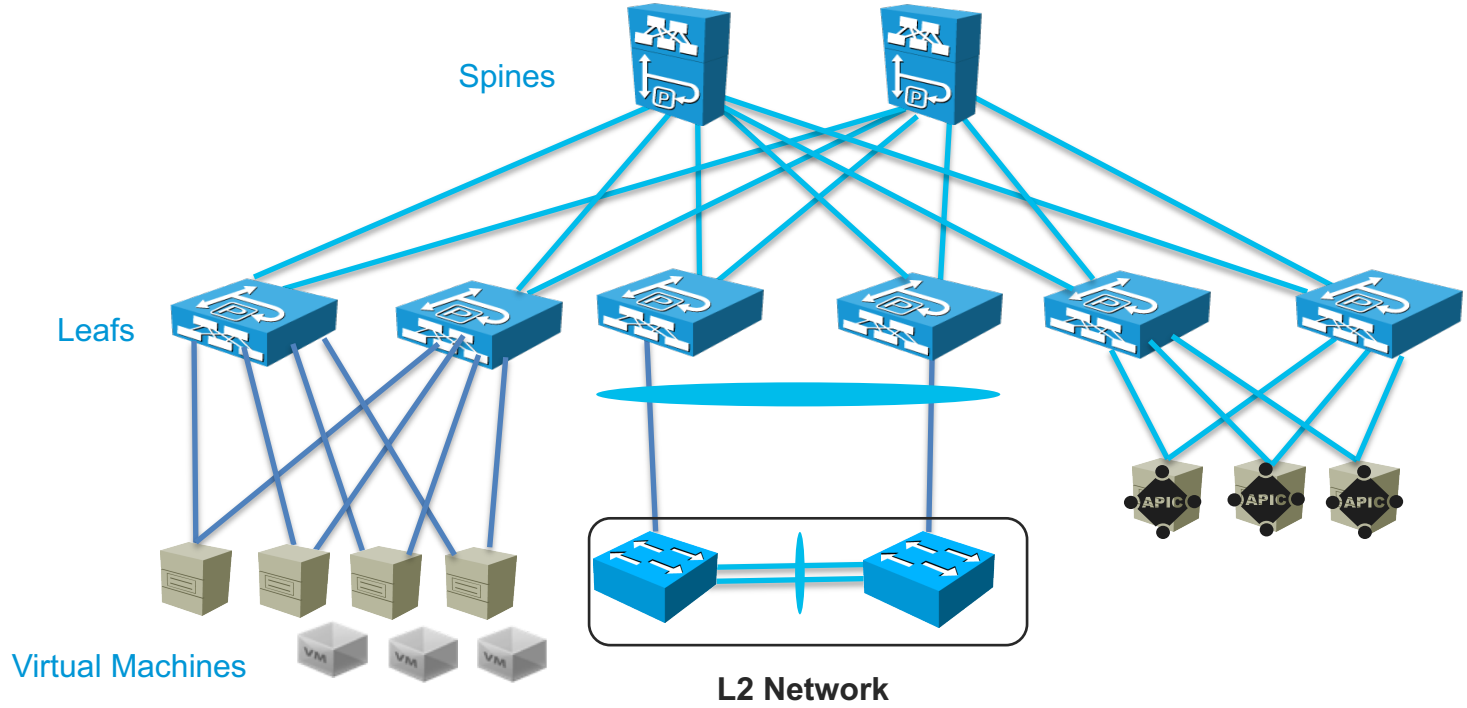
Fabric BD and EP settings

Configure Global Settings that are now considered best practices

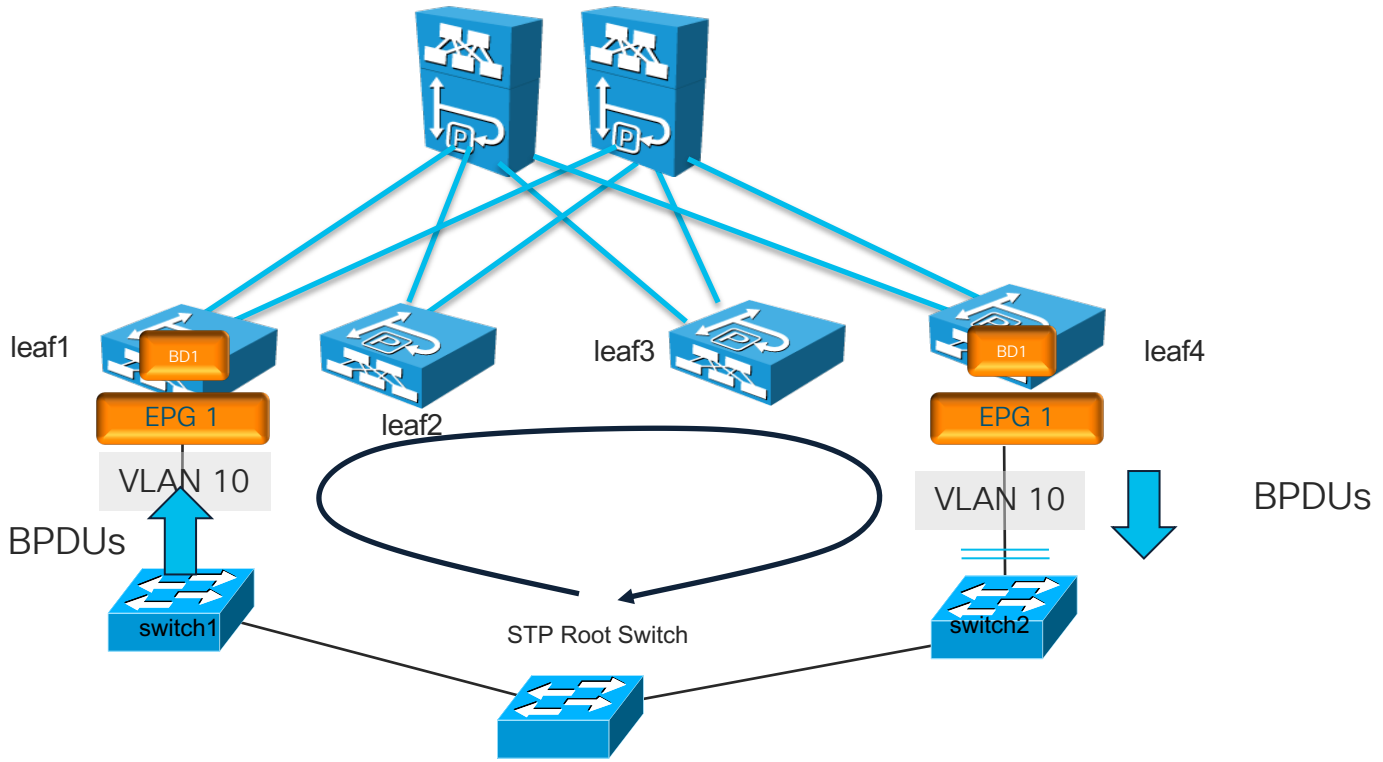
- Disable Remote EP Learn
- Enforce Subnet Check Globally
- IP Aging Enabled
- Rogue EP Protection (more details later)
- MCP per VLAN enabled
- Enable Domain Validation: this option can only be enabled because it is meant to be used always. You may have configurations that were working even if they were incorrect. Before enabling it make sure you verify the domain assignment to the EPGs and the associated AEPs.



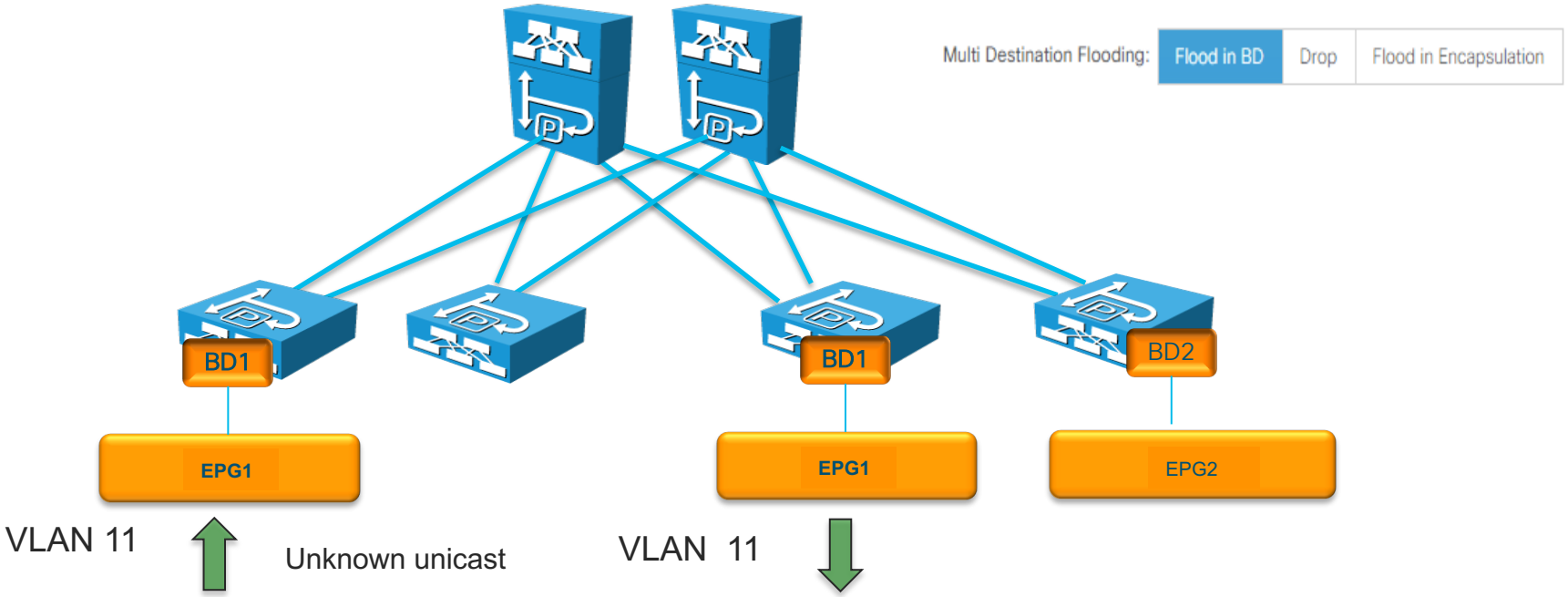
Very often the new ACI network connects to an existing network via L2, which could look like this…



ACI forwarding BPDUs allows the external switches which run STP to prevent loops

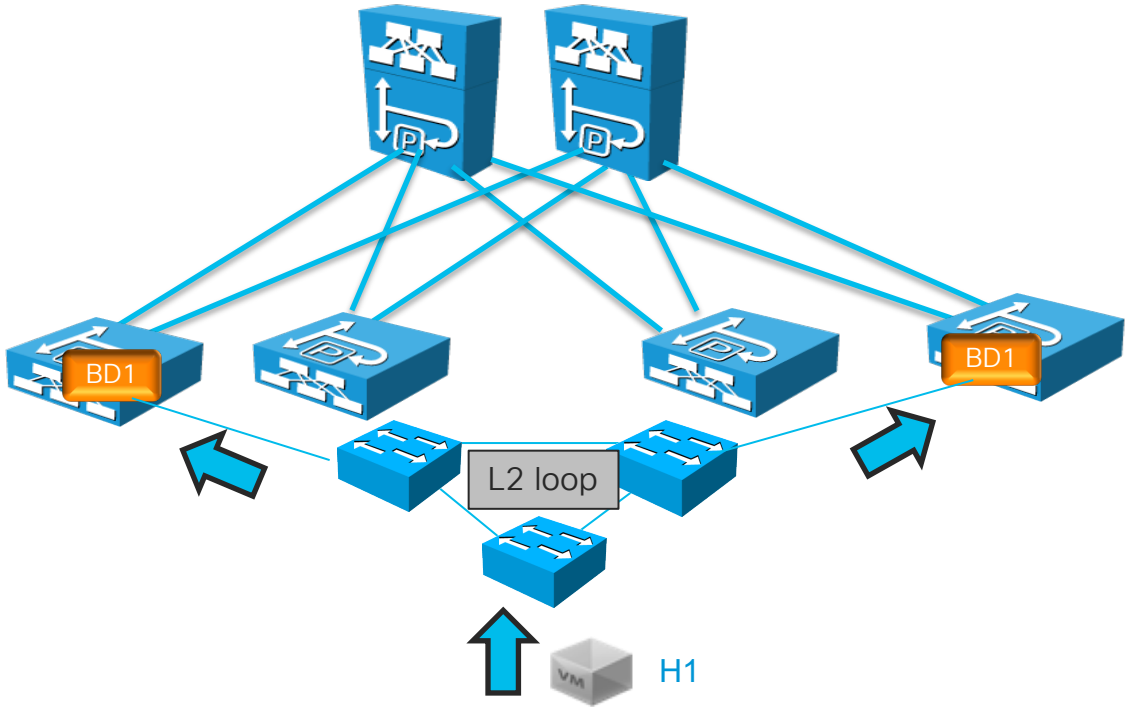


With Network Centric Deployments the flooded traffic is scoped by VLAN = BD, hence there is less risk to introduce L2 loops (compared to application centric deployments)



If a loop is occurring endpoints would be continuously flapping

- If a Loop occur like in this picture the host MAC keeps flapping between leafs like leaf1 and leaf4
- ACI has multiple way to detect loops with this type of scenario
- ACI can detect loops from EP moves and it can disable BD learning when too many EP moves occur, or it can disable the ports where the move is happening or it can quarantine the specific EP that are flapping.



EP move dampening, EP Loop Protection, *Rogue EP Detection*

	EP move dampening	EP Loop Protection	Rogue EP Detection
Scope	per BD	Global	Global
Detection	aggregate number of all moves per BD	number of moves of an individual endpoint in the specified interval	number of moves of an individual endpoint in the specified interval
Detects MAC and/or IP move	Detects MAC moves, IP moves	Detects MAC moves	Detects MAC moves, IP moves
Possible actions	BD learn disable per leaf	port disable or BD learn disable per leaf	Programs static entry to disable learning for the specific entry

What Rogue EP does and what it doesn't do

- Rogue EP protects the fabric from too many EP moves by quarantining the source
- If ACI manages to associate the EP to the correct port, then this doesn't disrupt the EP whose packet is looped from communicating

- Rogue EP doesn't prevent the replication of packets with the loop
- Packets continue to replicate when the EPs are quarantined
- MCP and BPDUs should still be used to prevent the loop

What tool to use for which purpose

- Rogue EP:
  - This is to protect the fabric against issues such as a flapping EP due to the wrong teaming configurations such as A/A TLB teaming when the BD is not configured correctly
  - This is to protect also the ACI control plane from having to manage too many EP moves which could be caused also by L2 Loops
- MCP:
  - Use MCP to protect the fabric and the external L2 network from loops.
  - MCP disables the link not the individual VLAN
- BPDUs:
  - Allow BPDUs if you want to protect the network from loops in a more granular level than MCP, i.e. per-VLAN instead of per-interface. BPDUs are forwarded within a site. Not forwarded between sites (ACI Multi-Site)

Hardening the ACI fabric to reduce the chance and/or impact of Layer 2 Loops

MCP globally enabled

MCP Instance Policy default

Admin State: Disabled Enabled

Rogue EP Control

Properties

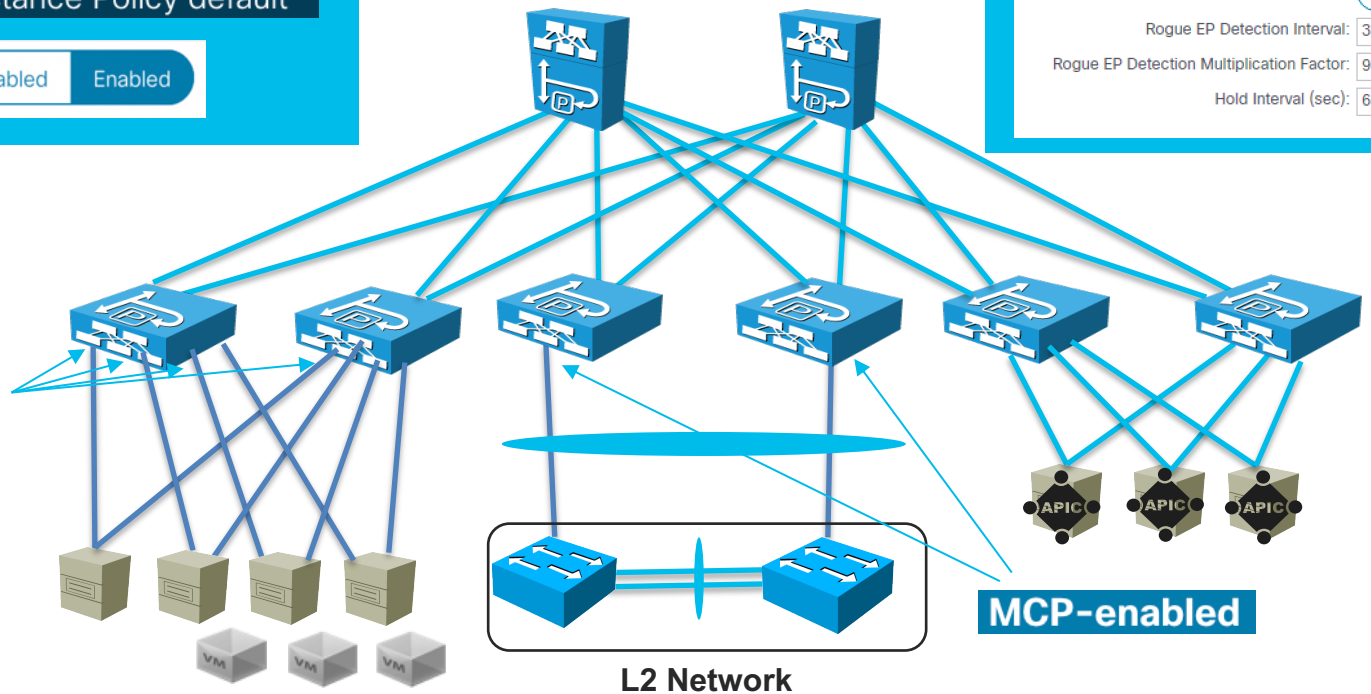
Administrative State: Disabled Enabled

Rogue EP Detection Interval: 30

Rogue EP Detection Multiplication Factor: 9

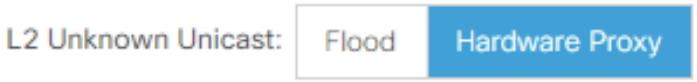
Hold Interval (sec): 600

BPDU guard



On the ACI BD, differently from traditional networks, you can optimize Unknown Unicast with Hardware-proxy

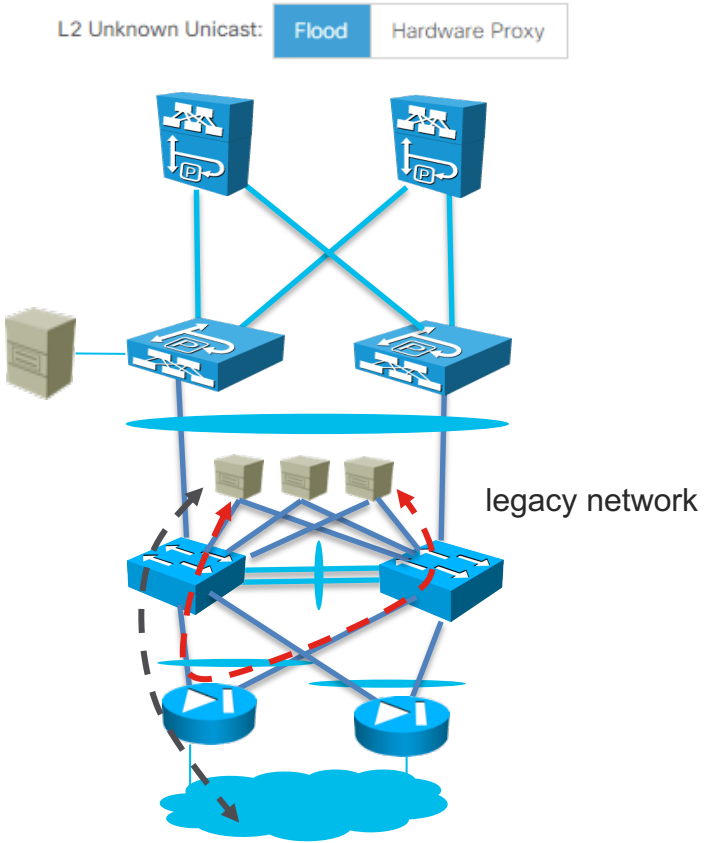
- Flooding on
- Unknown L2 unicast: flood in Bridge Domain
- If you don't enable ip routing because you want to use the BD as a pure L2 domain, then keep uucast flooding on



- "Hardware-proxy" option on
- Unknown L2 unicast:
  - If the destination is not known on the leaf, the leaf will send it to Spine proxy. If Spine also does not know the address, discard the packet.

Migration Scenarios: unknown unicast flooding makes the deployment easier

- Reason for using L2 unknown unicast flooding:
  - ACI may be used just as a L2 network and an external device provides the default gateway
  - Servers in the existing network may not send gratuitous ARPs thus ACI may not have their EP information
  - Or TCN frames from the legacy network may flush the mapping database
- But if later you need to change the BD from unknown unicast flooding to hw-proxy **make sure you have a script in place to ping the servers so that their traffic updates the mapping database**.



If ACI is used as the default gateway for the servers, then Hw-proxy on is preferred

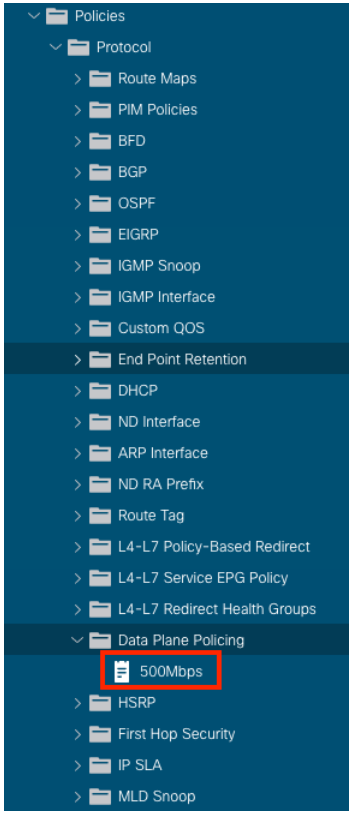
- The forwarding table on the leaf is heavily used for remote endpoints if the fabric has a lot of VMs and the servers are very chatty
- The table is programmed whenever there's an ARP flood or a flood or a unicast communication.
- Hw-proxy is recommended because **it makes the fabric scale better:**
  - If you use uucast flooding and the forwarding table is full and no remote entries can be stored, traffic to remote endpoints is flooded in the fabric
  - If you use hw-proxy the forwarding lookup happens in the spine-proxy rather than the table on the leaf, hence traffic is not flooded
- Similarly if you happen to disable dataplane learning for some reasons, hw-proxy makes the fabric forwarding optimal

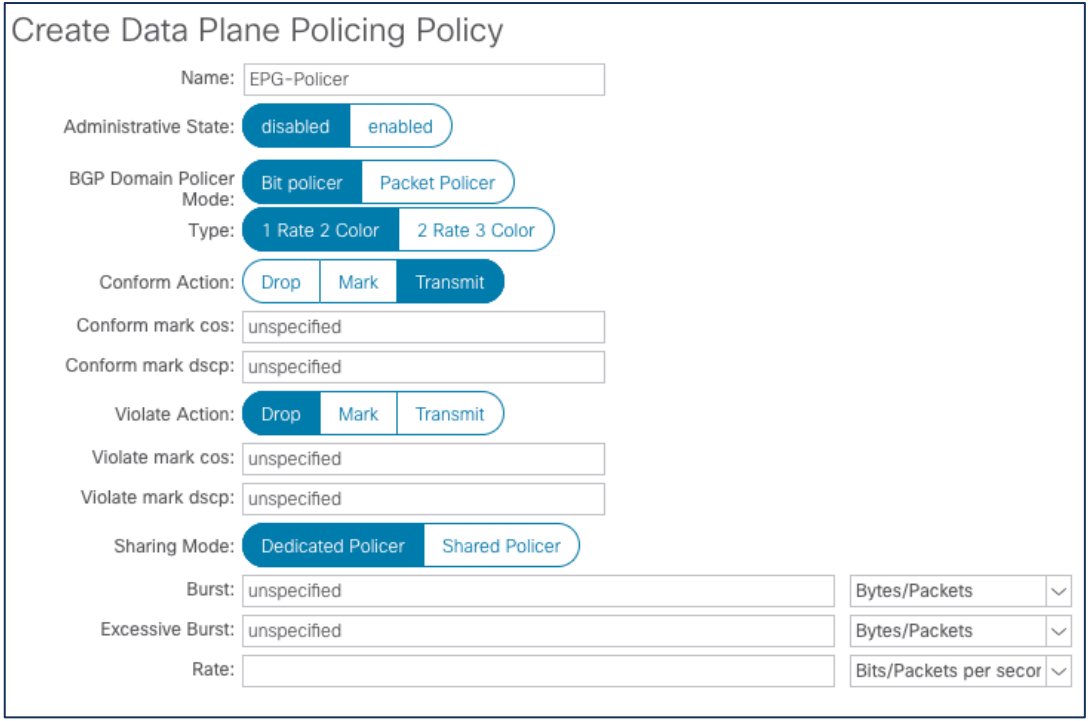
Data Plane Policing

Ingress Policing Per EPG

- Ability to limit traffic on a port at EPG granularity
- If a leaf interface is shared by multiple EPGs per EPG policing can prevent a single EPG from monopolizing the bandwidth of the link
- Policer policies defined at the tenant and can be used across EPGs. Enforcement is at the interface level

Data Plane Policing Policy





Quota Management

ACI Quota Management

ACI Quota for Managed Objects

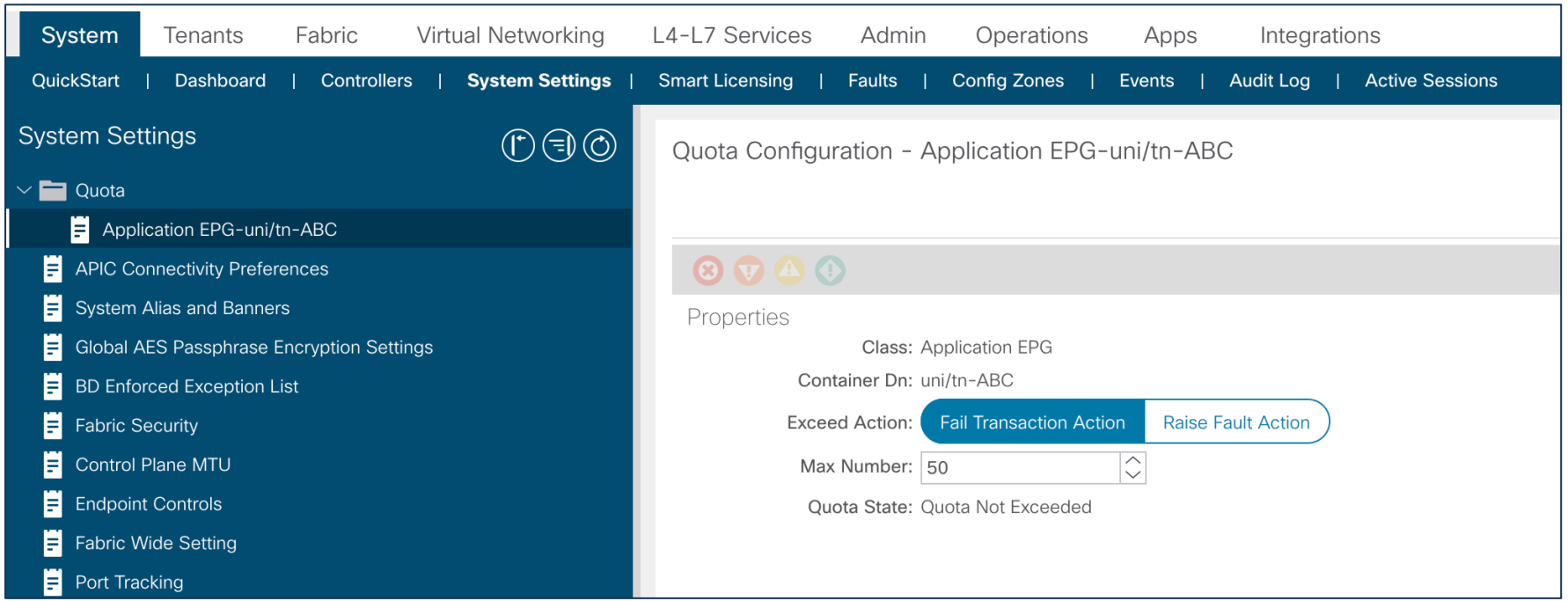
Not all configuration uses hardware resources when configured. The ACI quota management feature allows the administrator to specify quota limits for logical object configuration

- In ACI every configuration is a Managed Object (MO)
- This feature allows the admin to limit the number of MOs globally or at a subtree (example, Tenant)

Quota Management supported objects

- Application EPG
- Bridge Domain
- VRF
- Contract Consumer
- Contract Interface
- Taboo Contract Association
- Contract Provider
- Subnet
- External Network Instance Profile
- L3 Outside

ACI Quota Management Configuration




CISCO