Cisco

Cisco live! Jan 30 - Feb 3, 2012 • London, UK

Redundancy Mechanisms for Carrier Ethernet Networks and Layer 2 VPN Services

BRKSPG-2611

Housekeeping

- We value your feedback- don't forget to complete your online session evaluations after each session & the Overall Conference Evaluation which will be available online from Thursday
- Visit the World of Solutions and Meet the Engineer
- Visit the Cisco Store to purchase your recommended readings
- Please switch off your mobile phones
- After the event don’t forget to visit Cisco Live Virtual: www.ciscolivevirtual.com

Agenda

- Introduction
- Resiliency Fundamentals
- Access Resiliency Mechanisms
- Aggregation and Core Resiliency Mechanisms
- MAC Flushing Mechanisms
- Redundancy Solutions
- Summary

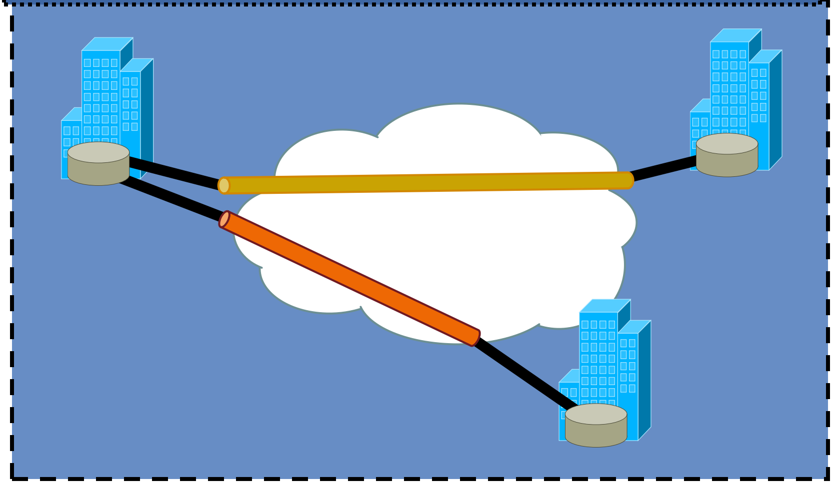
Learn. Connect. Collaborate. together.

Introduction

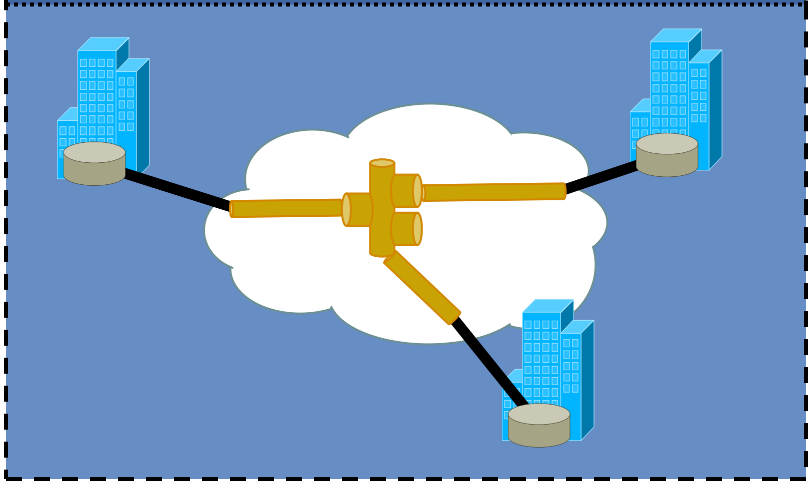
Carrier Ethernet Services

Metro Ethernet Forum (MEF) Service Visualization

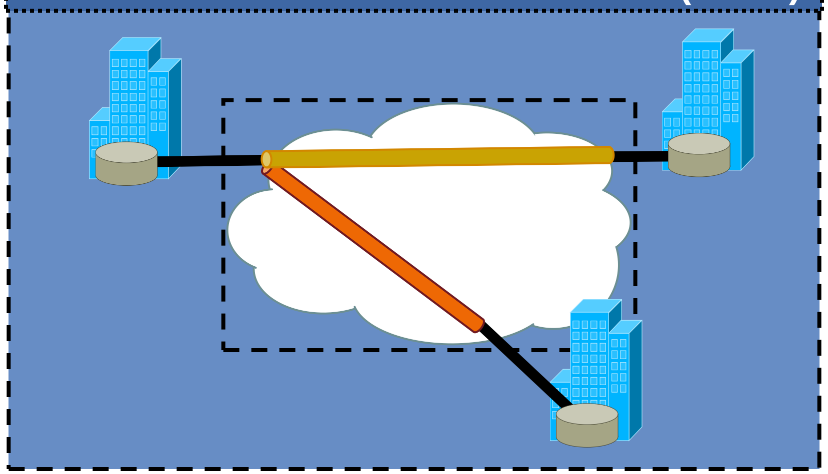
E-LINE: Ethernet Private Line (EPL)



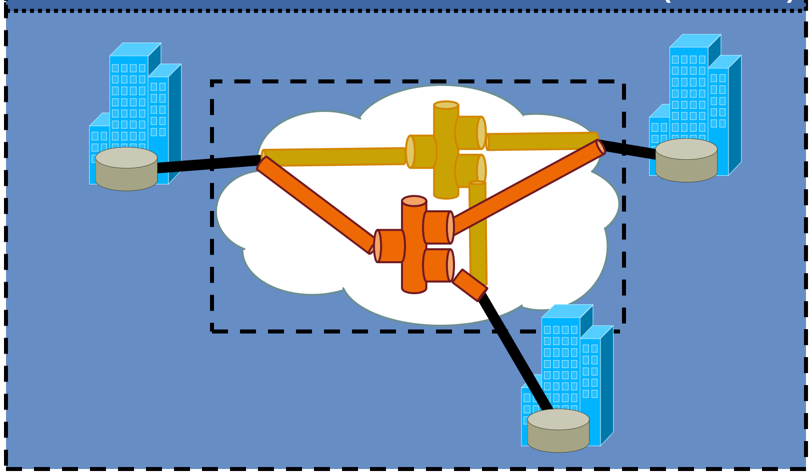
E-LAN: Ethernet Private LAN (EP-LAN)



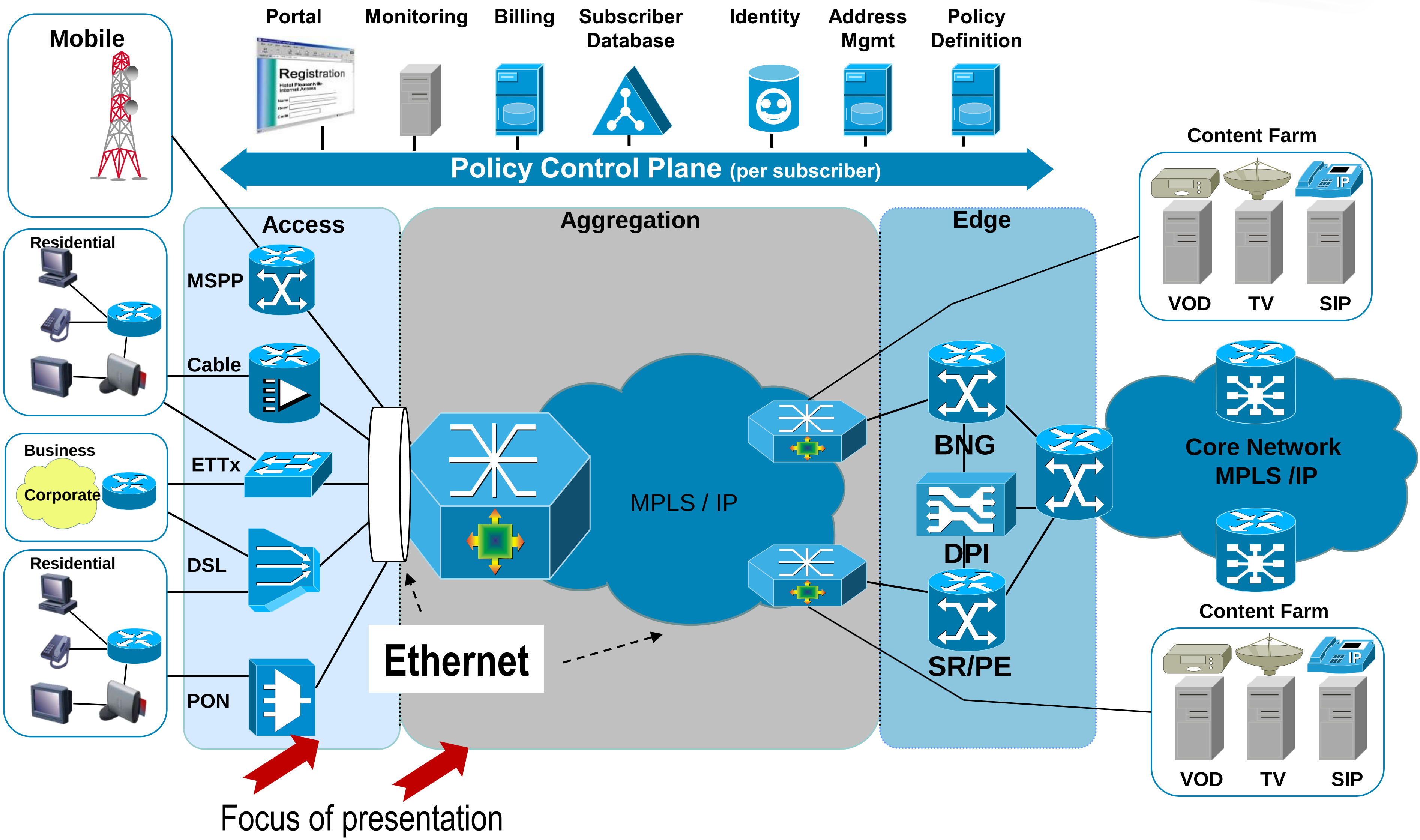
E-LINE: Ethernet Virtual Private Line (EVPL)



E-LAN: Ethernet Virtual Private LAN (EVP-LAN)



Carrier Ethernet Networks



Learn. Connect.
Collaborate. together.

Resiliency Fundamentals

Resiliency Fundamentals

- Resiliency definition from Metro Ethernet Forum:

  "A self-healing property of the network that allows it to continue to function with minimal or no impact to the network users upon disruption, outages or degradation of facilities or equipment in the MEN" [MEF-2]

- User's perspective

  SLA attributes such as:

  - Availability
  - Mean Time To Restore (MTTR)
  - Mean Time Between Failure (MTBF)

  Actual methods and mechanisms used by SP not relevant

- Provider's perspective

  Translation of SLAs to network protection requirements

  Selection of mechanisms / protocols to provide such protection

Ethernet-Aware Resiliency Mechanisms

Key Requirements

- MUST NOT allow data-plane loops

  Not even transient ones, as Ethernet header has no Time To Live (TTL) or equivalent field

- MUST ensure congruency of forward and reverse data-plane paths

  Prevent MAC moves in scenarios with Load Balancing

- MUST ensure a unique entry/exit point into an Ethernet segment

  Prevent delivery of duplicate packets - Designated Forwarder notion

- MUST ensure MAC-relearning after topology change notification

  Prevent black-holing of traffic - MAC address tables must be updated after re-convergence events

Ethernet-Aware Resiliency Mechanisms

Generic Requirements

- Failure type requirements
  - Link failures (hard and soft (degrade) conditions)
  - Node failures
- Failure detection requirements
- Failure notification requirements
- Protection switching requirements
  - Connectivity Restoration Time (i.e. Recovery Time)
  - SLS Restoration Time (i.e. Full Restoration Time)
- Protection resource allocation requirements
  - 1+1, 1:1, n:1, m:n, 1:n
- Topology requirements
  - Hub and spoke / rings
- Resource selection requirements
  - Revertive mode
  - Controls – manual switch / forced switch / lockout
- End-user transparency

Ethernet-Aware Resiliency Mechanisms

- Ethernet Virtual Circuits (EVC) implementing an Ethernet service usually traverse different transports
- End-to-end protection involves different resiliency mechanisms (sometimes even layered ones – layered protection)
  - The lower the layer, the faster the protection
  - The higher the layer the longer the path that can be protected
- This presentation covers different resiliency mechanisms used in the access and aggregation/core layers of a Carrier Ethernet Network and the interactions among them

Access Resiliency Mechanisms

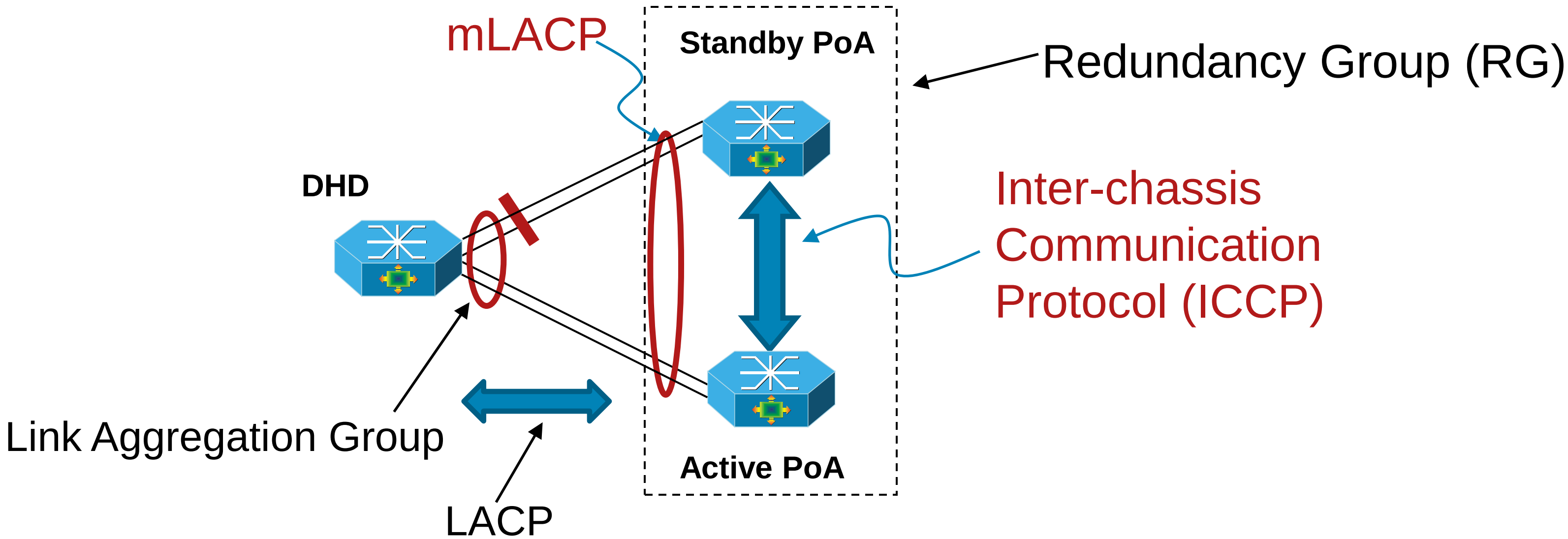
Learn. Connect. Collaborate. together.

Access Resiliency Mechanisms

Multi-Chassis LACP (mLACP) and Inter-Chassis Communication Protocol (ICCP)

Multi-Chassis LACP and ICCP Overview

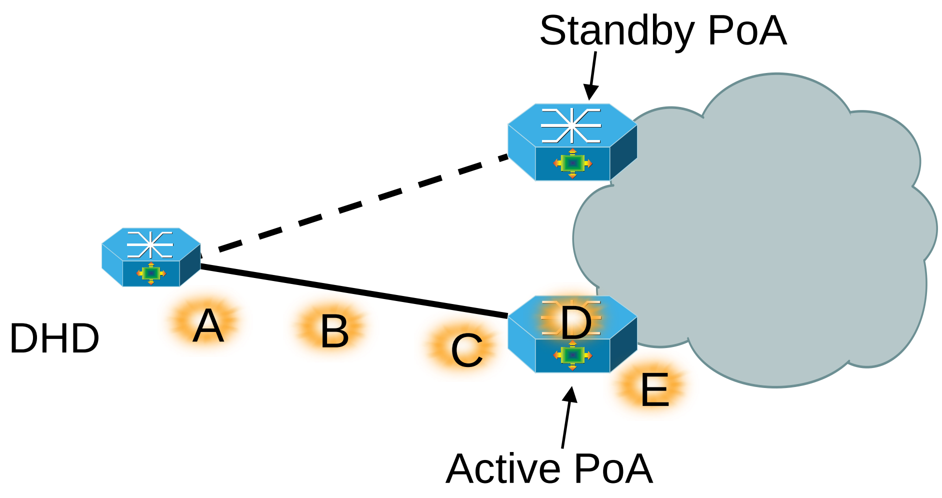
- mLACP & ICCP enable a switch/router to use standard Ethernet Link Aggregation for device dual-homing, with active/standby redundancy
- Dual-homed Device (DHD) operates as if it is connected to single virtual device and runs IEEE std. 802.1AX-2008 (LACP)
- Point of Attachment (PoA) nodes run Inter-chassis Communication Protocol (ICCP) to synchronize state & form a Redundancy Group (RG)



Protected Failure Points

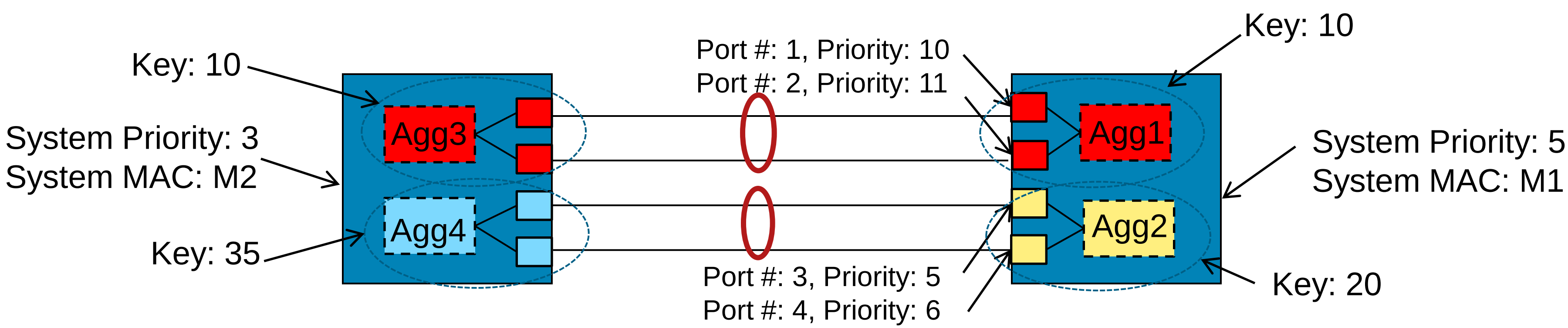
mLACP Offers Protection Against 5 Failure Points:

- A: DHD Port Failure
- B: DHD Uplink Failure
- C: Active PoA Port Failure
- D: Active PoA Node Failure
- E: Active PoA Isolation from Core Network



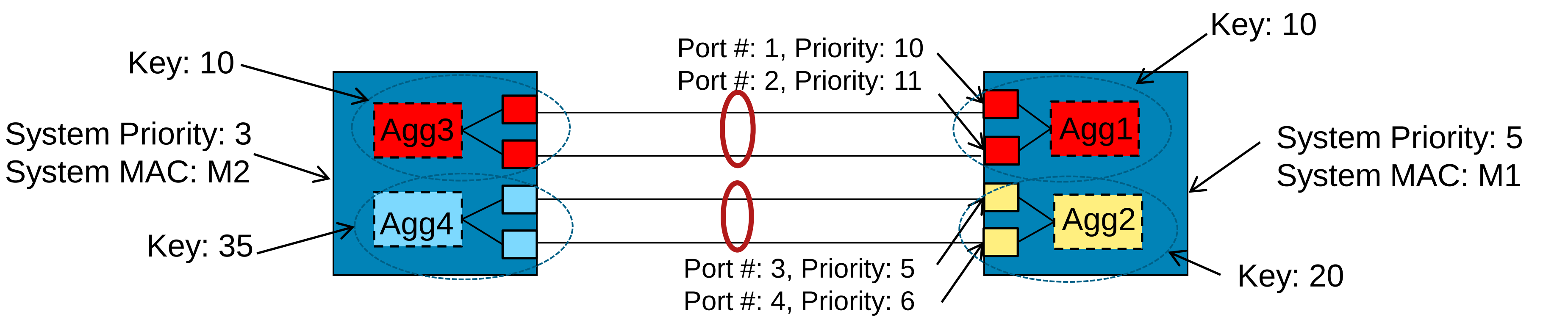
Background: Link Aggregation Control Protocol

- System attributes:
  - System MAC address: MAC address that uniquely identifies the switch
  - System priority: determines which switch's Port Priority values win
- Aggregator (bundle) attributes:
  - Aggregator key: identifies a bundle within a switch (per node significance)
  - Maximum links per bundle: maximum number of forwarding links in bundle – used for Hot Standby configuration
  - Minimum links per bundle: minimum number of forwarding links in bundle, when threshold is crossed the bundle is disabled



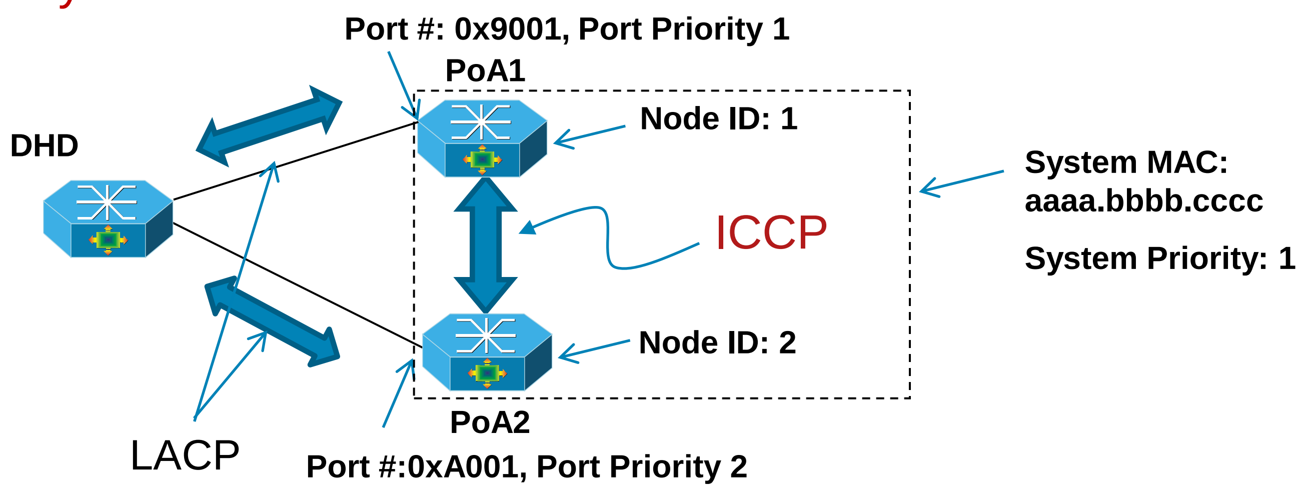
Background: Link Aggregation Control Protocol (Cont.)

- Port attributes:
  - Port key: defines which ports can be bundled together (per node significance)
  - Port priority: specifies which ports have precedence to join a bundle when the candidate ports exceed the Maximum Links per Bundle value
  - Port number: uniquely identifies a port in the switch (per node significance)



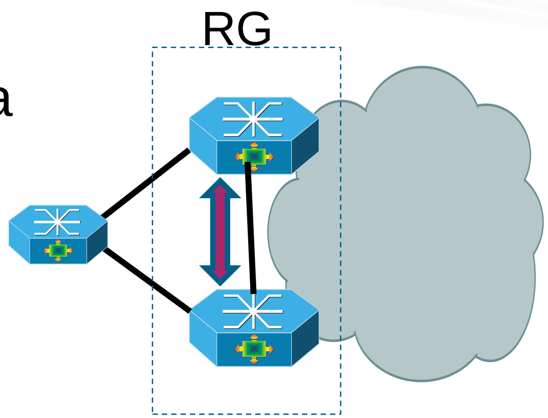
Extending LACP Across Multi-Chassis: mLACP

- mLACP uses ICCP to synchronize LACP configuration & operational state between PoAs, to provide DHD the perception of being connected to a single switch
- All PoAs use the same System MAC Address & System Priority when communicating with DHD
  - Configurable or automatically synchronized via ICCP
- Every PoA in the RG is configured with a unique Node ID (value 0 to 7). Node ID + 8 forms the most significant nibble of the Port Number
- For a given bundle, all links on the same PoA must have the same Port Priority

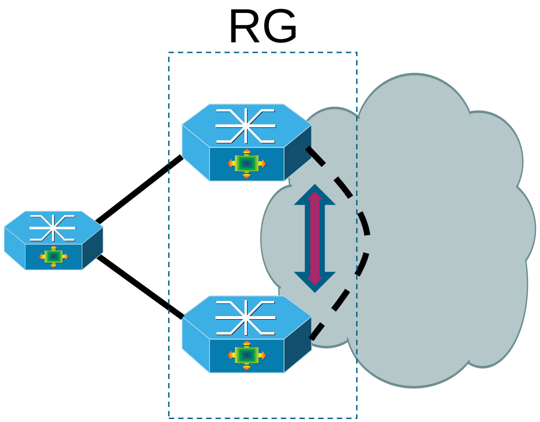


Inter-Chassis Communication Protocol

- ICCP allows two or more devices to form a ‘Redundancy Group’
- ICCP provides a control channel for synchronizing state between devices
- ICCP uses TCP/IP as the underlying transport
  - ICCP rides on targeted LDP session, but MPLS need not be enabled
- Various redundancy applications can use ICCP:
  - mLACP
  - Pseudowire redundancy
- Under standardization in IETF:
  - draft-ietf-pwe3-iccp-05.txt



ICCP over Dedicated Link



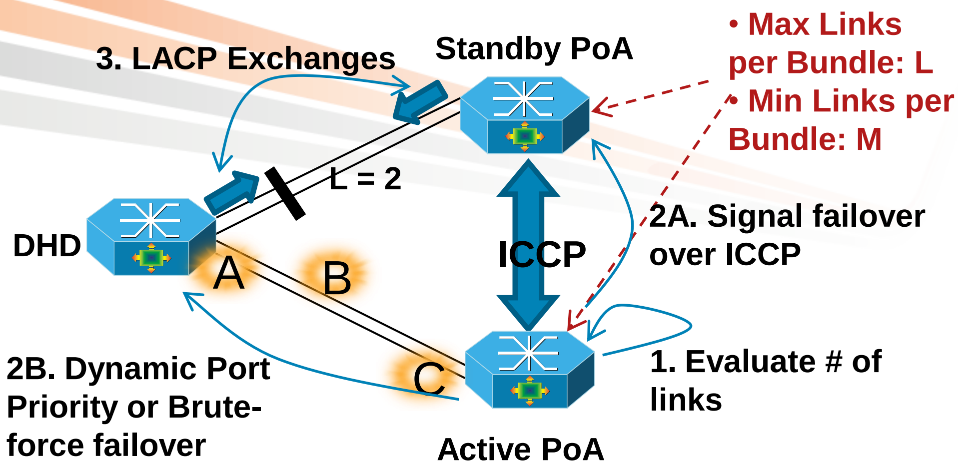
ICCP over Shared Network

Operational Variants

Variant	DHD Configuration	PoA Configuration	Advantages	Trade-Offs
DHD-based Control	Limits Max. No. of Links per Bundle (LM)	Limits Min. No. of Links per Bundle (must be set to LM)	Handle split-brain condition	Failover time depends on DHD implementation
PoA-based Control		Limits Max. No. of Links per Bundle	•Fast switchover •Flexible Min. Link policy on PoA	Susceptible to split brain problem if ICCP transport is not protected
Shared Control	Limits Max. No. of Links per Bundle	Limits Max. No. of Links per Bundle	•Handle split-brain condition •Flexible Min. Link policy on PoA	Failover time depends on DHD implementation

Failover Operation

Port/Link Failures



Step 1 – For port/link failures, active PoA evaluates number of surviving links (selected or standby) in bundle:

- If > M, then no action
- If < M, then trigger failover to standby PoA

Step 2A – Active PoA signals failover to standby PoA over ICCP

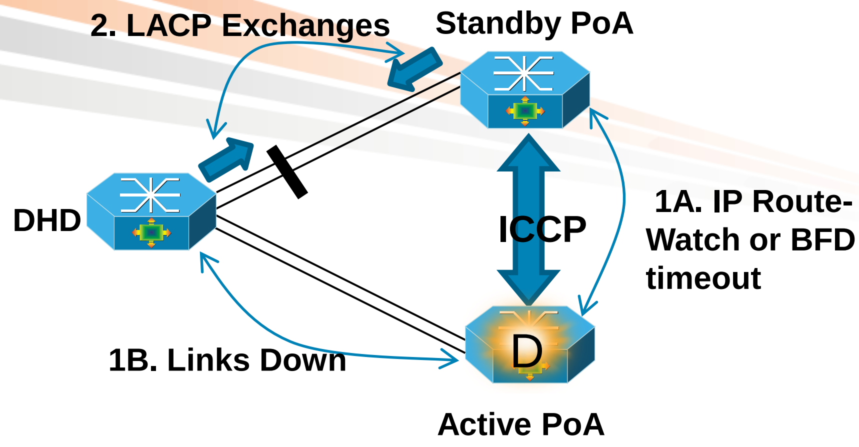
Step 2B – Failover is triggered on DHD by one of:

- Dynamic Port Priority Mechanism: real-time change of LACP Port Priority on active PoA to cause the standby PoA links to gain precedence
- Brute-force Mechanism: change the state of the surviving links on active PoA to admin down

Step 3 – Standby PoA and DHD bring up standby links per regular LACP procedures

Failover Operation

Node Failure



Step 1A – Standby PoA detects failure of Active PoA via one of:

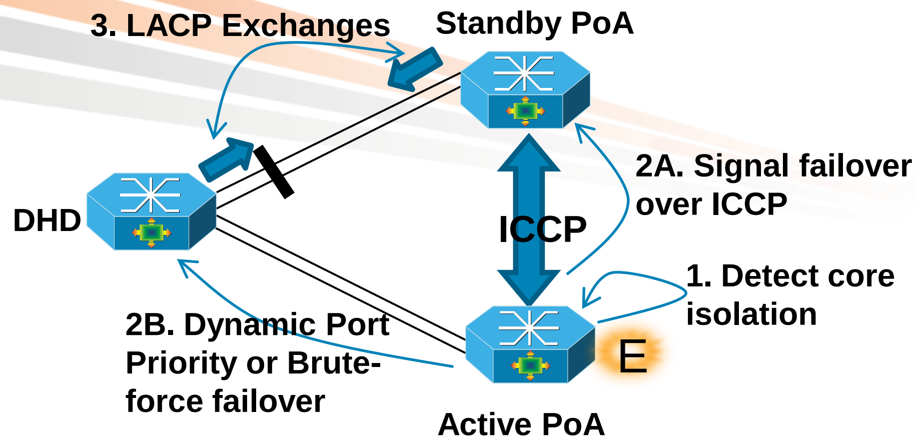
- IP Route-watch: loss of IP routing adjacency
- BFD: loss of BFD keepalives

Step 1B – DHD detects failure of all its uplinks to previously active PoA

Step 2 – Both Standby PoA and DHD activate their Standby links per regular LACP procedures

Failover Operation

PoA Isolation from Core



Step 1 – Active PoA detects all designated core interfaces are down

Step 2A – Active PoA signals standby PoA over ICCP to trigger failover

Step 2B – Active PoA uses either Dynamic Port Priority or Brute-force Mechanism to signal DHD of failover

Step 3 – Standby PoA and DHD bring up standby links per regular LACP procedures

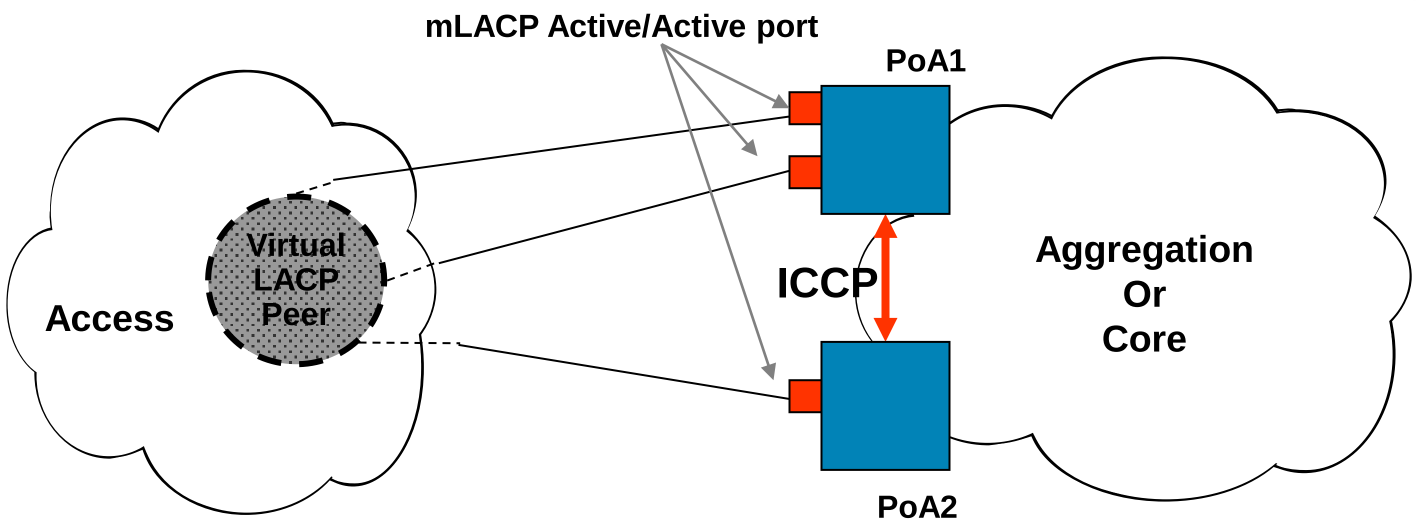
mLACP/ICCP Advantages

- Allow dual-homing of access node that doesn't support spanning-tree (e.g. Router CE or DSLAM)
- Support co-located and geo-redundant PEs
- Support revertive and non-revertive operation
- Standards based solution using IEEE 802.1AX and draft-ietf-pwe3-iccp

Access Resiliency Mechanisms

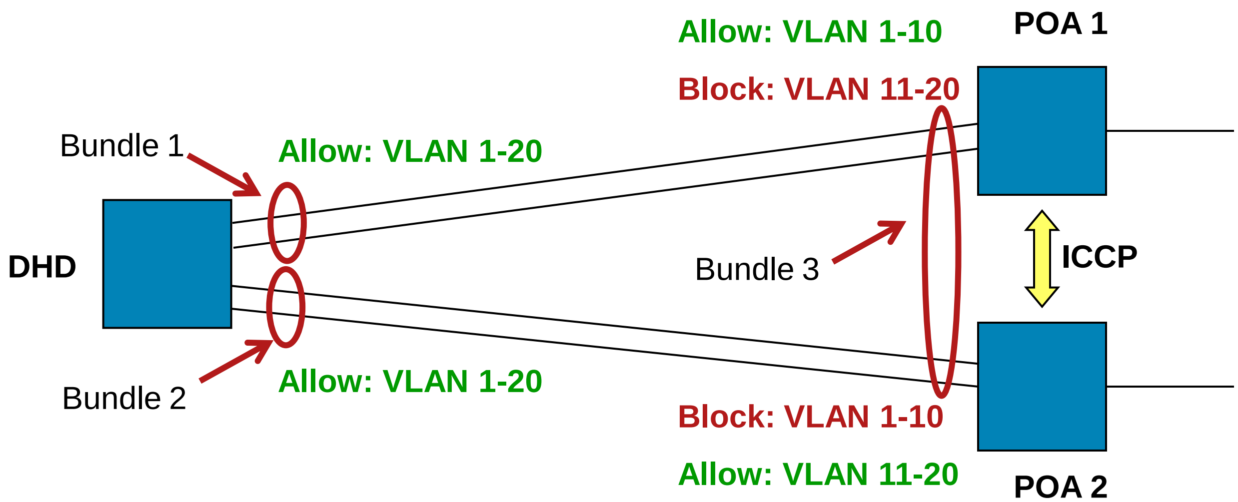
mLACP Active/Active (per VLAN Load-Balancing)

Conceptual Model



- PoA ports are configured to assume mLACP Active/Active (mLACP-AA) role:
  - Ports act as if connected to a virtual device over an MC-LAG with mLACP
  - Ports placed in Active/Active Mode with manual VLAN load-balancing
- Access node(s) perceive the ports/links as being independent.
- Supports Dual Homed Device (DHD)

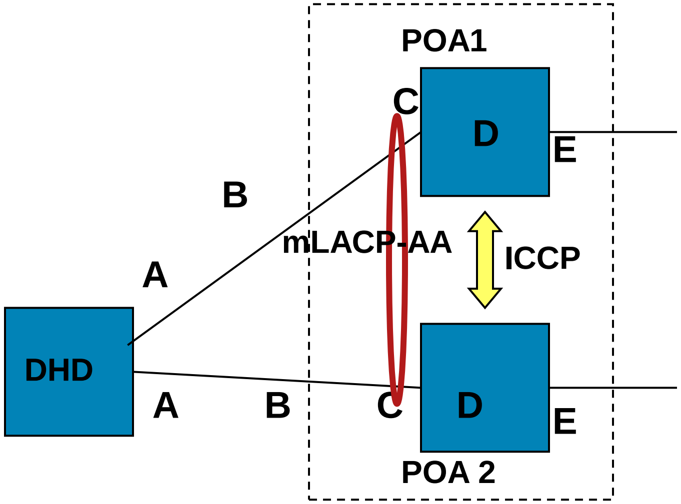
Setup



- DHD configures all uplinks towards a single POA in a bundle (LAG)
  - Links towards different POAs belong to different bundles
- DHD enables all VLANs on both bundles to PoAs
- POAs configured to allow certain VLANs and block others
  - A given VLAN can be active on a single PoA at a time
  - Per VLAN load-balancing
- Traffic from DHD to core initially flooded to both PoAs until DHD learns which bundle is active for what VLANs

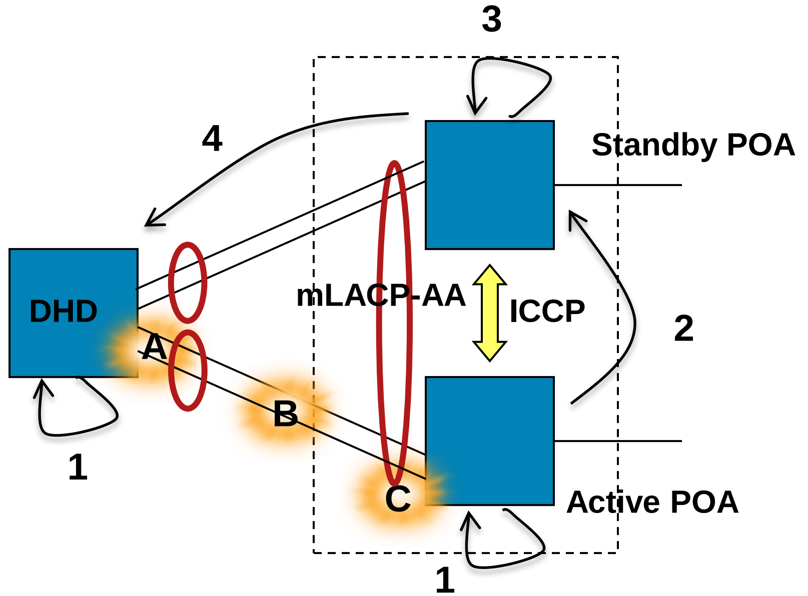
Fault Protection Points

- Provide Protection Against 5 Failure Points:
  - –A: DHD Uplink Port Failure
  - –B: DHD Uplink Failure
  - –C: POA Downlink Port Failure
  - –D: POA Node Failure
  - –E: POA Isolation from core network



Failure Procedures

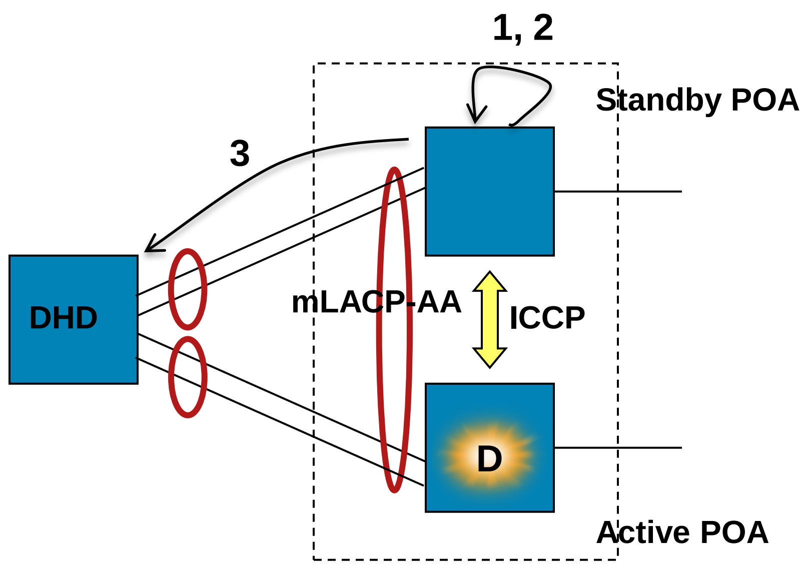
For Failure Points A, B, and C



1. DHD & Active POA detect port down
2. Active POA signals switchover to Standby via ICCP
3. Standby unblocks affected VLANs over downlink and flushes its MAC tables
4. Standby triggers Multiple VLAN Registration Protocol (MVRP) 'new' declaration towards DHD to induce MAC flushing

Failure Procedures

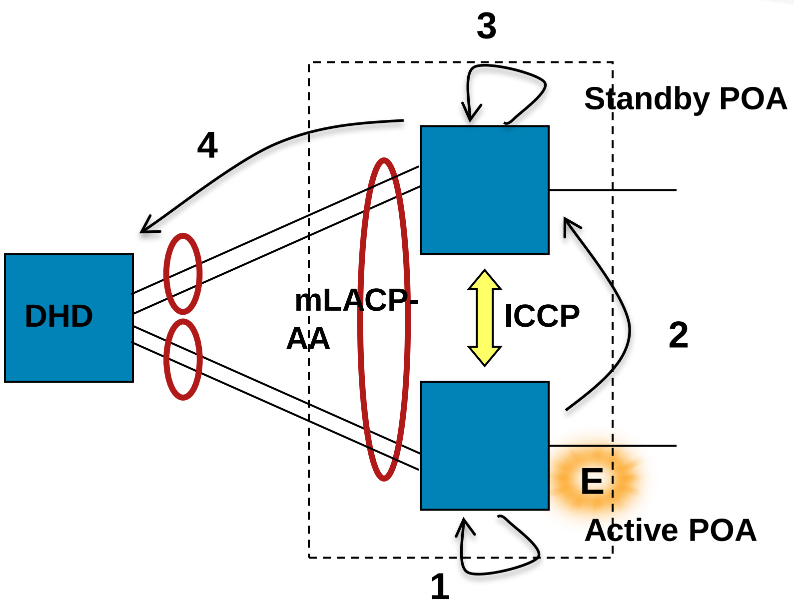
For Failure D



1. Standby POA detects failure of active POA via IP Route-Watch or BFD
2. Standby POA unblocks affected VLANs over downlink
3. Standby POA flushes its MAC tables & triggers MVRP MAC flush notification towards DHD

Failure Procedures

For Failure E



1. Active POA detects isolation from core, blocks its previously active VLANs
2. Active POA informs standby POA of need to failover via ICCP
3. Standby POA activates (unblocks) affected VLANs on downlink and flushes its MAC tables
4. Standby POA triggers MVRP registrations with ‘new” bit set (for affected VLANs) towards DHD to trigger MAC flushing.

Access Resiliency Mechanisms

Ethernet Ring Protection (ITU-T G.8032)

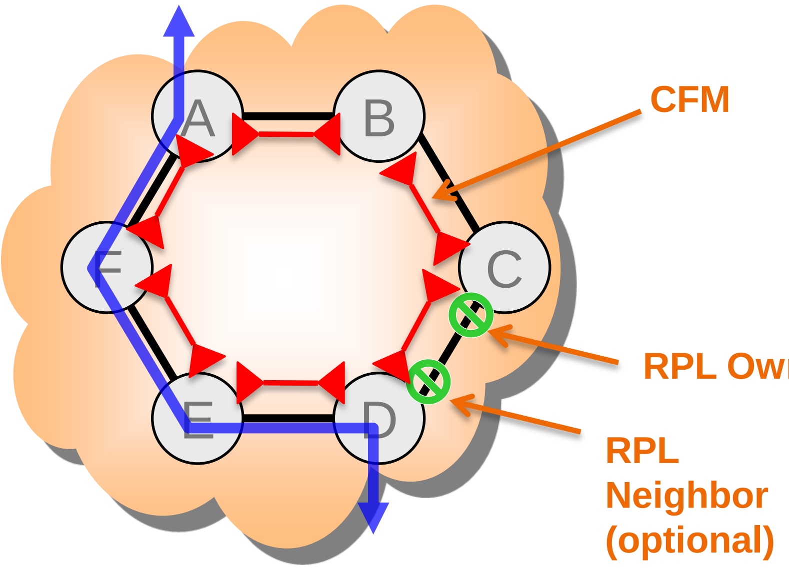
Overview

- Protection switching at Ethernet layer
  - Fast convergence (50 ms) with HW support
- Leverage Ethernet CFM (ITU-T Y.1731) for
  - Fault Detection (IEEE 802.1ag Continuity Check Message - CCM)
  - Control Channel (R-APS)
- Topology Support
  - Closed Ring
  - Open Ring (G.8032 v.2)
  - Cascaded Rings (Ladder Network) (G8032 v.2)
- Load Balancing (multi-instance support) (G.8032 v.2)
- Administrative Tools (G.8032 v.2)
  - Manual Switchover
  - Forced Switchover

Setup and Basic Operation

Setup

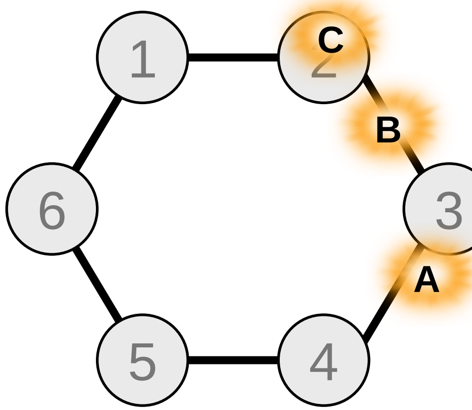
- Map VLANs into Ethernet Ring Protection (ERP) Instances
- Select Ring Protection Link (RPL) per instance and configure ports as RPL owners
- Optionally: Configure RPL Neighbor ports
- Use CFM Down MEPs to monitor link faults via CCMs



Normal Operation

- When no faults, RPL Owner (and neighbor) are blocked.
- RPL Owner (& neighbor) send R-APS message with No Request/Link Blocked every 5 sec.

Protected Failure Points

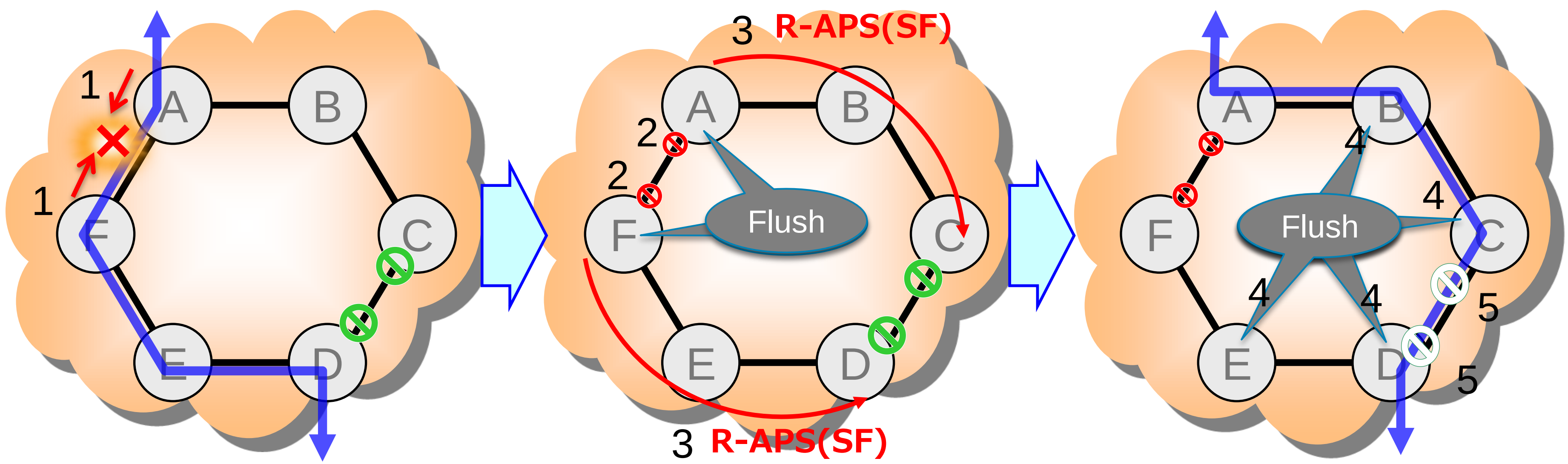


G.8032 protects against any single Link, Port or Node failure within a ring

- A: Failure of a port within the ring
- B: Failure of a link within the ring
- C: Failure of a node within the ring

Failure Handling

1. Switches detect link failure via:
   - Link Down Event (PHY based)
   - Loss of CFM CCMs
2. Switches block ports connected to failed link & flush MAC tables
3. Send R-APS messages with Signal Fail (SF) code on other ring port
4. Switches receiving R-APS (SF) flush their MAC forwarding tables
5. RPL Owner (and neighbor) unblock their ports



Administrative Tools

- Forced Switch (FS)
  - Allows operator to block a particular ring port
    - Effective even if there is existing SF condition
    - Multiple FS commands supported per ring
    - May be used to allow immediate maintenance operations
- Manual Switch (MS)
  - Allows operator to block a particular ring port
    - Not effective if existing FS or SF condition
    - Overridden by new FS and SF conditions
    - New MS commands are ignored
- Clear
  - Cancels an existing FS/MS command on the ring port
    - May be used (at RPL Owner Node) to trigger reversion

R-APS — Control Channel

- R-APS message format based on ITU-T Y.1731

  Opcode = 40 (R-APS)

- Sent to well-known multicast MAC address

  MAC DA = 01-19-A7-00-00-[Ring ID]

  For time being, only Ring ID = 0x01 is allowed per standard

- R-APS messages for different ERP instances must use different VLANs

<table>
<tr><td></td><td colspan="8">1</td><td colspan="8">2</td><td colspan="8">3</td><td colspan="8">4</td></tr>
<tr><td></td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td></tr>
<tr><td>1</td><td colspan="3">MEL</td><td colspan="5">Version (1)</td><td colspan="8">OpCode (R-APS = 40)</td><td colspan="8">Flags (0)</td><td colspan="8">TLV Offset (32)</td></tr>
<tr><td rowspan="2">5</td><td colspan="4" rowspan="2">Request /State</td><td colspan="4" rowspan="2">Sub-code</td><td colspan="8">Status</td><td colspan="16" rowspan="2">Node ID (6 octets)</td></tr>
<tr><td>RB</td><td>DNF</td><td>BPR</td><td colspan="5">Status Reserved</td></tr>
<tr><td>9</td><td colspan="32">Node ID</td></tr>
<tr><td>13</td><td colspan="32">Reserved 2 (24 octets)</td></tr>
<tr><td>…</td><td colspan="32">…</td></tr>
<tr><td>37</td><td colspan="32">[optional TLV starts here; otherwise End TLV]</td></tr>
<tr><td>last</td><td colspan="24"></td><td colspan="8">End TLV (0)</td></tr>
</table>

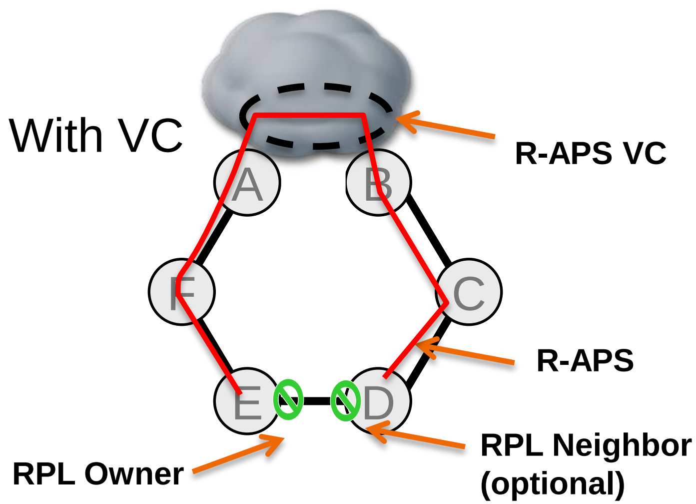
Indicates whether eastbound or westbound port is blocked

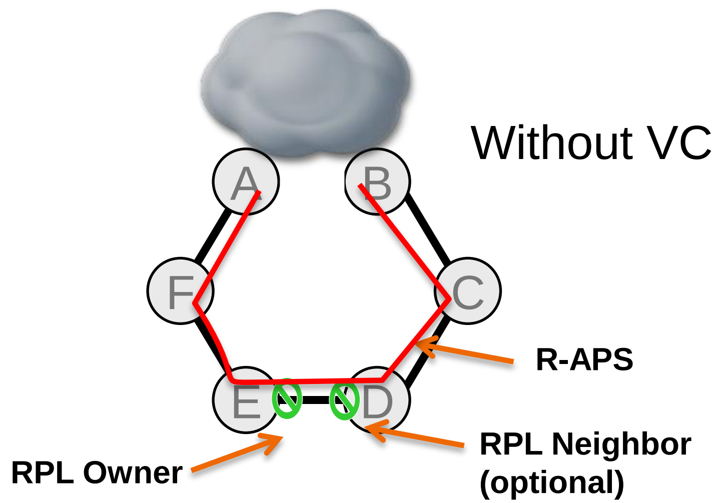
MAC Address to uniquely identify the transmitting switch

Open Ring Support

Two Solutions:

- Open ring with R-APS Virtual Channel (VC)
  - R-APS messages flow over a virtual channel supplied by another network to close the ring control channel
  - Ring is closed from control perspective but open from data perspective
- Open ring without R-APS Virtual Channel (VC)
  - Special handling of R-APS on the ring: R-APS control messages can pass over the RPL to reach all nodes
  - Requires independent blocking of control vs. data channels on RPL owner/neighbor





Interconnecting Rings

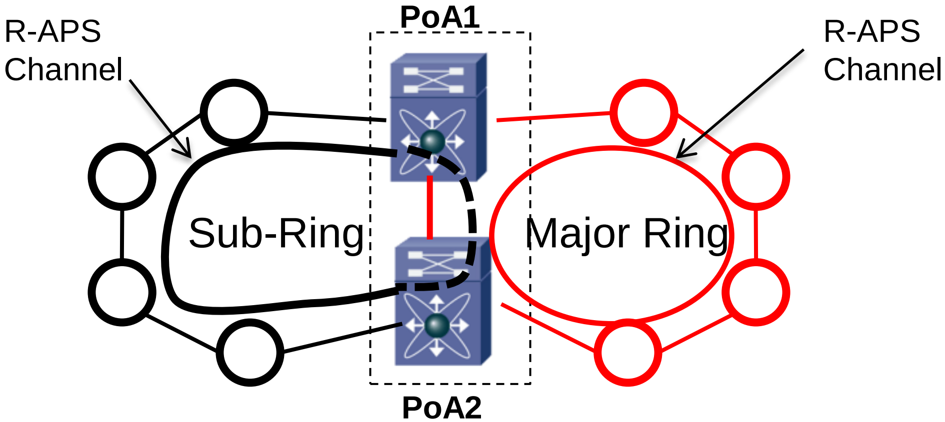
- Networks can be constructed out of closed and open rings

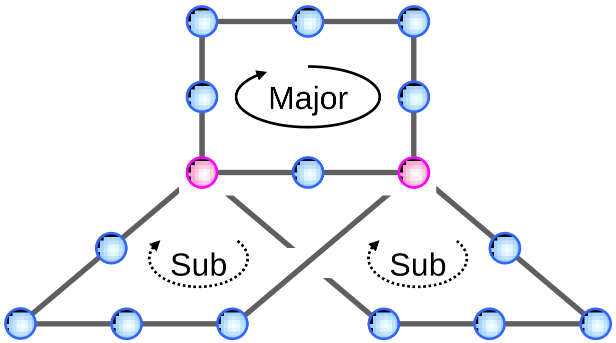
  Rule: a given link must belong to a single ring

  1 Major ring (closed) and multiple Sub-rings (open)

- R-APS Event Message to signal 'MAC flushing notification' from one ring to another interconnected ring

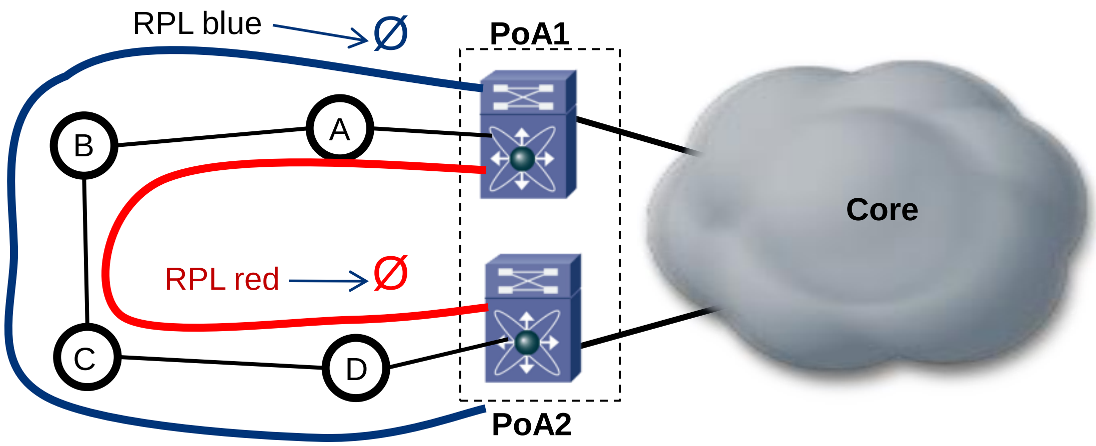
- If one ring provides R-APS VC for a subtended ring, the R-APS channels for the two rings must be in different VLANs for correct operation





Ring Instances

- G.8032 v.2 supports multiple ERP instances over a ring
- Disjoint VLANs are mapped into instances
- Every ERP instance can have a different RPL
  - Enables load-balancing over the ring



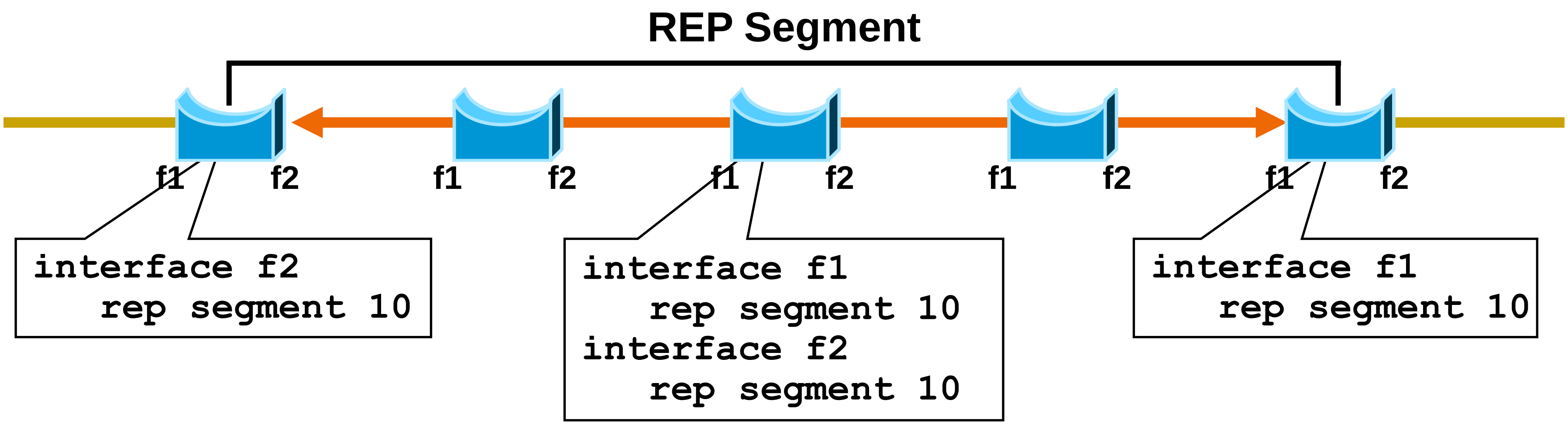
Learn. Connect. Collaborate. together.

Access Resiliency Mechanisms

Resilient Ethernet Protocol (REP)

REP Protocol Basics

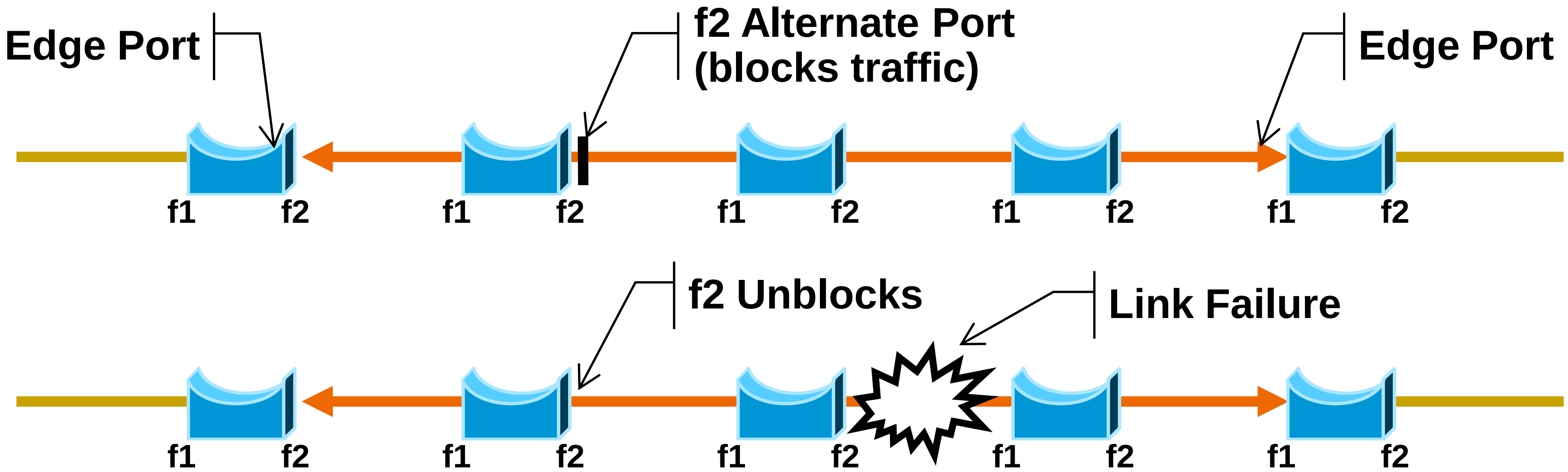
A Segment Protocol



- REP operates on chain of bridges called segments
- A port is assigned to a unique segment using:
  (config-if)# **[no] rep segment {*id*}**
- A segment can have up to two ports on a given bridge

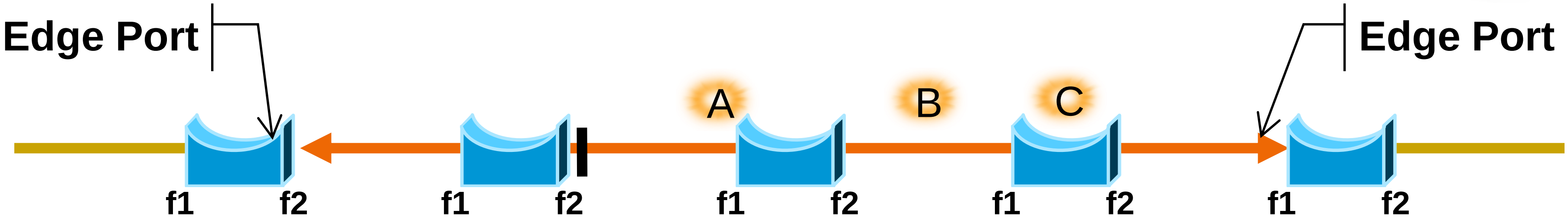
REP Protocol Basics

Blocked Port



- When all links are operational, a unique port blocks the traffic on the segment
  - No connectivity between edge ports over the segment
- If any failure occurs within the segment, the blocked port goes forwarding

Protected Failure Points

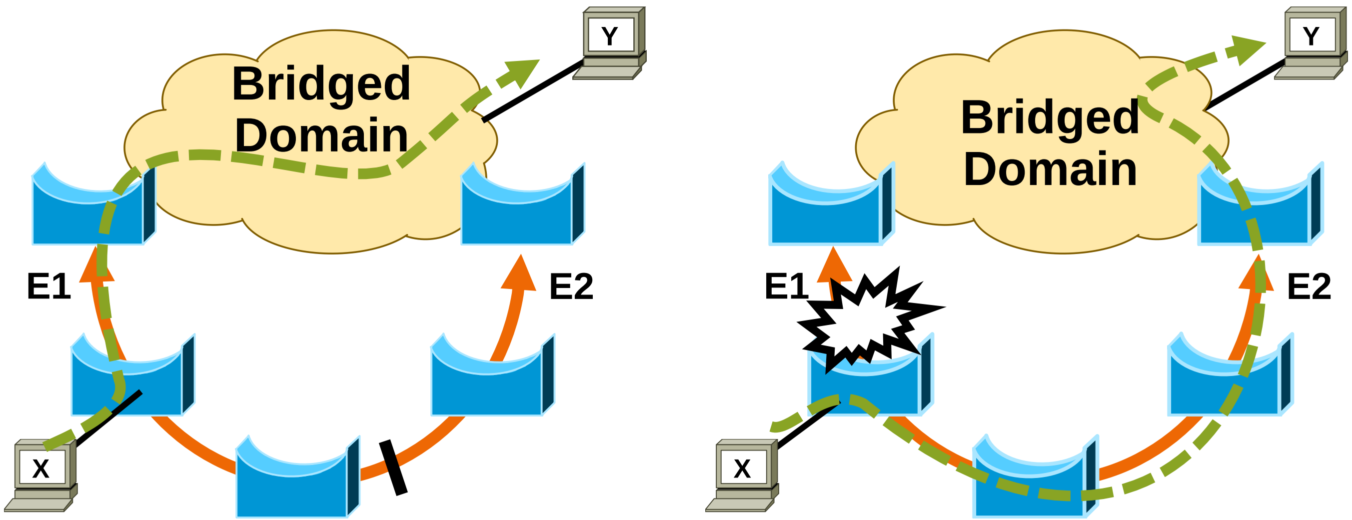


REP Protects Against Any Single Link, Port or Node Failure Within a Segment

- A: Failure of a port within the segment
- B: Failure of a link within the segment
- C: Failure of a node within the segment

REP Protocol Basics

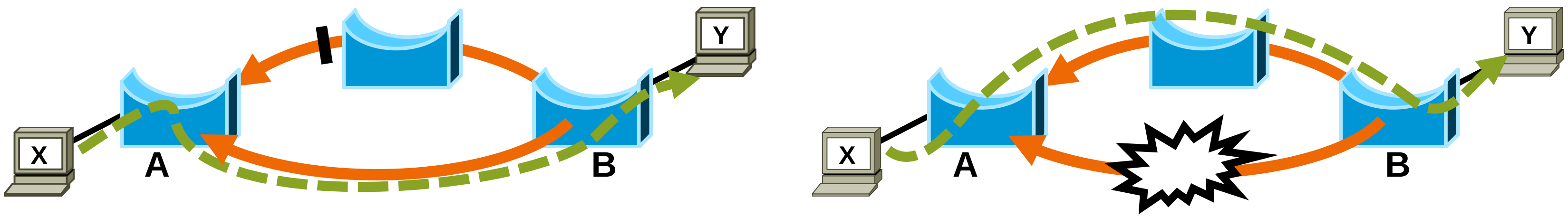
REP Provides Two Redundant Gateways



- The segment provides one level of redundancy
- Hosts on the segment can reach the rest of the network through either edge port, as necessary

REP Protocol Basics

REP Creates a Redundant Link



- Segments can be wrapped into a ring
- Can be seen as a redundant link in that case
- Identification of edge ports requires additional configuration in that case

REP Advantages

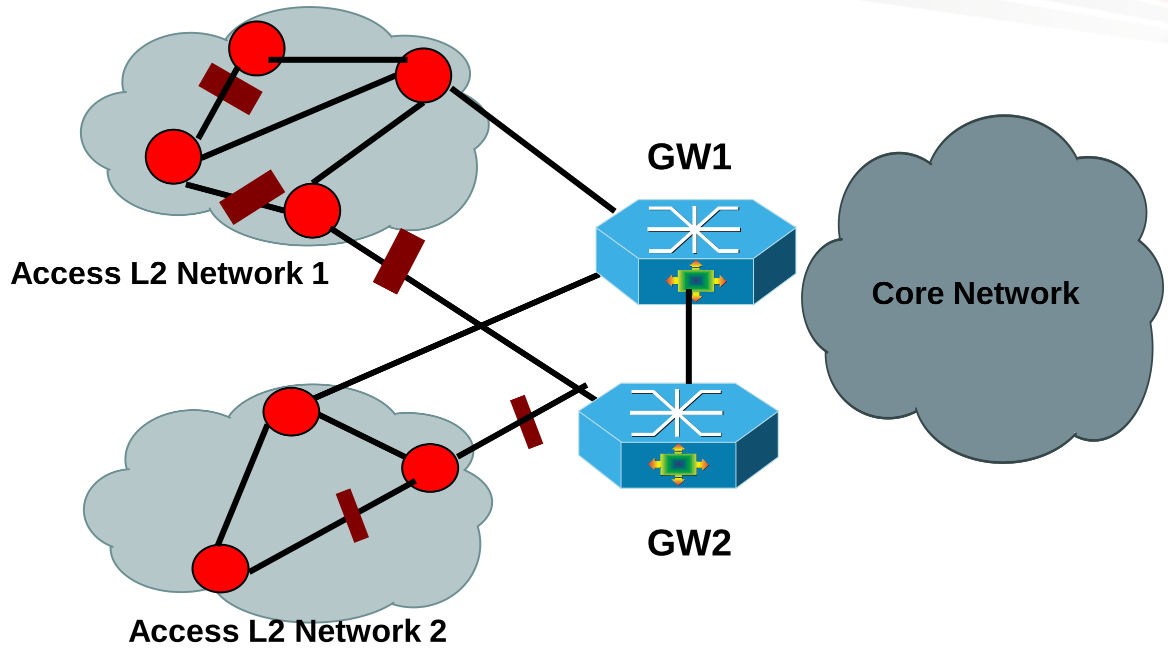
- Fast and predictable convergence
  - Convergence time: 50 to 250ms
  - Fast failure notification even in large rings with high number of node
  - Manual configuration for predictable failover behavior
- Co-existence with spanning tree
  - STP is deactivated on REP interfaces
  - Limit the scope of spanning tree
  - Topology changes notification from REP to STP
- Optimal bandwidth utilization
  - VLAN load balancing
- Easy to configure and troubleshoot
  - Topology archiving for easy troubleshooting
  - Known fixed topology
  - Simple mechanism to setup the alternate port (blocking)

Learn. Connect. Collaborate. together.

Access Resiliency Mechanisms

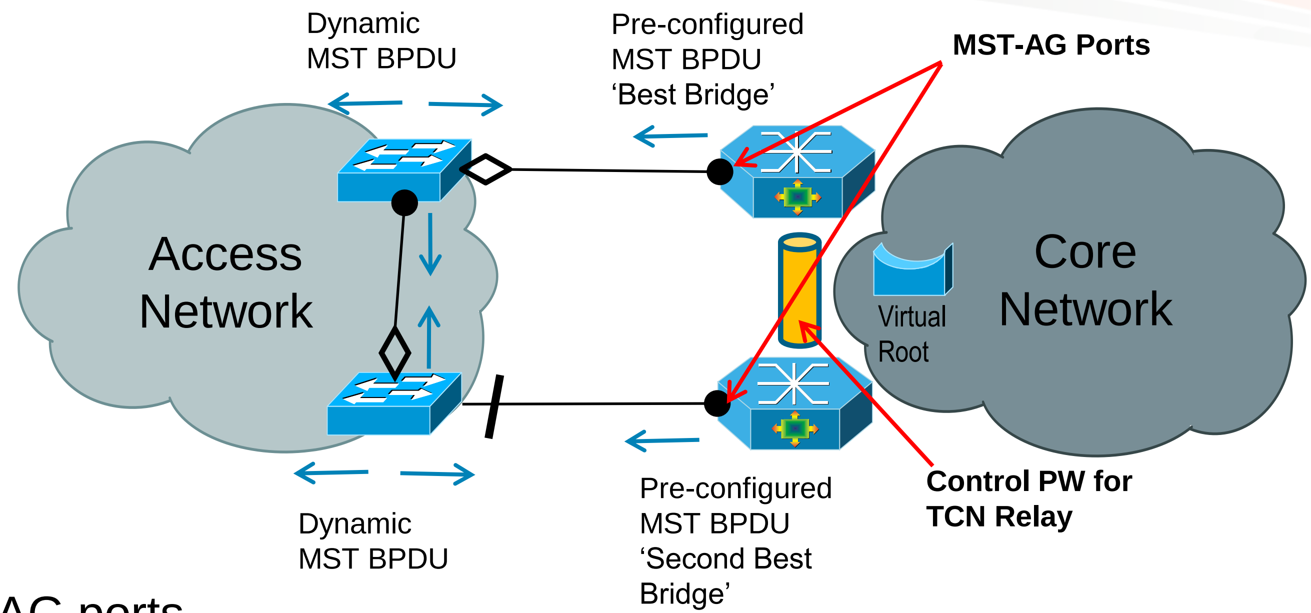
MST Access Gateway (MST-AG)
(a.k.a. Reverse Layer 2 Gateway Ports (R-L2GP))

Motivation for MST-AG

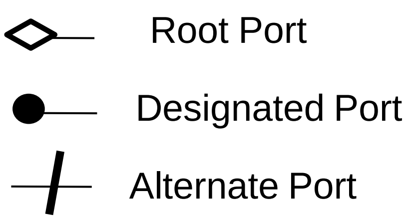


- Terminate multiple Ethernet access networks into same pair of 'Gateway' nodes
- Each access network maintains independent topology (control plane isolation)
- Fast convergence in all cases
- Access nodes run standard MST
- Gateway nodes act as root bridges

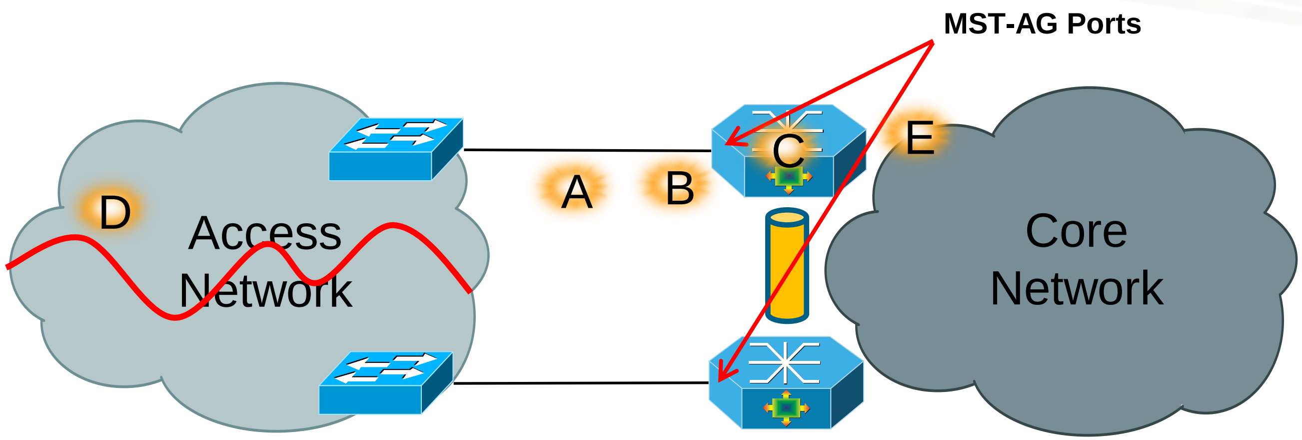
MST-AG Overview



- MST-AG ports
  - Send pre-configured BPDUs advertising "virtual root" by best and second best bridge
  - Ignore incoming BPDUs from access network, except for TCN
  - Always in Designated Forwarding state
- React and relay TCN over a special control pseudowire
- L2 access network
  - Can have arbitrary topology (e.g. ring or mesh)
  - Runs standard MST protocol
  - Handles port blocking/unblocking



Protected Failure Points

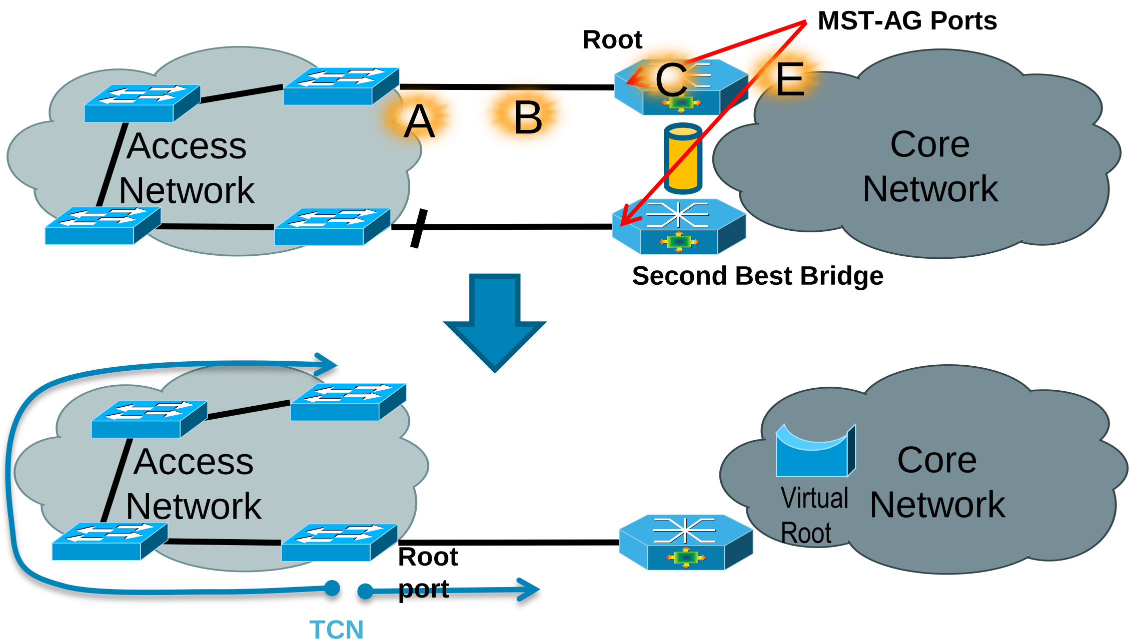


MST-AG Provides Protection Against Any of the Following Failure Points:

- A: Failure of link connecting access network to gateway
- B: Failure of gateway access-facing port
- C: Gateway node failure
- D: Failure within access network, including access network total split
- E: Isolation of the gateway from core network (via Link State Tracking feature)

Failure Scenarios

Gateway Direct Failures



- Access switches detect failure

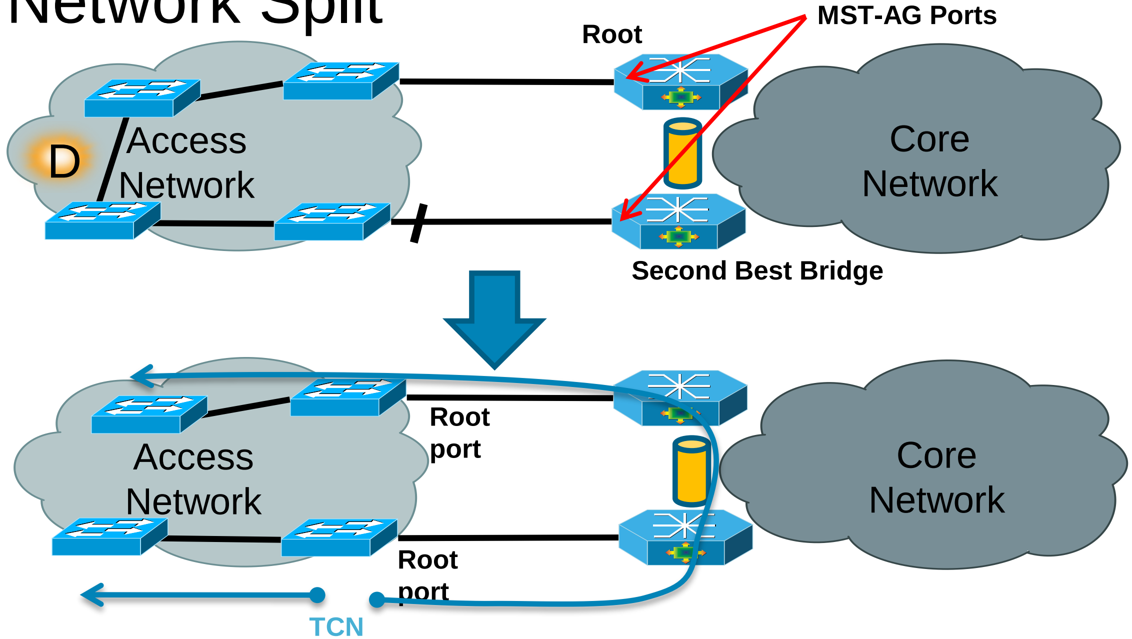
  Note: for Failure E, gateway brings down line-protocol on link to access

- MST re-converges in access network, choosing path through second Gateway to reach the root

- TCN propagated all the way to new root

Failure Scenarios

Access Network Split

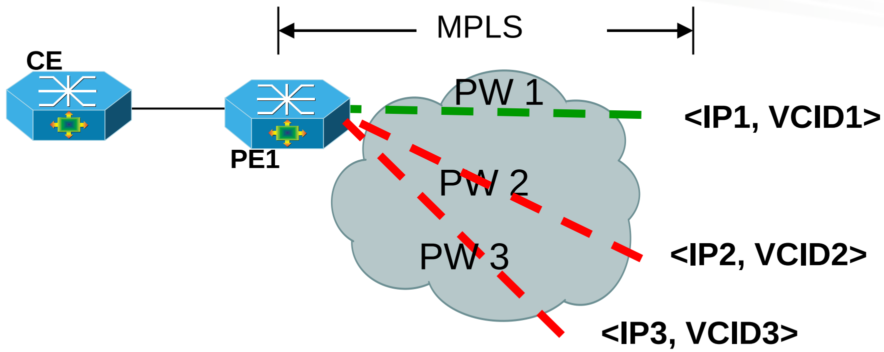


- Access network completely partitioned
- Sub-network isolated from original root selects path through second Gateway
- TCN is propagated to new root, relayed over control PW and into the other sub-network

Aggregation and Core Resiliency Mechanisms

Pseudowire Redundancy with LDP

Background



- Designate Pseudowires as either primary or backup

  Primary Pseudowire used for traffic forwarding, and backup takes over in case of failure (1:1 or N:1 protection)

- Signaling Redundant/Backup Pseudowires in targeted LDP session

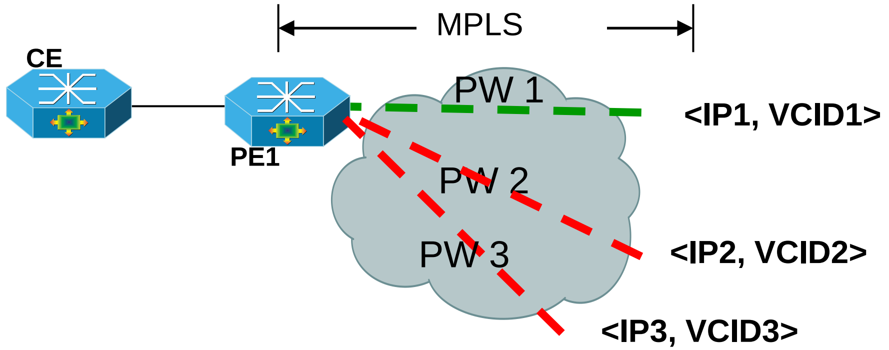
  Cold Redundancy: Backup PWs not signaled until required to take over

  Warm Redundancy: Backup PWs signaled up in the control-plane but held down in the data-plane. Use AC Fault code-point in LDP Status Message to indicate a backup PW

  Hot Redundancy: Backup PWs signaled up in the control-plane (use PW Preferential Forwarding Status Bit) and data-plane programmed

One-Way Pseudowire Redundancy

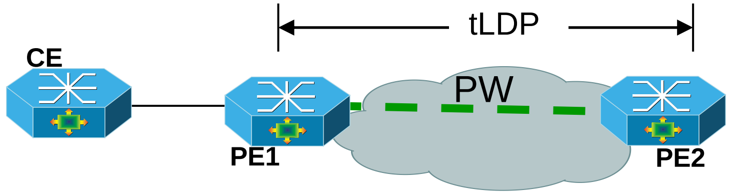
Overview



- Allows dual-homing of one local PE to two or more remote PEs
- Two pseudowires: primary and backup provide redundancy for a single AC (1:1 Protection)
- Multiple backup PWs (different priorities) can be defined (N:1)
- Alternate LSPs (TE Tunnels) can be used for additional redundancy
- Upon primary PW failure, failover is triggered after a configurable delay (seconds)
- Configurable Revertive / Non-Revertive upon recovery

Pseudowire Redundancy with LDP

PW Status Signaling



0x00000000 - Pseudowire forwarding (clear all failures)

0x00000001 - Pseudowire Not Forwarding

0x00000002 - Local Attachment Circuit (ingress) Receive Fault

0x00000004 - Local Attachment Circuit (egress) Transmit Fault

0x00000008 - Local PSN-facing PW (ingress) Receive Fault

0x00000010 - Local PSN-facing PW (egress) Transmit Fault

RFC 4447

0x00000020 - PW Preferential Forwarding Status

0x00000040 – PW Request Switchover Status

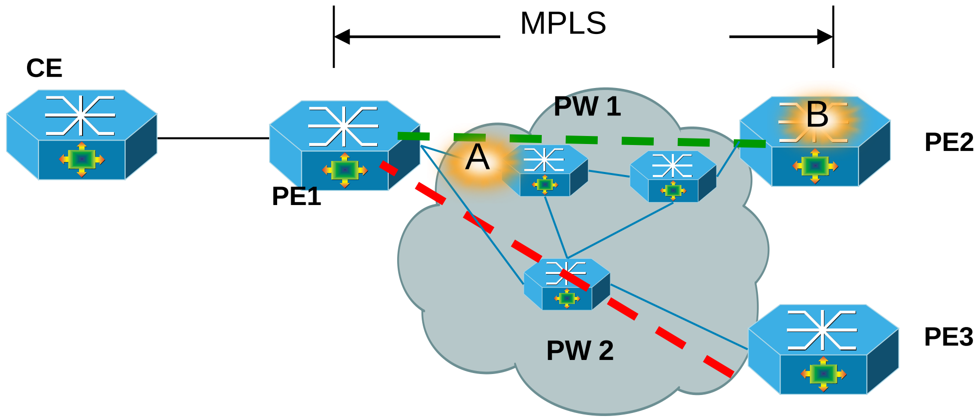
draft-ietf-pwe3-redundancy-bit

Whent set == PW fwd Standby; when cleared == PW fwd Active

Only this bit is required/used (with help of ICCP)

One-Way Pseudowire Redundancy

Failure Protection Points



A. Loss of next hop P node as notified by IGP

PW failover is delayed to allow IGP chance to restore

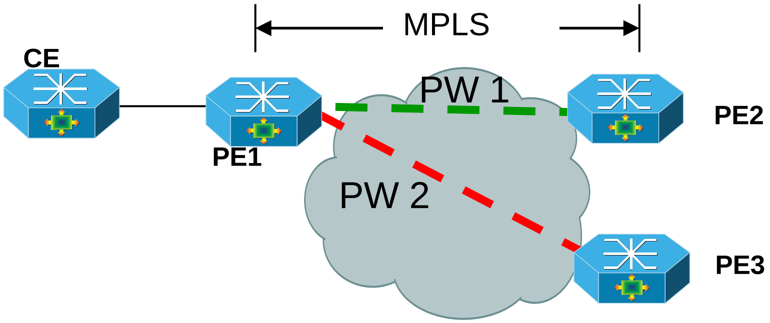
B. Loss of Remote PE

LDP session timeout

BFD timeout

One-Way Pseudowire Redundancy

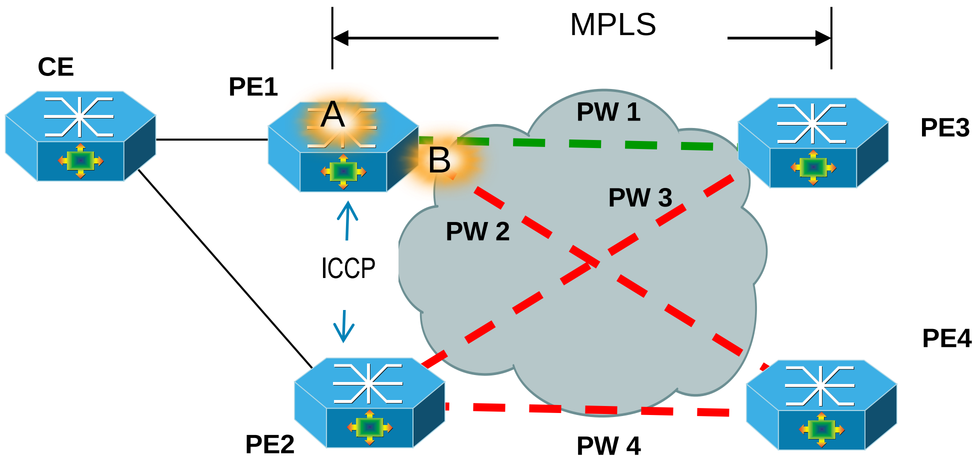
Operation



- Control is on dual-homed PE side, via static configuration
- Signaling:
  - If PEs support LDP PW status per RFC4447, backup PW is signaled (up in control-plane, down in data-plane)
  - If PEs do not support PW status, backup PW is not signaled in the control-plane
- Failover operation:
  - Upon primary PW failure, failover is triggered after a configurable delay (seconds)
  - Upon recovery, system reverts to primary PW after configurable delay (seconds)

Two-Way Pseudowire Redundancy

Overview

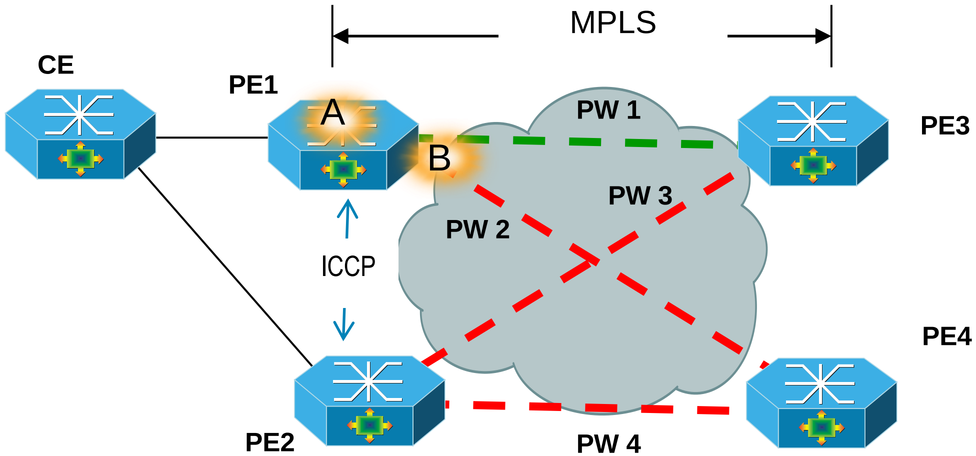


- Allows dual-homing of two local PEs to two remote PEs
- PW Preferential Forwarding Status determined by ICCP application (e.g. mLACP)

  Four pseudowires: 1 primary and 3 backup provide redundancy for a dual-homed device

Two-Way Pseudowire Redundancy

Failure Protection Points



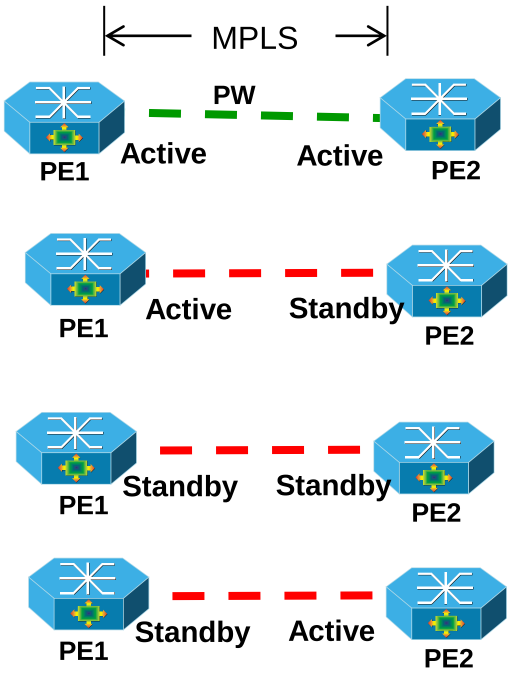
A. Failure of primary PE node

B. Isolation of primary PE node from the MPLS core

Two-Way Pseudowire Redundancy

Independent Operation Mode

- Every PE decides the local forwarding status of the PW: Active or Standby
- A PW is selected as Active for forwarding if it is declared as Active by both local and remote PEs
- A PW is selected as Standby for forwarding if it is declared as Standby by either local or remote PE



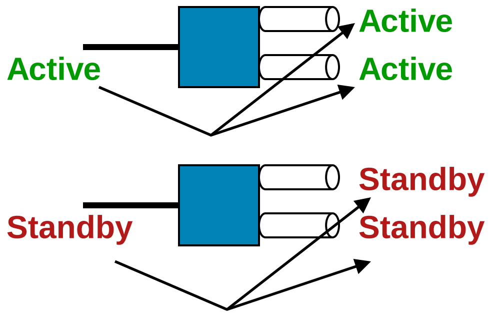
Two-Way Pseudowire Redundancy

Determining Pseudowire State

- VPWS / H-VPLS – two-way coupled:

  When AC changes state to Active[1], both PWs will advertise Active

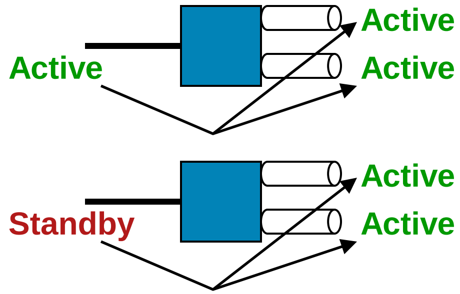
  When AC changes state to Standby[1], both PWs will advertise Standby



- H-VPLS – two-way decoupled:

  Regardless from AC state, Primary PW and Backup PWs will advertise Active state

- For H-VPLS, all PWs in VFI (at nPE) are Active simultaneously, for both access and core PWs



(1) Active / Standby AC states determined for example by mLACP

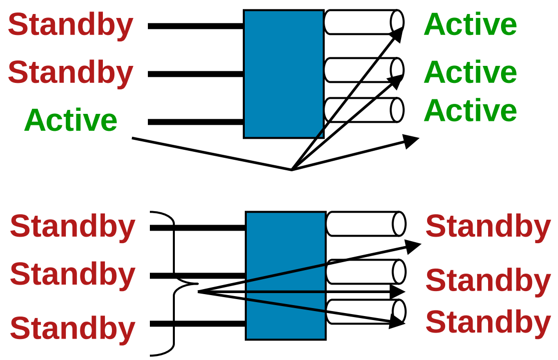
Two-Way Pseudowire Redundancy

Determining Pseudowire State (Cont.)

- VPLS – Two-way Coupled:

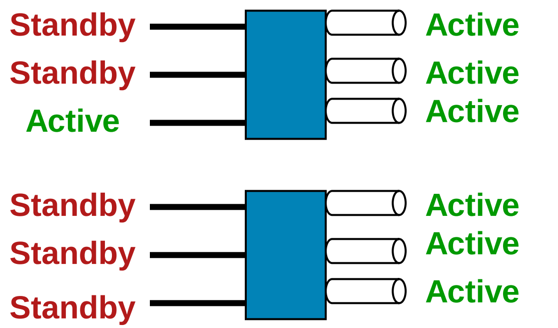
  When at least 1 AC in VFI changes state to Active, all PWs in VFI will advertise Active

  When all ACs in VFI change state to Standby, all PWs in VFI will advertise Standby mode



- VPLS – Two-way Decoupled:

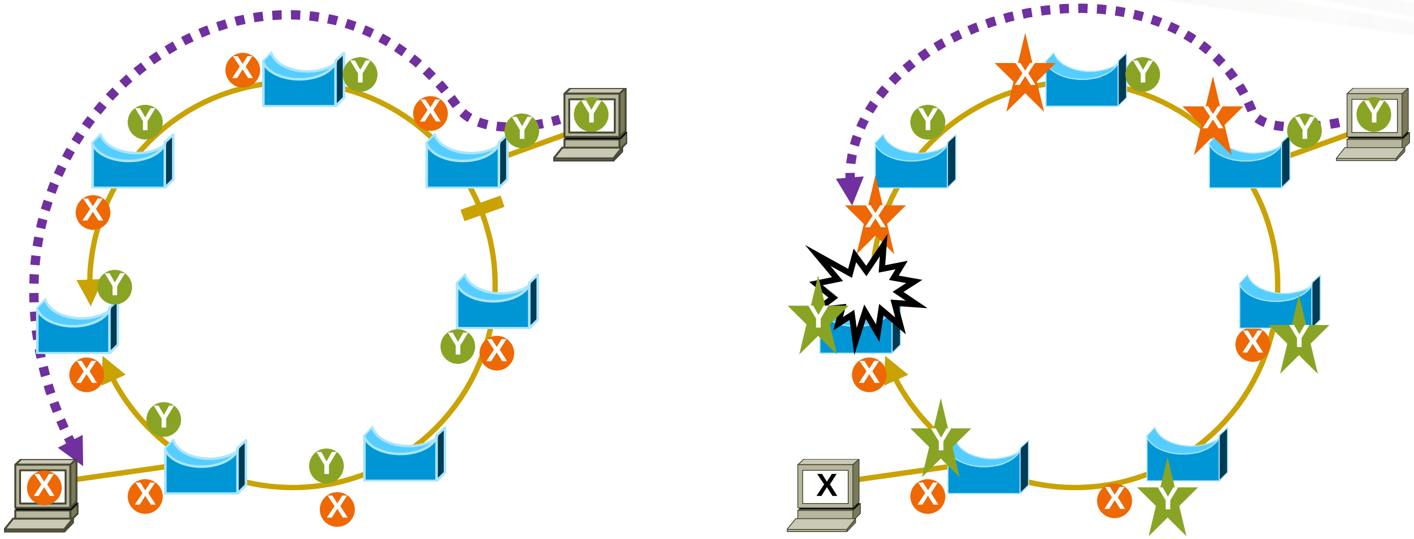
  Regardless from AC states, all PWs in VFI will advertise Active state



MAC Flushing Mechanisms

Why MAC Flushing Is Needed?

Topology Changes



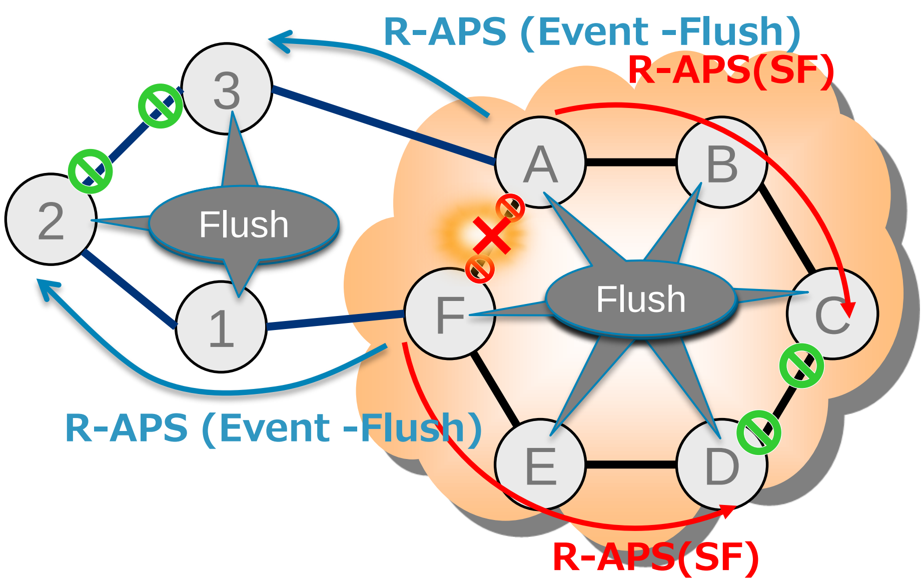
Filtering Entries Populated from Conversation X-Y

After a Change in the Topology, “Starred” Entries Are Incorrect

- Bridges learn the location of the stations from the traffic they forward
- Mac-addresses are added to a filtering table
- After a failure, the filtering table must be updated

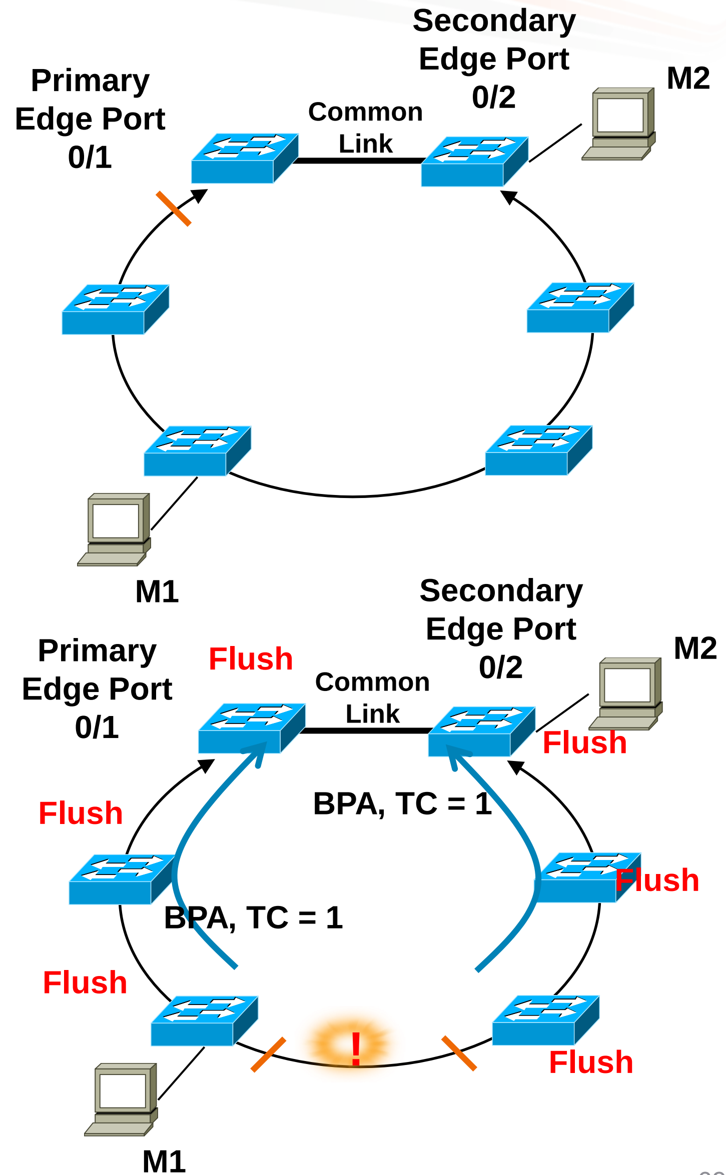
G.8032 MAC Flushing Notification

- Nodes evaluate every R-APS message received. If the message indicates that the location of blocking has moved (via Node ID and BPR), then flushing is triggered.
- A specific R-APS Event Message with Flush indication is used to trigger a burst of 3 flushes from one ring to another in case of cascaded rings



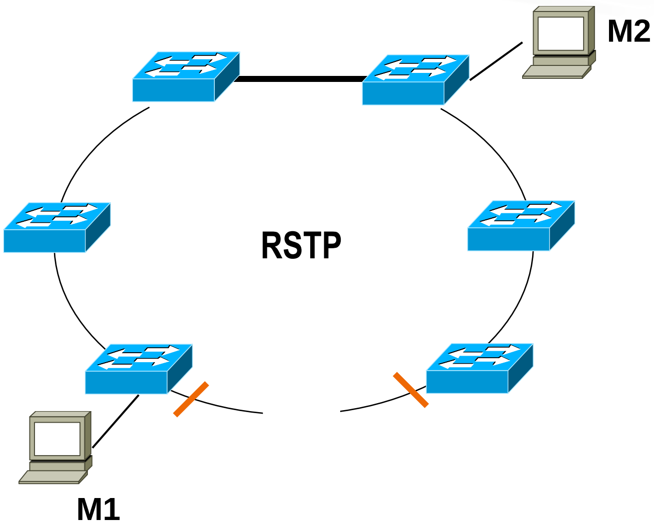
REP Topology Change Notification

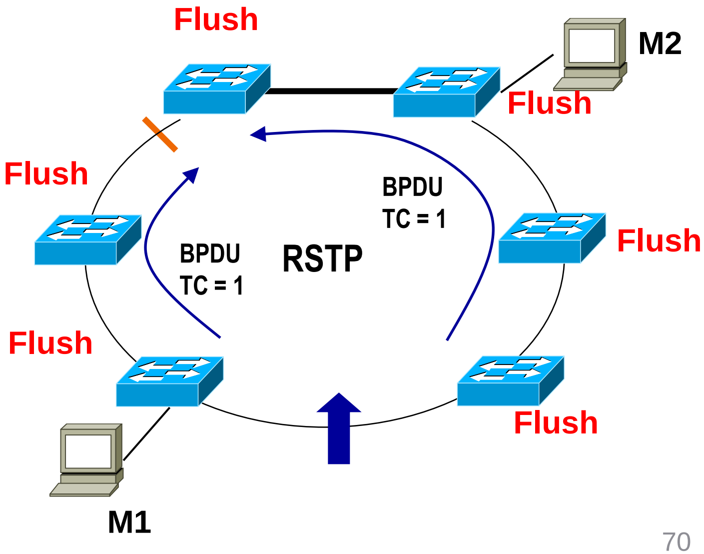
- On topology change, nodes next to fault send Blocked Port Advertisement (BPA) with Topology Change (TC) bit set to 1
- Nodes react to this by flushing their MAC tables for affected VLAN(s)
- Topology changes not propagated beyond segment except by explicit configuration



RSTP Topology Change Notification

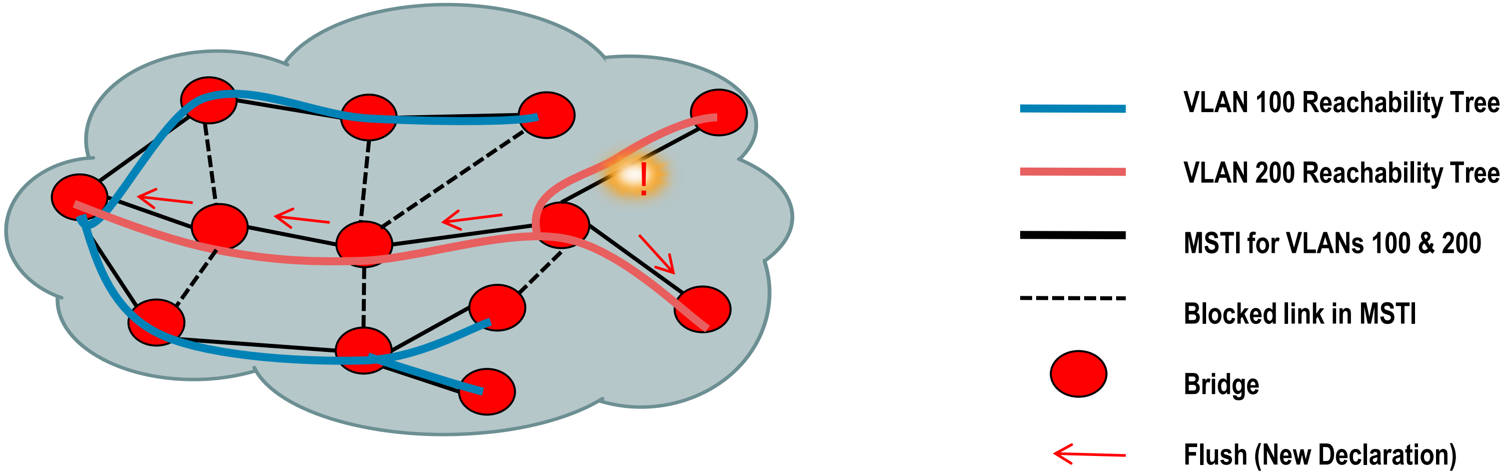
- Rapid STP (IEEE 802.1D-2004) introduced new Topology Change Notification mechanism (from IEEE 802.1D-1998)
- Detection — Transitions from blocking to forwarding state cause topology change
  - i.e., only increase in connectivity is TC
  - Link Down events no longer trigger TCN
  - Edge ports (port-fast) are not flushed
- Notification — via TCN Flag in configuration BPDU
  - TCN BPDU no longer used; no ack required (TCA flag not used)
- “Broadcasted” on the network by the initiator (not by the Root bridge as in IEEE 802.1D-1998)



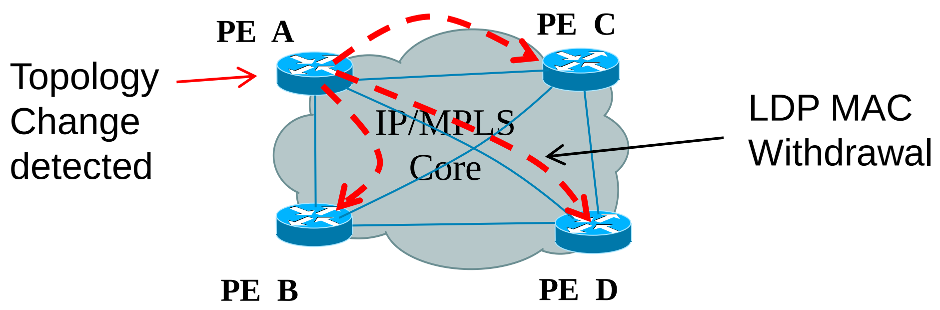


Multiple VLAN Registration Protocol (MVRP)

- Application of IEEE 802.1ak Multiple Registration Protocol (MRP)
- Builds dynamic VLAN reachability trees within a spanning tree instance
    - Enables source pruning of floods
- Defines new declaration messages as a replacement for TCNs
    - Sent in addition to existing STP TC messages
    - Generated by ports declaring a given VID on bridges that detect a topology change
- Net effect — only VLANs active in the area of the network that is actually affected by the topology change are flushed
    - VLANs not present in that part of the network are unaffected
    - VLANs that are affected are only flushed in the affected sub-tree



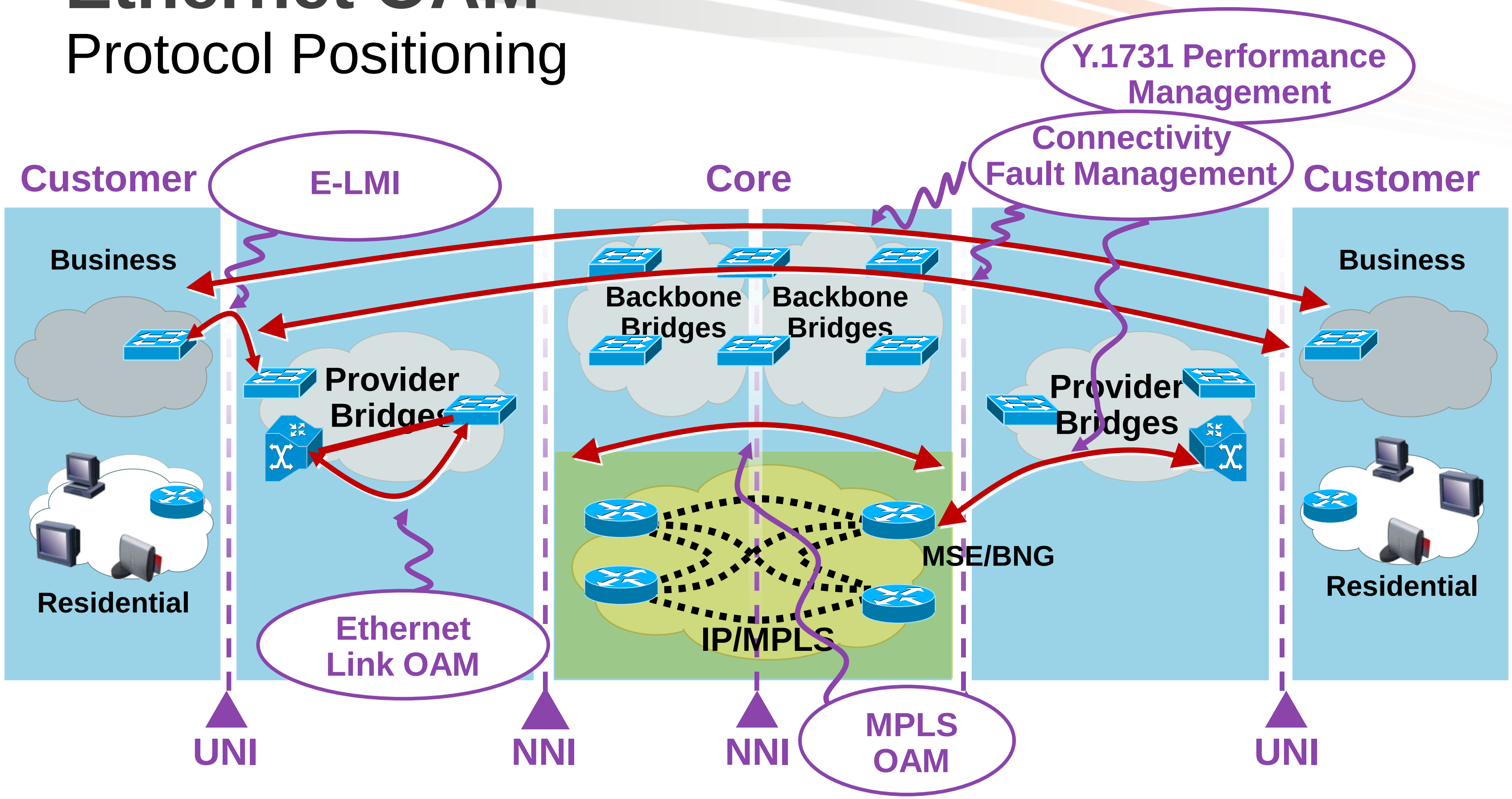
LDP MAC Address Withdrawal



- Transmitted by a VPLS PE that detects a topology change to all other PEs in the VPLS instance
- Out of band indication
- Optionally may contain a list of MAC addresses to be flushed
    - If MAC list is empty → flush all addresses except those learnt from transmitting PE
    - If specific MAC → remove specified MAC address(es)
- Defined in RFC4762

Ethernet OAM Resiliency Triggers

Ethernet OAM
Protocol Positioning



- E-LMI - User to Network Interface (UNI)
- Link OAM - Any point-to-point 802.3 link
- CFM / Y.1731 - End-to-End UNI to UNI
- MPLS OAM - within MPLS cloud

CFM Overview

- Family of protocols that provides capabilities to detect, verify, isolate and report end-to-end ethernet connectivity faults
- Employs regular Ethernet frames that travel in-band with the customer traffic
  - Devices that cannot interpret CFM Messages forward them as normal data frames
- CFM frames are distinguishable by Ether-Type (0x8902) and dMAC address (for multicast messages)
- Standardized by IEEE in 2007
  - IEEE std. 802.1ag-2007

CFM Protocols

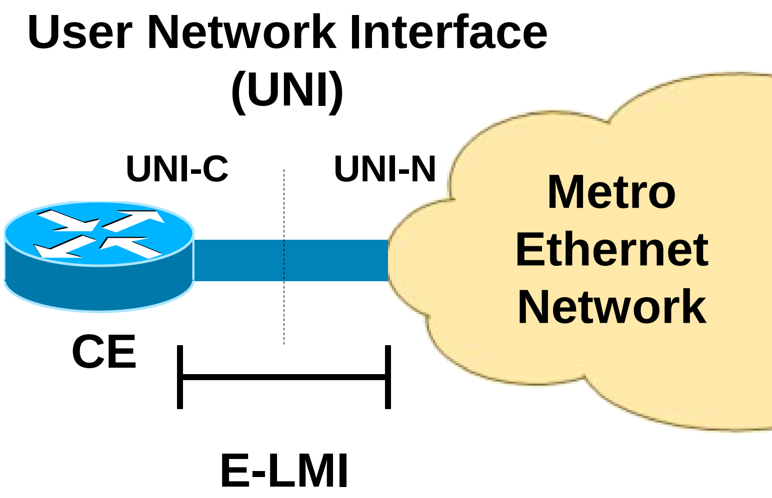
- There are three (3) protocols defined by CFM
- Continuity Check Protocol
  - Fault Detection
  - Fault Notification
  - Fault Recovery
- Loopback Protocol
  - Fault Verification
- Linktrace Protocol
  - Path Discovery and Fault Isolation

Ethernet LMI
Overview

- Provides protocol and mechanisms used for:
  - Notification of EVC addition, deletion or status (Active, Not Active, Partially Active) to CE
  - Communication of UNI and EVC attributes to CE (e.g. CE-VLAN to EVC map)
  - CE auto-configuration
  - Notification of Remote UNI name and status to CE

Cisco Enhancement

- Asymmetric protocol based on Frame Relay LMI, mainly applicable to the UNI (UNI-C and UNI-N)
- Specification completed by MEF: http://www.metroethernetforum.org/PDFs/Standards/MEF16.doc



Interworking Scenarios

Main Examples Supported by Cisco IOS / IOS-XR

CFM	➔	**E-LMI**
Link OAM	➔	**CFM**
MPLS PW OAM	➔	**E-LMI**

End-to-End Redundancy Solutions

End-to-End Redundancy Solutions

Service Type	Transport Enabler	Access Redundancy	Protocol / Feature
E-LINE	VPWS	Hub and Spoke (Active / Backup)	mLACP + 2-way PW Red. (coupled mode)
E-LINE	VPLS	Ring	MST + MST-AG
E-LINE	VPLS	Ring	G.8032 / REP
(*) E-LAN	VPLS	Hub and Spoke (Active / Backup)	mLACP + 2-way PW Red. (decoupled mode)
(*) E-LAN	H-VPLS	Hub and Spoke (Active / Backup)	mLACP + 2-way PW Red. (decoupled mode)
(*)E-LAN	VPLS	Ring	REP

(*) See Appendix Section

Learn. Connect. Collaborate. together.

E-LINE Availability Models

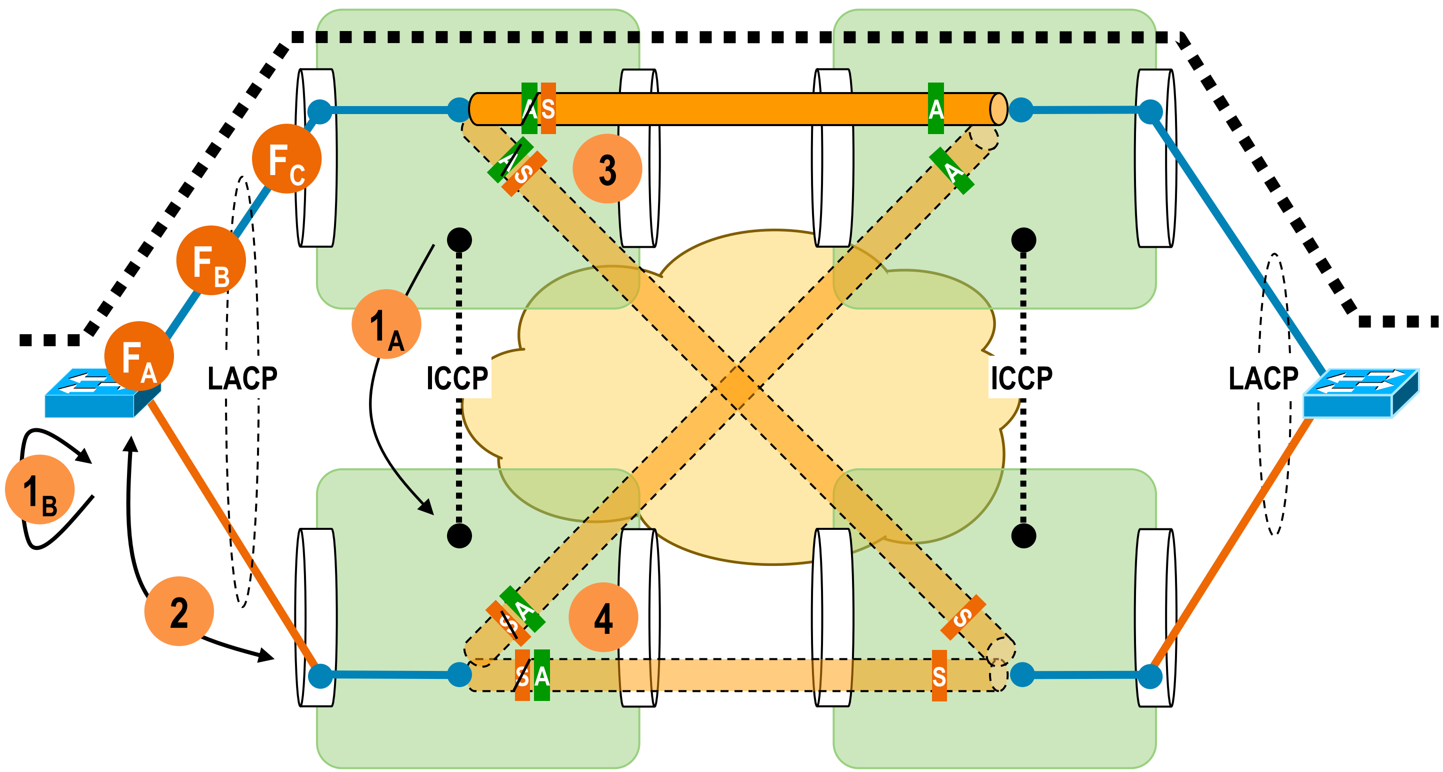
Active/Backup Access Node Redundancy (mLACP)

E-LINE Availability Model

VPWS

Active / Backup Access Node Redundancy (mLACP)

- Port / Link Failures



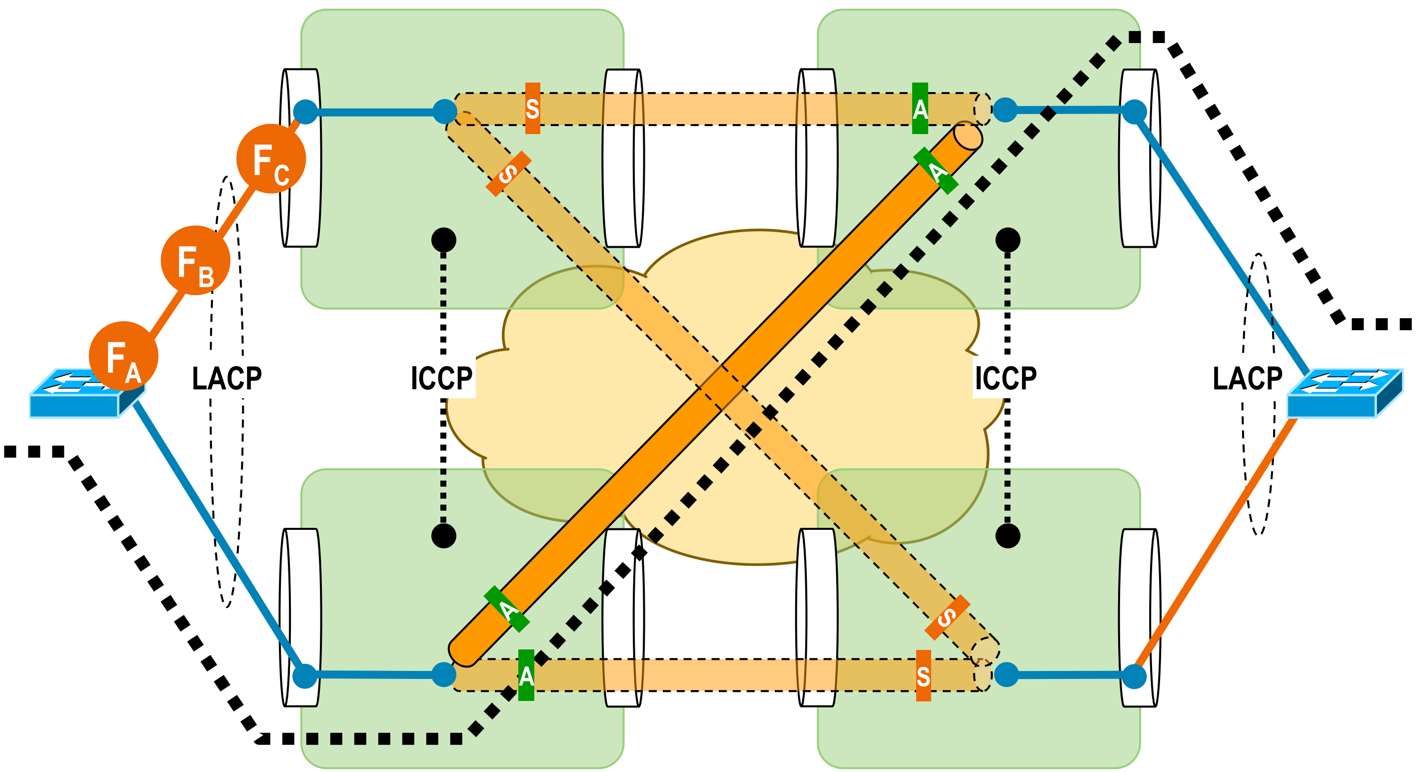
Events	
I	Initial state
$F_{A\text{-}C}$	Port / Link Failures
1_A	Active PoA detects failure and signals failover over ICCP
1_B	Failover triggered on DHD
2	Standby link brought up per LACP proc.
3	Active PoA advertises “Standby” state on its PWs
4	Standby PoA advertises “Active” state on its PWs

- For VPWS Coupled Mode, attachment circuit (AC) state (Active/Standby) drives PW state advertised to remote peers

Forwarding EoMPLS PW — Non-Forwarding EoMPLS PW

E-LINE Availability Model
Active / Backup Access Node Redundancy (mLACP)

- Port / Link Failures (cont.)



Events	
I	Initial state
F_{A-C}	Port / Link Failures
1_A	Active PoA detects failure and signals failover over ICCP
1_B	Failover triggered on DHD
2	Standby link brought up per LACP proc.
3	Active PoA advertises "Standby" state on its PWs
4	Standby PoA advertises "Active" state on its PWs
E	End State

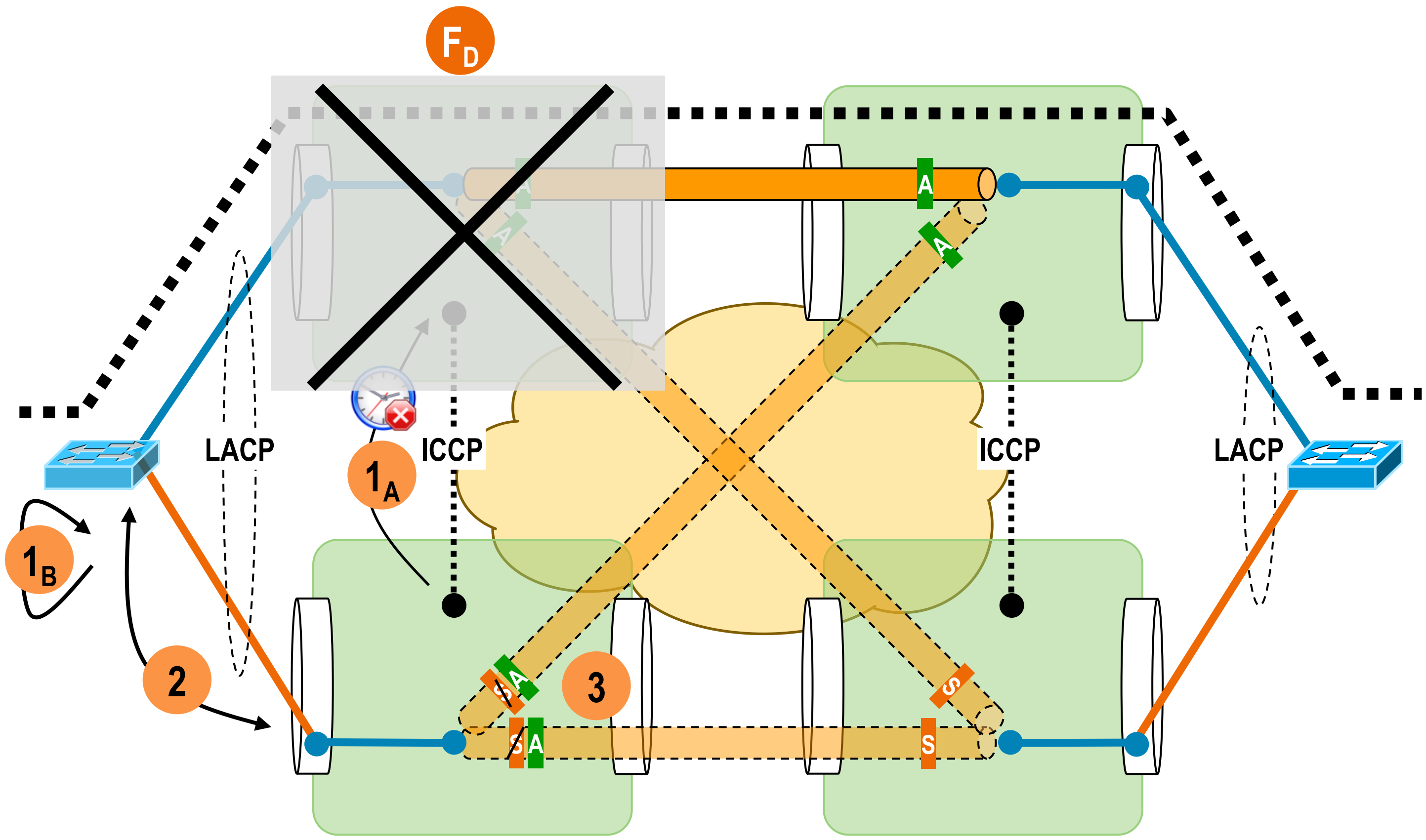
- Local site access failure does not trigger LACP failover at remote site (i.e. control-plane separation between sites)

E-LINE Availability Model

Active / Backup Access Node Redundancy (mLACP)

VPWS

- PoA Node Failure



Events	
I	Initial state
F_D	Active PoA Node Failure
1_A	Standby PoA detects node failure (BFD timeout or IP route-watch)
1_B	Failover triggered on DHD
2	Standby link brought up per LACP proc.
3	Standby PoA advertises "Active" state on its PWs

- PoA node failures detected by BFD (session timeout) or IP route-watch (loss of routing adjacency)

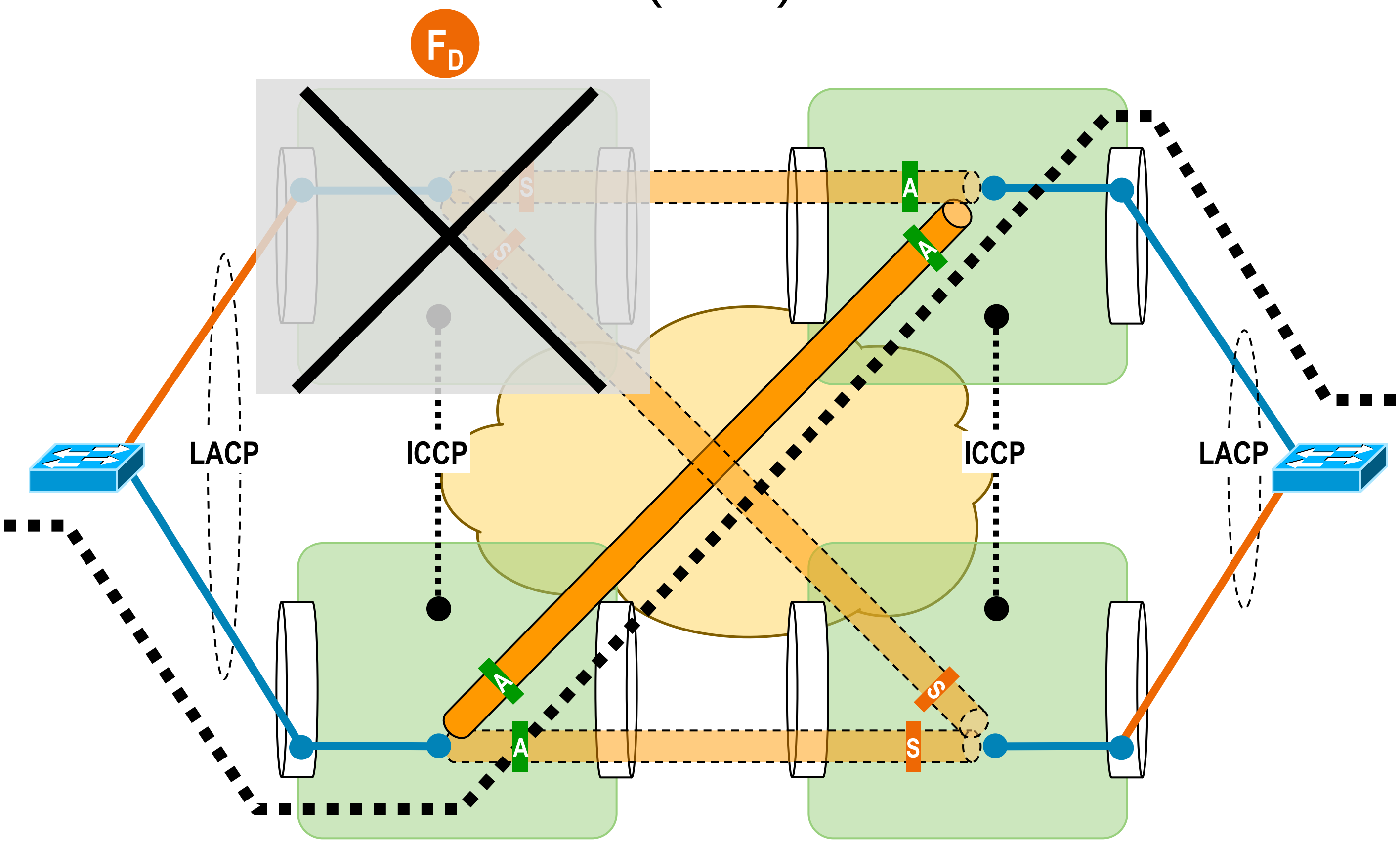
Forwarding EoMPLS PW — Non-Forwarding EoMPLS PW

E-LINE Availability Model

Active / Backup Access Node Redundancy (mLACP)

VPWS

- PoA Node Failure (cont.)



Events	
I	Initial state
F_D	Active PoA Node Failure
1_A	Standby PoA detects node failure (BFD timeout or IP route-watch)
1_B	Failover triggered on DHD
2	Standby link brought up per LACP proc.
3	Standby PoA advertises “Active” state on its PWs
E	End State

- No remote LACP switchover even if remote PoAs detect loss of PW before local LACP switchover is performed

Forwarding EoMPLS PW

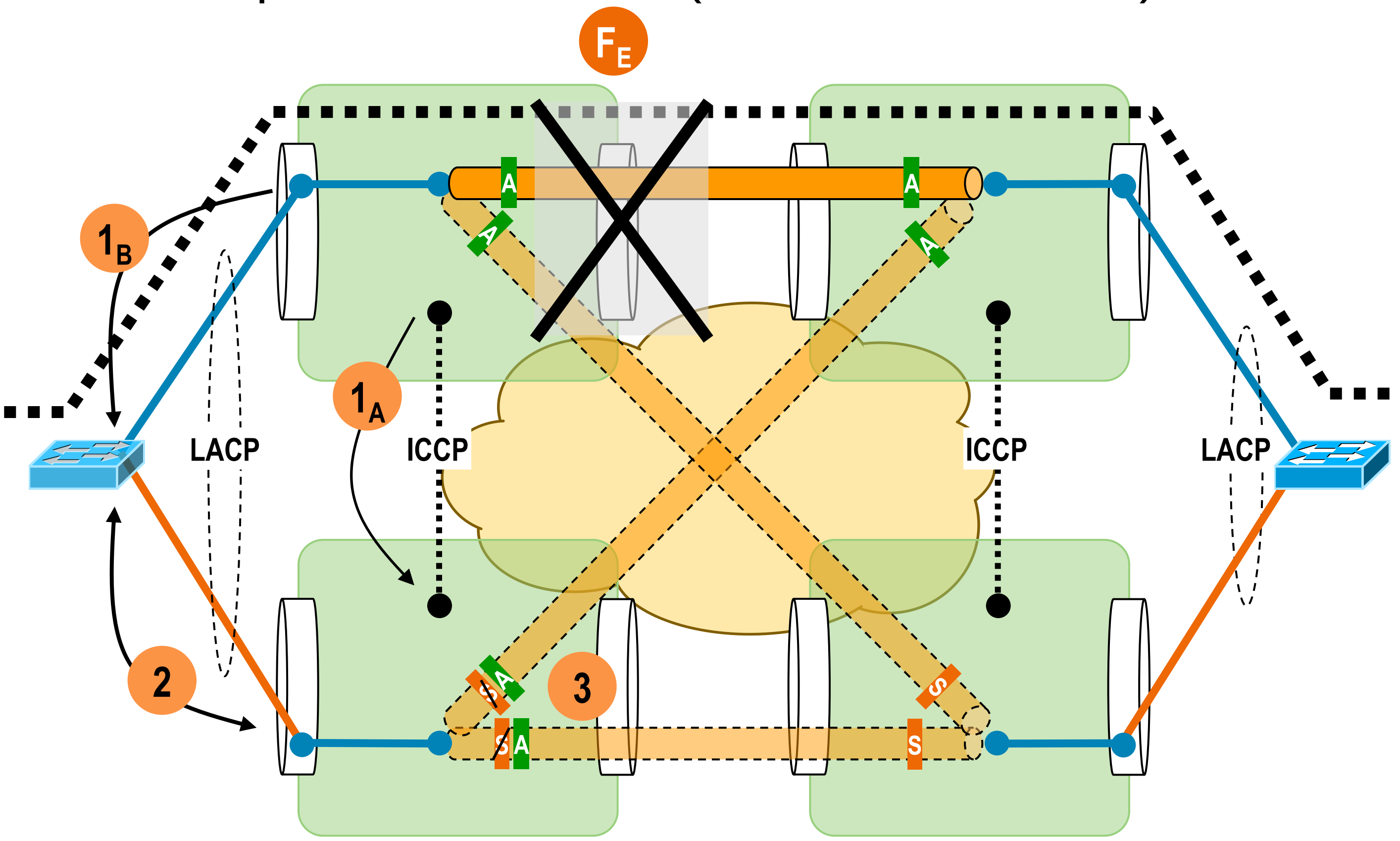
Non-Forwarding EoMPLS PW

E-LINE Availability Model

Active / Backup Access Node Redundancy (mLACP)

VPWS

- Uplink Core Failure (PoA Core Isolation)



Events	
I	Initial state
F_E	Core Isolation
1_A	Active PoA detects core isolation and signals failover over ICCP
1_B	Active PoA signals failover to DHD (dynamic port priority changes / bruteforce)
2	Standby link brought up per LACP proc.
3	Standby PoA advertises "Active" state on its PWs

- Link and Node failures in the Core are handled by IP routing and/or MPLS FRR – do not trigger LACP switchover

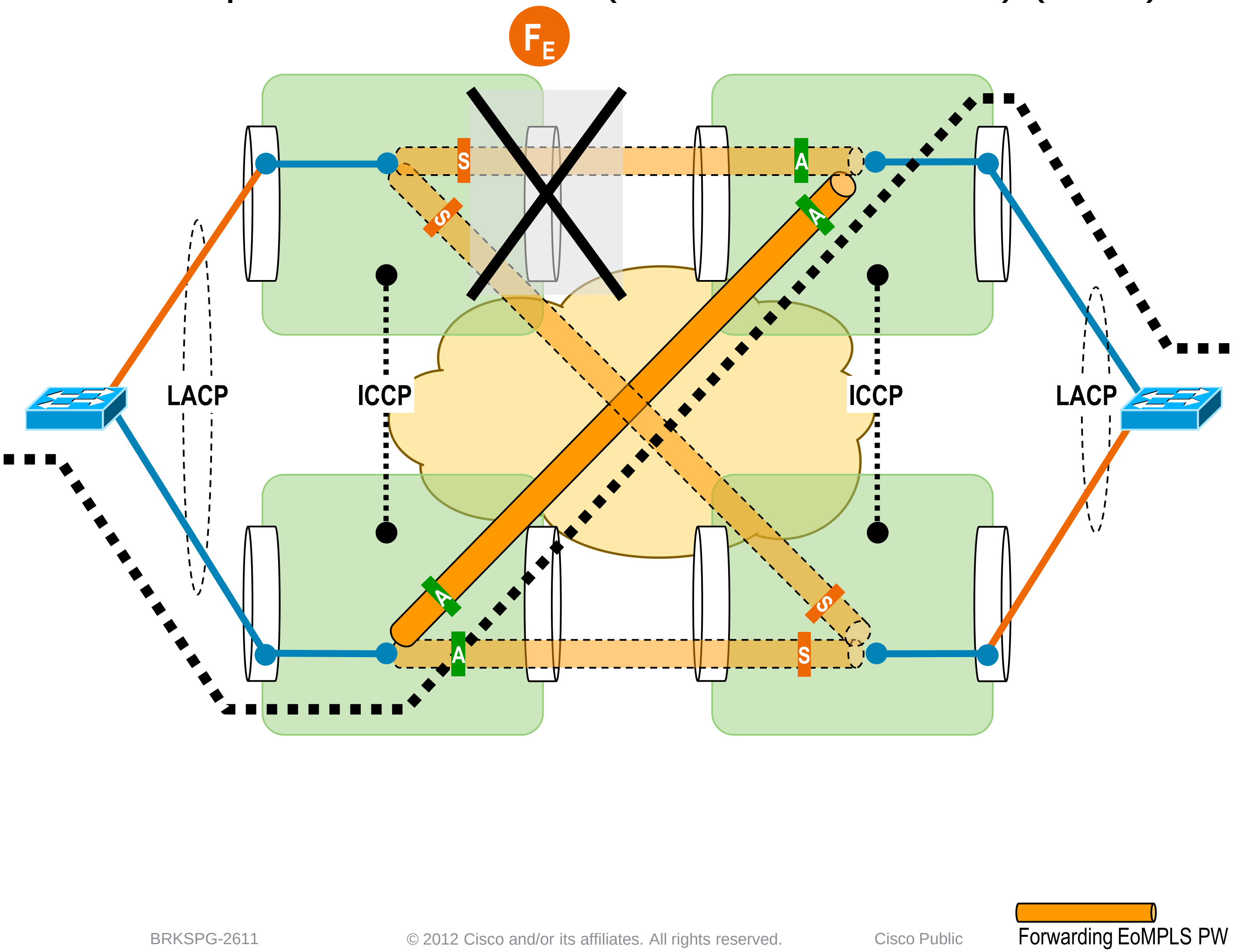
Forwarding EoMPLS PW
Non-Forwarding EoMPLS PW

E-LINE Availability Model

Active / Backup Access Node Redundancy (mLACP)

VPWS

- Uplink Core Failure (PoA Core Isolation) (cont.)



Events	
I	Initial state
F_E	Core Isolation
1_A	Active PoA detects core isolation and signals failover over ICCP
1_B	Active PoA signals failover to DHD (dynamic port priority changes / bruteforce)
2	Standby link brought up per LACP proc.
3	Standby PoA advertises “Active” state on its PWs
E	End State

E-LINE Availability Models

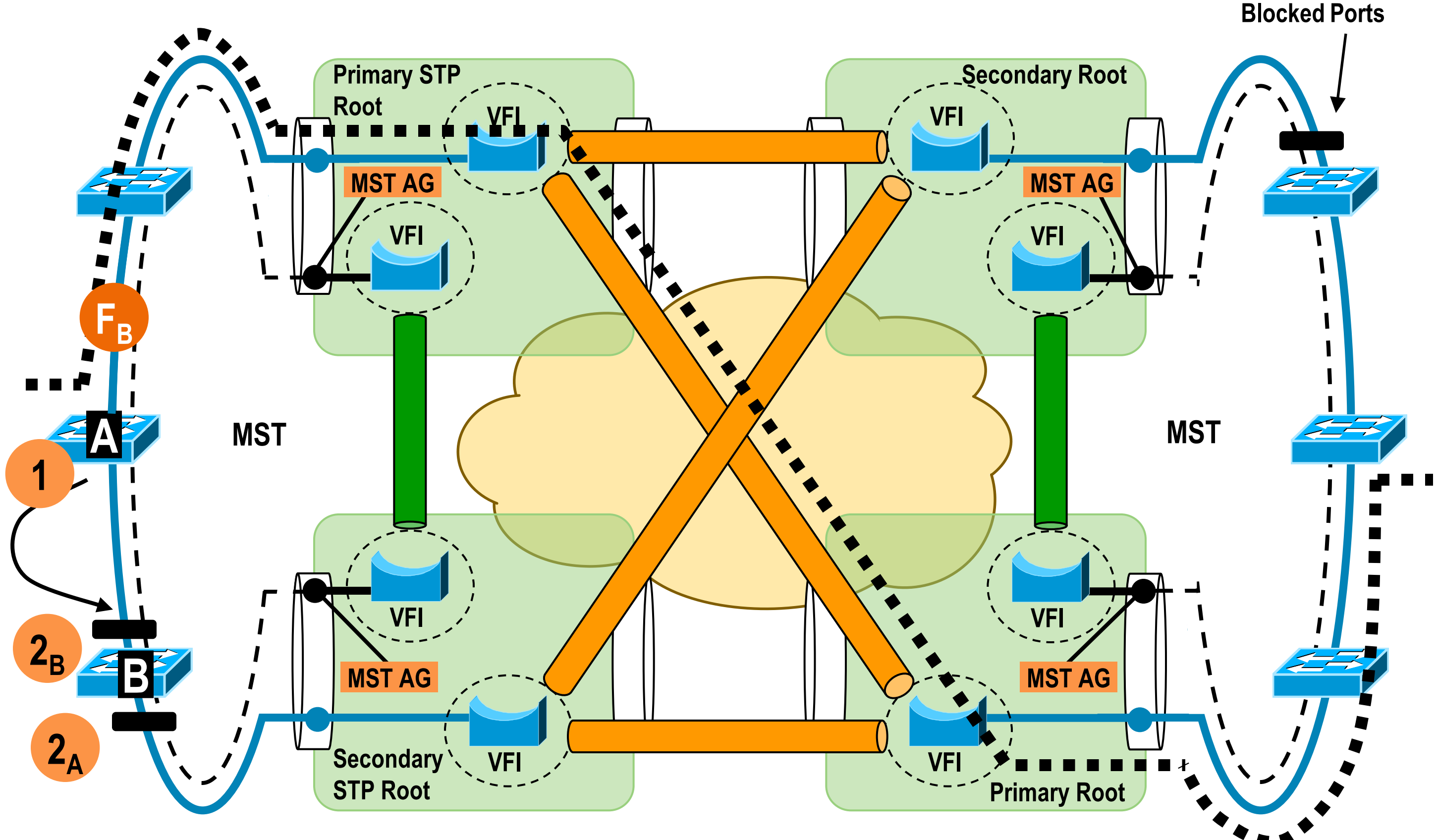
Ring Access Node Redundancy (MST)

E-LINE Availability Model

Ring Access Node Redundancy (MST)

VPLS

- MST Ring Span Failure



Events	
I	Initial state
F_B	Ring Span failure
1	Access switch "A" detects link failure (looses root port), blocks failed port and sends root proposal to "B"
2_A	"B" selects bottom AGG as new root (unblocks port towards it)
2_B	"B" blocks port towards "A"

- MST isolation; not carried over MPLS cloud
- MST Access Gateway (MST-AG) on Aggregation Nodes transmits statically configured BPDUs

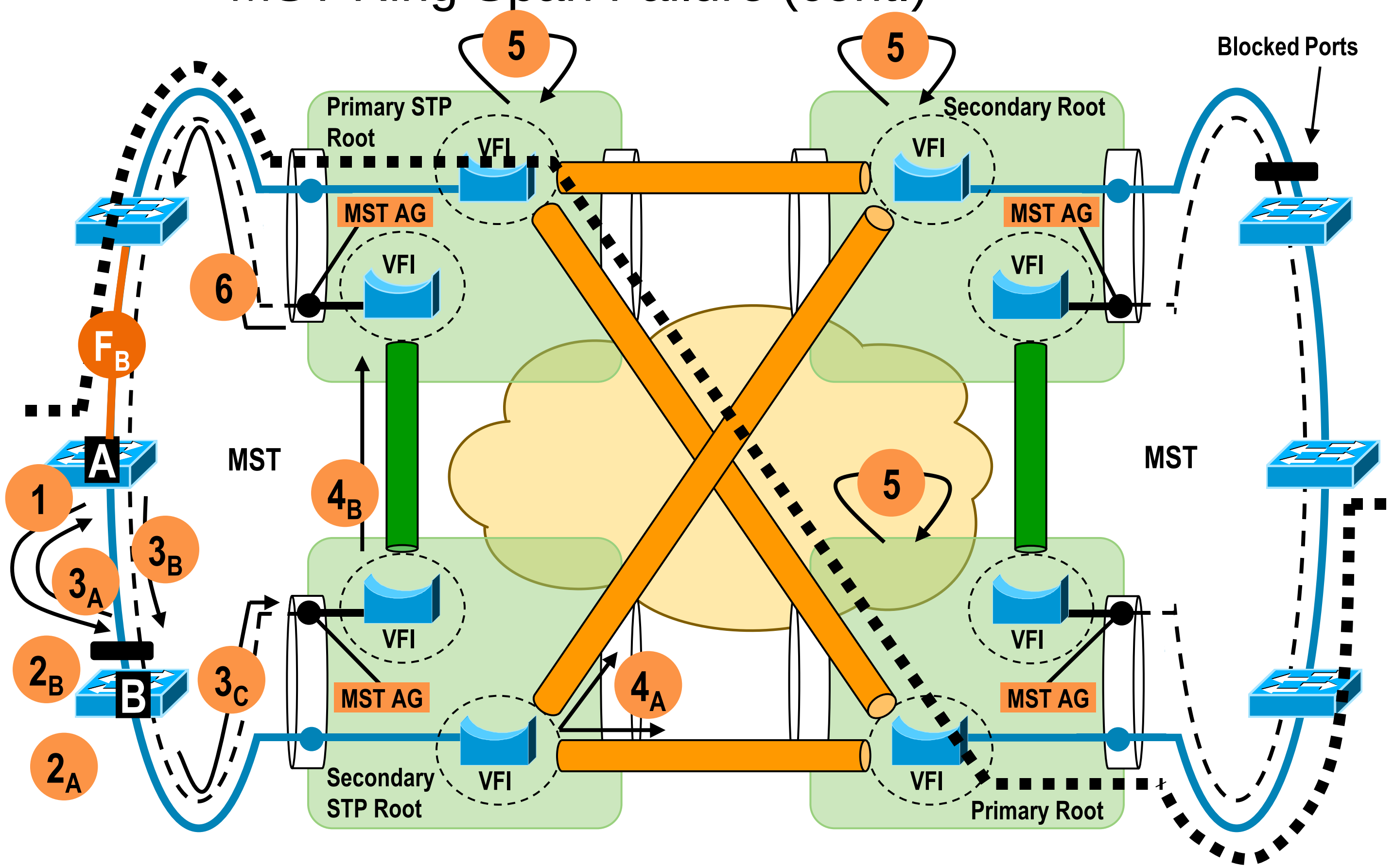
STP TCN EoMPLS PW
Forwarding EoMPLS PW
Non-Forwarding EoMPLS PW
MST Native Vlan

E-LINE Availability Model

Ring Access Node Redundancy (MST)

VPLS

- MST Ring Span Failure (cont.)



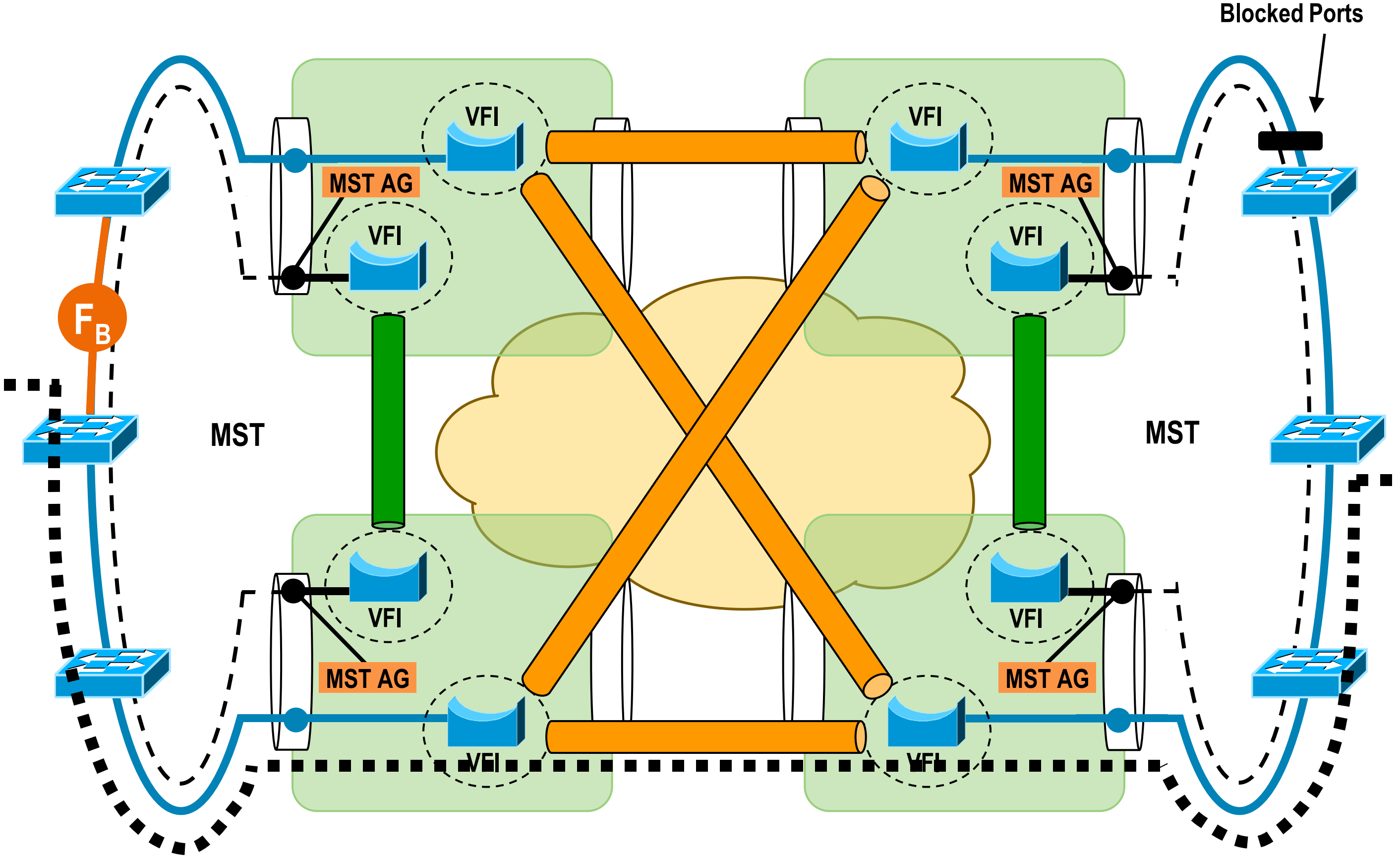
- Special VFI between AGG nodes to relay TCN BPDUs used to trigger MAC flushes after a Topology Change (TC)

Events	
3_{A-B}	Proposal / Agreement handshake between "B" and "A". "B" unblocks port towards "A"
3_C	"B" flushes MAC table. Signals Topology Change (TC) to AGG device
4_A	AGG flushes MAC table. Triggers LDP MAC add. withdrawal to VPLS peers
4_B	AGG device propagates TCN over BPDU PW
5	AGG (local and remote) flush MAC tables
6	Top AGG generates TCN on local ring

E-LINE Availability Model

Ring Access Node Redundancy (MST)

- MST Ring Span Failure (cont.)



Events	
6	Top AGG generates TCN on local ring
E	End State

- Each ring on unique TCN domain for control plane isolation
- Two MST instances for VLAN load balancing over ring

E-LINE Availability Models

Ring Access Node Redundancy (REP)*

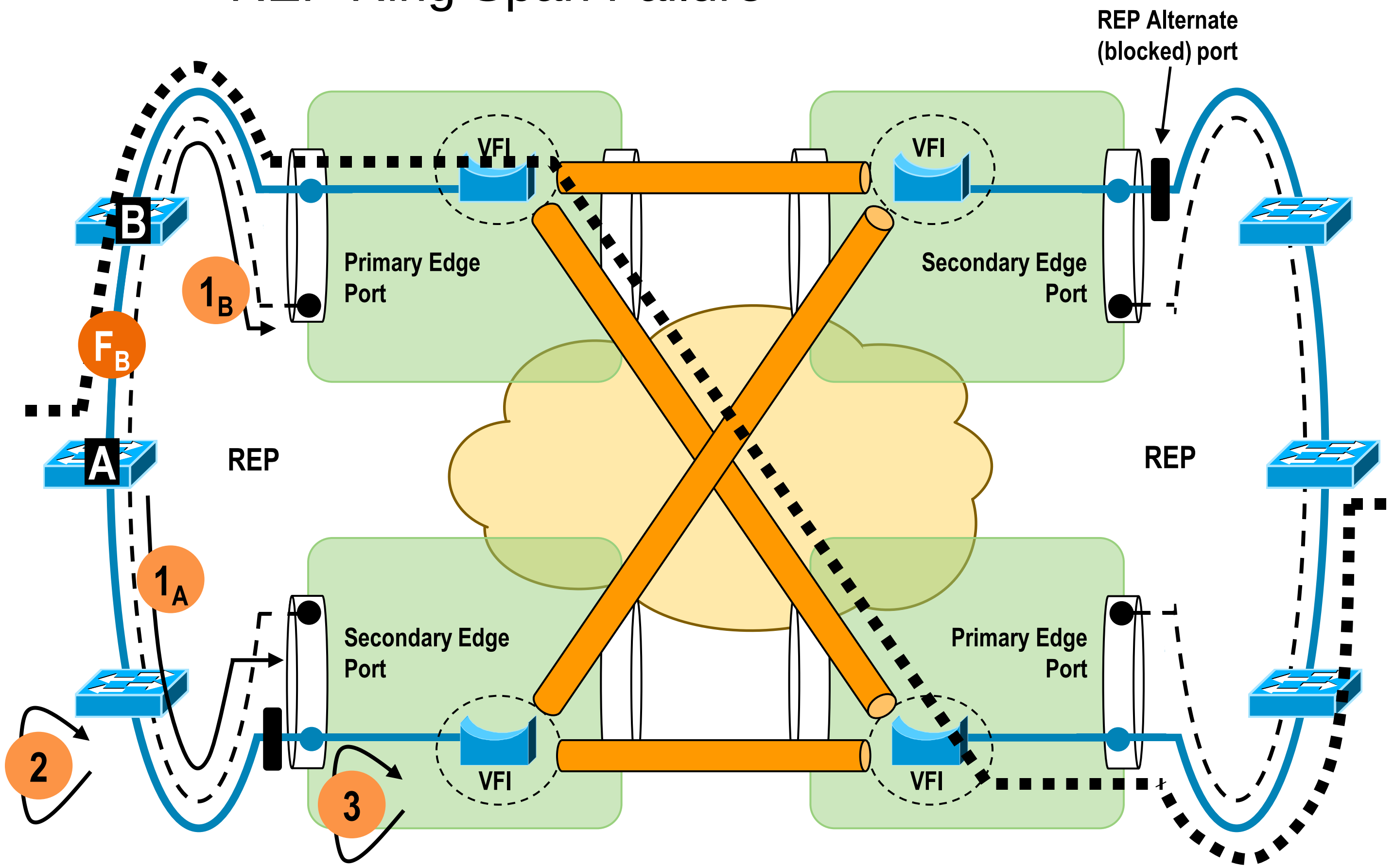
(*) – same principle applies to ITU-T G.8032

E-LINE Availability Model

Ring Access Node Redundancy (REP)

VPLS

- REP Ring Span Failure



Events	
I	Initial state
F_B	Ring Span failure
$1_{A\text{-}B}$	Access switches “A” and “B”detect link failure. Send Blocked Port Advertisement (BPA) with TC bit set on the segment
2	Access nodes in the ring flush MAC tables and propagate BPA
3	AGG node receives BPA and unblocks alternate port

- REP enabled segment with Edge Ports on Aggregation Nodes
- VLAN load balancing using Alternate Port configured on Secondary Edge Port

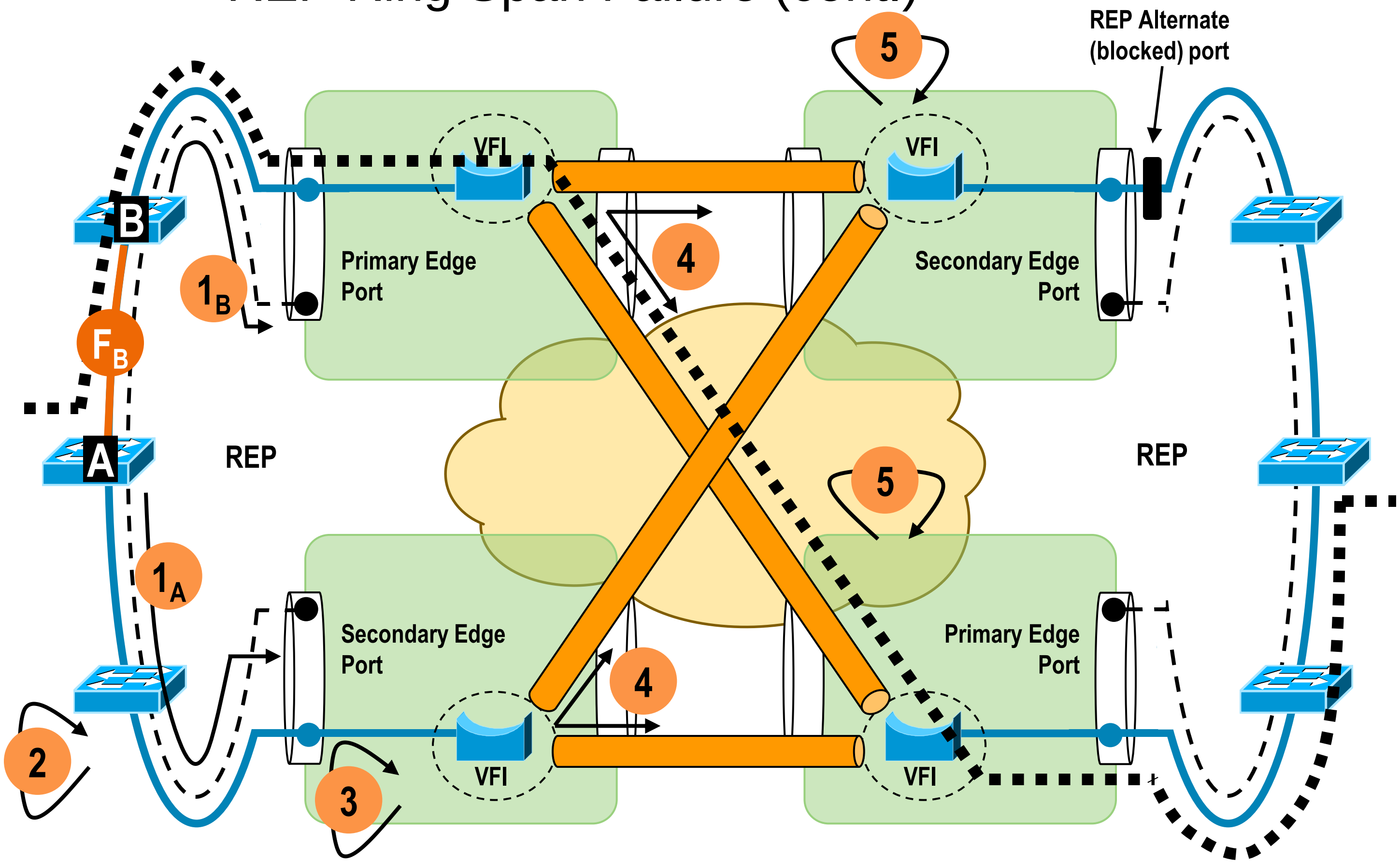
Forwarding EoMPLS PW
Non-Forwarding EoMPLS PW
REP Admin vlan

E-LINE Availability Model

Ring Access Node Redundancy (REP)

VPLS

- REP Ring Span Failure (cont.)

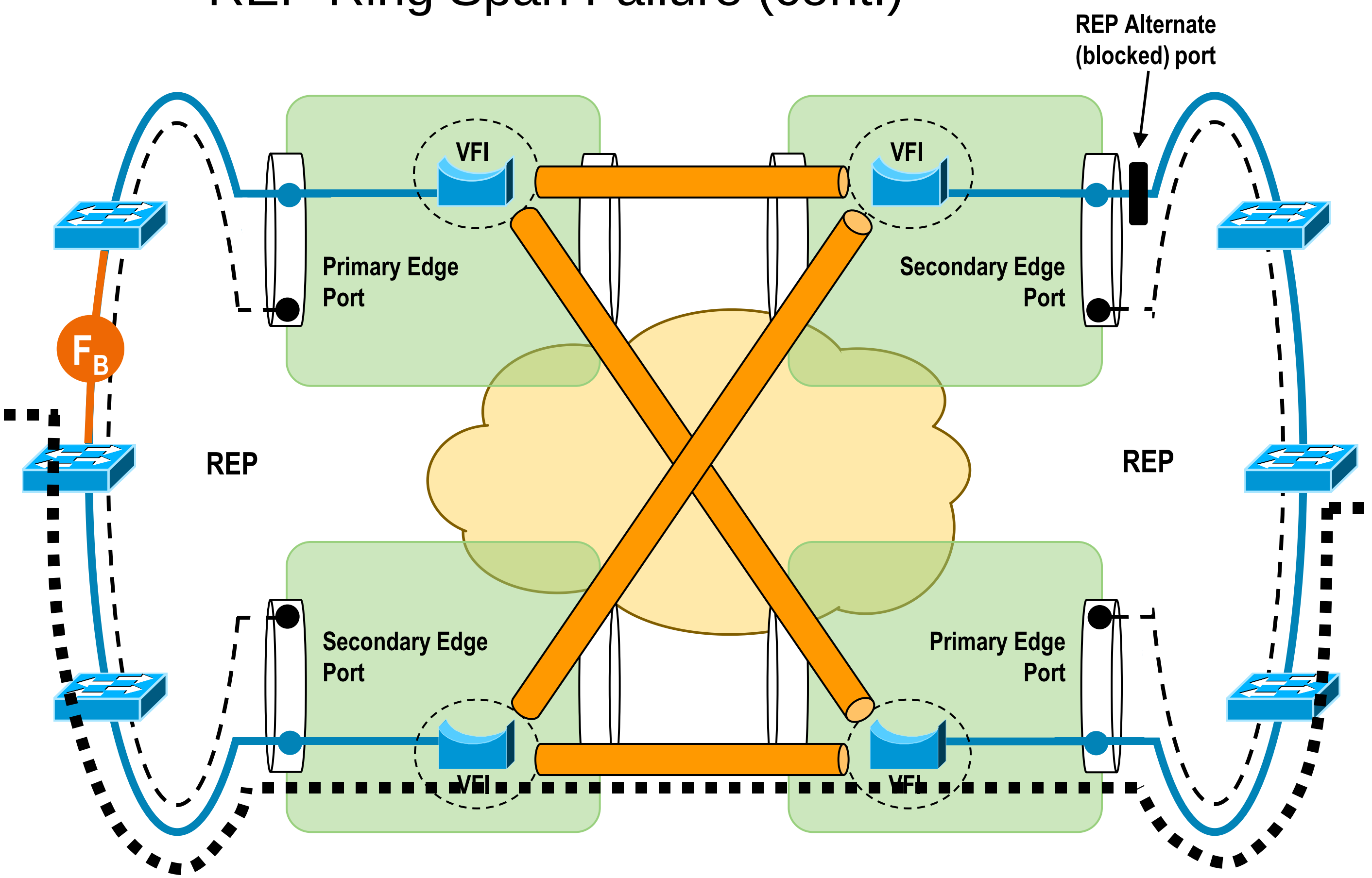


Events	
3	AGG node receives BPA and unblocks alternate port
4	AGG nodes flush MAC tables. Trigger LDP MAC add withdrawal to VPLS peers
5	Remote peers flush MAC tables

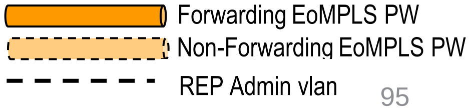
E-LINE Availability Model
Ring Access Node Redundancy (REP)

VPLS

- REP Ring Span Failure (cont.)

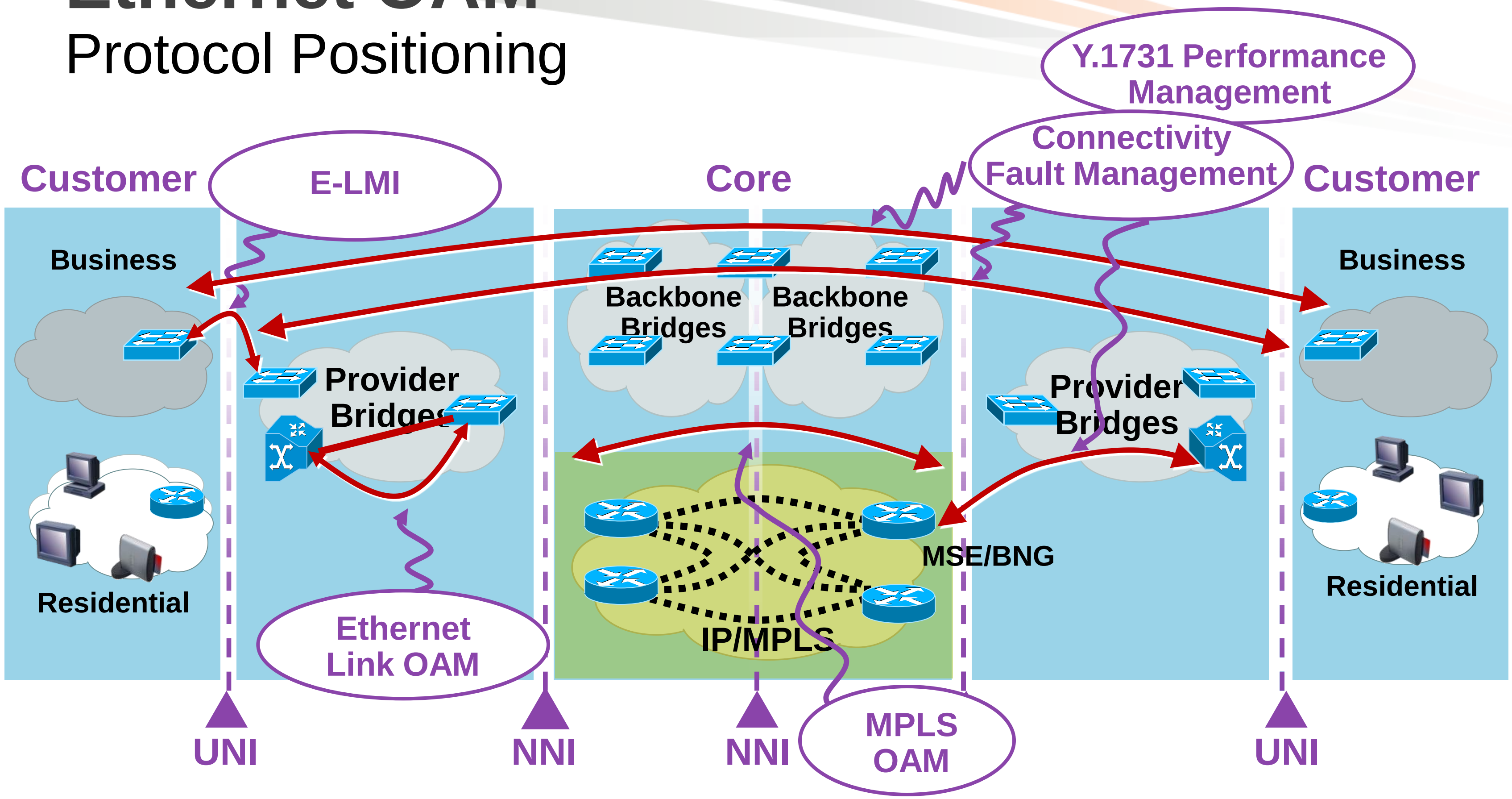


Events	
5	Remote peers flush MAC tables
E	End State



Ethernet-OAM Scenarios

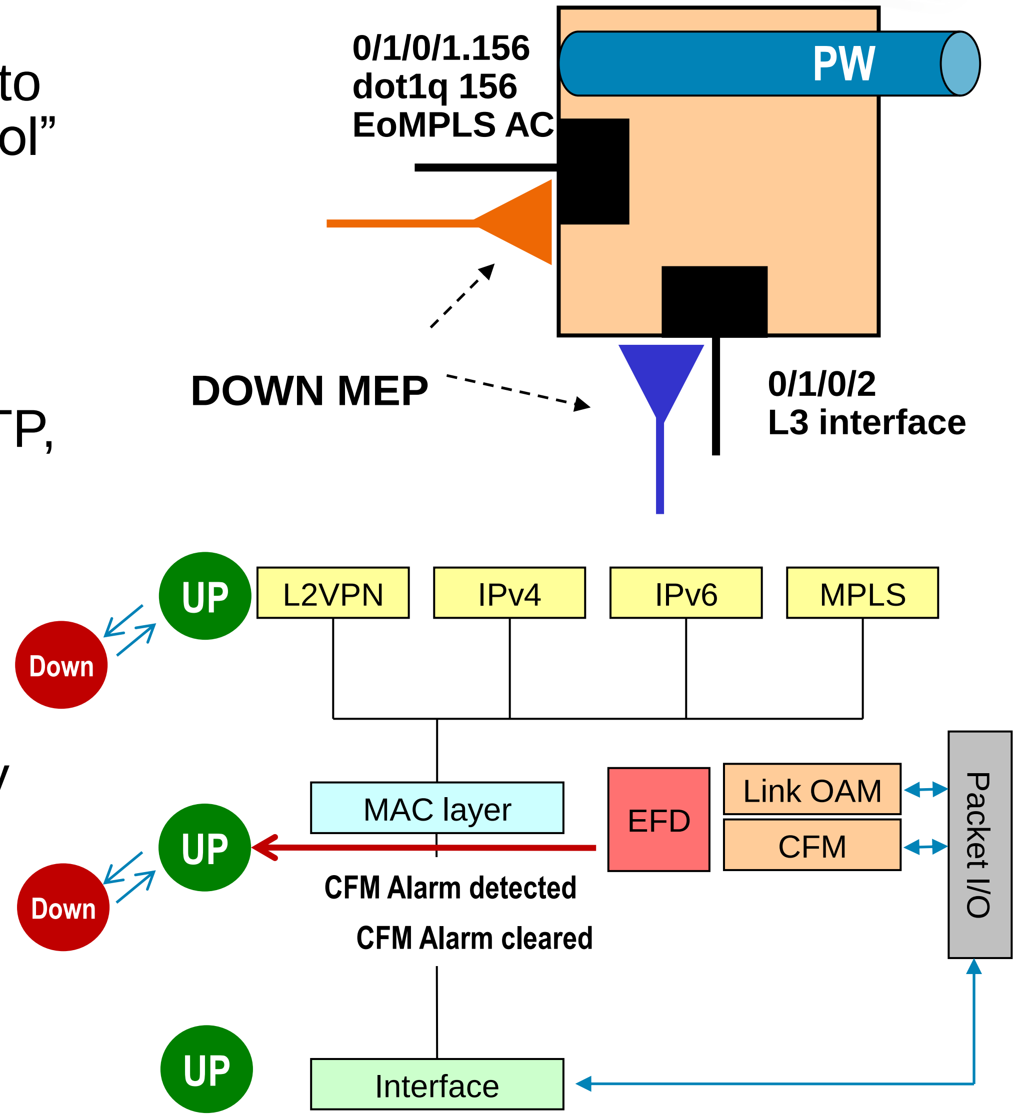
Ethernet OAM
Protocol Positioning



- E-LMI - User to Network Interface (UNI)
- Link OAM - Any point-to-point 802.3 link
- CFM / Y.1731 - End-to-End UNI to UNI
- MPLS OAM - within MPLS cloud

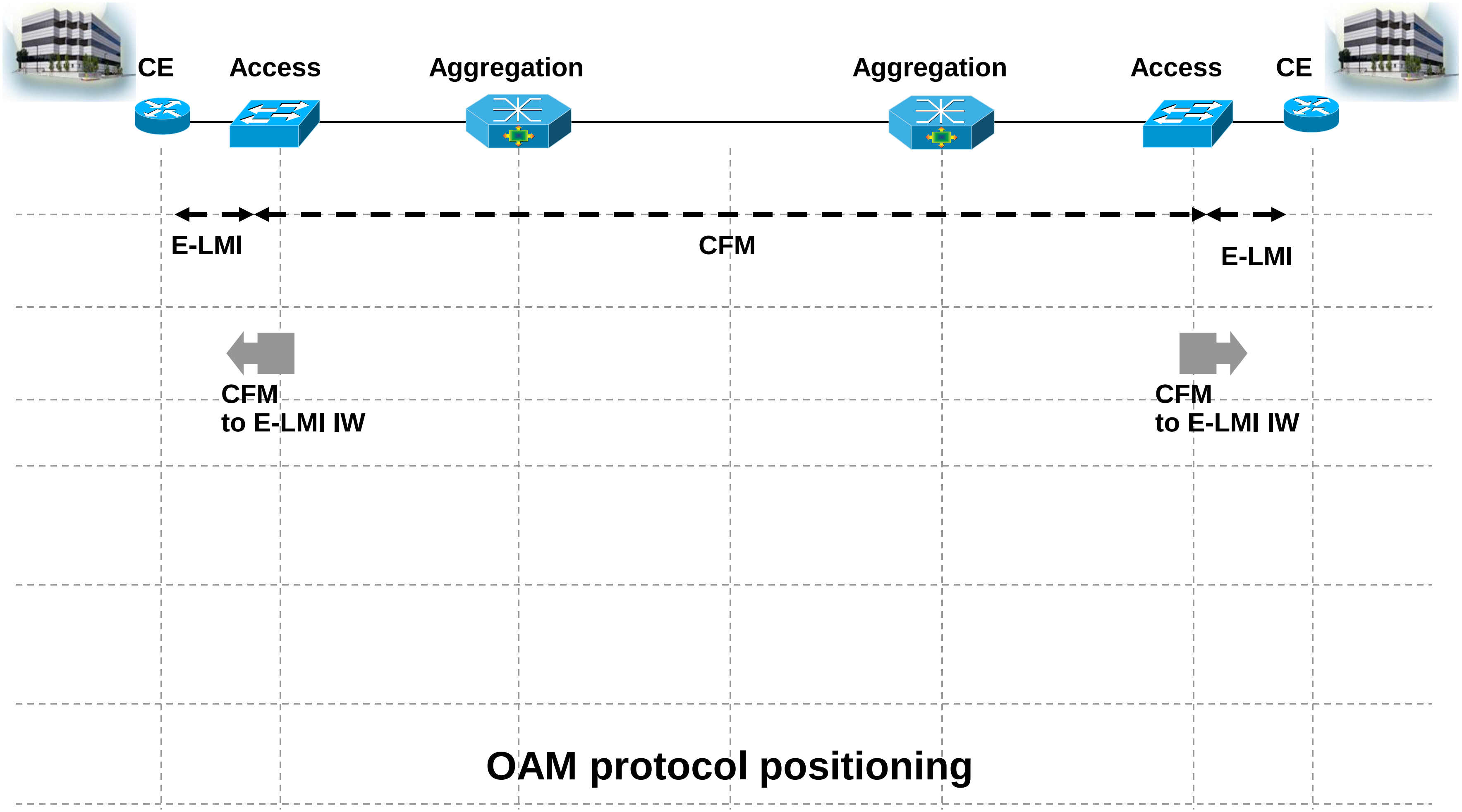
Ethernet Failure Detection (EFD)

- Mechanism for E-OAM protocol to bring down interface “line protocol” state when a defect is detected
- No service frame traffic flows
- (Sub)-Interface is “down” to routing/switching protocols (MSTP, ARP, IGPs, BGP) – will trigger reconvergence
- E-OAM protocol continues to operate
- Brings interface up automatically when defect is resolved



Deploying Carrier Ethernet OAM
Ethernet Layer 2 VPN Services

CFM to ELMI IW scenario

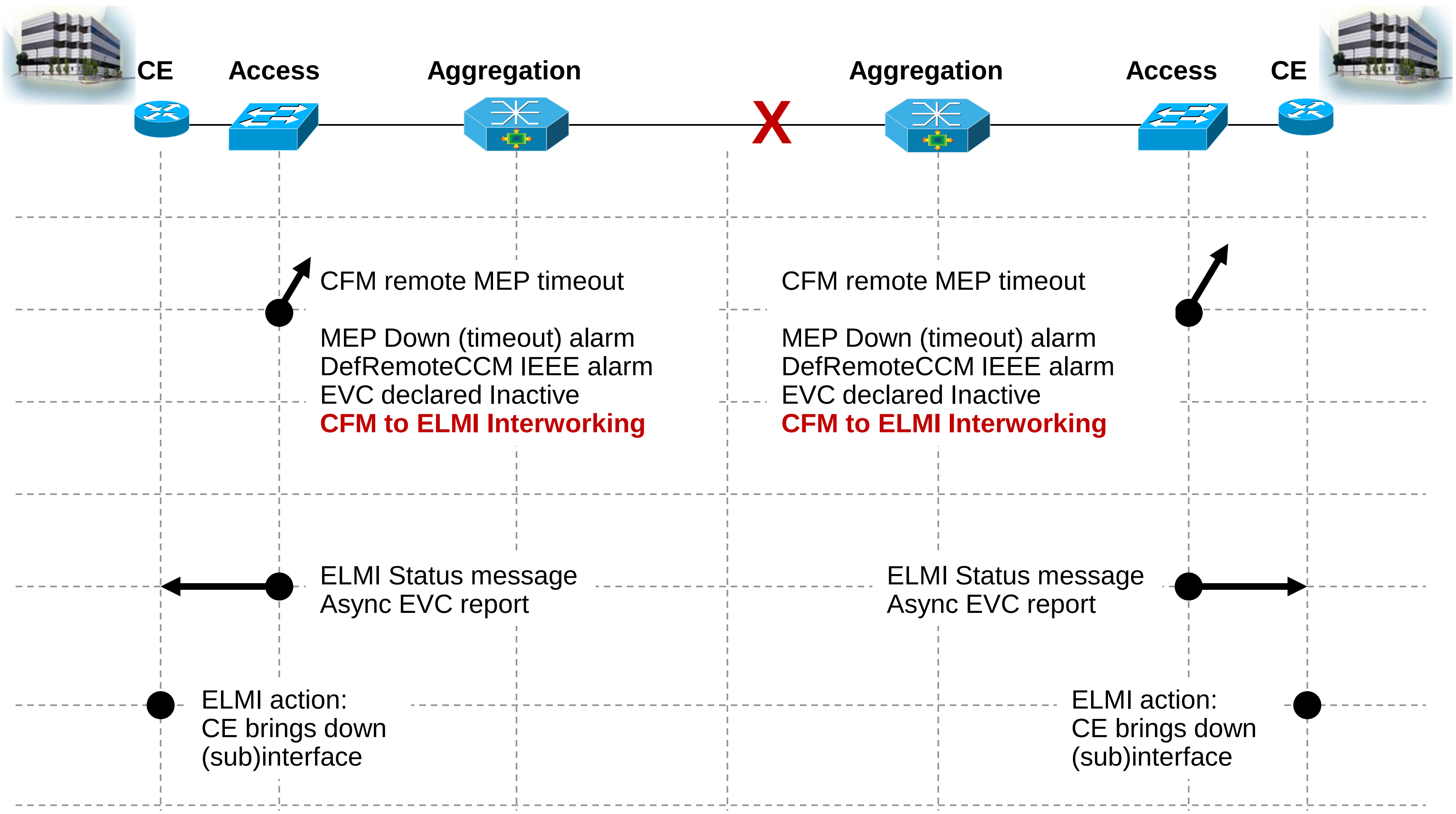


OAM protocol positioning

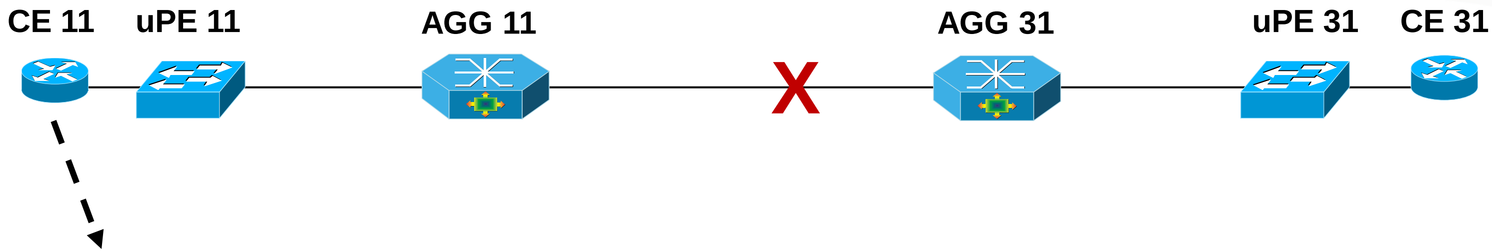
Deploying Carrier Ethernet OAM

Ethernet Layer 2 VPN Services

Failure Scenario: Network Failure



Deploying Carrier Ethernet OAM
Ethernet Layer 2 VPN Services



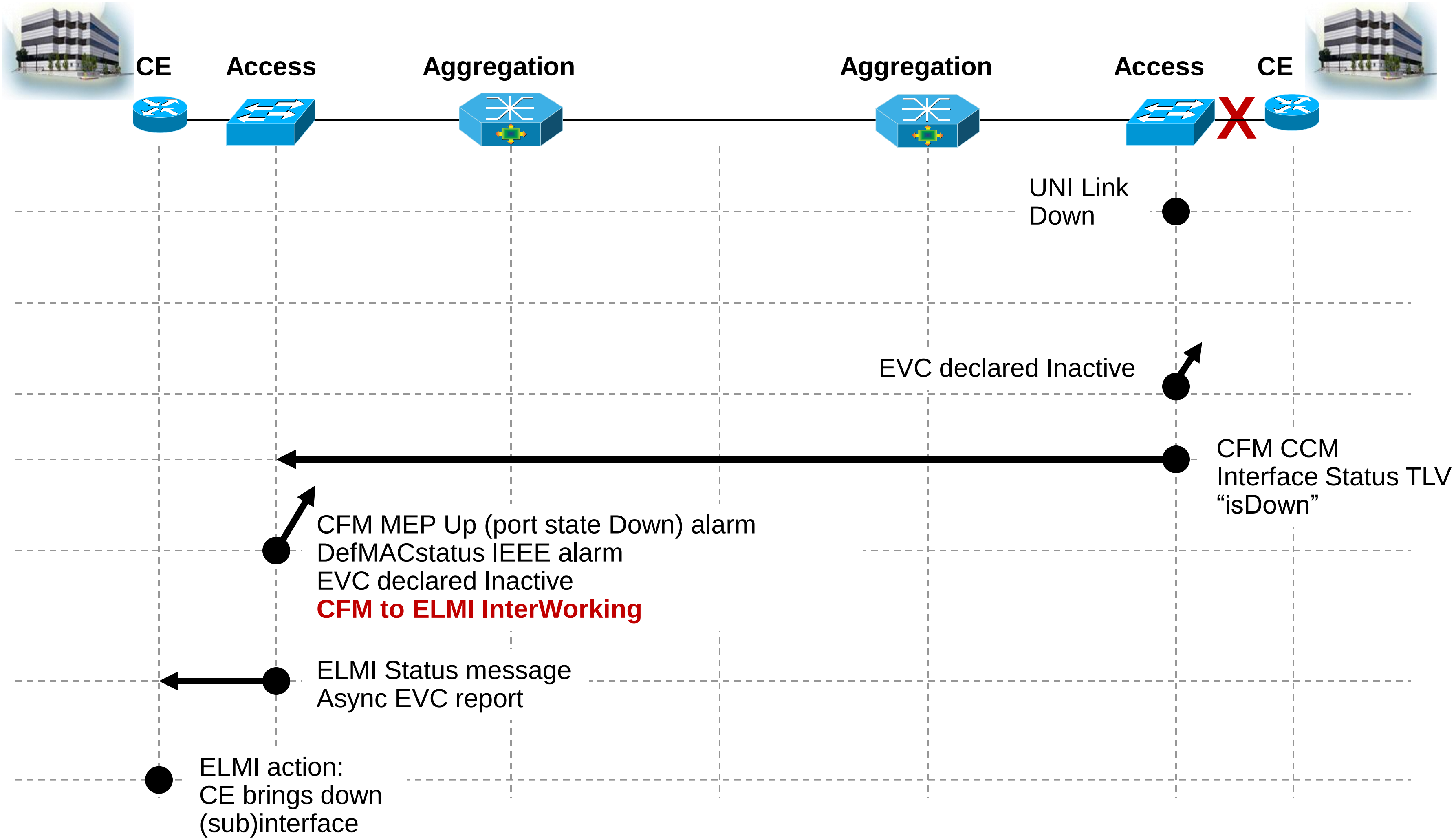

```
CE11#

*Apr  8 04:33:44.991: %LINEPROTO-5-UPDOWN: Line protocol on Interface Ethernet0/0.100,
changed state to down


CE11#show ethernet lmi evc detail EVC_P2P_100
EVC Id: EVC_P2P_100
interface Ethernet0/0
  Time since Last Full Report: 00:01:13
  Ether LMI Link Status: Up
  UNI Status: Up
  UNI Id: CE11_UNI
  CE-VLAN/EVC Map Type: Service Multiplexing with no bundling
  VLAN: 100

  EVC Status: Inactive
  EVC Type: Point-to-Point
  Remote UNI Count: Configured = 1, Active = 0

  UNI Id                                    UNI Status    Port
  ------                                    ----------    ----
  CE31_UNI                                  Unreachable   Remote
```

Network Failure:
Remote UNI shows
UNREACHABLE

Deploying Carrier Ethernet OAM

Ethernet Layer 2 VPN Services

Failure Scenario: UNI Link Down



Deploying Carrier Ethernet OAM

Ethernet Layer 2 VPN Services

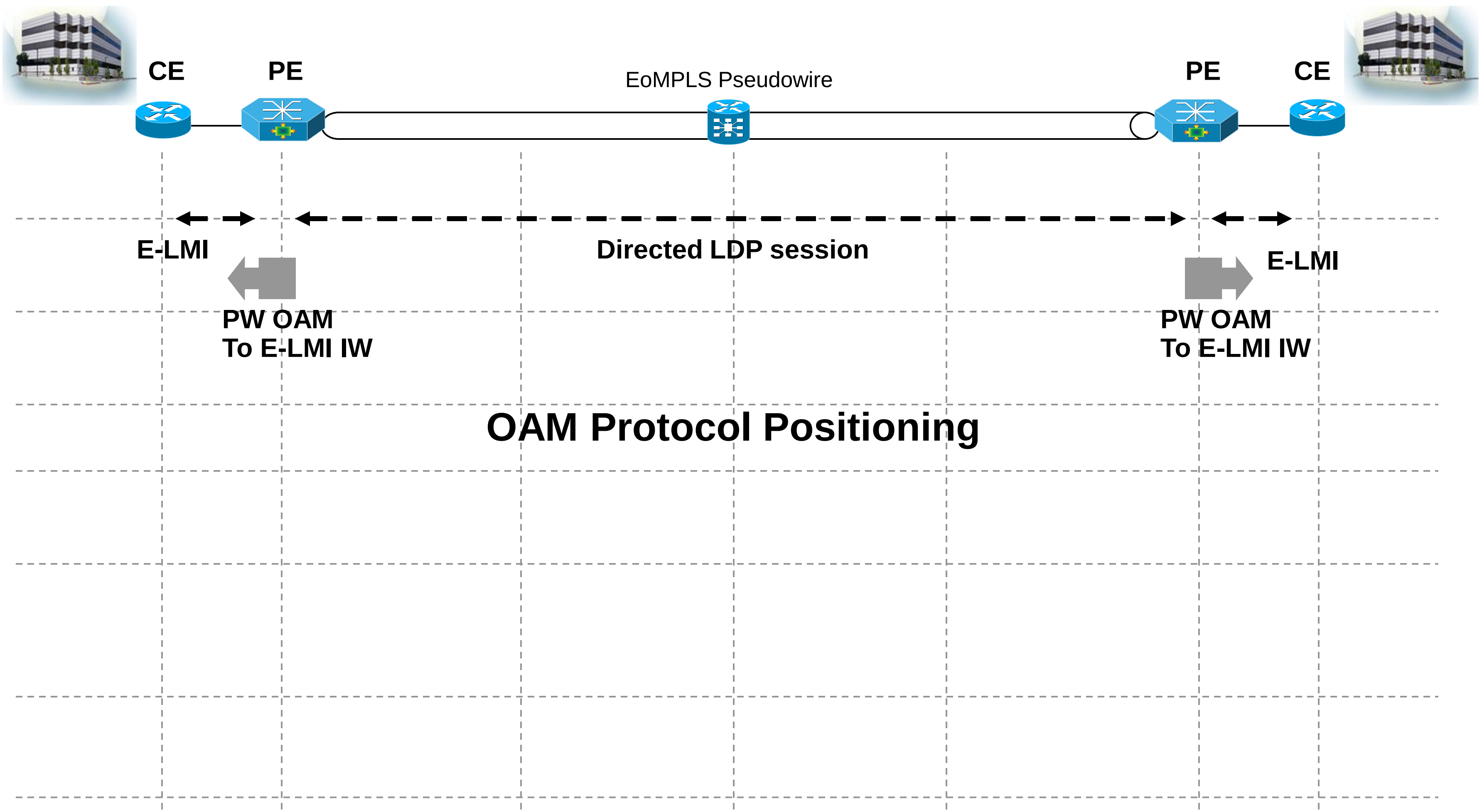

```
CE11#

*Apr  8 04:41:54.907: %LINEPROTO-5-UPDOWN: Line protocol on Interface Ethernet0/0.100,
changed state to down

CE11#show ethernet lmi evc detail EVC_P2P_100
EVC Id: EVC_P2P_100
interface Ethernet0/0
  Time since Last Full Report: 00:01:07
  Ether LMI Link Status: Up
  UNI Status: Up
  UNI Id: CE11_UNI
  CE-VLAN/EVC Map Type: Service Multiplexing with no bundling
  VLAN: 100

  EVC Status: Inactive
  EVC Type: Point-to-Point
  Remote UNI Count: Configured = 1, Active = 0

  UNI Id                                    UNI Status    Port
  ------                                    ----------    ----
  CE31_UNI                                  Down          Remote
```

UNI Failure:
Remote UNI shows DOWN

Deploying Carrier Ethernet OAM

Ethernet Layer 2 VPN Services

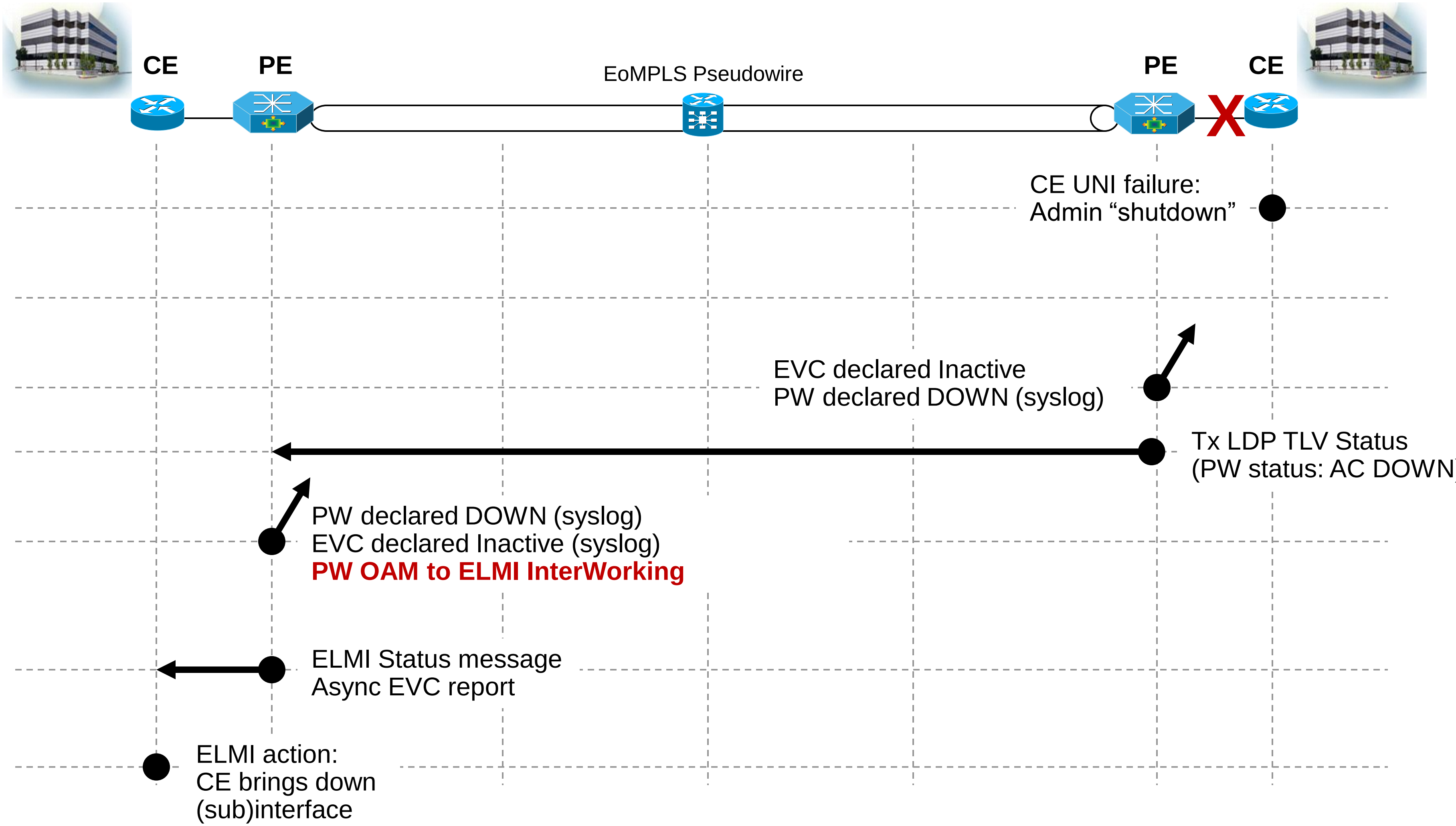
PW OAM to ELMI IW scenario



Deploying Carrier Ethernet OAM

Ethernet Layer 2 VPN Services

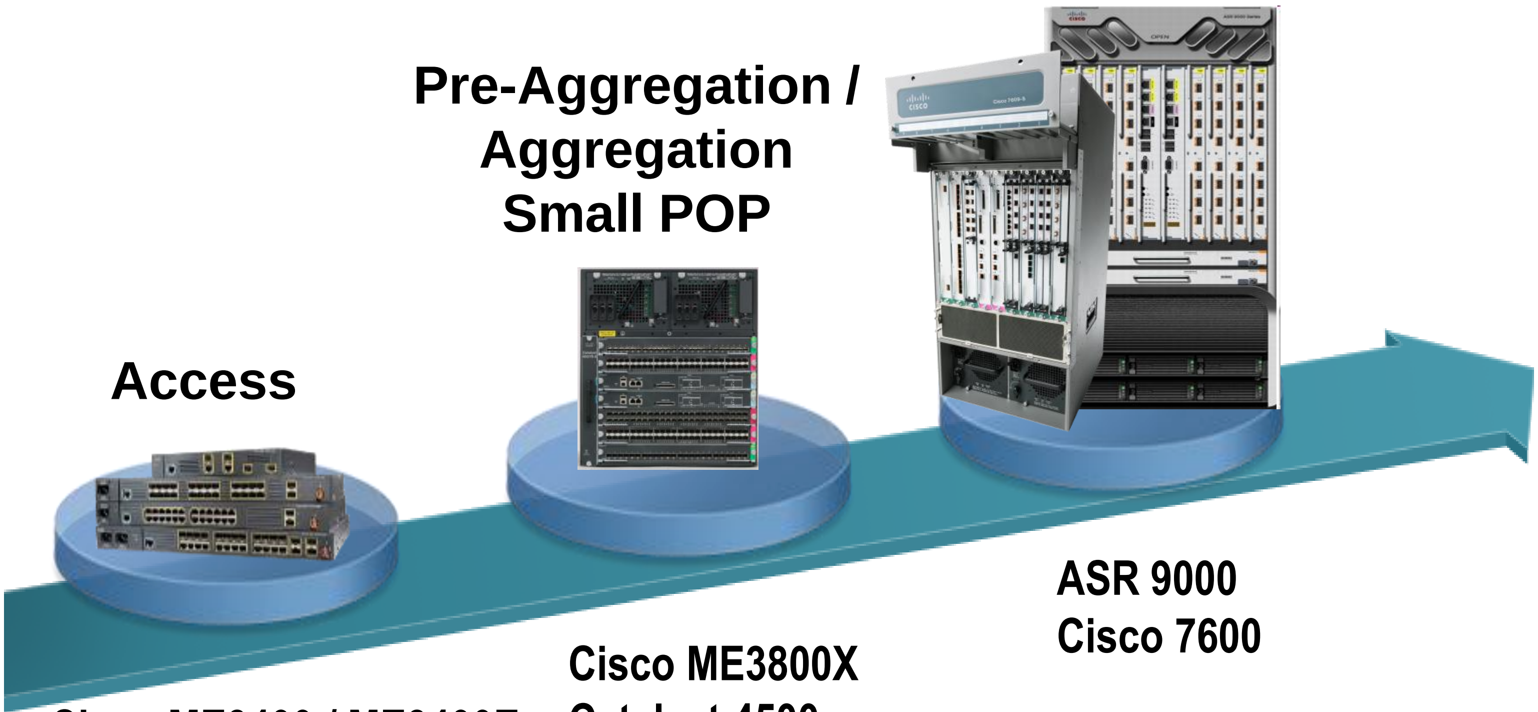
Failure Scenario: UNI Failure



Learn. Connect. Collaborate. together.

Platform Support

Carrier Ethernet Portfolio
Cisco Platform Support



Access

Cisco ME3400 / ME3400E
Cisco ME3600X
Cisco MWR 2941
Cisco ME4924
Catalyst 4900
Cisco ASR 901

Pre-Aggregation / Aggregation Small POP

Cisco ME3800X
Catalyst 4500
Cisco ASR 903

Aggregation Large POP

ASR 9000
Cisco 7600

Summary

Summary

- Various access redundancy mechanisms are available, which enable node as well as network multi-homing:
  - Multichassis LACP (mLACP)
  - MST Access Gateway (MST-AG)
  - REP Access Gateway
- Aggregation/core redundancy mechanisms operating at the pseudowire layer primarily protect against PE node failures:
  - One-way Pseudowire Redundancy
  - Two-way Pseudowire Redundancy
- Above mechanisms can interwork to provide comprehensive end-to-end resiliency solutions for E-Line and E-LAN services

Recommended Reading

Learn. Connect.
Collaborate. together.

Please visit the Cisco Store for suitable reading.

References

- Cisco IOS — L2VPN Pseudowire Redundancy

  http://www.cisco.com/en/US/docs/ios/wan/configuration/guide/wan_l2vpn_pw_red_ps6922_TSD_Products_Configuration_Guide_Chapter.html

- Cisco IOS — Multichassis LACP Configuration Guide

  http://www.cisco.com/en/US/docs/ios/cether/configuration/guide/ce_mlacp.html

- Cisco ME 3400 / 3400E — REP Configuration Guide

  http://www.cisco.com/en/US/docs/switches/metro/me3400e/software/release/12.2_55_se/configuration/guide/swrep.html

- Cisco 7600 — ES+ Layer 1 and Layer 2 features (covering MST / REP on EVC, Two-way PW redundancy, ICCP, mLACP, MST-AG)

  http://www.cisco.com/en/US/docs/routers/7600/install_config/ES40_config_guide/es40_chap4.html

- Cisco 7600 — H-VPLS N-PE Redundancy for QinQ and MPLS Access (covering MST on nPE, LDP MAC Address Withdrawal)

  http://www.cisco.com/en/US/docs/ios/mpls/configuration/guide/mp_hvpls_npe_red.html

- Cisco 7600 — Link State Tracking

  http://www.cisco.com/en/US/docs/routers/7600/ios/15S/configuration/guide/lst.html

References (Cont.)

- Cisco ASR 9000 — Configuring Link Bundles (covering Multichassis LACP)

  http://www.cisco.com/en/US/docs/routers/asr9000/software/asr9k_r4.0/lxvpn/configuration/guide/lesc40lbun.html

- Cisco ASR 9000 — L2VPN and Ethernet Services Configuration Guide (covering MST, MST-AG, PW Redundancy, LDP MAC Address Withdrawal)

  http://www.cisco.com/en/US/docs/routers/asr9000/software/asr9k_r4.0/lxvpn/configuration/guide/lesc40.html

Acronyms—IP and MPLS

Acronym	Description
AC	Attachment Circuit
AS	Autonomous System
BFD	Bidirectional Failure Detection
CoS	Class of Service
ECMP	Equal Cost Multipath
EoMPLS	Ethernet over MPLS
FRR	Fast Re-Route
H-VPLS	Hierarchical VPLS
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
LDP	Label Distribution Protocol
LER	Label Edge Router
LFIB	Labeled Forwarding Information Base
LSM	Label Switched Multicast
LSP	Label Switched Path
LSR	Label Switching Router
MPLS	Multi-Protocol Label Switching
NLRI	Network Layer Reachability Information
PSN	Packet Switch Network

Acronym	Description
PW	Pseudo-Wire
PWE3	Pseudo-Wire End-to-End Emulation
QoS	Quality of Service
RD	Route Distinguisher
RIB	Routing Information Base
RR	Route Reflector
RSVP	Resource Reservation Protocol
RSVP-TE	RSVP based Traffic Engineering
RT	Route Target
TE	Traffic Engineering
tLDP	Targeted LDP
VC	Virtual Circuit
VCID	VC Identifier
VFI	Virtual Forwarding Instance
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VPWS	Virtual Private Wire Service
VRF	Virtual Route Forwarding Instance
VSI	Virtual Switching Instance

Acronyms—Ethernet/Bridging

Acronym	Description
ACL	Access Control List
BD	Bridge Domain
BPA	Blocked Port Advertisement (REP PDU)
BPDU	Bridge Protocol Data Unit
BRAS	Broadband Access Server
CE	Customer Equipment (Edge)
C-VLAN / CE-VLAN	Customer / CE VLAN
CoS	Class of Service
DHD	Dual Homed Device
DSLAM	DSL Access Modulator
E-LAN	Ethernet LAN service (multipoint)
E-Line	Ethernet Line service (point-to-point)
E-Tree	Ethernet Tree service (rooted multipoint)
EFP	Ethernet Flow Point
EPL	Ethernet Private Line
EP-LAN	Ethernet Private LAN
EVC	Ethernet Virtual Connection
EVPL	Ethernet Virtual Private Line

Acronym	Description
EVP-LAN	Ethernet Virtual Private LAN
ICCP	Inter-Chassis Communication Protocol
IEEE	Institute of Electrical and Electronics Engineers
IPoETV	TV on IP over Ethernet
IPTV	Television over IP
L2GP	Layer 2 Gateway Ports
LACP	Link Aggregation Control Protocol
LAN	Local Area Network
MEF	Metro Ethernet Forum
MEN	Metro Ethernet Network
MIRP	Multiple I-Tag Registration Protocol
mLACP	Multi-Chassis LACP
MRP	Multiple Registration Protocol
MST / MSTP	Multiple Instance STP
MSTG-AG	MST Access Gateway

Acronyms—Ethernet/Bridging (Cont.)

Acronym	Description
MSTi	MST Instances
MTBF	Mean Time Between Failures
MTTR	Mean Time To Recover
MVRP	Multiple VLAN Registration Protocol
OAM	Operations, Administration and Maintenance
PE	Provider Edge device
PoA	Point of Attachment
Q-in-Q	VLAN tunneling using two 802.1Q tags
QoS	Quality of Service
R-L2GP	Reverse L2GP
REP	Resilient Ethernet Protocol
REP-AG	REP Access Gateway
RG	Redundancy Group
SLA	Service Level Agreement
SLS	Service Level Specification
STP	Spanning Tree Protocol
SVI	Switch Virtual Interface (interface vlan)
S-VLAN	Service VLAN (Provider VLAN)
TC	Topology Change

Acronym	Description
TCN	Topology Change Notification
UNI	User to Network Interface
VID	VLAN Identifier
VLAN	Virtual LAN
VoD	Video on Demand
VoIP	Voice over IP

Acronyms—
Provider Backbone Bridging

Acronym	Description
B-BEB	B-Component BEB
BCB	Backbone Core Bridge
B-DA	Backbone Destination Address
BEB	Backbone Edge Bridge
B-MAC	Backbone MAC Address
B-SA	Backbone Source Address
B-Tag	B-VLAN Tag
B-VLAN	Backbone VLAN
C-DA	Customer Destination Address
CE	Customer Equipment (Edge)
C-MAC	Customer MAC Address
C-SA	Customer Source Address
80	C-VLAN Tag
C-VLAN / CE-VLAN	Customer / CE VLAN
DA	Destination MAC Address
FCS	Frame Check Sequence
IB-BEB	Combined I-Component & B-Component BEB

Acronym	Description
I-BEB	I-Component BEB
IEEE	Institute of Electrical and Electronics Engineers
I-SID	Instance Service Identifier
I-Tag	I-SID Tag
MAC	Media Access Control
N-PE	Network-facing Provider Edge device
PB	Provider Bridge
PBB	Provider Backbone Bridge / Bridging
PBBN	Provider Backbone Bridging Network
PBN	Provider Bridging Network
PE	Provider Edge device
Q-in-Q	VLAN tunneling using two 802.1Q tags
SA	Source MAC Address
S-Tag	S-VLAN Tag
S-VLAN	Service VLAN (Provider VLAN)
UNI	User to Network Interface
U-PE	User-facing Provider Edge device
VLAN	Virtual LAN

Thank you.

Learn. Connect. Collaborate. together.

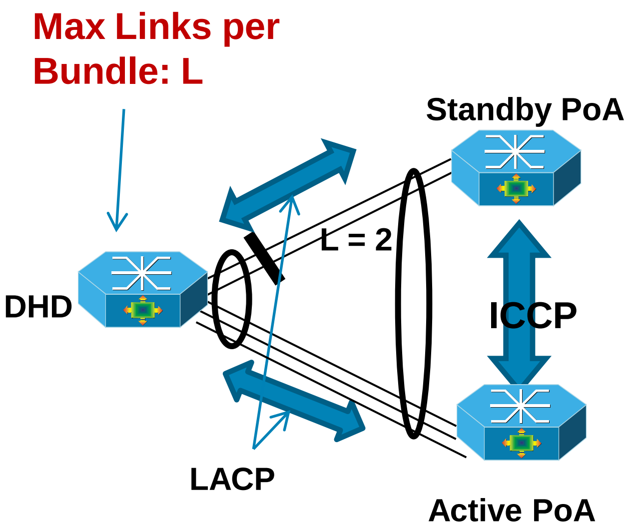
Appendix

Multi-Chassis LACP (mLACP) and Inter-Chassis Communication Protocol (ICCP)

Operational Variants

DHD-Based Control

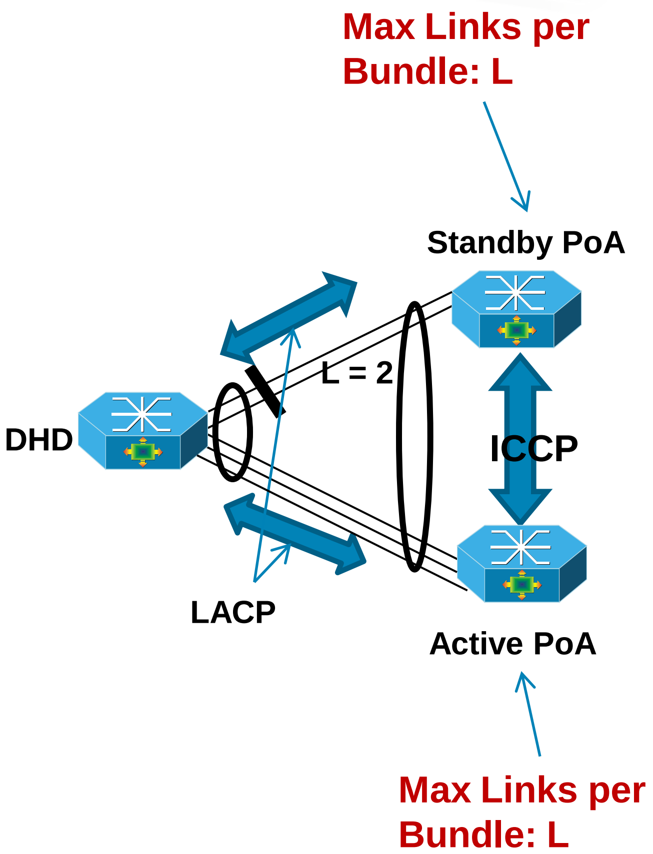
- DHD is configured to limit the maximum number of links per bundle
  - Limit must be set to L, where L is the minimum number of links from DHD to any single PoA
- PoAs must be configured with Minimum Links per Bundle policy set to L as well
  - This prevents unsupported scenario where uplinks from DHD to both PoAs attempt to go active
- Selection of active/standby links is the responsibility of the DHD
- Advantages: Split Brain condition can be easily detected
- Disadvantages: If DHD does not support LACP fast failover, then failover time will be standard



Operational Variants

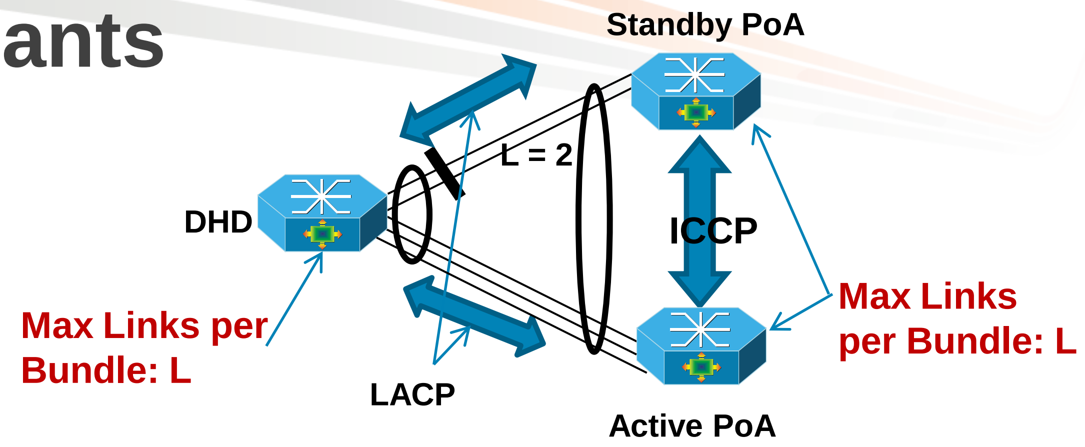
PoA-Based Control

- Each PoA is configured to limit the maximum number of links per bundle
  - Limit must be set to L, where L is the minimum number of links from DHD to any single PoA
- Selection of active/standby links is the responsibility of the PoAs
- Advantages: Faster switchover times compared to other variants, and Minimum Link policy on PoA can be flexible
- Disadvantage: If ICCP transport is lost, Split Brain condition would occur



Operational Variants

Shared Control



- DHD and PoAs are configured to limit the maximum number of links per bundle
  - Limit must be set to L, where L is the minimum number of links from DHD to any single PoA
- Selection of active/standby links is the responsibility of DHD and PoAs combined
- Advantages: Split brain condition can be detected, and Minimum Link policy on PoA can be flexible
- Disadvantages: If DHD does not support LACP fast failover, then failover time will be standard

Appendix
End-to-End Redundancy Solutions (Cont.)

E-LAN Availability Models

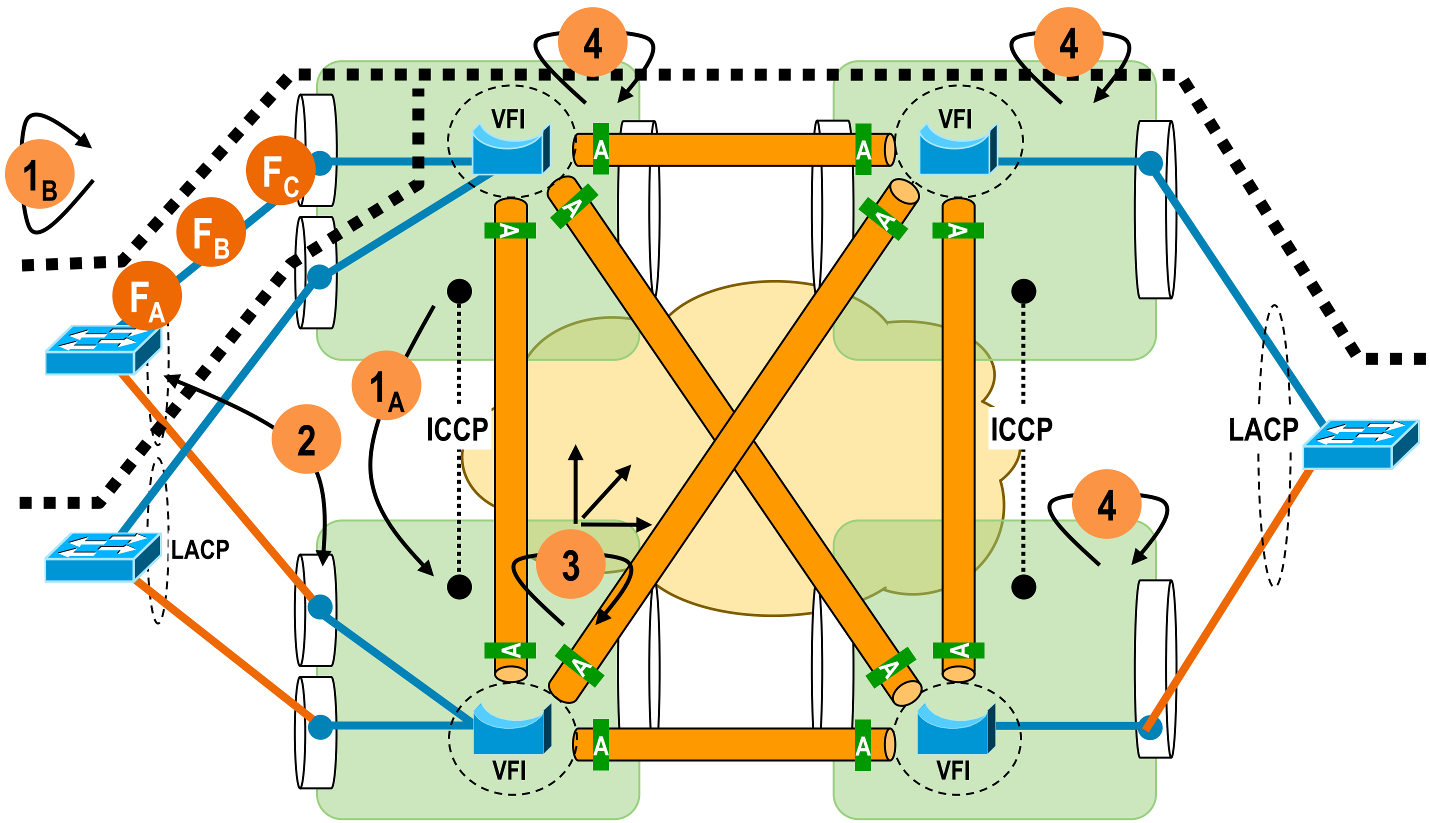
Active/Backup Access Node Redundancy (mLACP)

E-LAN Availability Model

Active / Backup Access Node Redundancy (mLACP)

VPLS

- Port / Link Failures



Events	
I	Initial state
F_{A-C}	Port / Link Failures
1_A	Active PoA detects failure and signals failover over ICCP
1_B	Failover triggered on DHD
2	Standby link brought up per LACP proc.
3	Standby PoA flushes MAC table and triggers LDP MAC add. withdrawal to remote peers
4	Remote PEs flush MAC addresses

- For VPLS Decoupled Mode, VFI's PWs always advertised in Active state, regardless of AC state

Forwarding EoMPLS PW   Non-Forwarding EoMPLS PW

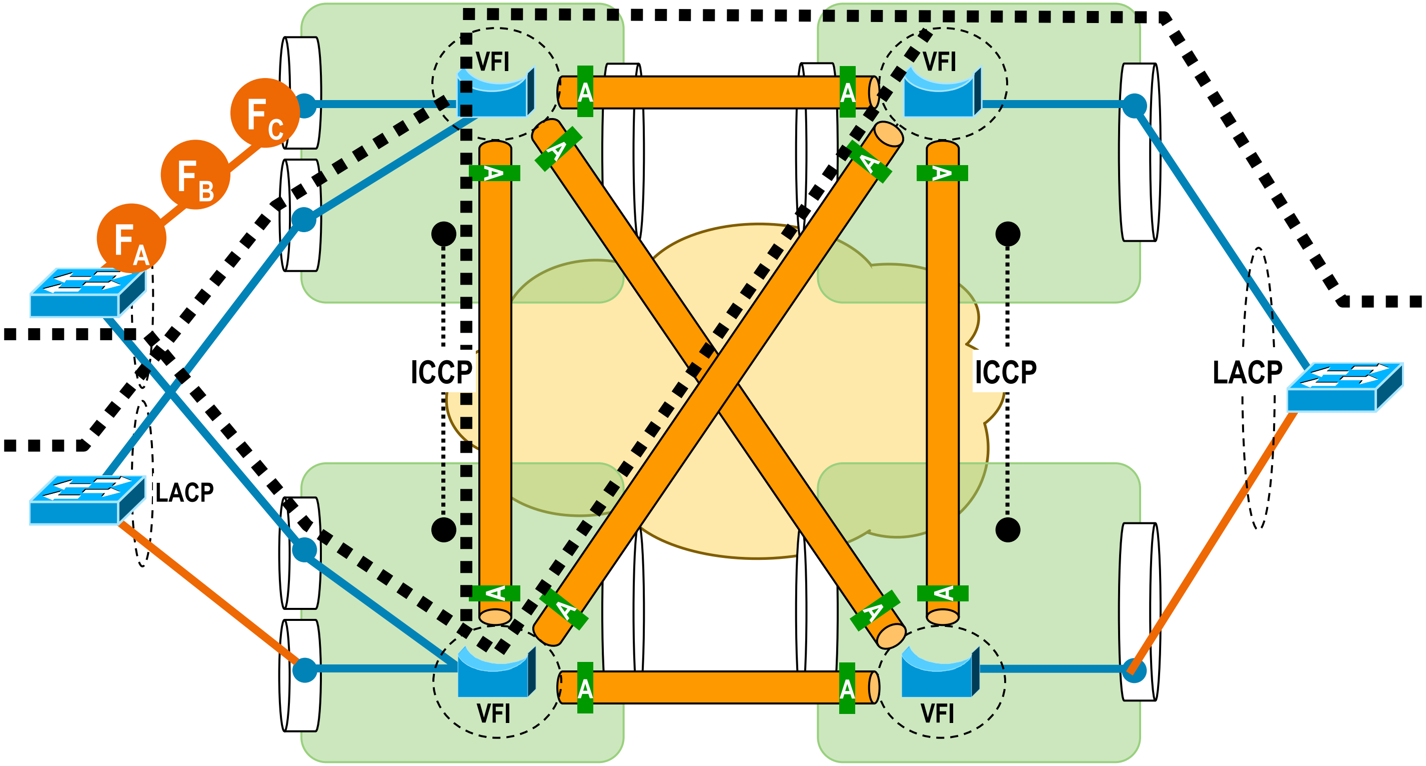
E-LAN Availability Model

Active / Backup Access Node Redundancy (mLACP)

VPLS

- Port / Link Failures (cont.)

Events	
4	Remote PEs flush MAC addresses
E	End State



  Cisco Public Forwarding EoMPLS PW Non-Forwarding EoMPLS PW

Learn. Connect. Collaborate. together.

E-LAN Availability Models

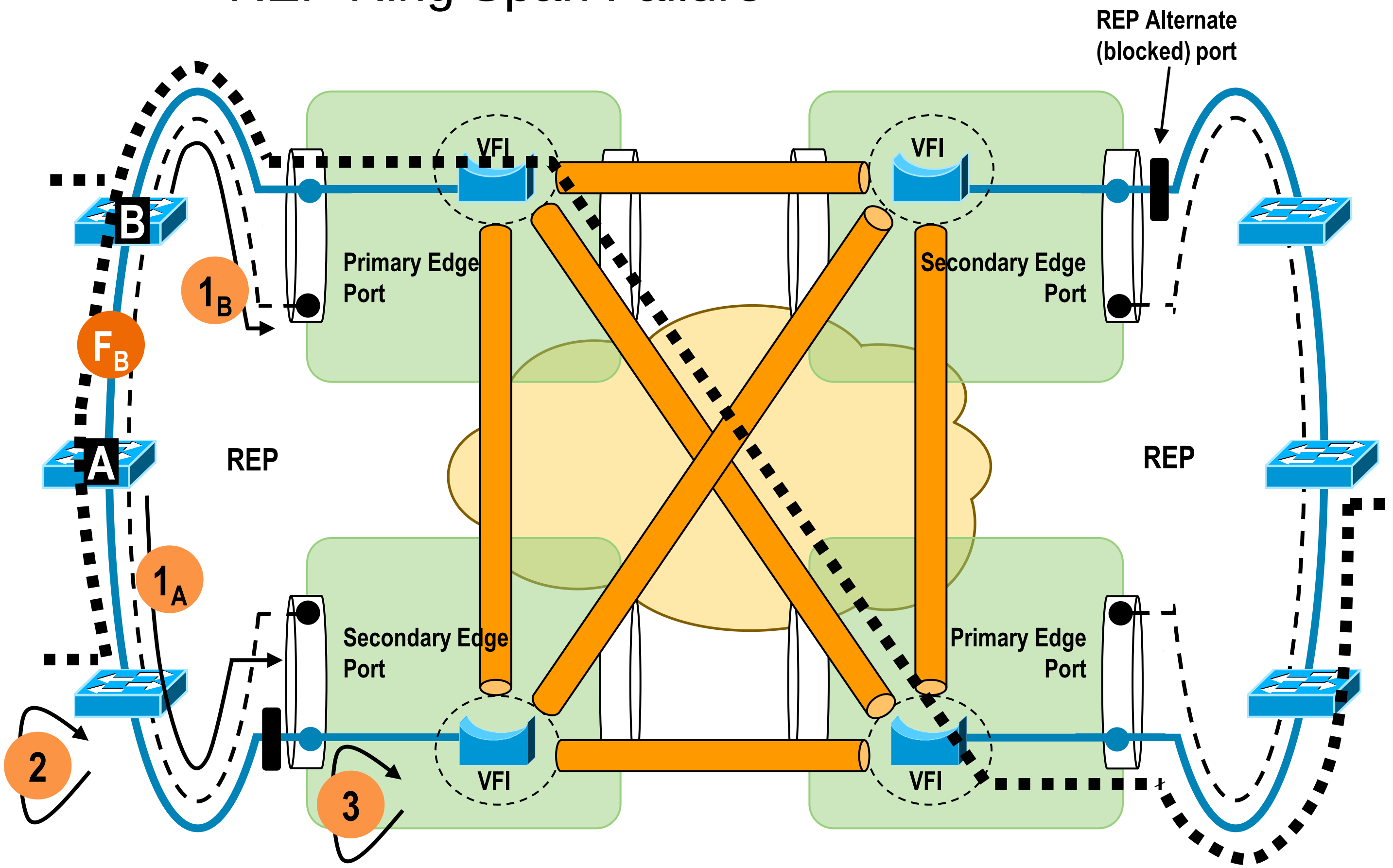
Ring Access Node Redundancy (REP)

VPLS

E-LAN Availability Model

Ring Access Node Redundancy (REP)

- REP Ring Span Failure



Events	
I	Initial state
F_B	Ring Span failure
1_{A-B}	Access switches "A" and "B"detect link failure. Send Blocked Port Advertisement (BPA) with TC bit set on the segment
2	Access nodes in the ring flush MAC tables and propagate BPA
3	AGG node receives BPA and unblocks alternate port

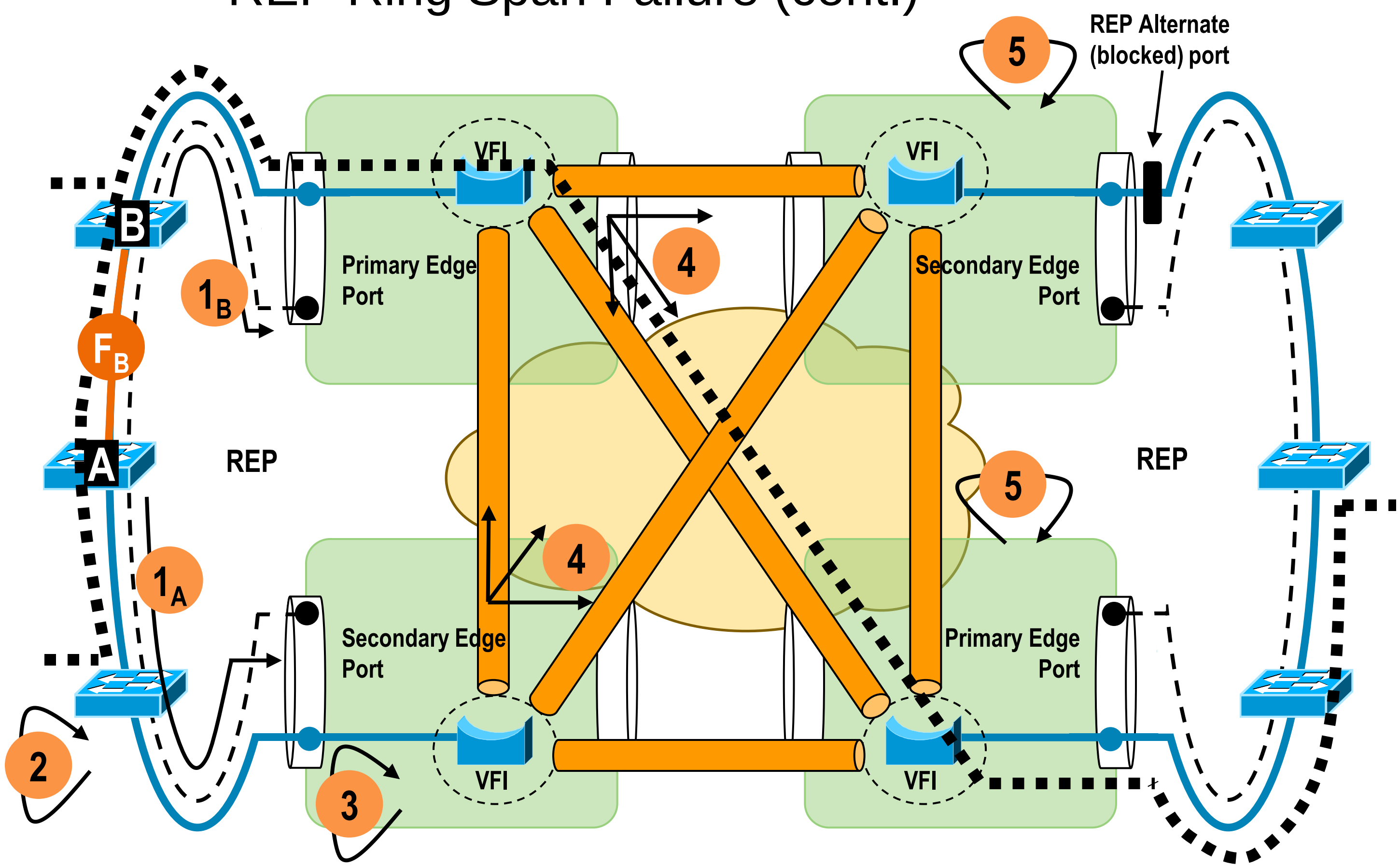
- REP enabled segment with Edge Ports on Aggregation Nodes
- VLAN load balancing using Alternate Port configured on Secondary Edge Port

Forwarding EoMPLS PW
Non-Forwarding EoMPLS PW
REP Admin vlan

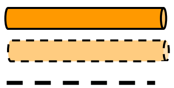
E-LAN Availability Model

Ring Access Node Redundancy (REP)

- REP Ring Span Failure (cont.)



Events	
3	AGG node receives BPA and unblocks alternate port
4	AGG nodes flush MAC tables. Trigger LDP MAC add withdrawal to VPLS peers
5	Remote peers flush MAC tables

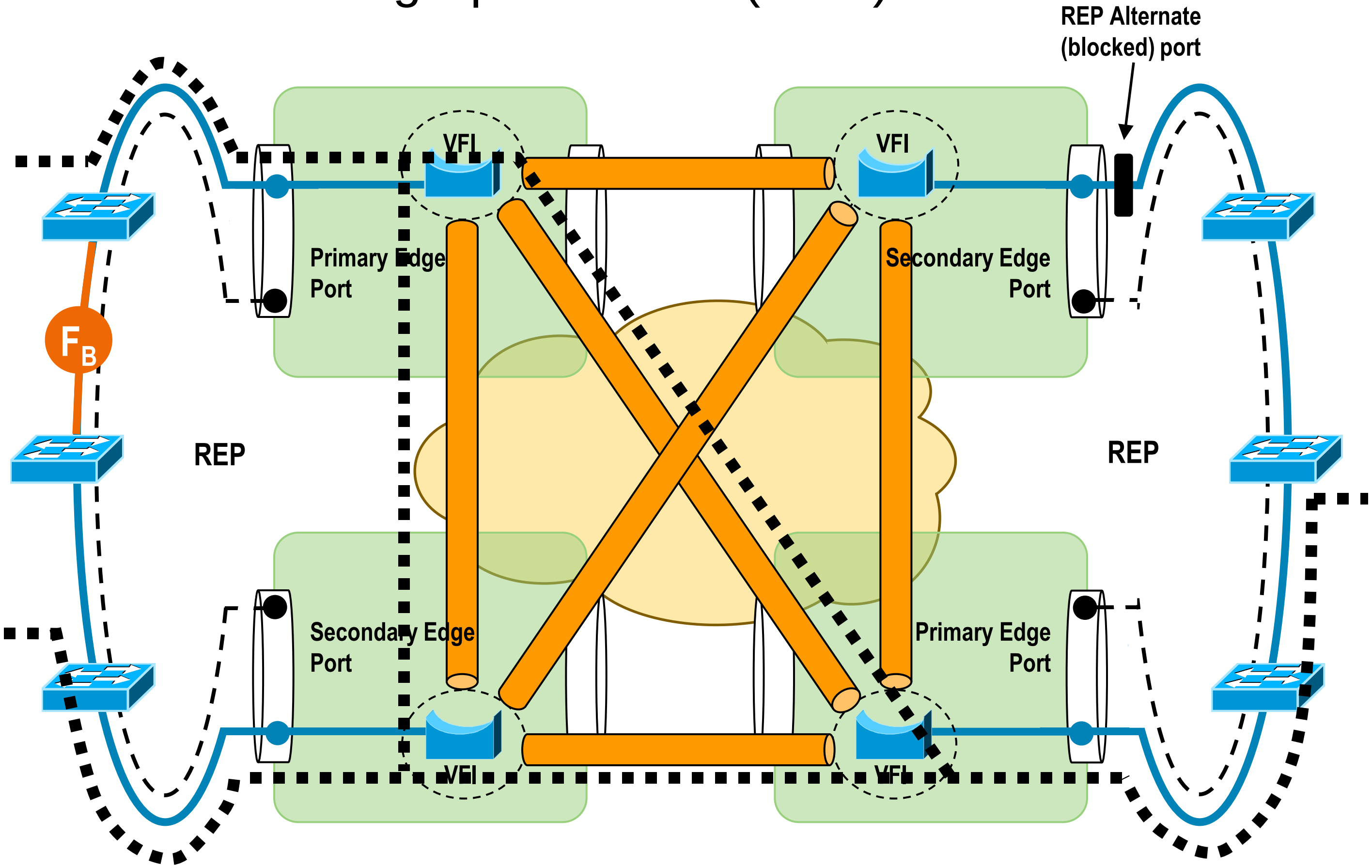


E-LAN Availability Model

Ring Access Node Redundancy (REP)

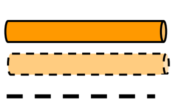
VPLS

- REP Ring Span Failure (cont.)



Events	
5	Remote peers flush MAC tables
E	End State

- Topology depicted shows full mesh VPLS but can also be implemented using H-VPLS with Active/Standby PWs



E-LAN Availability Models

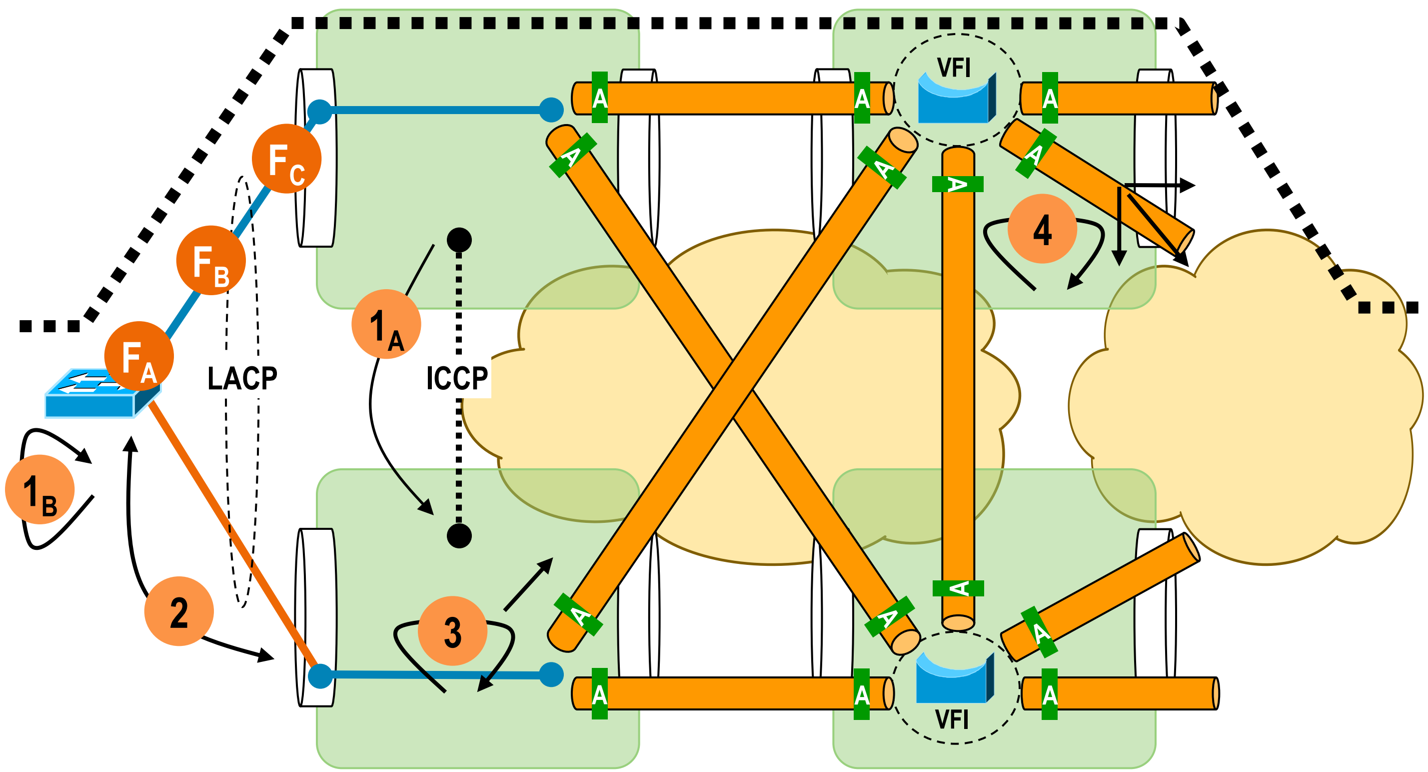
H-VPLS (MPLS Access)
Active/Backup Access Node Redundancy (mLACP)

E-LAN Availability Model

H-VPLS

Active / Backup Access Node Redundancy (mLACP)

- Port / Link Failures



- For H-VPLS Decoupled Mode, Primary/Backup PW in active/active states respectively, regardless of AC state

Events	
I	Initial state
$F_{A\text{-}C}$	Port / Link Failures
1_A	Active PoA detects failure and signals failover over ICCP
1_B	Failover triggered on DHD
2	Standby link brought up per LACP proc.
3	Standby PoA flushes MAC table and triggers LDP MAC add. withdrawal to VPLS hub PE
4	Hub PE flushes MAC addresses and triggers LDP MAC address withdrawal to other hub PEs

Forwarding EoMPLS PW

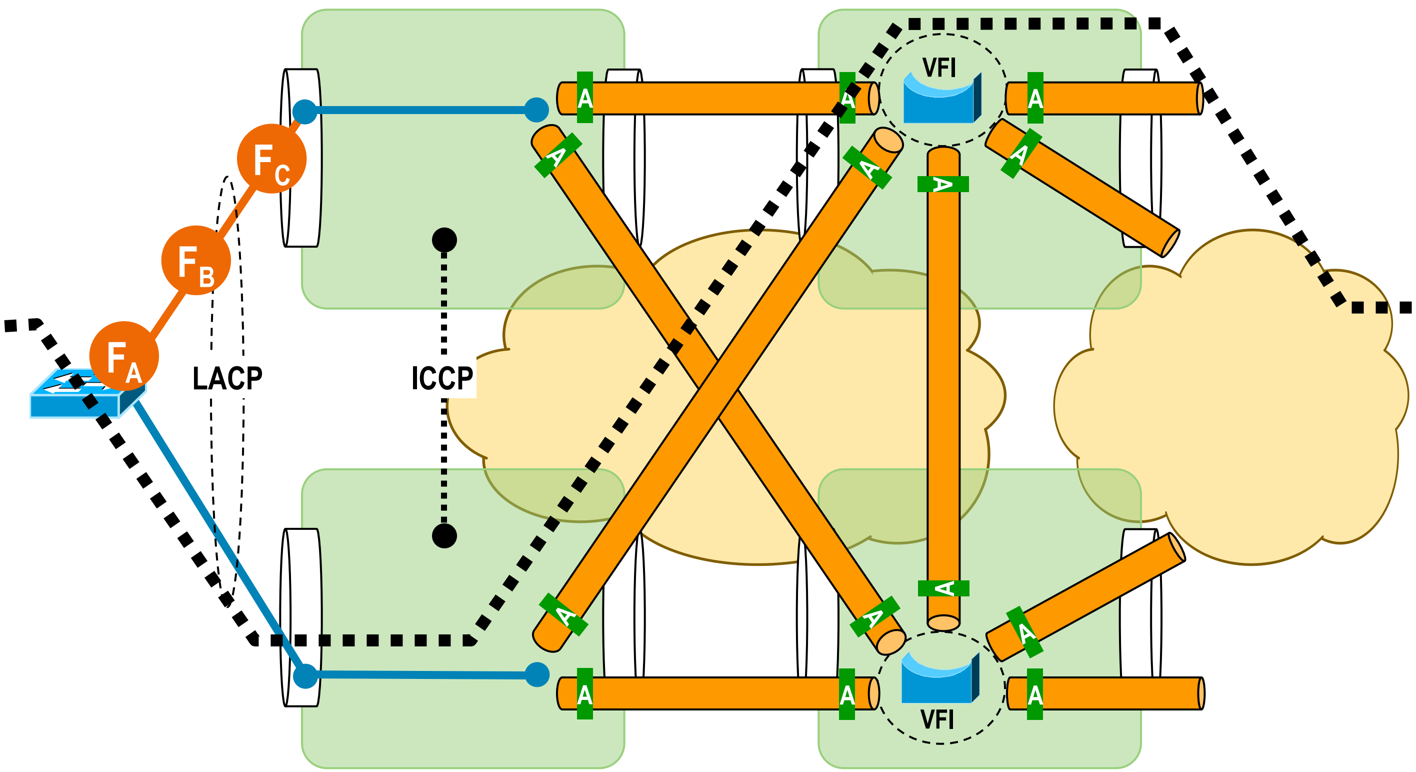
Non-Forwarding EoMPLS PW

E-LAN Availability Model

Active / Backup Access Node Redundancy (mLACP)

H-VPLS

- Port / Link Failures



Events	
4	Hub PE flushes MAC addresses and triggers LDP MAC address withdrawal to other hub PEs
E	End State

- Failure of VPLS Hub PE (detected by loss of routing adjacency (IP route-watch)), triggers failover to backup PW – No LACP switchover performed

Forwarding EoMPLS PW   Non-Forwarding EoMPLS PW