

# MPLS Architectural Considerations for a Transport Profile

**ITU-T - IETF Joint Working Team  
Dave Ward, Malcolm Betts, ed.**

**April 18, 2008**

# Table of Contents

- Executive Overview
  - Recommendation
- Introduction and Background Material
- High Level Architecture
- OAM Requirements
- OAM Mechanisms and Baseline Use Cases
- Associated Channel Level (ACH)
- Forwarding and OAM
  - LSP/PW OAM
  - Use Case Scenario and Label Stack Diagrams
  - Use of TTL for MIP OAM alert
  - Packet Context
- Control Plane
- Survivability
- Network Management
- Summary

# Executive Summary

# Recommendation

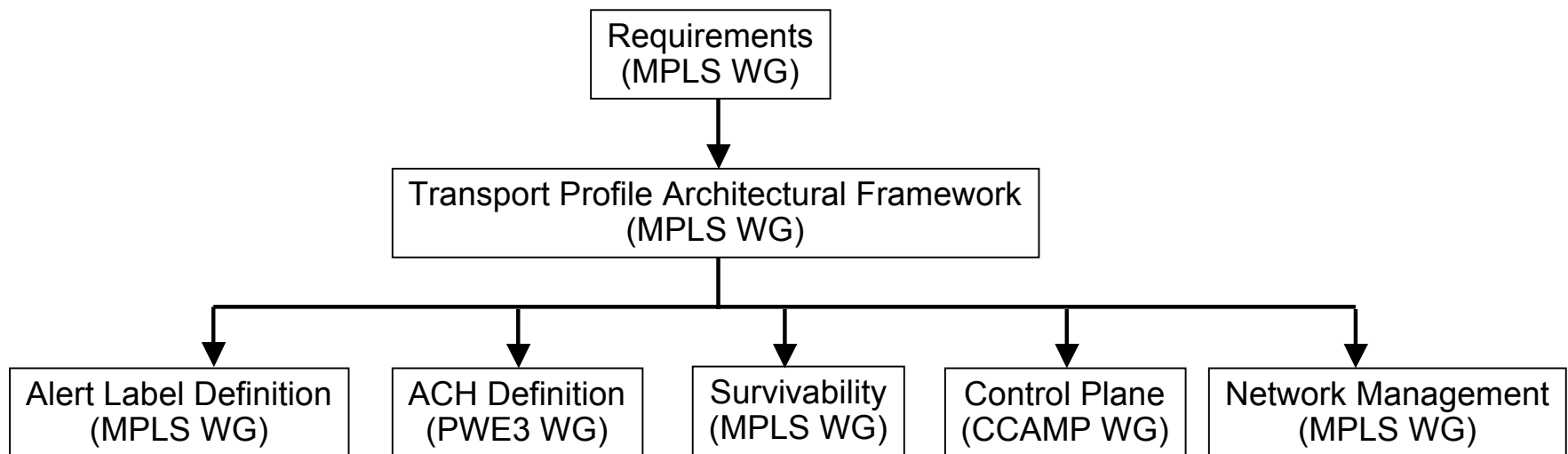
- Consensus on recommendation of Option 1
  - Jointly agree to work together and bring transport requirements into the IETF and extend IETF MPLS forwarding, OAM, survivability, network management and control plane protocols to meet those requirements through the IETF Standards Process
  - The Joint Working Team believes this would fulfill the mutual goal of improving the functionality of the transport networks and the internet and guaranteeing complete interoperability and architectural soundness
  - Refer to the technology as the Transport Profile for MPLS (MPLS-TP)
  - Therefore, we recommend that future work should focus on:
    - In the IETF: Definition of the MPLS “Transport Profile” (MPLS-TP)
    - In the ITU-T:
      - Integration of MPLS-TP into the transport network
      - Alignment of the current T-MPLS Recommendations with MPLS-TP and,
      - Terminate the work on current T-MPLS
- The technical feasibility analysis demonstrated there were no “show stopper” issues in the recommendation of Option 1 and that the IETF MPLS and Pseudowire architecture could be extended to support transport functional requirements
  - Therefore the team believed that there was no need for the analysis of any other option

## Future inter-SDO organizational structure

- It is proposed that the MPLS interop design team, JWT and ad hoc T-MPLS groups continue as described in SG15 TD515/PLEN with the following roles:
  - Facilitate the rapid exchange of information between the IETF and ITU-T
  - Ensure that the work is progressing with a consistent set of priorities
  - Identify gaps/inconsistencies in the solutions under development
    - Propose solutions for consideration by the appropriate WG/Question
  - Provide guidance when work on a topic is stalled or technical decision must be mediated
- None of these groups has the authority to create or modify IETF RFCs or ITU-T Recommendations
  - Any such work will be progressed via the normal process of the respective standards body
  - Direct participation in the work by experts from the IETF and ITU-T is required

# Role for the IETF MPLS Interoperability Design Team

- The IETF MPLS Interoperability Design Team should be chartered to produce an MPLS-TP architectural documentation hierarchy
  - All documents would progress in appropriate IETF WGs according to the normal process
  - The list of specific documents to be written in the IETF will be created by the Design Team
    - To allow rapid development of the architectural foundation documents no additional work on MPLS-TP will be taken on until the architectural foundation RFCs have progressed into WG LC
  - The Design Team is the group sponsored by the Routing Area Directors to facilitate rapid communication and coherent and consistent decision making on the Transport Profile for MPLS
  - An example of such a tree (by functional area) is below:



# Development of RFCs on MPLS-TP

- Work areas will be assigned to the appropriate IETF Working Groups to develop the RFCs
  - Working group charters and milestones will be updated to reflect the new work
    - Expected to be completed before IETF 72 (July 2008)
    - This will include the list of documents in the architectural hierarchy
  - WGs will appoint authors and where appropriate form design teams to develop the RFCs
    - It is assumed that ITU-T participants will be active members of these design teams
    - The draft will be reviewed by the ITU-T prior to completion of WG last call
      - ITU-T review will be by correspondence, the results of this review will be conveyed via a liaison statement
        - » Review by correspondence will avoid delaying WG last call to align with an ITU-T SG/experts meeting
        - » Early communication via liaisons and the JWT should allow us to avoid major comments on the final documents
    - Apply for early allocation of RFC numbers and IANA codepoints once a document has completed IESG review

# Development of ITU-T Recommendations on MPLS-TP

- The normative definition of the MPLS-TP that supports the ITU-T transport network requirements will be captured in IETF RFCs
- The ITU-T will:
  - Develop Recommendations to allow MPLS-TP to be integrated with current transport equipment and networks
    - Including in agreement with the IETF, the definition of any ITU-T specific functionality within the MPLS-TP architecture
      - Via the MPLS change process (RFC 4929)
    - Revise existing Recommendations to align with MPLS-TP
  - It is anticipated that following areas will be in scope. The actual Recommendations will be identified by the questions responsible for the topic areas.
    - Architecture (e.g. G.8110.1)
    - Equipment (e.g. G.8121)
    - Protection (e.g. G.8131, G.8132)
    - OAM (e.g. G.8113, G.8114)
    - Network management (e.g. G.7710, G.7712, G.8151, ...)
    - Control plane (e.g. G.7713, G.7715, ...)
  - ITU-T Recommendations will make normative references to the appropriate RFCs



# Development of ITU-T Recommendations on MPLS-TP - 2

- Work areas will be assigned to the Questions as defined in COM 15 - C1 (Questions allocated to SG15)
  - Work will be progressed in each question
    - Direct participation by interested parties from the IETF is strongly encouraged
    - Draft versions of Recommendations will be provided to the IETF for review via a liaison to a WG and/or via the JWT
  - It is anticipated that approval will be using AAP as defined in Recommendation A.8
    - Interim WP meetings may be required to allow timely consent of Recommendations that rely on normative references to RFCs
    - Final text for consent will be provided to the IETF for review
      - Initiation of the AAP process should be timed such that members can base AAP comments on an appropriate IETF WG consensus review of the consented text
      - Early communication via liaisons and the JWT should allow us to avoid major comments on the final documents
        - » e.g. the draft Recommendation for consent should be sent to the IETF for review prior to the SG meeting

# Documentation schedule

- First draft of the Transport Profile Architectural Framework
  - IETF 72 (July 2008)
  - WG last call completion Q2/2009
- Draft to request new reserved label for MPLS TP alert
  - IETF 72 (July 2008)
- RFCs on Alert Label and ACH definition
  - WG last call completion Q2/2009
- Updated ITU-T Recommendations
  - Q2/2009 (may need to schedule experts meeting/WP plenary to avoid delaying consent to the October 2009 meeting of SG 15)

A significant amount of work is required to achieve these milestones

- We need to start immediately (May 2008)
- Need a commitment from interested parties to edit and drive the drafts

# Introduction and Background Material

# What am I reading?

- This presentation is a collection of assumptions, discussion points and decisions that the combined group has had during the months of March and April, 2008
  - This represents the agreed upon starting point for the technical analysis of the T-MPLS requirements from the ITU-T and the MPLS architecture to meet those requirements
- The output of this technical analysis is the recommendation given to SG 15 on how to reply to the IETF's liaison of July 2007
  - IETF requested decision on whether the SDOs work together and extend MPLS aka "option 1: or
  - ITU-T choose another ethertype and rename T-MPLS to not include the MPLS moniker aka "option 2"
- The starting point of the analysis is to attempt to satisfy option 1 by showing the high level architecture, any showstoppers and the design points that would need to be addressed after the decision has been made to work together.
  - Option 1 was stated as preferred by the IETF and if it can be met; Option 2 will not be explored

# Some contributors to this architecture

- BT
- Verizon
- ATT
- NTT
- Comcast
- Acreo AB
- Alcatel-Lucent
- Cisco
- Ericsson
- Huawei
- Juniper
- Nortel
- Old Dog Consulting

# How is the effort organized?

1. In ITU-T
  - TMPLS ad hoc group
2. In IETF
  - MPLS interoperability design team
3. DMZ between the SDOs: Joint Working Team
  - Segmented into groups looking at
    1. Forwarding
    2. OAM
    3. Protection
    4. Control Plane
    5. Network Management
  - Goal: Produce a technical analysis showing that MPLS architecture can perform functionality required by a transport profile.
    - Compare w/ ITU-T requirements and identify showstoppers
    - Find any obvious design points in MPLS architecture that may need extensions

# MPLS - Transport Profile: What are the problems?

- Desire to statically configure LSPs and PWEs via the management plane
  - Not solely via control (routing/signaling) plane
  - If a control plane is used for configuration of LSPs/PWEs failure and recovery of the control plane must not impact forwarding plane (a la NSR/NSF)
  
- Transport OAM capabilities don't exist for LSP and PWE independent of configuration mechanism (management plane or GMPLS or PWE control plane)
  - Full transport FCAPS - AIS, RDI, Connection verification (aka connectivity supervision in G.806), loss of connectivity (aka continuity supervision in G.806), support of MCC and SCC etc
  - Recent drafts to IETF demonstrate some issues
  
- Service Providers are requesting consistent OAM capabilities for multi-layered network and interworking of the different layers/technologies (L2, PWE, LSP)
  - Include functionality of Y.1711 and Y.1731 into one architecture

# MPLS -TP: What are the problems? 2

- Service Providers want to be able to offer MPLS LSPs and PWEs as a part of their transport offerings and not just associated with higher level services (e.g. VPNs)
- Service Providers want LSPs/PWEs to be able to be managed at the different nested levels seamlessly (path, segment, multiple segments)
  - aka Tandem Connection Monitoring (TCM), this is used for example when a LSP/PWE crosses multiple administrations
- Service Providers want additional protection mechanisms or clear statements on how typical “transport” protection switching designs can be met by the MPLS architecture
- Service Providers are requesting that OAM and traffic are congruent
  - Including scenarios of LAG or ECMP
  - Or create LSP/PWEs that don't traverse links with LAG/ECMP



# MPLS - TP Requirements Overview

- Meet functional requirements stated earlier by service providers
- No modification to MPLS forwarding architecture
- Solution Based on existing Pseudo-wire and LSP constructs
- Bi-directional congruent p2p LSPs
- No LSP merging (e.g. no use of LDP mp2p signaling in order to avoid losing LSP head-end information)
- Multicast is point to multipoint not MP2MP

# MPLS - TP Requirements Overview .2

- OAM function responsible for monitoring the LSP/PWE
  - Initiates path recovery actions
- IP forwarding is not required to support of OAM or data packets
  - OOB management network running IP is outside scope of feasibility study
- Can be used with static provisioning systems or with control plane
  - With static provisioning, no dependency on routing or signaling (e.g. GMPLS or, IGP, RSVP, BGP, LDP)
- Mechanisms and capabilities must be able to interoperate with existing MPLS and PWE control and forwarding planes

# MPLS-TP Major Solution Constructs

NOTE: These two constructs were used as the basis for the Technical Feasibility study performed by the ad hoc team, JWT and IETF MPLS Interoperability Design Team

## 1. Definition of MPLS-TP alert label (TAL) and a Generic Associated Channel (GE ACH)

Allows OAM packets to be directed to an intermediated node on a LSP/PWE

Via label stacking or proper TTL setting

Define a new reserved label (13 is suggested):

It is believed that Label 14 cannot be reused at this point

## 2. Generic Associated Channel (GE ACH) functionality supports the FCAPS functions by carrying OAM, APS, ECC etc. packets across the network

Use of PWE-3 Associated Channel to carry OAM packets

GE ACH are codepoints from PWE ACH space but, not necessarily, for PWE purposes

GE ACH would be present for OAM of all LSPs

# MPLS-TP Major Solution Observations

1. Bringing ACH functionality into LSPs begins to blur the architectural line between an MPLS LSP and an MPLS Pseudowire
  - The functional differences between an MPLS LSP and MPLS PW must be retained in the architecture
2. The same OAM mechanism (e.g. ACH) can be unified for LSPs and PWE
  - Enabling the same functionality for both and ease of implementation
  - Avoid breaking anything (e.g. ECMP)
  - There may be specific differences that are discovered in design phase
  - ACH functionality for LSPs should be limited to only OAM, APS & ECC management channel data
3. A great deal of IETF protocol, design and architectural reuse can be employed to solve the requirements
  - No fundamental change to the IETF MPLS architecture was found to be necessary

# MPLS-TP Alert Label Observations - 1

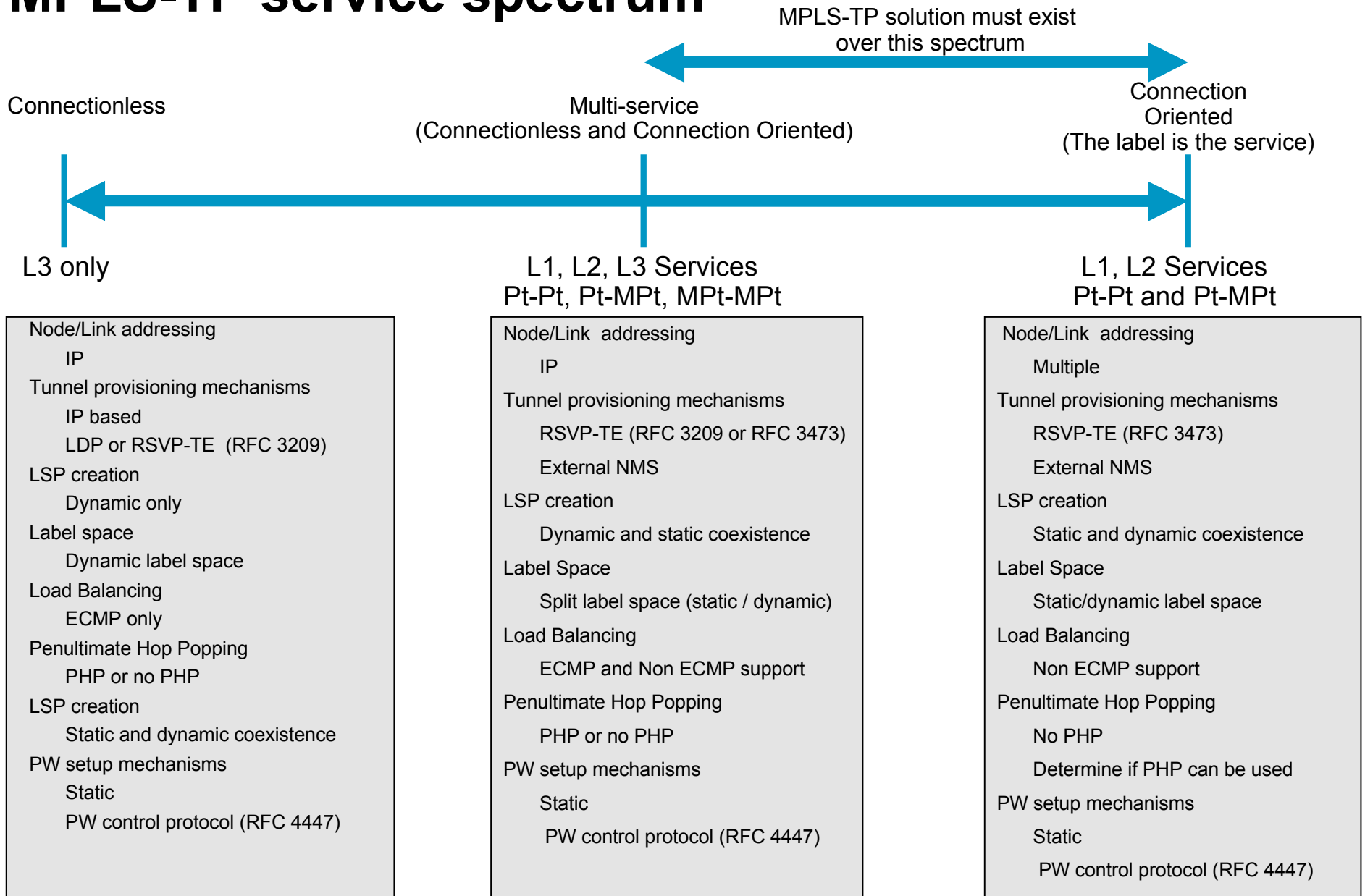
- The JWT has established that to create an MPLS-TP there is a need for an associated channel that shares fate and coexists with data
- One possibility would be to use the OAM Alert Label (label 14) to establish this channel but:
- IETF WGs and ITU-T SGs were polled to find out the state of implementation and deployment of Y.1711 and RFC3429
  - The conclusion was that there are enough implementations and deployments so that it is not possible to immediately deprecate Y.1711 and RFC3429

# MPLS-TP Alert Label Observations - 2

- The JWT has concluded that a new reserved label may be needed for the MPLS TP alert
- This label would be requested from the pool of un-allocated reserved MPLS labels
  - Label 13 has been suggested.
- The suggested roadmap is to gradually move all OAM functionality defined by label 14 over to the new reserved label
- The specification of the new OAM channel must be accompanied with a decision to stop further extension of OAM based on label 14
  - Only maintenance operations continue

# High Level Architecture

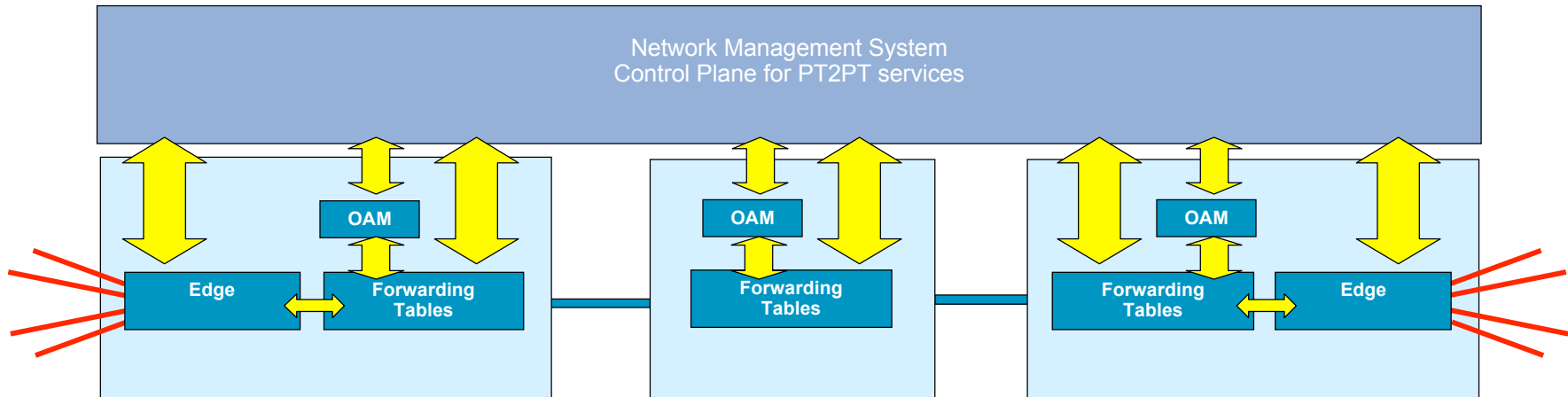
# MPLS-TP service spectrum



- IMPERATIVE MPLS-TP MUST BE ABLE TO INTEROPERATE IN AN L3 NETWORK
- MPLS-TP MUST ALSO SUPPORT AND CO-EXIST WITH EXISTING PWE-3 SOLUTIONS

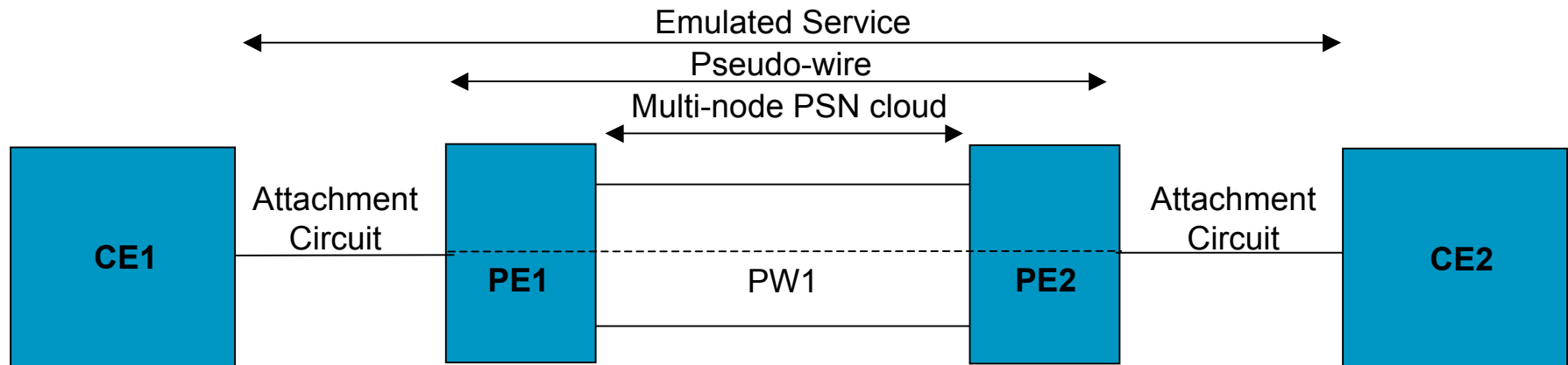


# MPLS+TP Static Provisioning



- Static provisioning and dynamic control plane
  - Requirements state that the solution must include static only provisioning
  - Any dynamic Control plane will be based on IETF solutions (GMPLS, IP/MPLS)
- Control Plane responsible for:
  - End to End, Segment LSPs and PWE-3 application labels (programming the LFIB)
  - Determining and defining primary and backup paths
  - Configuring the OAM function along the path
  - Others : Defining the UNI etc
- OAM responsible for monitoring and driving switches between primary and backup paths for the end to end path and path segments

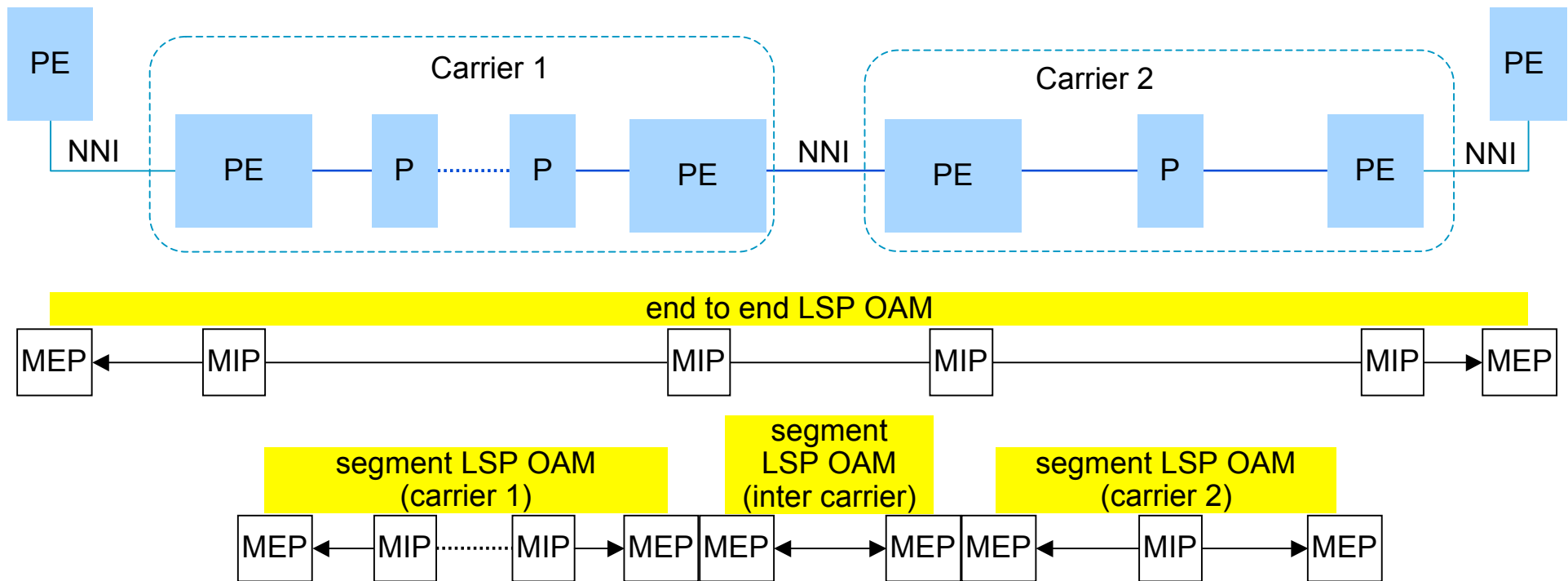
# MPLS Transport Profile - Terminology



- Definition of an MPLS Transport Profile (TP) within IETF MPLS standards
  - Based on PWE3 and LSP forwarding architecture
  - IETF MPLS architecture concepts
- The major construct of the transport profile for MPLS are LSPs
  - PW are a client layer

# LSP example

- end to end and per carrier monitoring

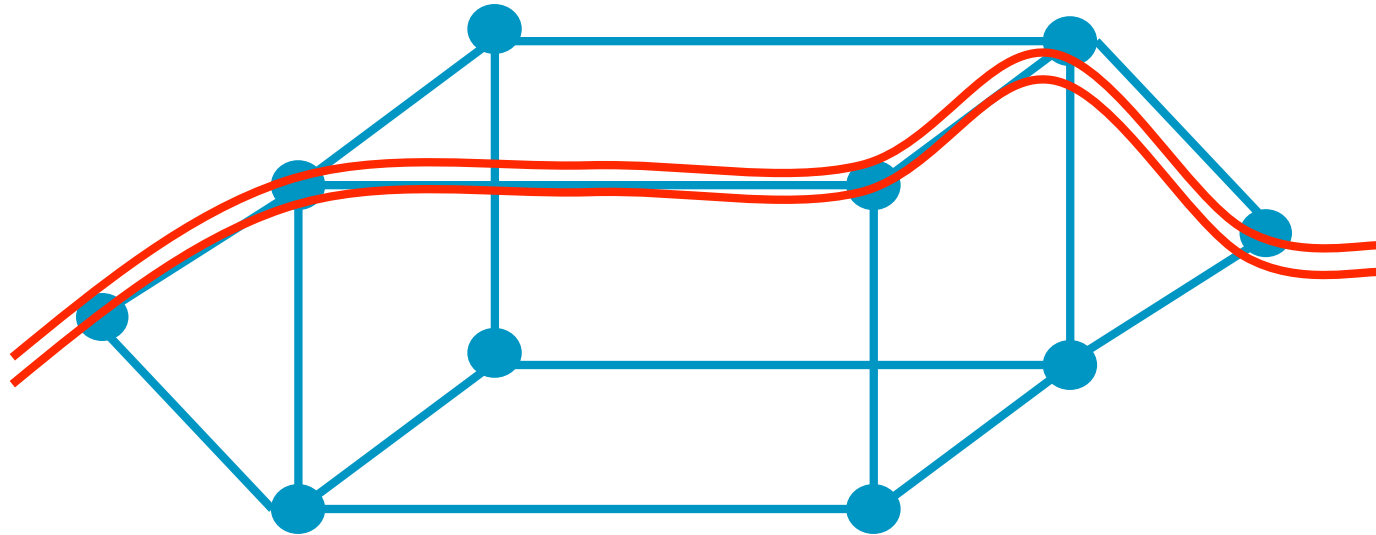


- A segment is between MEPs
- OAM is end to end or per segment
  - In SDH/OTN and Ethernet segment OAM is implemented using Tandem Connection Monitoring (TCM)
- The OAM in each segment is independent of any other segment
- Recovery actions (Protection or restoration) are always between MEPs i.e. per segment or end to end

Note: A policing function (traffic management/shaping) is normally co located with a MEP at a business boundary (UNI/NNI)

MEP: Maintenance End Point  
MIP: Maintenance Intermediate Point

# Bidirectional Paths



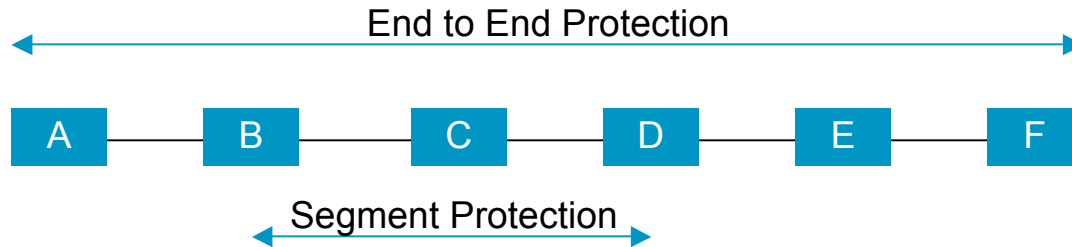
- External Static Provisioning
  - NMS responsible for configuration and ensuring bi-direction congruency
- If Dynamic Control Plane
  - GMPLS bidirectional RSVP for LSP path establishment

# OAM requirements

# OAM Requirements

- Must be able to monitor LSP, PWE3
  - Inter layer fault correlation
  - Failure indication propagation across multiple segments
  - Monitoring of Physical layer, layer 1, layer 2 is out of scope
- Packet loss rather than bit error based measurements/metrics for L2, LSP, PWE3
- Per segment (aka tandem connection) and end to end
  - Fault detection/isolation
  - Recovery - protection switch or restoration
- A security architecture

# What is segment recovery?



- End to End recovery:
  - Fault detection and recovery of the end to end pseudo-wire
  - Fault detection and recovery of the end to end LSP Segment recovery:
- Fault detection and recovery of a segment
  - The recovery mechanism used in a segment is independent of other segments
- Segment constructs
  - Hierarchical nested LSP: Existing construct
  - MS-PW segment: Currently defined construct in PWE3
  - Stacked TCM label (mapped 1:1 with corresponding LSP/PW)

# Node identification

- Will need to work through identification requirements
  - What about algorithmically derived label from the IP identifier
  - What IP identifier if we do not need IP to support forwarding or OAM?
  - Need to be able to rearrange the DCC without disturbing the forwarding/OAM?

A node has multiple identifiers including the following:

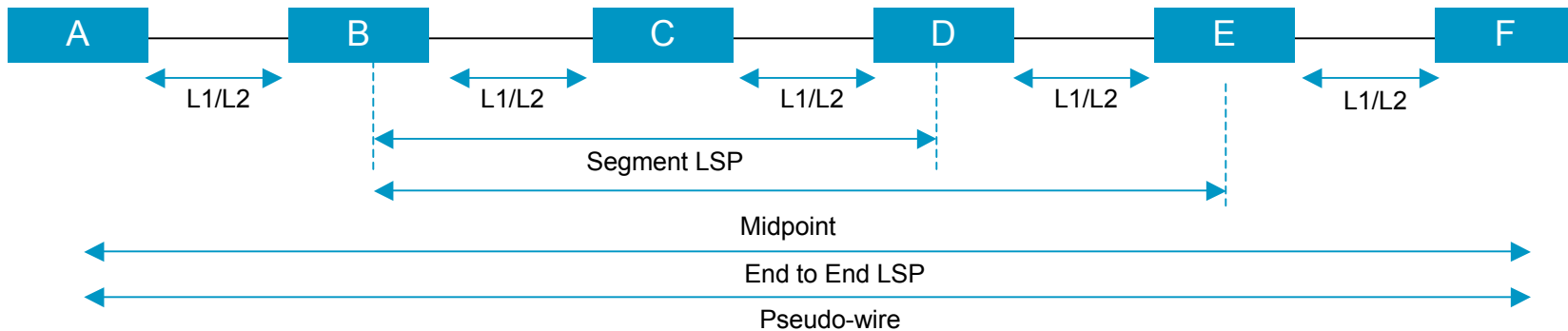
- Management identifier – normally user friendly, based on the location
- MEP/MIP identifier
- DCC address - how do management messages reach this node
- Control plane identifiers - how are the various control components identified
- Forwarding plane identifier - end points and intermediate points - e.g. NNIs

*These are design issues, no “show stoppers” found*



# OAM mechanisms

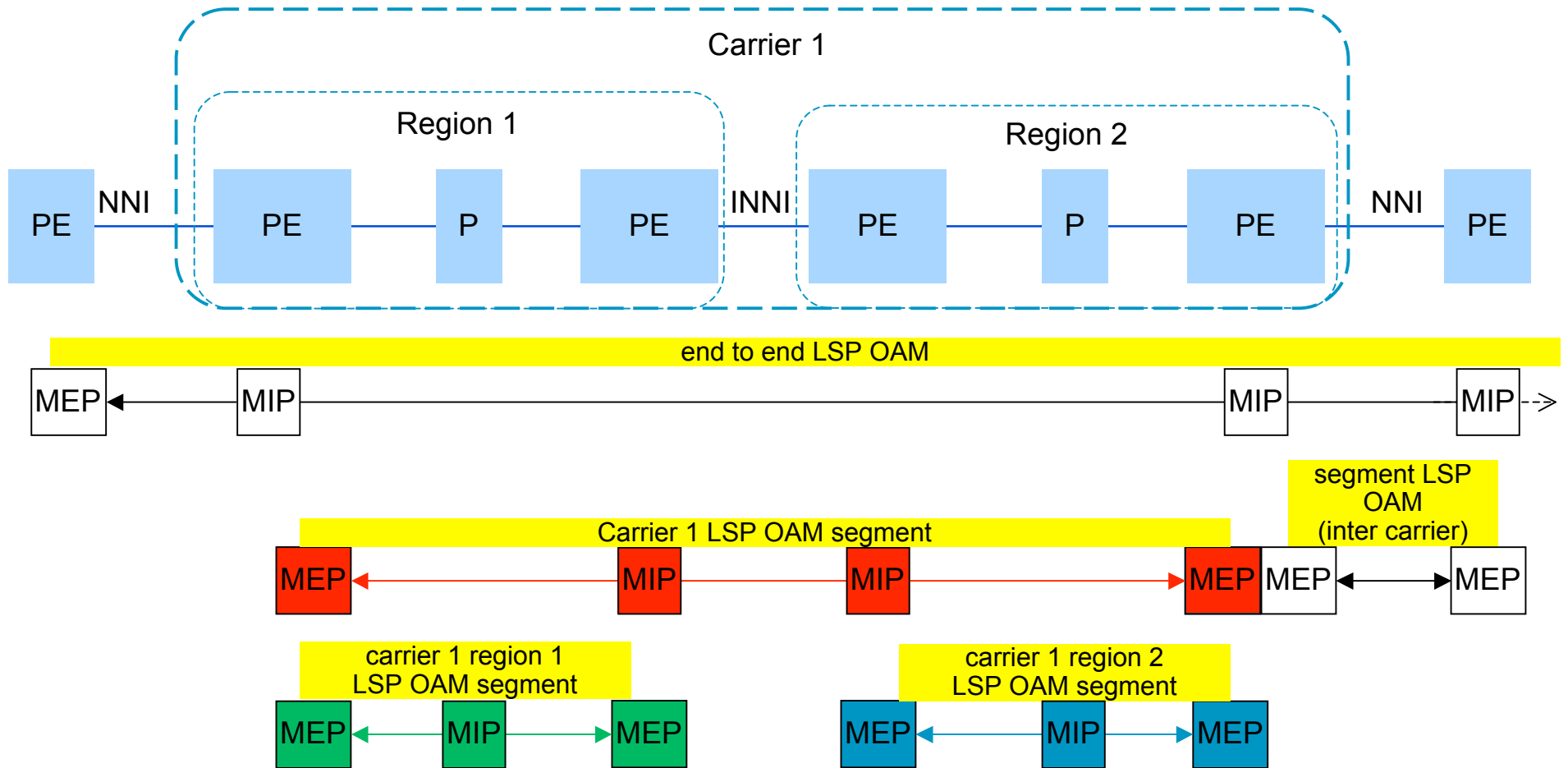
# Overview: OAM hierarchy and mechanisms



- L0/L1 : Loss of Light; G.709, SONET/SDH LoS, LoF, ES, SES (NOT DISCUSSED)
- Non MPLS L2 connectivity : Native L2 solution 802.1ag (Not Discussed) , Non IP BFD  
Failure propagation across layers is supported by this architecture
- General LSPs : Generic Exception Label and Generic Associated Channel  
Includes End to End and segment LSPs  
Used to carry a variety of OAM, Mgmt, signalling protocols.
- Pseudo-wires : PWE3 Associated Channel

# LSP monitoring example

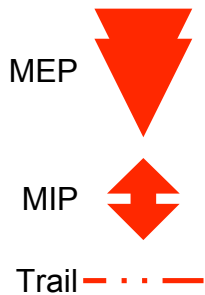
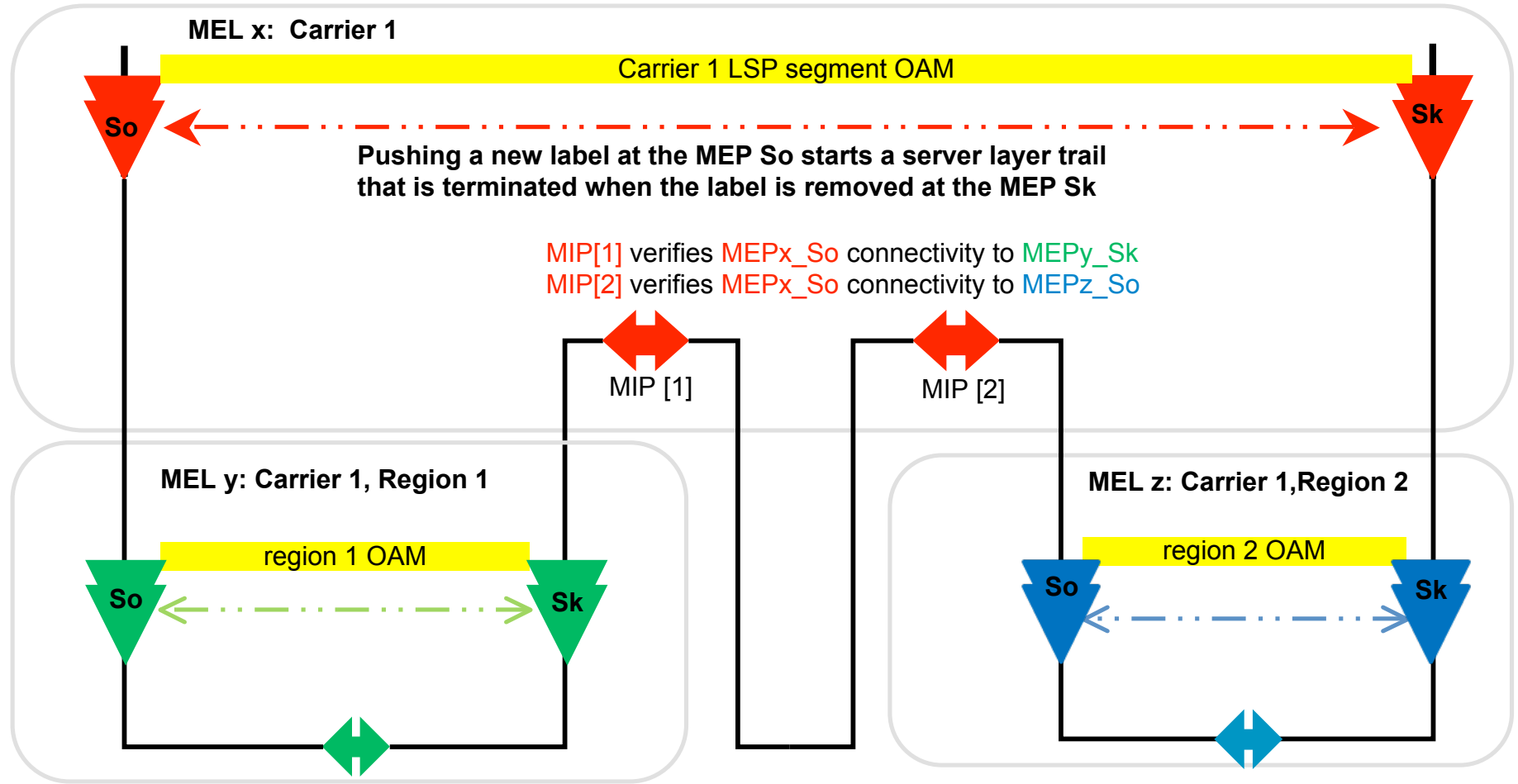
- monitoring within carrier 1



### 3 LSP OAM levels + PW OAM

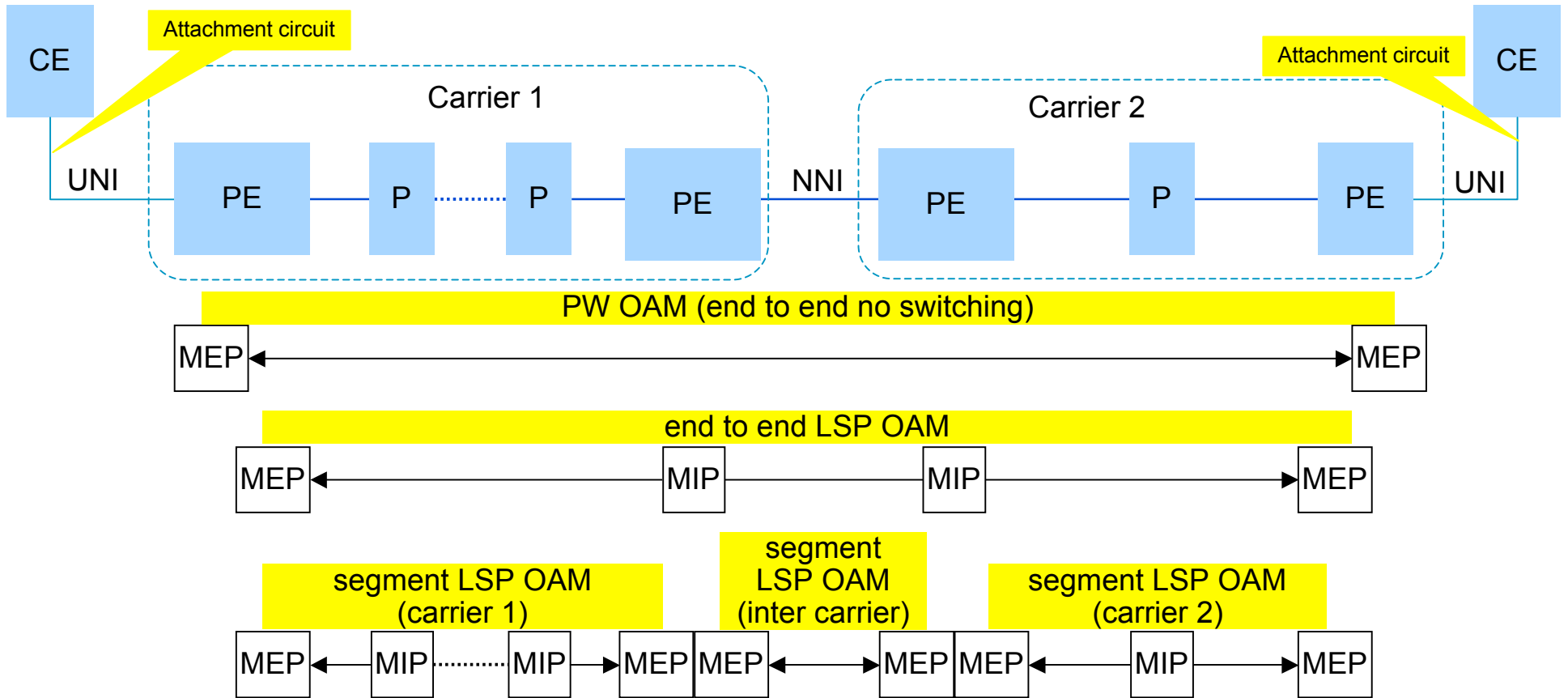
- end to end LSP + 2 nested segment LSP levels (Carrier 1 + regions 1/2)
- Nested segments are supported by Tandem Connection Monitoring (TCM) in SDH/OTN and Y.1731

# Carrier 1 example MEPs/MIPs relationships



**A MIP must support monitoring on the ingress port (logically before the label swap)**  
**An implementation may optionally support a second MIP to monitor the egress port**  
*How will this MIP be addressed*

# PW over LSP monitoring example

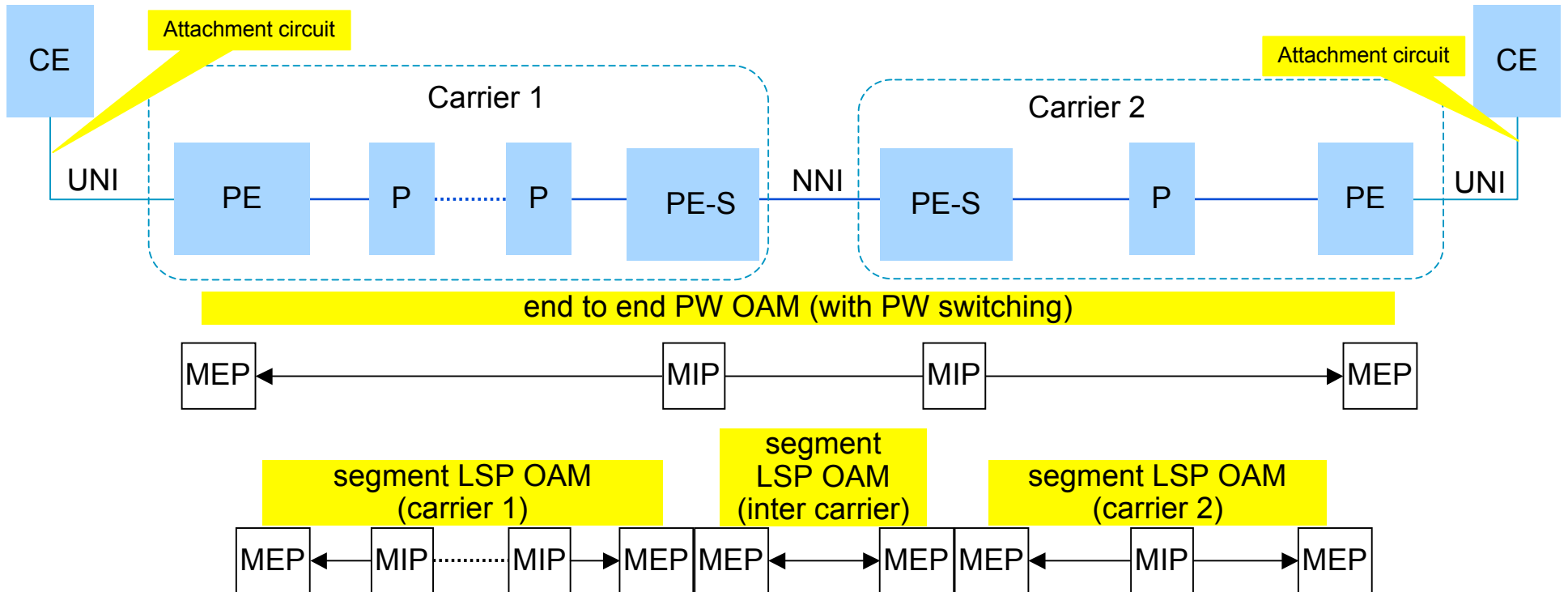


- end to end LSP OAM is used since PW OAM cannot create MIPs at the inter carrier boundary without a PW switching function

Note: A policing function (traffic management/shaping) is normally co located with a MEP at a business boundary (UNI/NNI)

MEP: Maintenance End Point  
MIP: Maintenance Intermediate Point

# PW over LSP example with PW switching



- end to end LSP OAM is not required since the PW switching points can support a MIP

Note: A policing function (traffic management/shaping) is normally co-located with a MEP at a business boundary (UNI/NNI)

MEP: Maintenance End Point  
MIP: Maintenance Intermediate Point

# Associated Channel Level (ACH)

# Associated Channel Level ACH: Overview

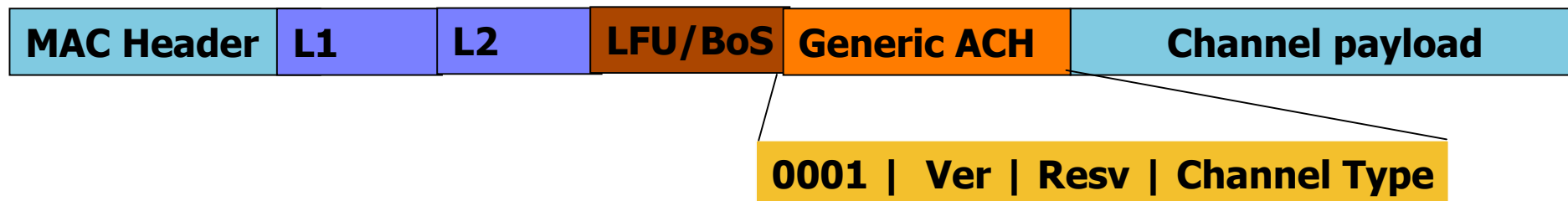
- Generalised mechanism for carrying management / OAM information
  - OAM capabilities : Connectivity Checks (CC) and “Connectivity Verification” (CV)
  - Management information: Embedded Control Channel (ECC)
    - To support the Data Communications Network (DCN) and the Signalling Communication Network (SCN) – see G.7712
  - APS information
- Associated Channel Capabilities
  - Multiple channels can exist between end points
  - Channel Type Indicates what protocol that is carried
  - To service an MPLS-TP network new channel types will need to be defined
- Management and Control Plane Information (DCN and SCN connectivity)
  - Via ECC where IP is not configured
- Generic ACH contains a “channel Type” field
  - Need for a registry of protocols
  - This needs to be blocked for different functions (IP-Free BFD is currently 7)
  - We may want to define a vendor specific and experimental range

**No Showstoppers found**



## LSP monitoring and alarming

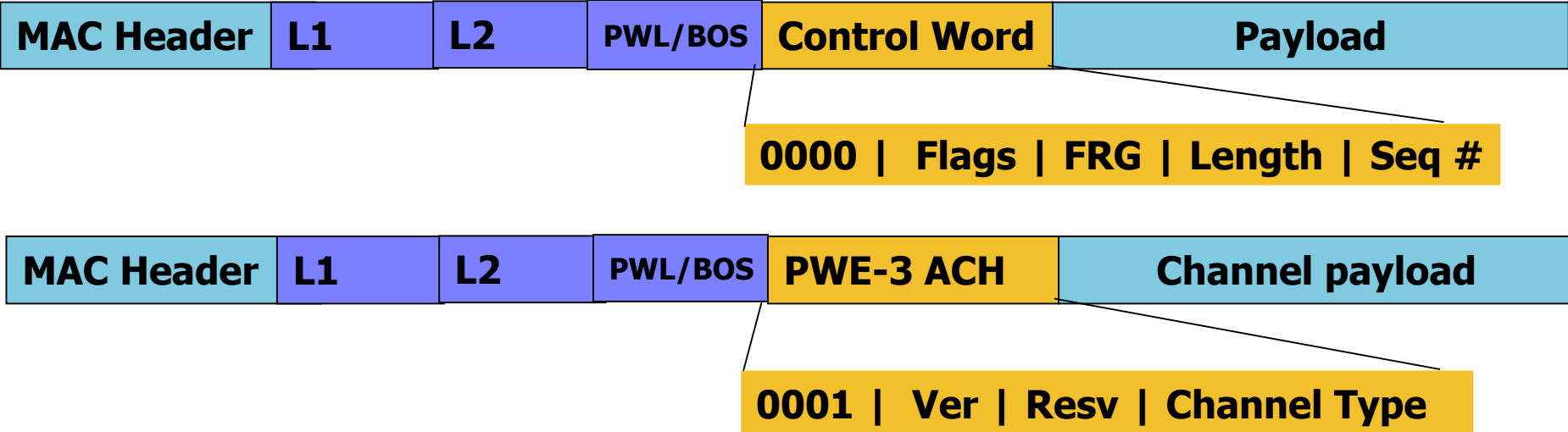
### Generic Exception Label and Generic Associated Channel Proposal



- Assign a Transport Alert Label as a Label For yoU (LFU) from reserved label space:
  - Label 13 has been proposed because,
  - Label 14 has been allocated to Y.1711
  - Y.1711 arch fits within “ACH” architecture
- Bottom of Stack is always set on LFU in the transport profile
- Define a Generic Associated Channel function
  - Similar to the PWE-3 Associated Channel but doesn't have to be associated with a PW
  - Important the first nibble tells system not to load balance (so not 06 or 04)
- Generic Associated Channel is always under a Generic Exception Label if endpoint (MEP)
- Generalised Associated Channel defines what packet function using “channel type” field
  - Examples : What OAM function is carried, DCC, etc

# Pseudo-wire monitoring and alarming

## PWE-3 Control Word and PW-Associated Channel



This is a representation of what is in RFC 4385

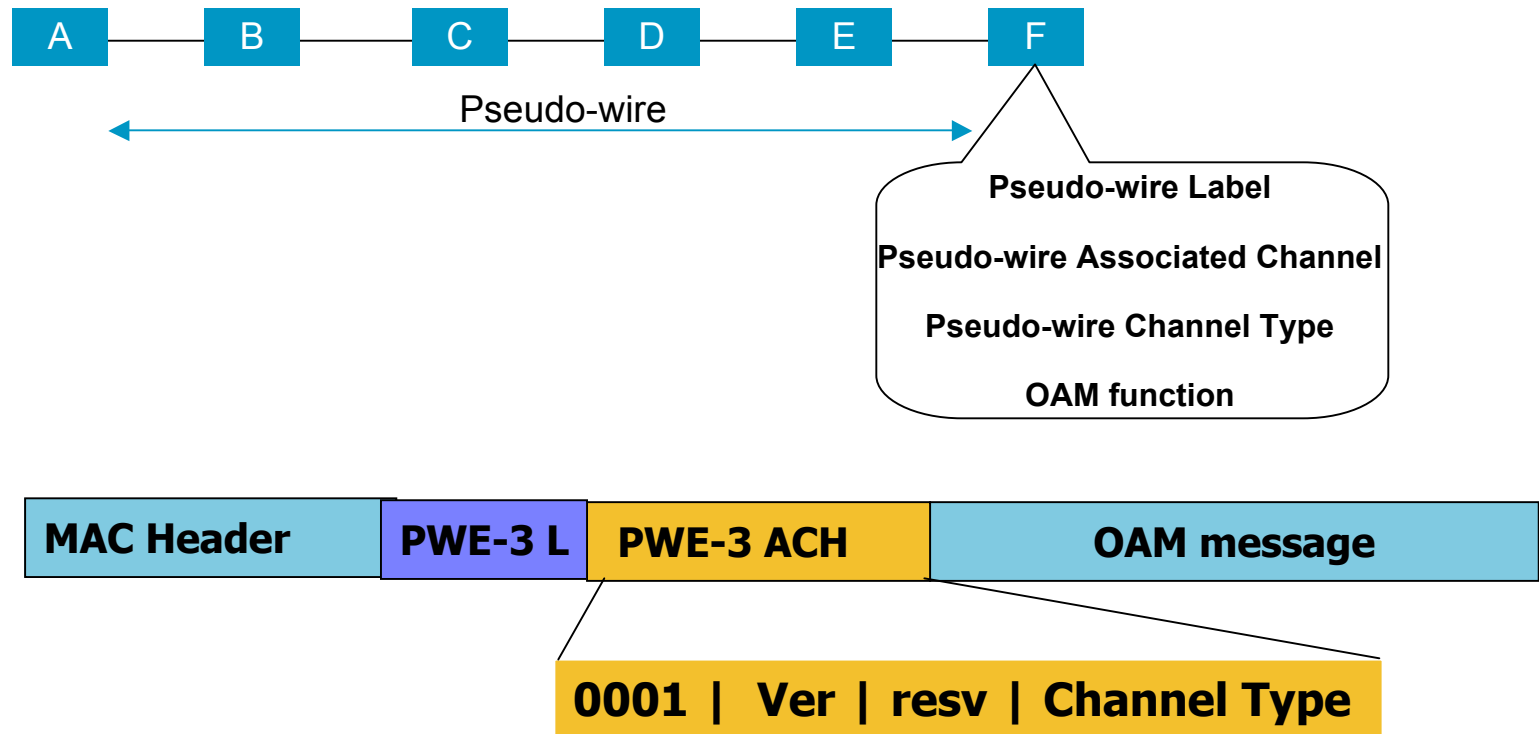
# Required Functionality demarked by Associated Channel

- CV : Connectivity Verification (detection of configuration errors)
- PM: Performance of the path
- AIS: Alarm suppression
- CC : Continuity Check : Is the path present (may reuse vanilla BFD here)
  - Light weight
  - Role is as a CC protocol, it is not a CV protocol
  - Not a connectivity verification protocol
  - VCCV-BFD provides capabilities over pseudo-wire
- ECC
  - OSS and control plane communication
- APS
  - Protection switching coordination
- Accounting/Billing information
- Security exchange
- Extra codepoint space to define new or use existing protocols for other functions

# Associated Channel Functionality Observations

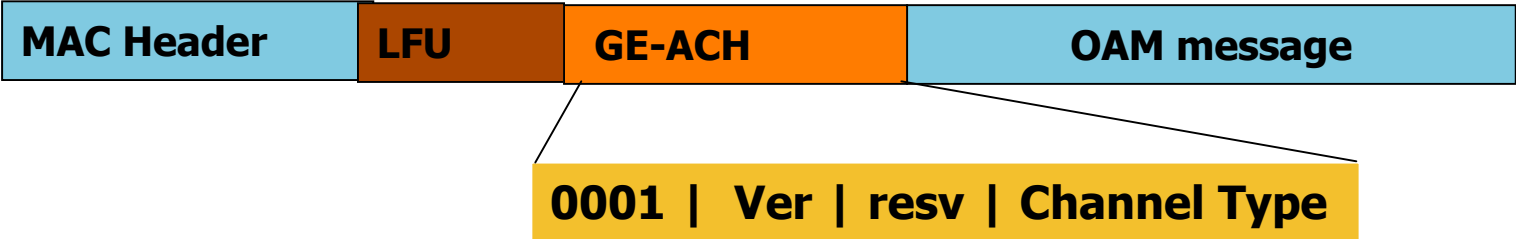
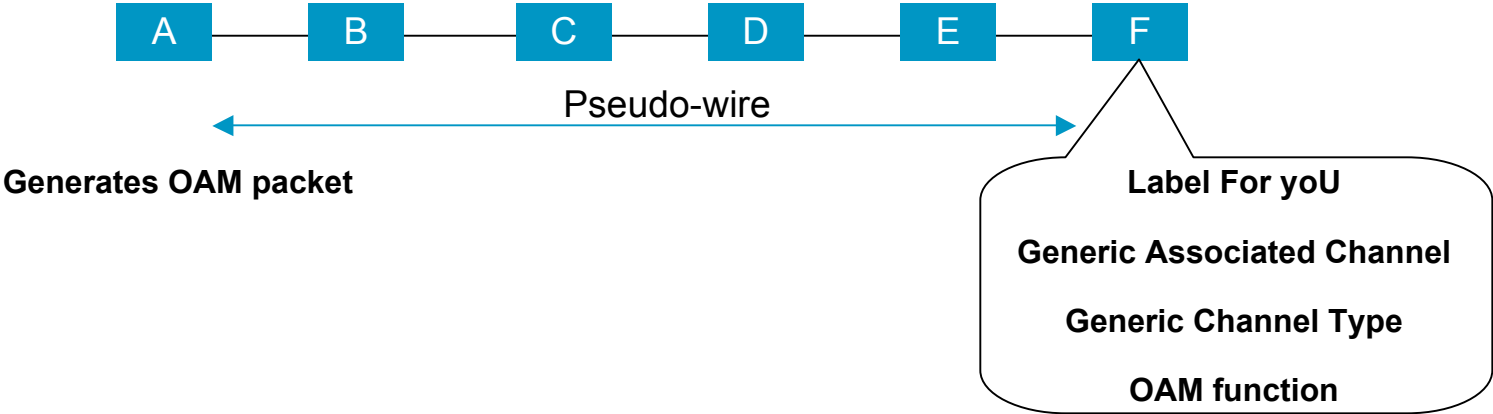
- Existing MPLS LSP OAM uses an IP based control channel and could be used for some OAM functions in transport networks
  - e.g. CC/CV
  - The new Alert label based control channel should be able to co-exist with the existing MPLS LSP OAM functions and protocols
- OAM message formats and protocol details carried in the OAM channel will be discussed in the design phase
  - We must figure out what the OAM messages/protocols should be used for the new requirements
  - Decide whether LSP-Ping or BFD can or should be tweaked or not

# Pseudo-wire OAM processing



- Processed by the pseudo-wire function on the end-points
  - End point or Pseudo-wire stitch point
- Verifies the operational status of the pseudo-wire
- Working with the native attachment circuit technology
  - An inter-working function with the native attachment circuit OAM.
  - Transport and act upon native attachment circuit OAM technology

# LSP End Point processing



- Label For yoU with Generic Channel Association
- Processed by the LSP end point
  - End to End LSP or Segment LSP
- Verifies the operational status of the LSP
  - Many options including Non IP BFD is an option encapsulation of Y.1731 pdu

Forwarding and OAM:  
LSPs / PW  
OAM and Label Stacks

# Scope of next slides

- Slides cover on MEP to MEP and MEP to MIP monitoring
  - Detailed OAM packet walkthrough not yet covered in this slide-set
  - For MIP monitoring traceroute or loopback is executed and TTL set accordingly
- Introduce concept of LSP/PW TCM label:
  - This is a label to indicate a tandem monitoring session context
  - Label is stacked above label of LSP or PW being monitored
  - 1 for 1 mapping between an LSP / PW and its TCM session. i.e. no multiplexing
  - Need mechanism to bind TCM label to underlying LSP or PW being monitored
- MEP to MIP
  - MEP sets the TTL of the LSP, TCM or PW label so that it will expire when the target MIP is reached
- PHP




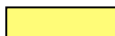
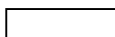


**No Showstoppers found**



# Notation and color conventions

- [Destination][[(using label provided by)]] [optionalFEC]/[StackBit]
- Thus D(E)/0 means Destination is D, using label provided by (E) - i.e. c is the tunnel next hop and the Sbit is 0 - i.e. not bottom of stack.
- Thus E(E)p/1 means Destination is E, using label provided by (E) the FEC is a pseudowire and the Sbit is 1, i.e. bottom of stack
- Special Labels and terms
  - LFU = Label For yoU - OAM alert label
  - Ach = Associated Channel Header
  - CW = Control Word
  - P = PW FEC

## Color Conventions

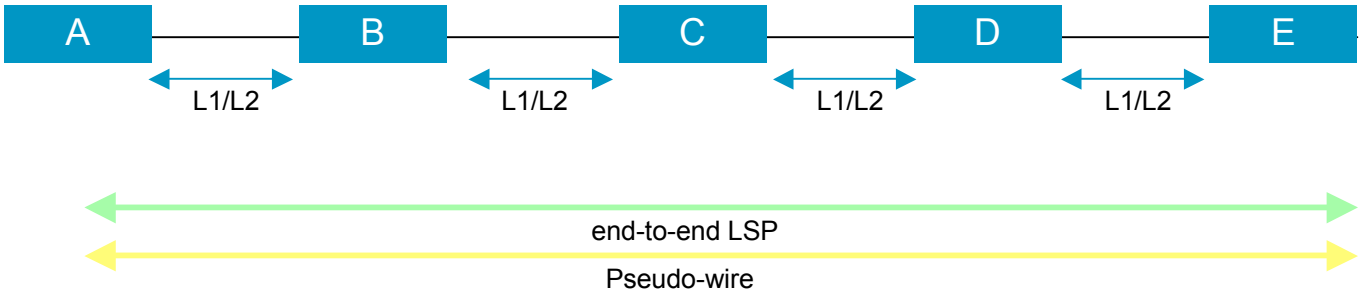
	LSP tandem OAM label
	LSP label
	PW tandem OAM label
	PW label
	PW control word
	Label For yoU
	ACH

# Scenarios

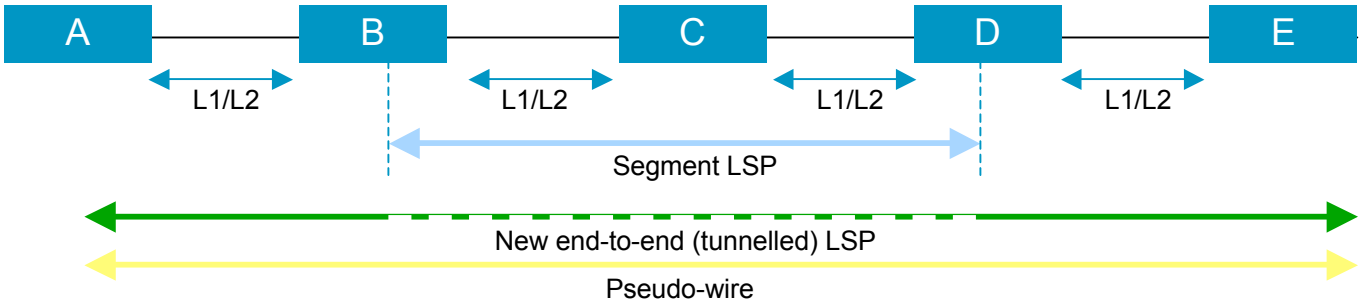
- SS-PW over intra-domain LSP
  - No TCM OAM
  - TCM-LSP OAM
- SS-PW over inter-domain LSP
  - LSP, TCM LSP & PW OAM
- Intra-domain MS-PW
  - MS-PW TCM OAM
- Intra-domain MS-PW
  - LSP OAM and TCM-MS-PW OAM
- Inter-provider MS-PW
  - PW E2Eand PW TCM OAM
- SS-PW over Intra-domain LSP
  - LSP MEP->MIP OAM using TTL
- Intra-domain MS-PW
  - MS-PW OAM: PW MEP-MIP, No TCM
- Intra-domain MS-PW
  - MS-PW OAM: TCM MEP->MIP, plus E2E PW

# Segment LSP setup

## Starting Point



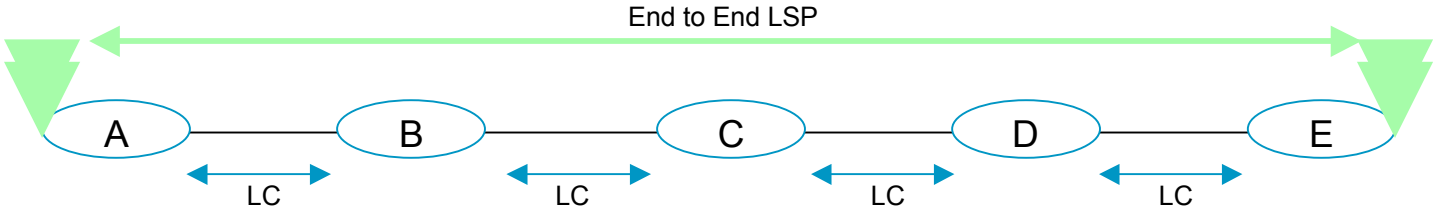
## Final Point



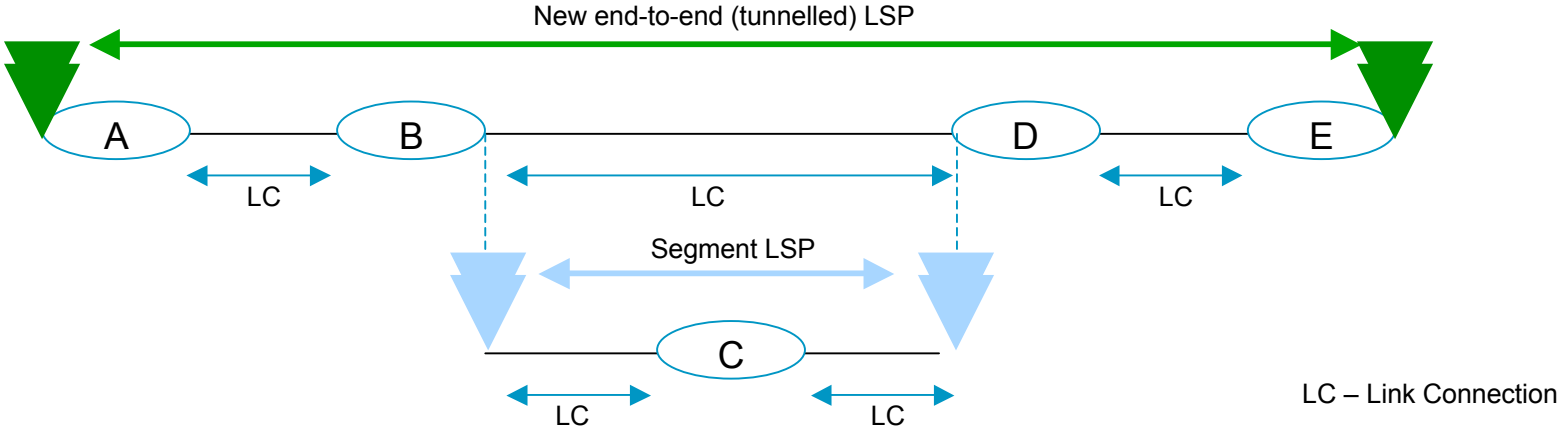
Objective:  
Use bridge-and-roll with make-before-break mechanism  
to ensure transition

# Segment LSP setup – G.805 view

## Starting Point



## Final Point

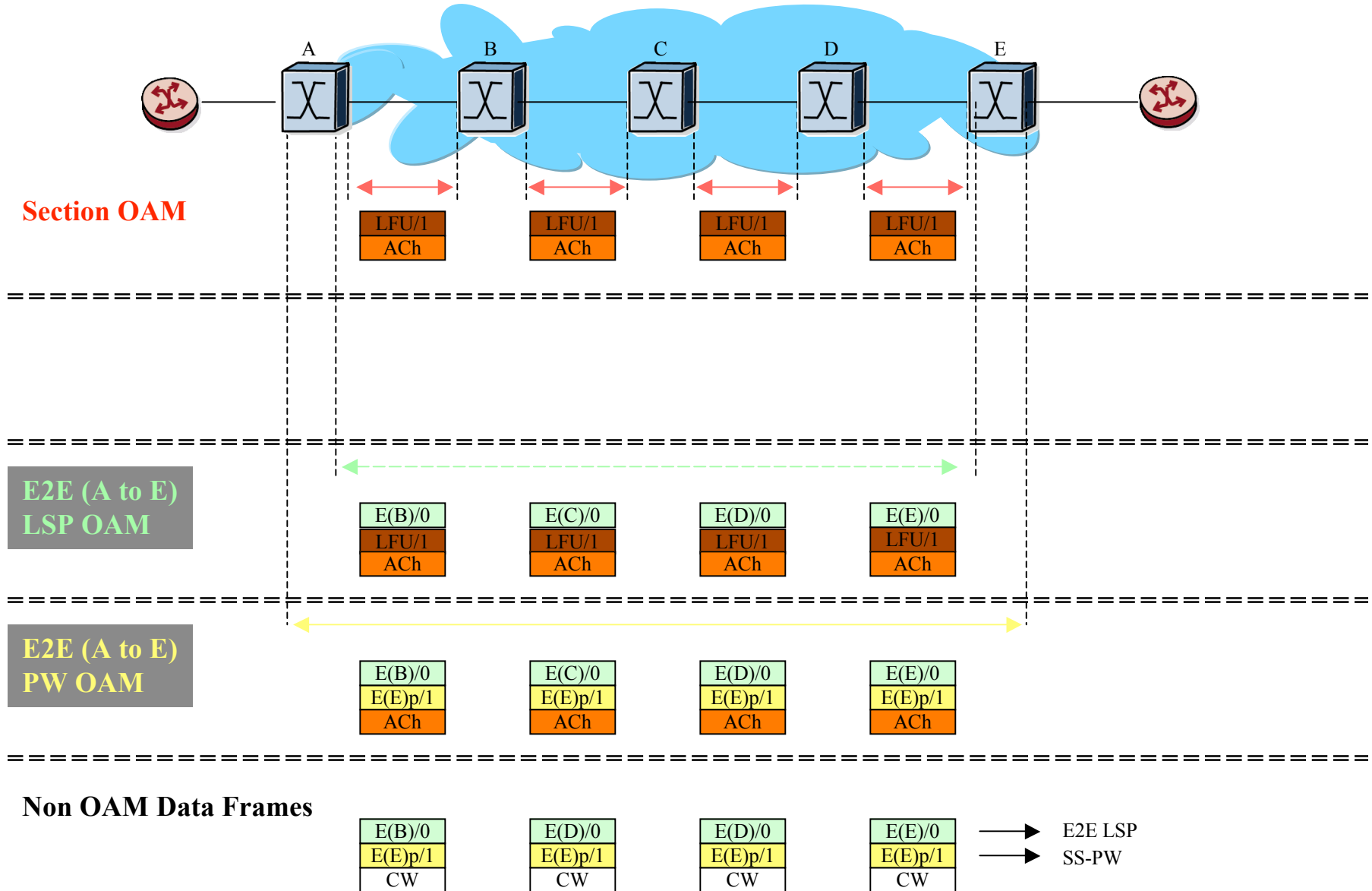


# Procedural Ordering Overview

- Step 1 : establish the **segment** LSP
  - Question : can segment LSP and existing end-to-end LSP share bandwidth?
- Step 2 : establish a **new** end-to-end LSP and which must be tunnelled in the **segment** LSP
  - Use MBB procedures (for sharing resources between **existing** and **new** end-to-end LSP).
- Step 3 : Perform switchover after Resv is received in A
  - ITU-T mechanisms rely on the creation of a Protection Group between the old and new (tunnelled) end-to-end LSP, the forcing of protection switching via APS and the tearing down of the Protection Group
- Step 4 : Tear down the **old** end-to-end LSP

# SS-PW, LSP OAM (no TCM)

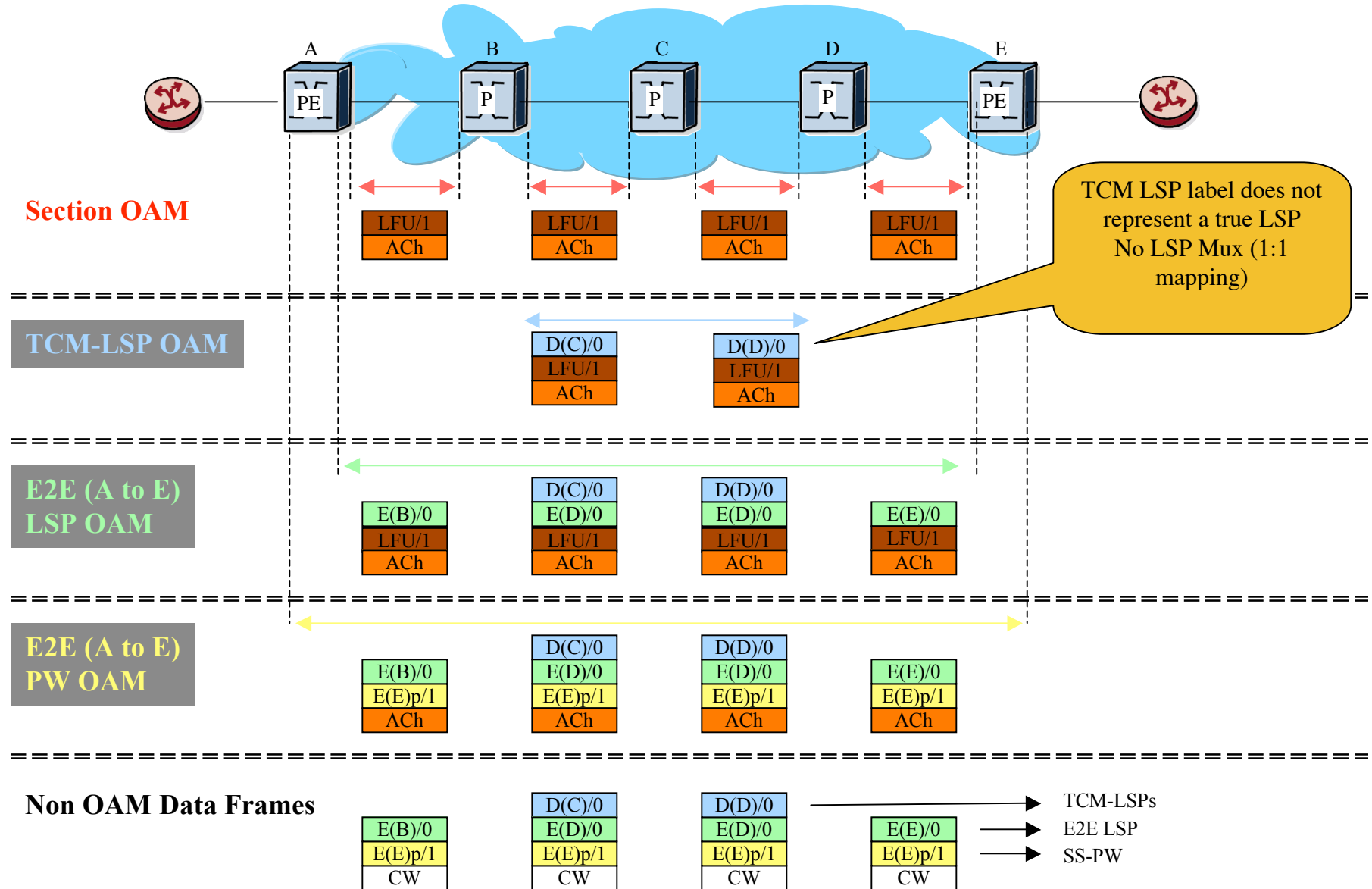
LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word



# SS-PW over intra-domain LSP

## LSP, TCM-LSP & PW OAM

LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word

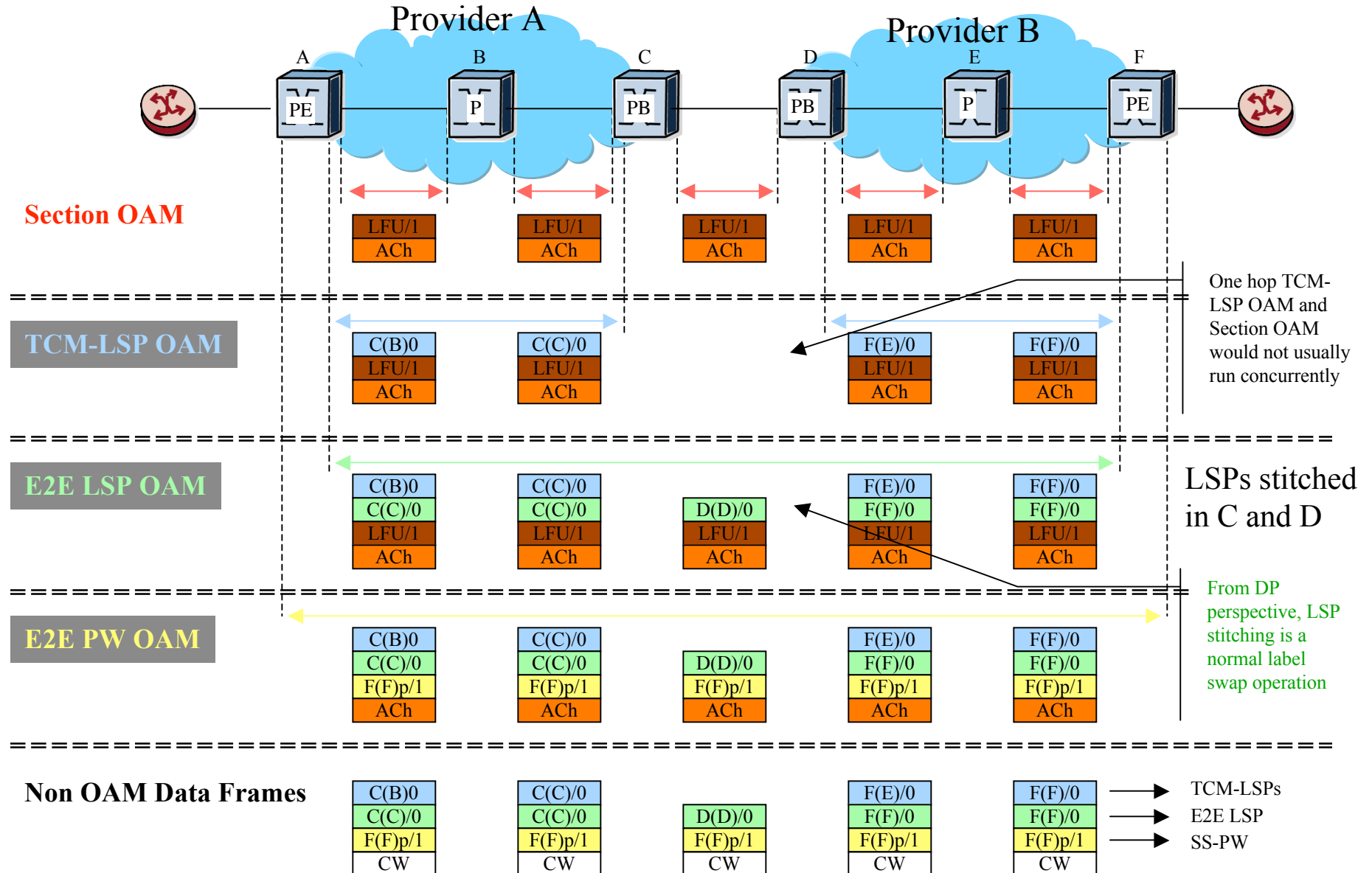


# SS-PW over inter-provider LSP

## LSP, TCM-LSP & PW OAM

LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word

PB = Provider Border LSR

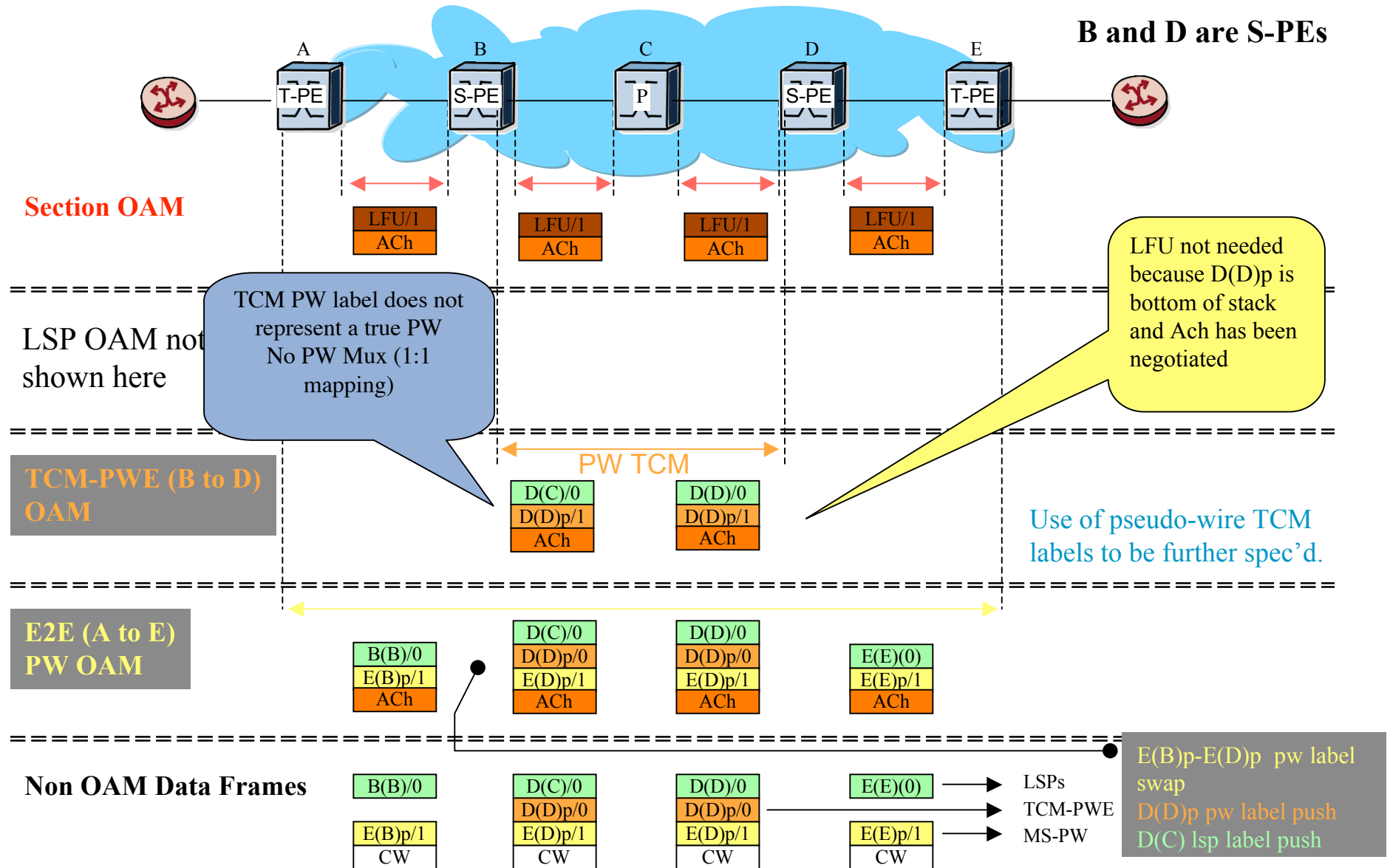




# Intra-domain MS-PW

## MS-PW & TCM-MS-PW OAM

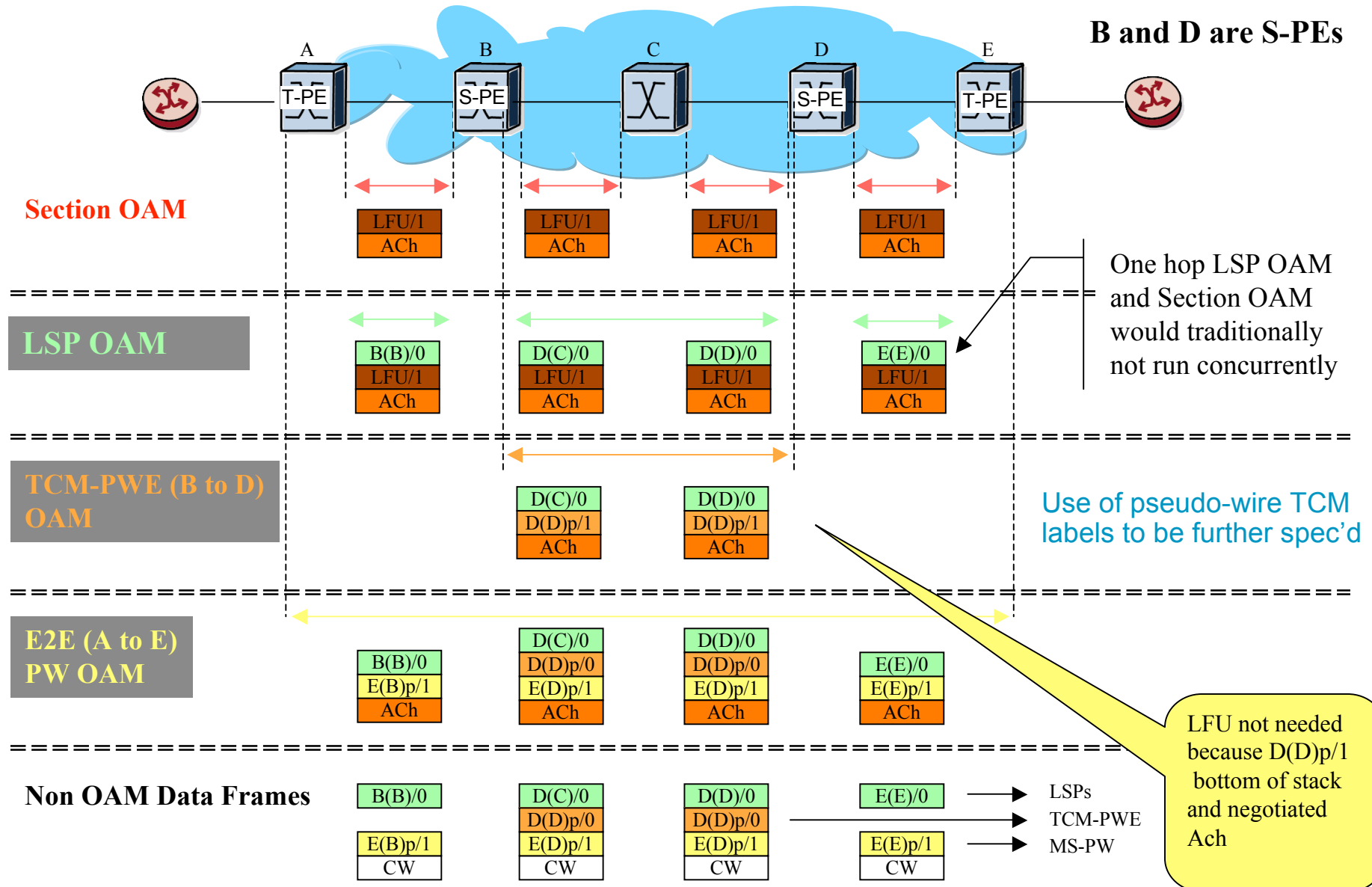
LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word



# Intra-domain MS-PW

## LSP, MS-PW & TCM-MS-PW OAM

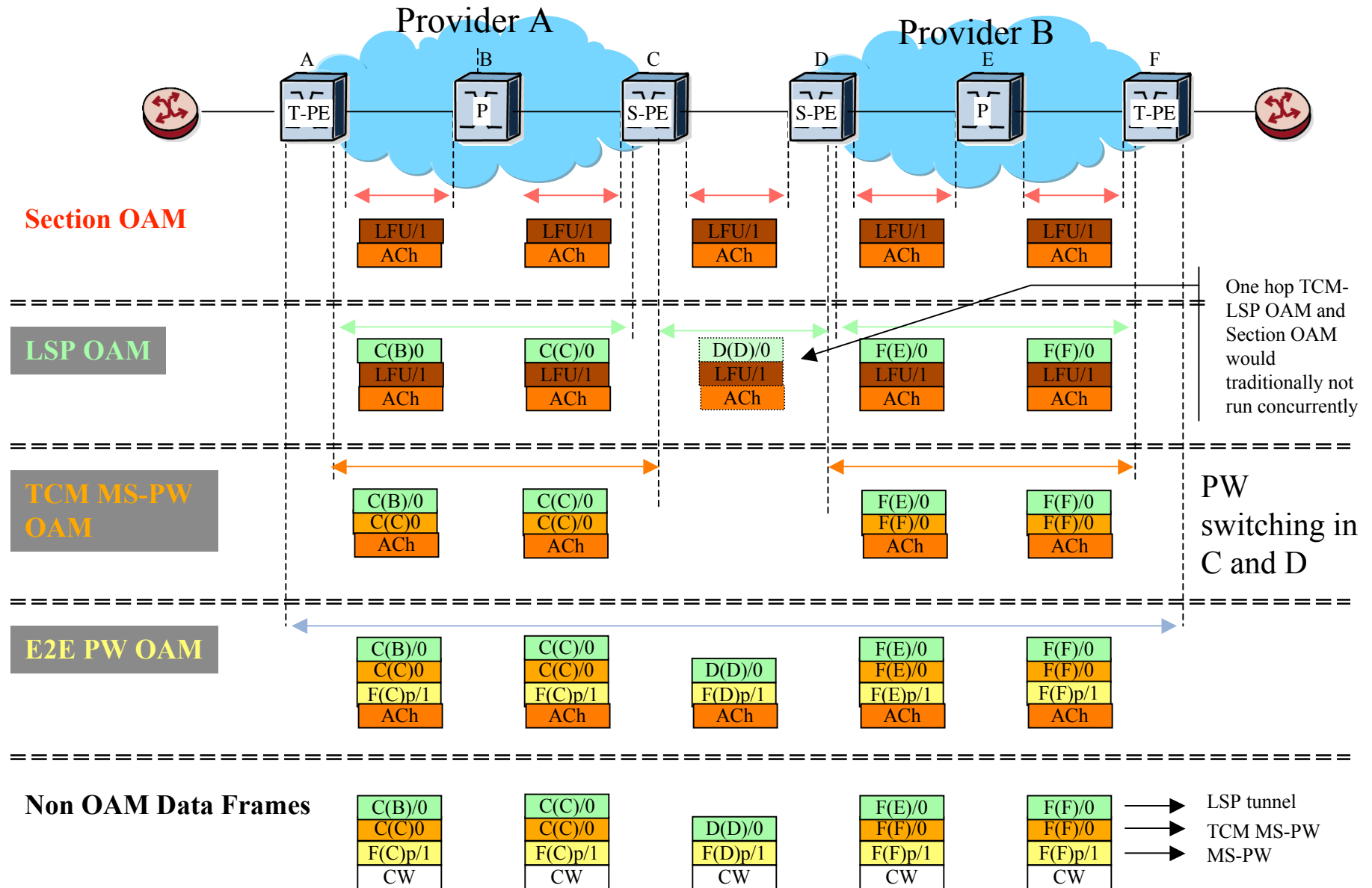
LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word



# Inter-provider MS-PW

## LSP, MS-PW & TCM-MS-PW OAM

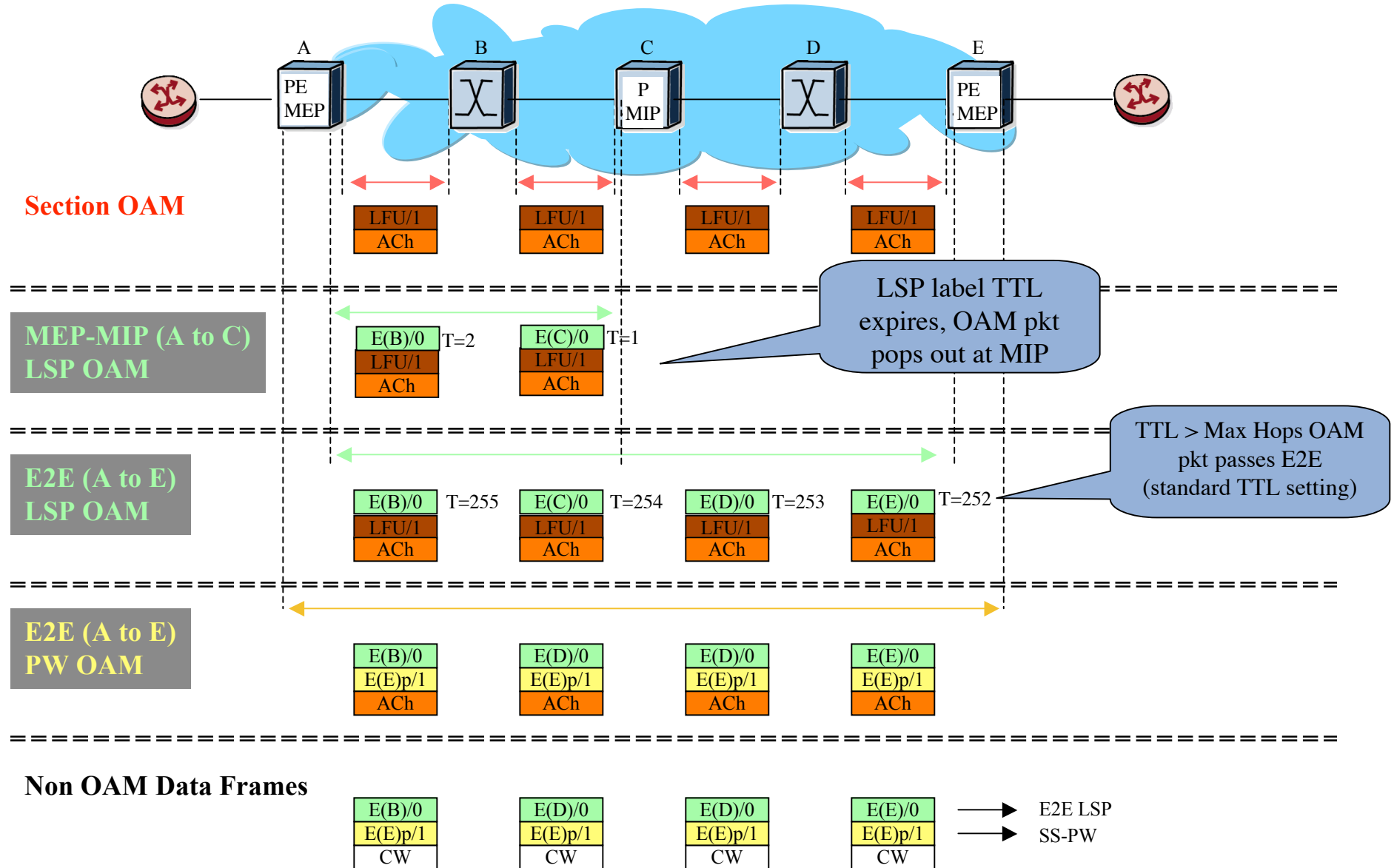
LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word



# SS-PW over Intra-domain LSP

## LSP MEP->MIP OAM using TTL

LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word  
 T = TTL



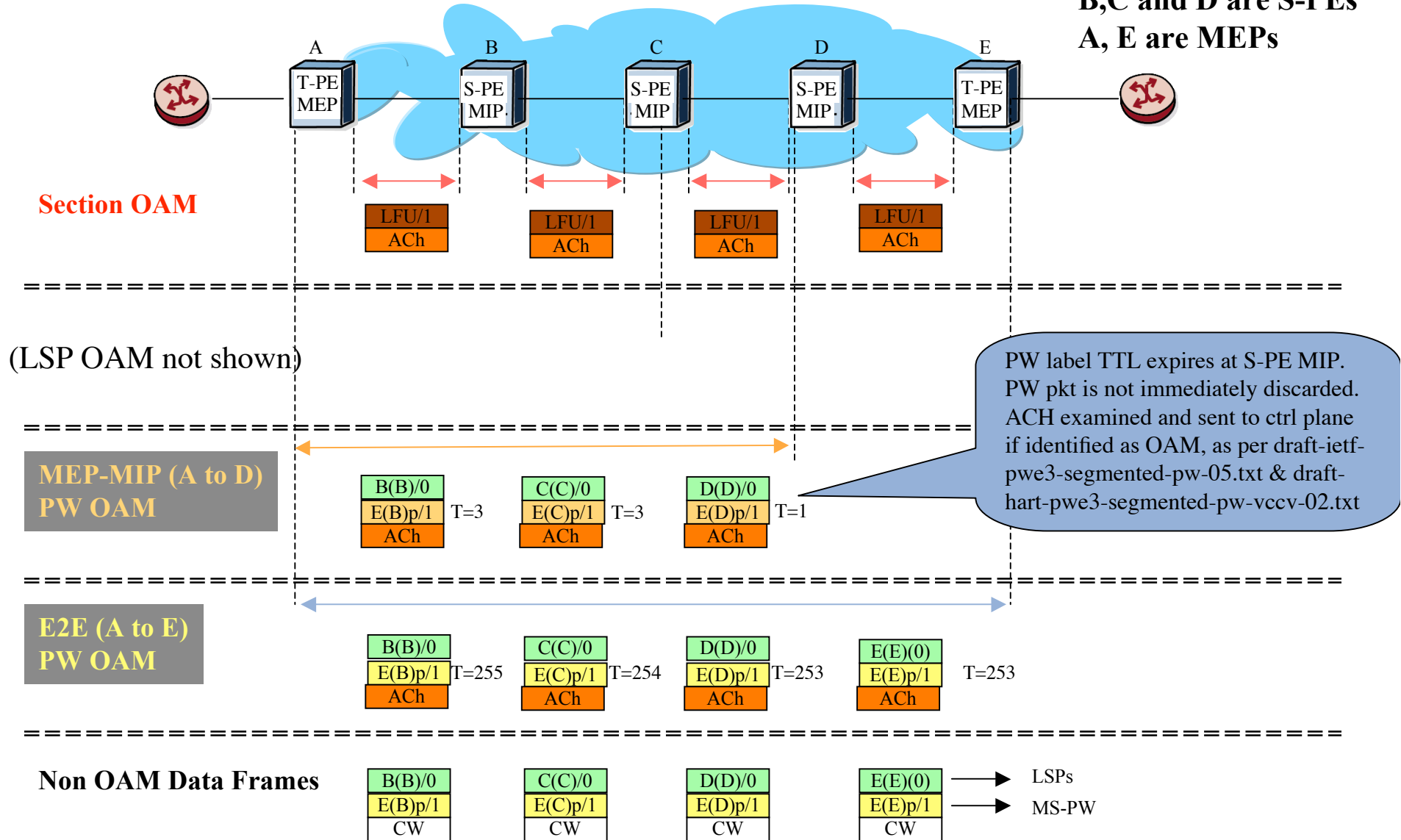
# Intra-domain MS-PW

## MS-PW MEP->MIP OAM using TTL (No TCM)

(See draft-ietf-pwe3-segmented-pw-)

LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word  
 T = TTL

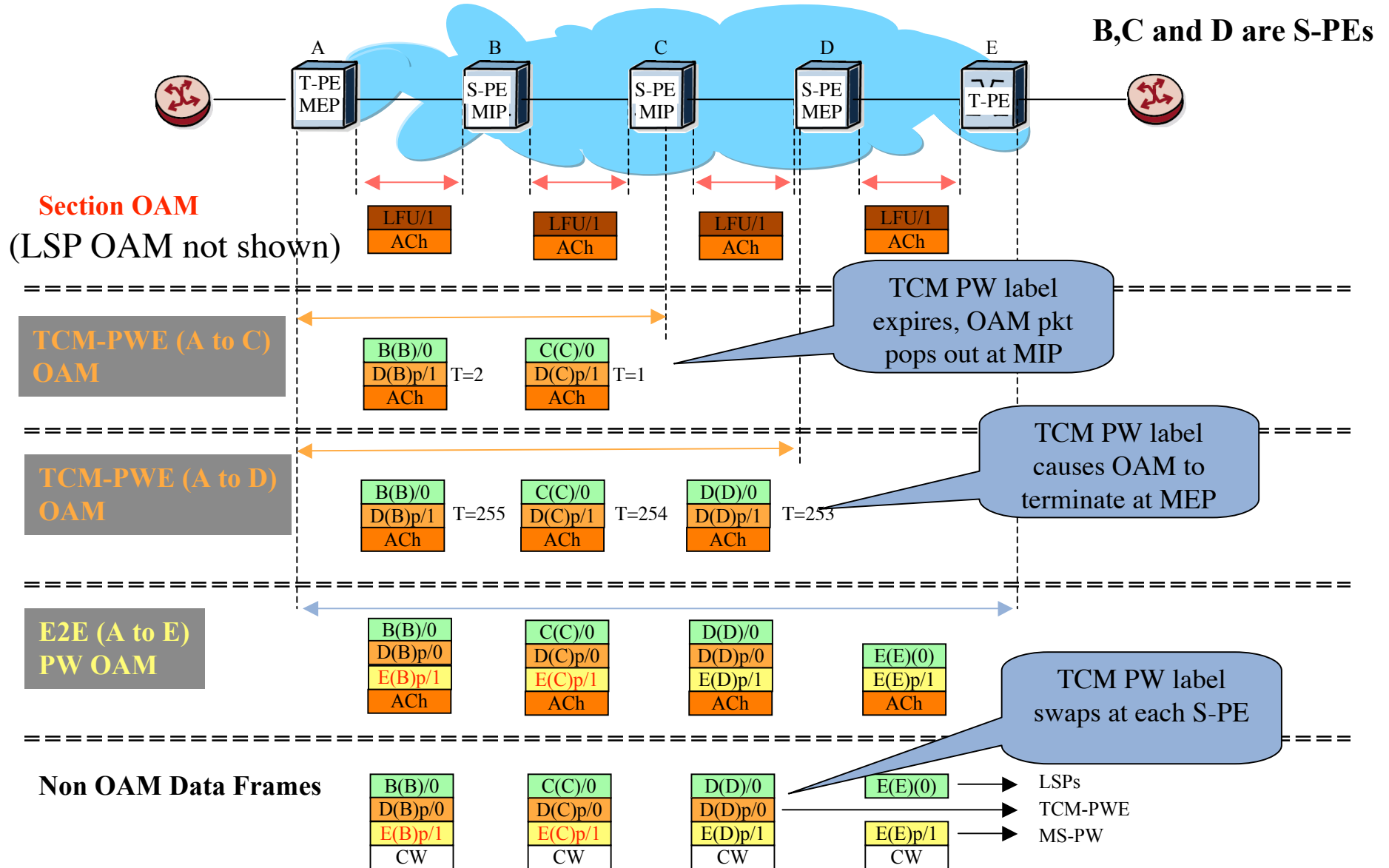
**B,C and D are S-PEs**  
**A, E are MEPs**



# Intra-domain MS-PW

## TCM-MS-PW MEP->MIP OAM using TTL

LFU – Label For You (label 13)  
 ACh – Associated Channel  
 CW – Control Word  
 T = TTL



# MEP to MIP OAM: TTL Processing for PWs and LSPs

- In order to maintain individual levels of OAM and path detection

Use pipe model per label level

TTL is not copied up the stack on a push

TTL is not copied down the stack on a pop

TTL is decremented on each swap and pop action

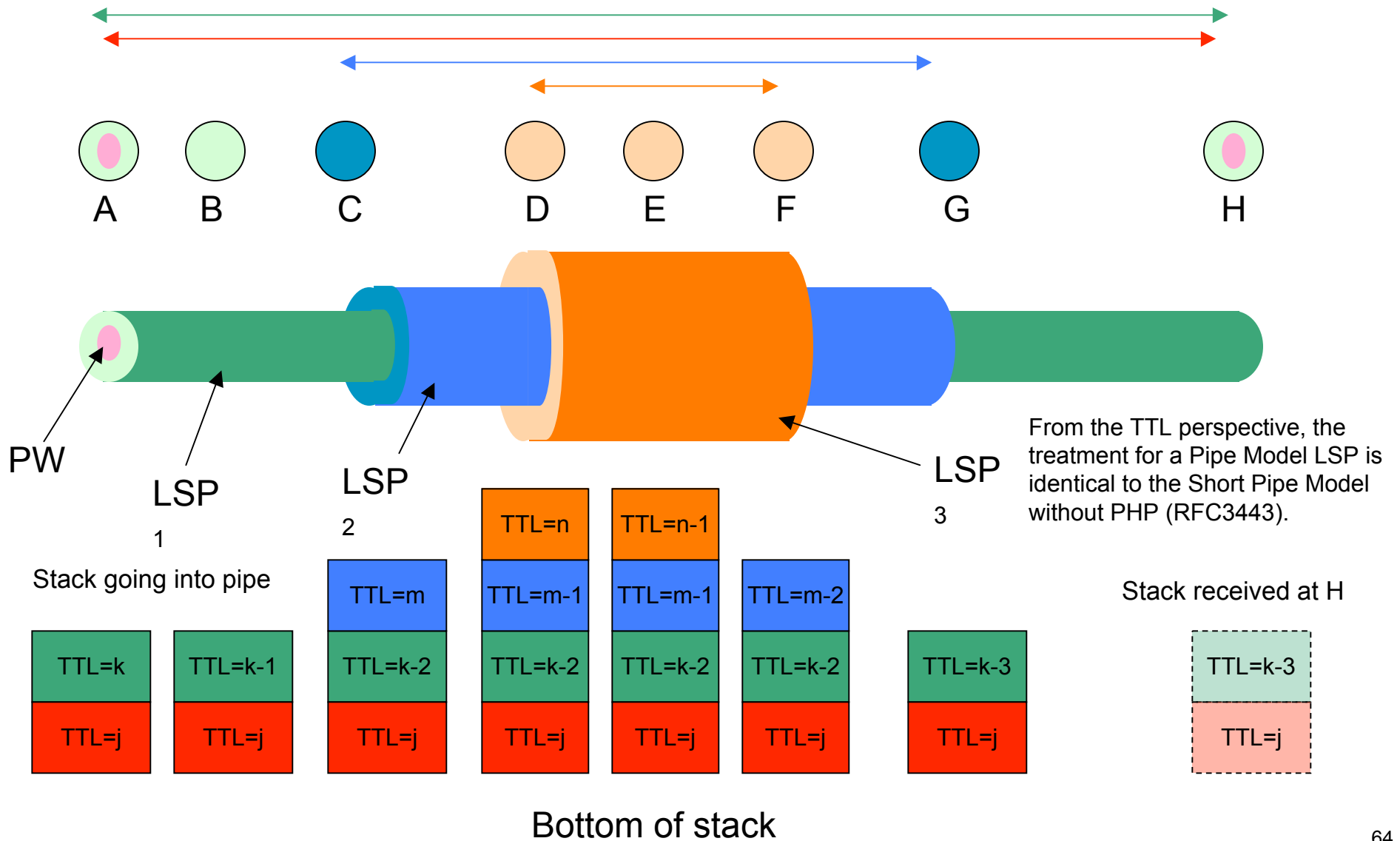
Traceroute for a level can be used to trap packets at each node that processes the label for that level in the label stack

*Scenarios to be added:*

*a) LSP on FRR path (both facility and detour)*

*b) PW with ACH processing (no need for LFU, so processing steps are slightly different from LSP processing)*

# Short Pipe Model with Nested TTL and No PHP Processing





## Nested LSP TTL Processing (1)

- The previous picture shows
  - PW: Pseudowire
  - LSP1: Level 1 LSP (PW is carried inside)
  - LSP2: Level 2 LSP (LSP1 is nested inside)
  - LSP3: Level 3 LSP (LSP2 is nested inside)
- TTL for each level is inserted by the ingress of the level
  - PW TTL is initialized to j at A
  - LSP1 TTL is initialized to k at A
  - LSP2 TTL is initialized to m at C
  - LSP3 TTL is initialized to n at D
- TTL for a particular level is decremented at each hop that looks at that level
  - PW TTL is decremented at H
  - LSP1 TTL is decremented at B, H
  - LSP2 TTL is decremented at G
  - LSP3 TTL is decremented at E, F

## Nested LSP TTL Processing (2) - pseudo code

If a packet arrives at a node with TTL != 1, then the TTL is decremented

If the LFIB action for this label is POP, then this node should be a MEP for this label level

If the packet has an LFU below the current label

The packet is passed to the control plane module for processing, including validating that the node is a MEP, the packet contents are consistent

The appropriate OAM actions, as described by the packet, are taken

A reply, if required, is returned to the MEP that originated this message

If the packet doesn't have an LFU below the current label

If the current label is not bottom of stack, continue processing label stack

If the current label is bottom of stack, forward the packet according to egress processing for this level

## Nested LSP TTL Processing (3) continued pseudocode

If a packet arrives at a node with TTL = 1, then the TTL is decremented and goes to 0

If the packet has no LFU below the current label, then the packet may be discarded

Statistics may be maintained for these packets

If the packet has an LFU just below the current label

If the LFIB action for this label is POP, then this node should be a MEP for this level

The packet is passed to the control plane module for processing, including validating that the node is a MEP, the packet contents are consistent

The appropriate OAM actions, as described by the packet, are taken

A reply, if required, is returned to the MEP that originated this message

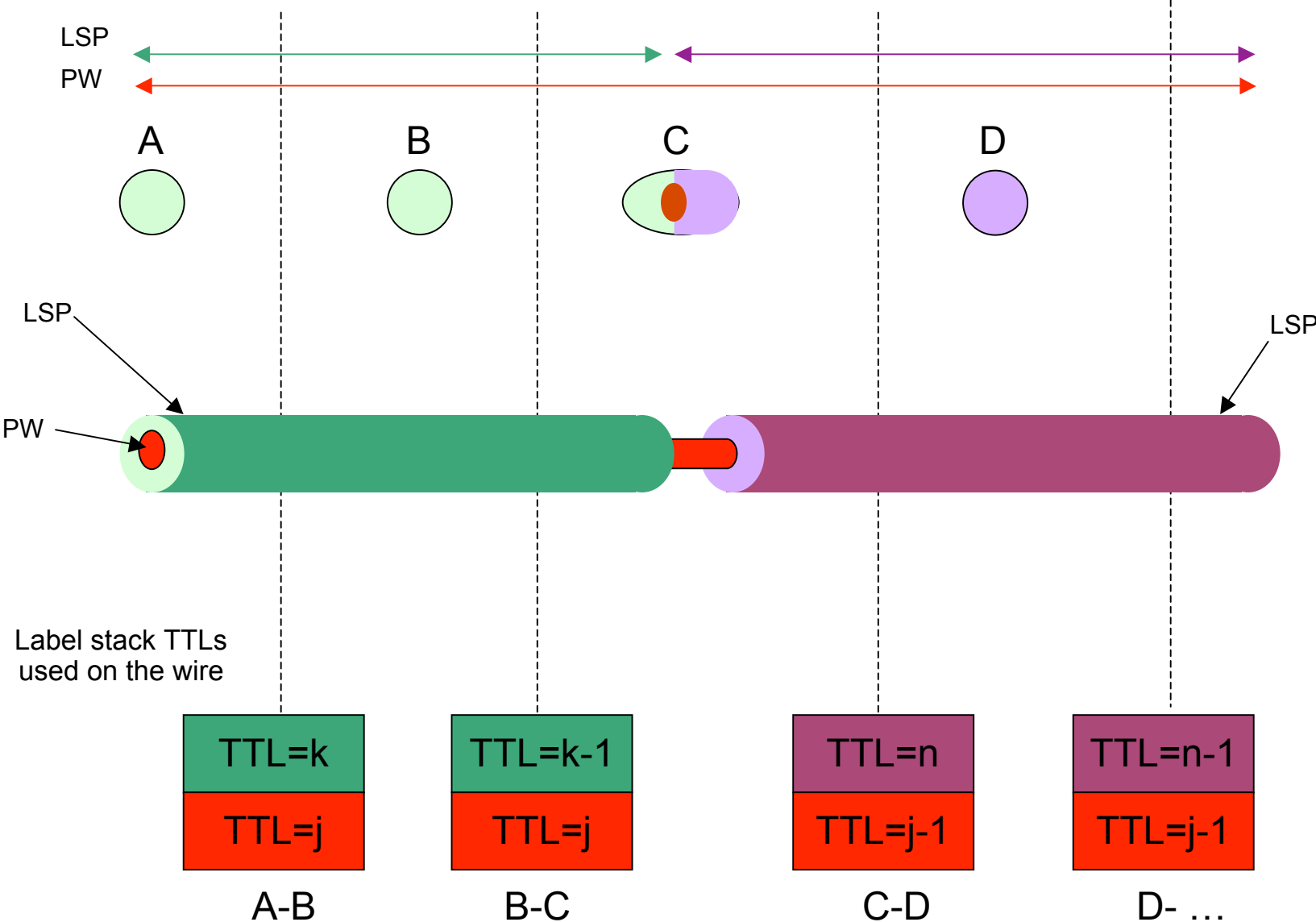
If the LFIB action for this label is SWAP, then this node should be a MIP for this level

The packet is passed to the control plane module for processing, including validating that the node is a MIP, the packet contents are consistent

The appropriate OAM actions, as described by the packet, are taken

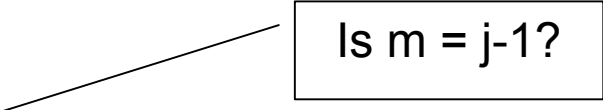
A reply, if required, is returned to the MEP that originated this message

# Multi-Segment PW TTL Processing



## Cascaded LSP TTL Processing

- The previous picture shows
  - PW1: Pseudowire
  - LSP1: Level 1 LSP (PW1 is carried inside)
  - PW2: Pseudowire (PW1 is stitched to PW2)
  - LSP2: Level 1 LSP (PW2 is carried inside)
- TTL for each level is inserted by the ingress of the level
  - PW1 TTL is initialized to  $j$  at A
  - LSP1 TTL is initialized to  $k$  at A
  - PW2 TTL is initialized to  $m$  at C
  - LSP2 TTL is initialized to  $n$  at C
- TTL for a particular level is decremented at each hop that looks at that level
  - PW1 TTL is decremented at C
  - LSP1 TTL is decremented at B, C
  - PW2 TTL is decremented at E
  - LSP2 TTL is decremented at D, E

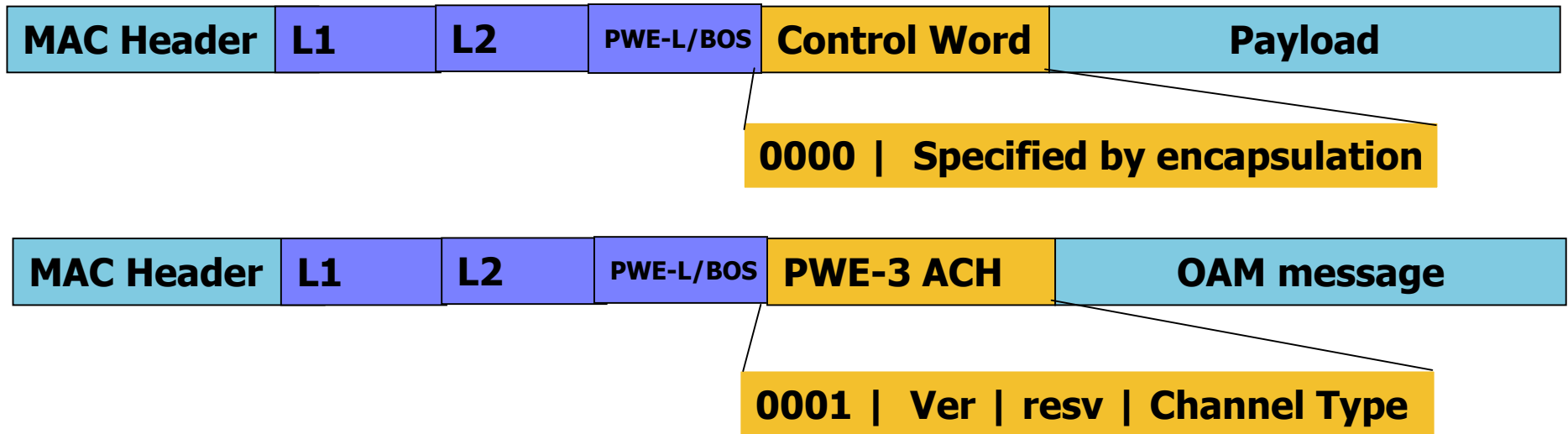


Is  $m = j - 1$ ?

# ECMP Considerations

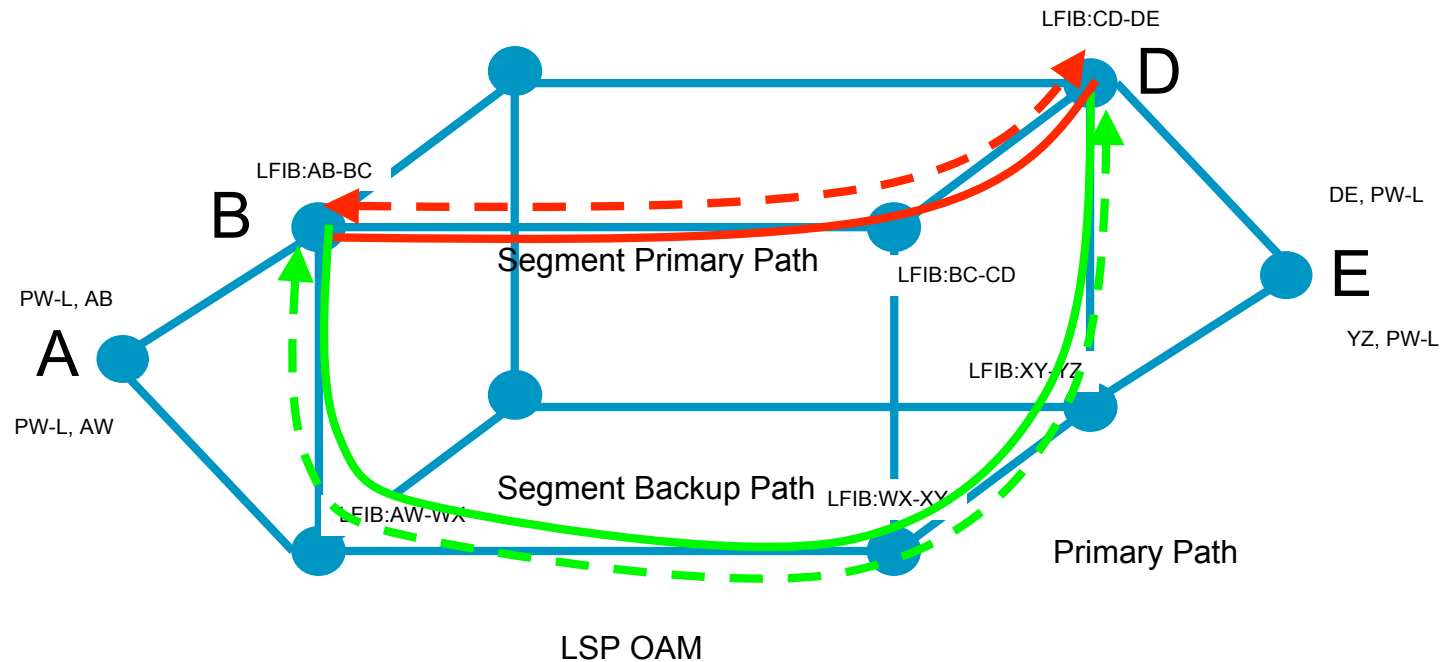
- OAM and Data MUST share fate.
- PW OAM fate shares with PW through the first nibble mechanism (RFC4928) and hence is fate shared over any MPLS PSN.
- Fate sharing is not assured for the MPLS Tunnel OAM/Data in the presence of ECMP.
- The current MPLS Transport Profile ensures OAM/Data fate sharing for the MPLS tunnel by excluding the use of MPLS ECMP paths (for example by only using RSVP or GMPLS signaled MPLS tunnels)
- There is a requirement to improve IETF MPLS OAM. This will require the problem of fate sharing in the presence of ECMP to be addressed.
- If the OAM/DATA fate sharing problem is solved for MPLS ECMP, then the Transport Profile may be extended to take advantage MPLS paths that employ ECMP.

# RFC4928 Mechanism



- Static Control Plane
  - Under the control of an external NMS therefore should not be an issue
  - Single discrete LSPs defined through static provisioning system
- Dynamic Control Plane environment
  - Routing protocols and LDP may set-up ECMP routes
  - Traffic Engineering can as well (auto-route)
- Recognized in IETF
  - RFC 4928 Avoiding Equal Cost Multipath Treatment in MPLS Networks : 0 or 1 in the first nibble of the payload
  - RFC 4385 PW3 Control Word for Use over an MPLS PSN : Defines “Generic PWE-3 control word” and “PW Associated Channel” formats
- A consistent approach required for MPLS with a transport profile
  - RFC 4928 implemented through use of control word and PWE-3 ACH
  - RFC 4385 for Control Word and PW associated Channel formats
  - NOTE: joint proposals to be made on “Load Balance” label technology in PWE3 WG

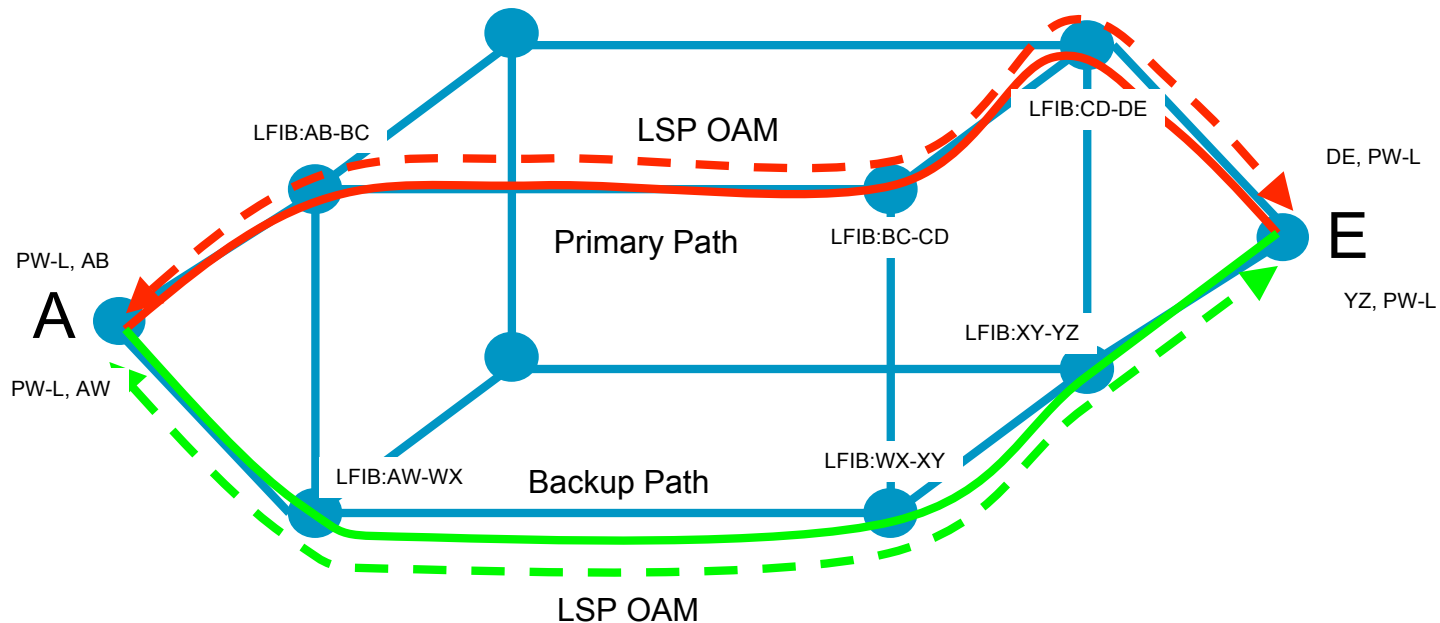
# Segment LSP operations



- Path diversity is not part of the OAM process. It is the responsibility of the Control or Management Plane
- OAM function uses LFU with Generic Channel Association
- Pre-provisioned segment primary and backup paths
- LSP OAM running on segment primary and back-up paths (using a nested LSP)
- OAM failure on backup path → Alert NMS
- OAM failure on primary path results in B and D updating LFIB to send traffic labelled for BD via segment backup path
- End to End traffic labelled for BD now pushed onto segment backup path



# End to End LSP operations



- Path diversity is not part of the OAM process. It is the responsibility of the Control Plane
- OAM function uses LFU with Generic Channel Association
- Pre-provisioned primary and backup paths
- LSP OAM running on primary and back-up paths
- OAM failure on backup path → Alert NMS
- OAM failure on primary path → A and E updating LFIB to send and receive PW-L traffic over backup path

# PHP

- It is believed that PHP may be able to be used in the transport profile.
- The issue is how do we maintain the packet context for both the data and OAM  
described on the following 3 slides
- One scenario follows:  
SS-PW, LSP and TCM-LSP

# Packet Context

- OAM operations require packet context.
- Work to date has proposed that this is supplied by the label value and hence precludes the use of PHP.
- Using the label as the identifier is a simple mechanism that can be applied to both OAM and data packets, but has a number of issues:
  - Precludes PHP which has cost and applicability implications for the OAM
  - Label errors may produce complex network issues
- Other context indicators may be available that allow the lifting of the PHP constraint (at least as an option).

# Alternative Context Indication

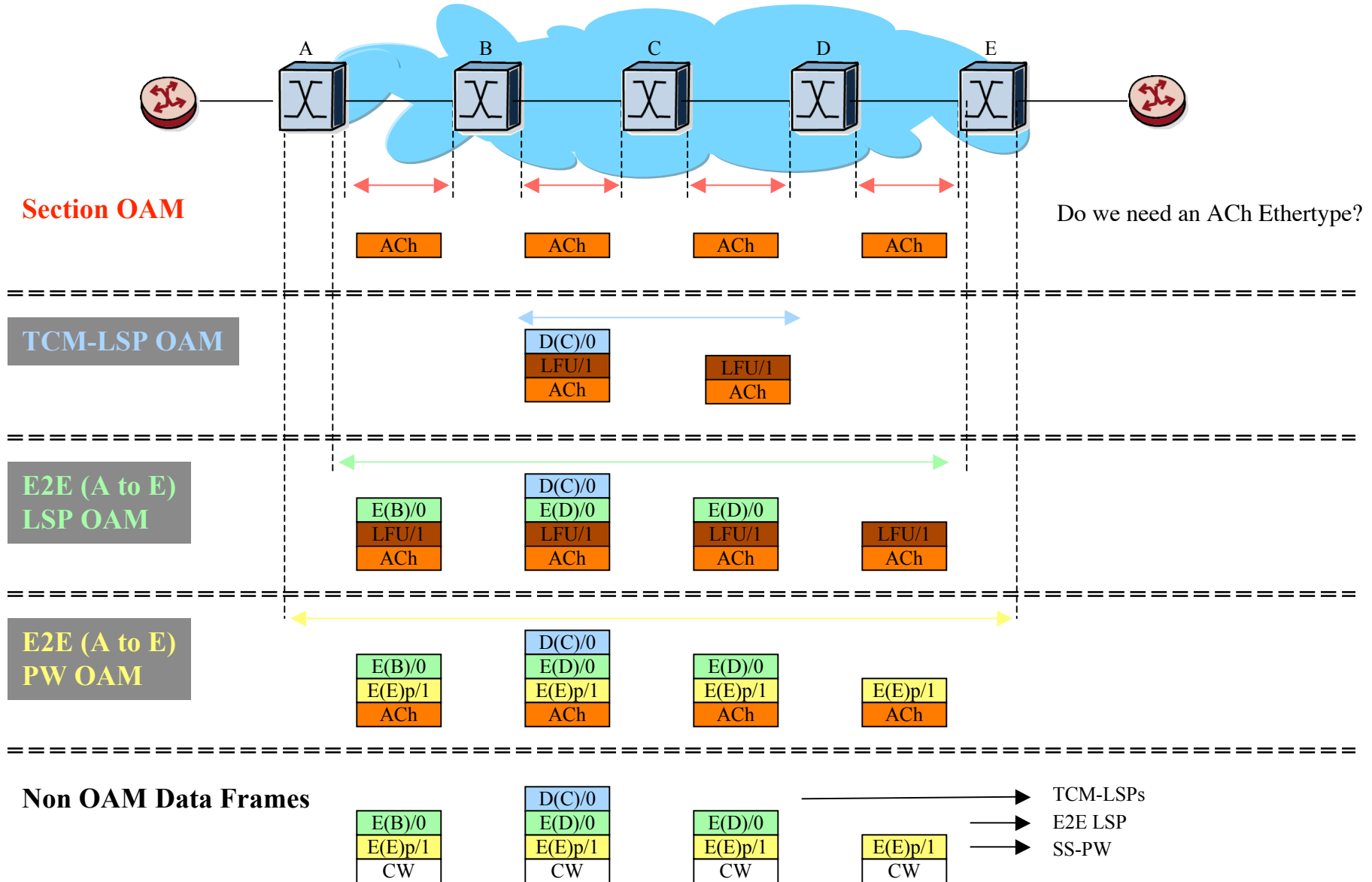
- In the case of IP the IP address provides context
- In the case of PW, the PW label provides context
- In the case of an OAM pkt, an identifier can provide context
- The issue are:
  - OAM and data must fate share;
  - Need to provide context identification for performance monitoring of data packets, or the need to provide an alternative mechanism that provides satisfactory performance information.

# Use of alternate context mechanisms

- The MPLS architecture supports label retention and hence we can proceed on the basis that this approach is available to the design team.
- There are costs to the prohibition of PHP that needs to be fully understood and accepted.
- During the design phase we need to:
  - Understand the costs, limitations, vulnerabilities and advantages of the PHP and non-PHP approaches
  - Either
    1. Confirm label as context identifier and hence confirm PHP restriction
    2. Propose an alternative mechanism that satisfies all needs and which permits PHP
    3. Propose the specification of a PHP and non-PHP method with appropriate applicability statements.

# SS-PW, LSP and TCM-LSP OAM - PHP

LFU – Label For You (label 13|14)  
 ACh – Associated Channel  
 CW – Control Word



# Control Plane

# Conclusions/Recommendations

- Control plane sub-team sees **no show-stoppers**
  - Existing IETF protocols can be used to provide required function
    - Transport network operation
    - DCN/SCN operation
  - IETF GMPLS protocols already applied to ASON architecture
  - Any protocol extensions needed will be easy to make
    - Configuration of MEPs/MIPs and activation of monitoring*
    - Support of bridge and roll capability*
      - Allows Tandem connection monitoring to be added to an existing LSP without disruption to the service*

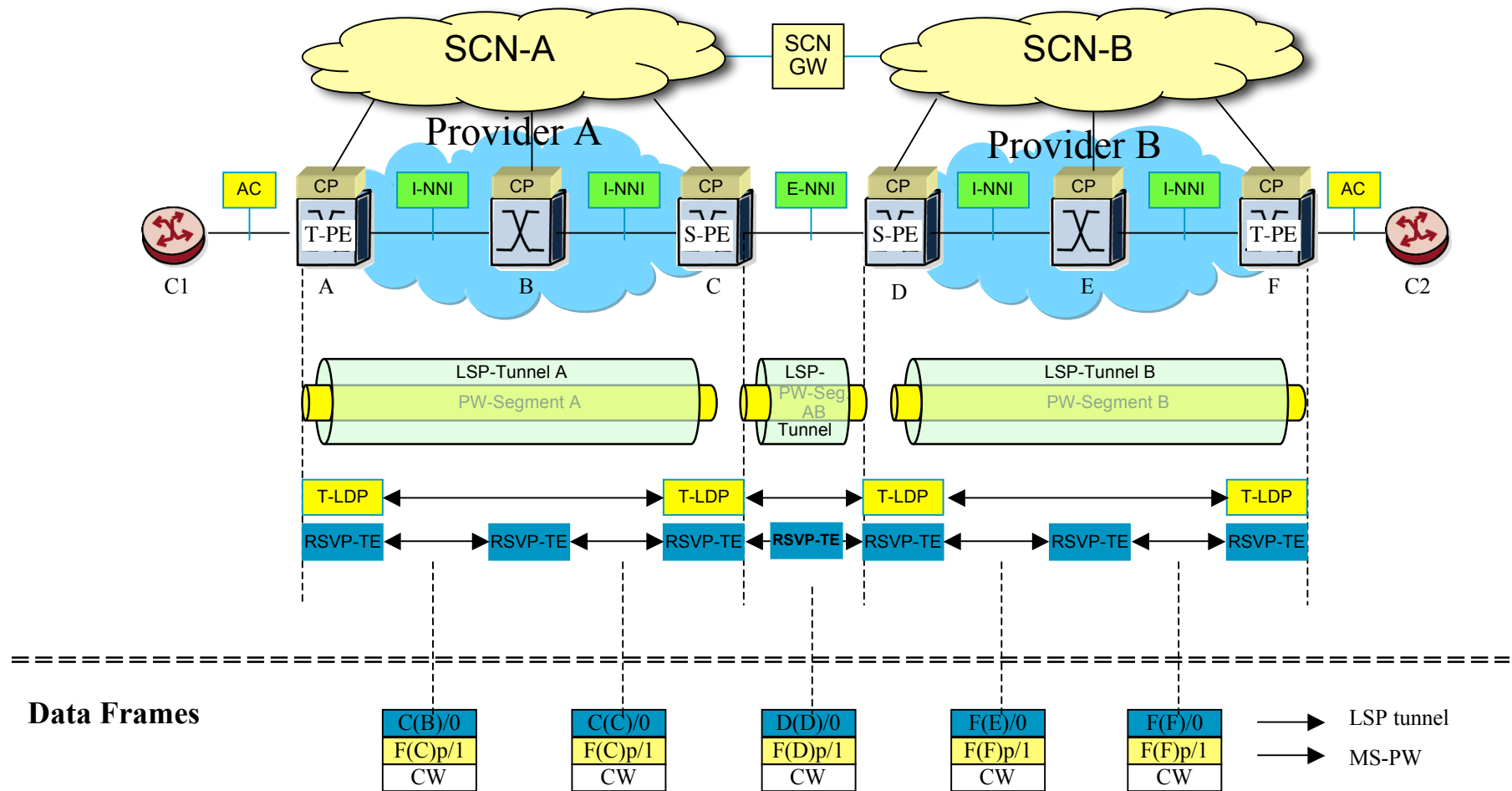


# Discussion

- Transport profile should meet the requirements of the ASON architecture
  - Use IETF protocol suite given it is used for ASON
    - GMPLS RSVP-TE for LSP signaling
    - GMPLS OSPF-TE and ISIS-TE for LSP TE information distribution
    - LDP will be used for PW setup (as part of client set up process)
- DCN/SCN
  - IP-based DCN/SCN
  - ACH defines ECC
    - Can have as many channels and protocols as necessary and therefore could support the SCN
  - Must have policing for DCN/SCN
  - IS-IS or OSPF running in DCN to provide DCN topology information
- Connectivity discovery and verification
  - Could use LMP if native mechanisms not adequate

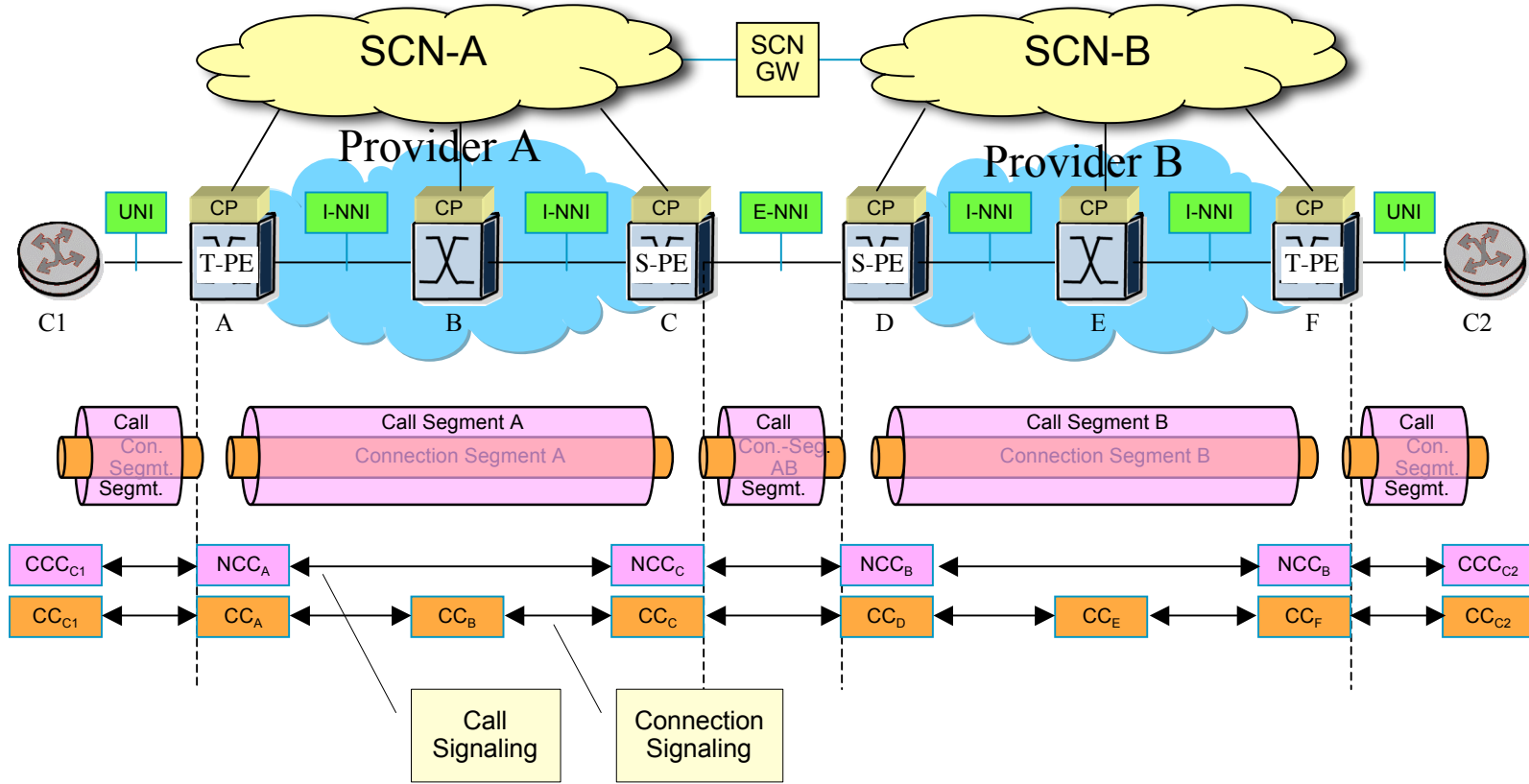
# Control Plane View of Inter-provider MS-PW

AC – Attachment Circuit  
 NNI – Network-Network Interface  
 I-NNI – Internal NNI  
 E-NNI – External NNI  
 SCN – Signaling Communication Network  
 SCN-GW Gateway  
 T-LDP – Targeted LDP



# ASON Call/Connection Model

CCC – Client Call Controller  
 NCC – Network Call Controller  
 CC – Connection Controller  
 UNI – User-Network-Interface  
 NNI – Network-Network Interface  
 I-NNI – Internal NNI  
 E-NNI – External NNI



# Survivability

# Advice

- Survivability sub team has not found any issues that prevent the creation of an MPLS transport profile
  - No showstoppers found*
- Therefore option 1 can be selected
- Summary of discussion
  - Three potential solutions have been identified
  - Each solutions has different attributes and advantages
  - Further work in the design phase should eliminate one or more of these options and/or provide an applicability statement

# Discussion

- Nested LSPs (potentially PWEs) provide levels of hierarchy to support per segment and path recovery
  - Must draw up PWE requirements*
- Most of the time intermediate nodes do not process the entire stack
- Each segment can act independently
  - Multiple potential solutions including
    - Native IETF mechanisms
    - Carry G.8131/G.8132 PDUs in an ACH

# Discussion - 2

- Native MPLS protection schemes, such as facility bypass and detours, can be used to provide ring protection in most, but not optimal in some scenarios
  - A single facility bypass LSP protects all LSPs over a specific link by wrapping traffic
  - A detour LSP can be used for optimal traffic delivery to the egress point (without wrapping)
    - A detour LSP is needed for every LSP to be protected.
  - Also can provide optimized exit preventing the 2x bandwidth in other wrapping repair technologies
  - Must add notion of DOWN and ADMINDOWN (e.g. standby bit)
- ITU-T G.8132 TM-SPRing defines a ring protection that includes additional capabilities to the MPLS protection schemes, by supporting coordinated protection in case of multiple failures (using single protection mechanism for all cases)
- MPLS ring protection strategies provide necessary functionality and option 1 can be recommended but, there appears to be cases where G.8132 may provide additional functionality that may be incorporated and specified

**We have found no showstoppers**

# Requirements summary - Rings

- MPLS-TP ring protection shall satisfy the following:
  - Less than 50 ms switching time
  - Protect p-t-p and p-t-mp connections
  - Support normal traffic and non-preemptable unprotected traffic
  - Provide hold-off timer and wait to-restore timer
  - Protect all traffic possible in case of single and multiple failures
    - Fiber, nodes or both
    - Failures that segment the ring
  - Support operator's commands
  - Support a priority scheme to arbitrate between switch requests from multiple faults and/or operator commands
    - Provide ability to coordinate multiple requests in the ring
  - Bi directional switching
  
- ITU-T References:
  - ETSI TS 101 009, Section 6.2.2
  - ITU-T G.841, Section 7.2.2
  - Telcordia GR-1230, Section 5
  - ITU-T Draft G.8132, Section 7



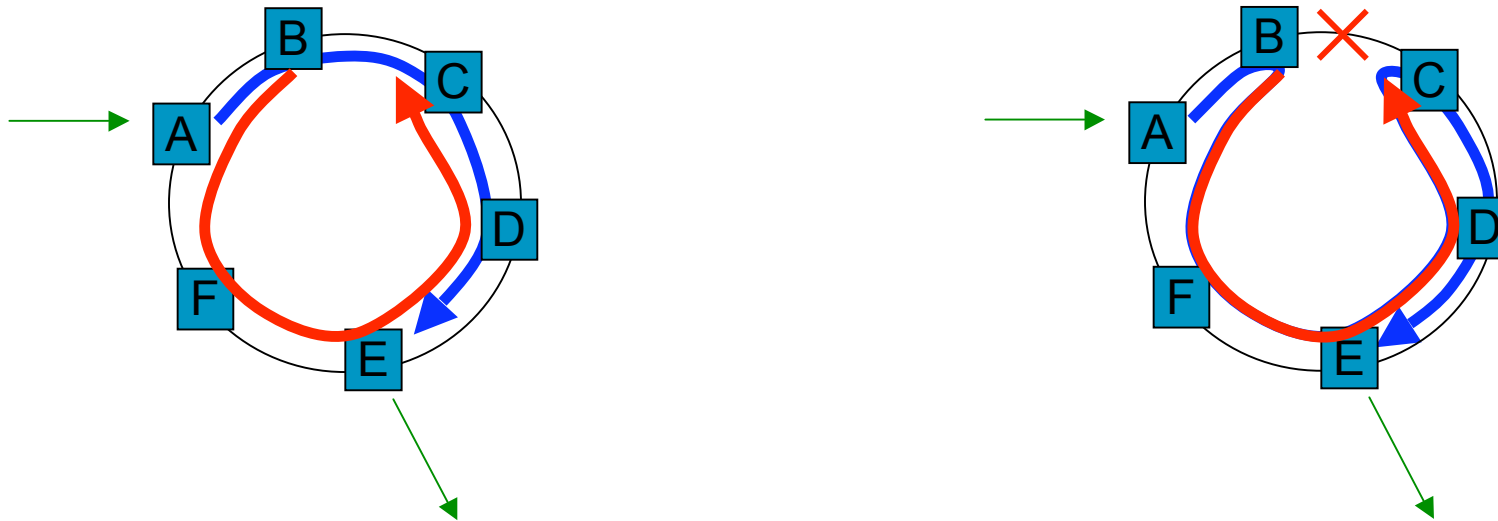
# Requirements summary - Linear

- MPLS-TP linear protection shall satisfy the following:
  - Less than 50 ms switching time
  - Protect p-t-p and p-t-mp connections
    - P-2-MP LSP protection based on detours is covered in RFC 4875, though an example is not included here
  - Support normal traffic and non-preemptable unprotected traffic
  - Provide hold-off timer and wait to-restore timer
  - Support operator's commands
  - Support a priority scheme to arbitrate between switch requests from multiple faults and/or operator commands
  - Bi directional switching
  - Revertive and non revertive operation
- ITU-T References:
  - G.808.1 – Generic linear protection
  - G.8131 T-MPLS linear protection
- Not addressed
  - Reuse (or simplify) the mechanism used for Ring protection?*

# Example Scenarios in the following slides

- Basic restoration in a ring
  
- MPLS protection scenarios
  - Facility Bypass
  - Restoration using detours
    - Sub-optimal
    - Optimized
  
- ITU-T G.8132 TM-SPRing protection overview
  - Label Allocation
  - OAM and APS messaging
  - P2P
  - P2MP
  - Multiple failures

# MPLS Facility Bypass Example



## Example:

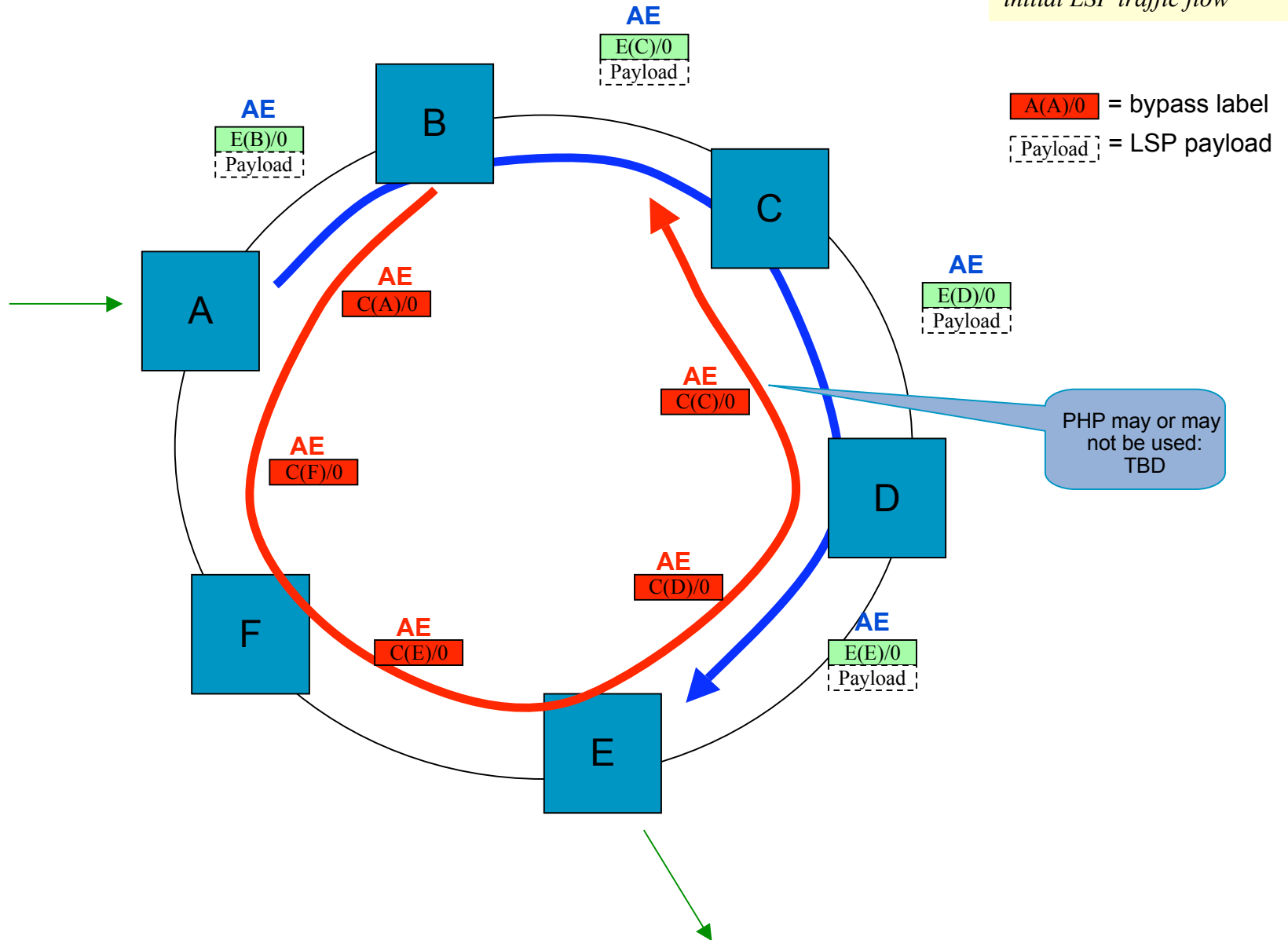
- Assume ingress to ring is at A and egress is at E
- Facility bypass (B-A-F-E-D-C) is established to protect link B-C
- Link B-C in the ring goes down
- Facility bypass protects failure of link B-C with the red path to the merge point (C)
- Emulates conventional optical ring failure recovery
- Requires two-label stack to redirect the LSP around the failure
- Scale issue:
  - One facility bypass provides protection for all LSPs over link B-C
  - One facility bypass for each link in the ring (shared by all LSPs on that link)

# MPLS Facility Bypass Label Stack .1

## Initial State, unidirectional LSP

**AE** = Initial clockwise ring  
**AE** = bypass for **AE**

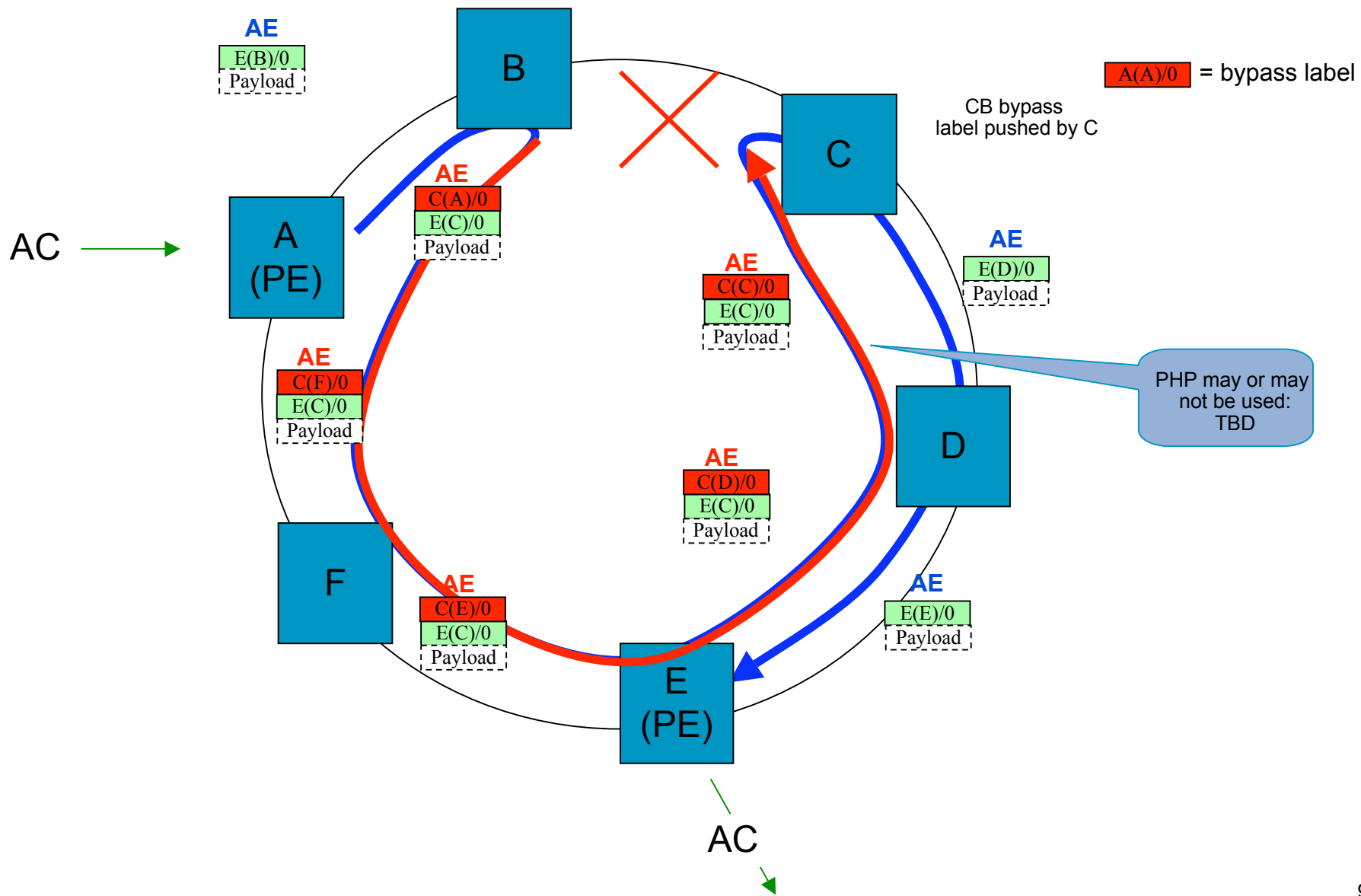
*Spin is relative to initial LSP traffic flow*



# MPLS Facility Bypass Label Stack .2

## Failure state, Unidirectional LSP

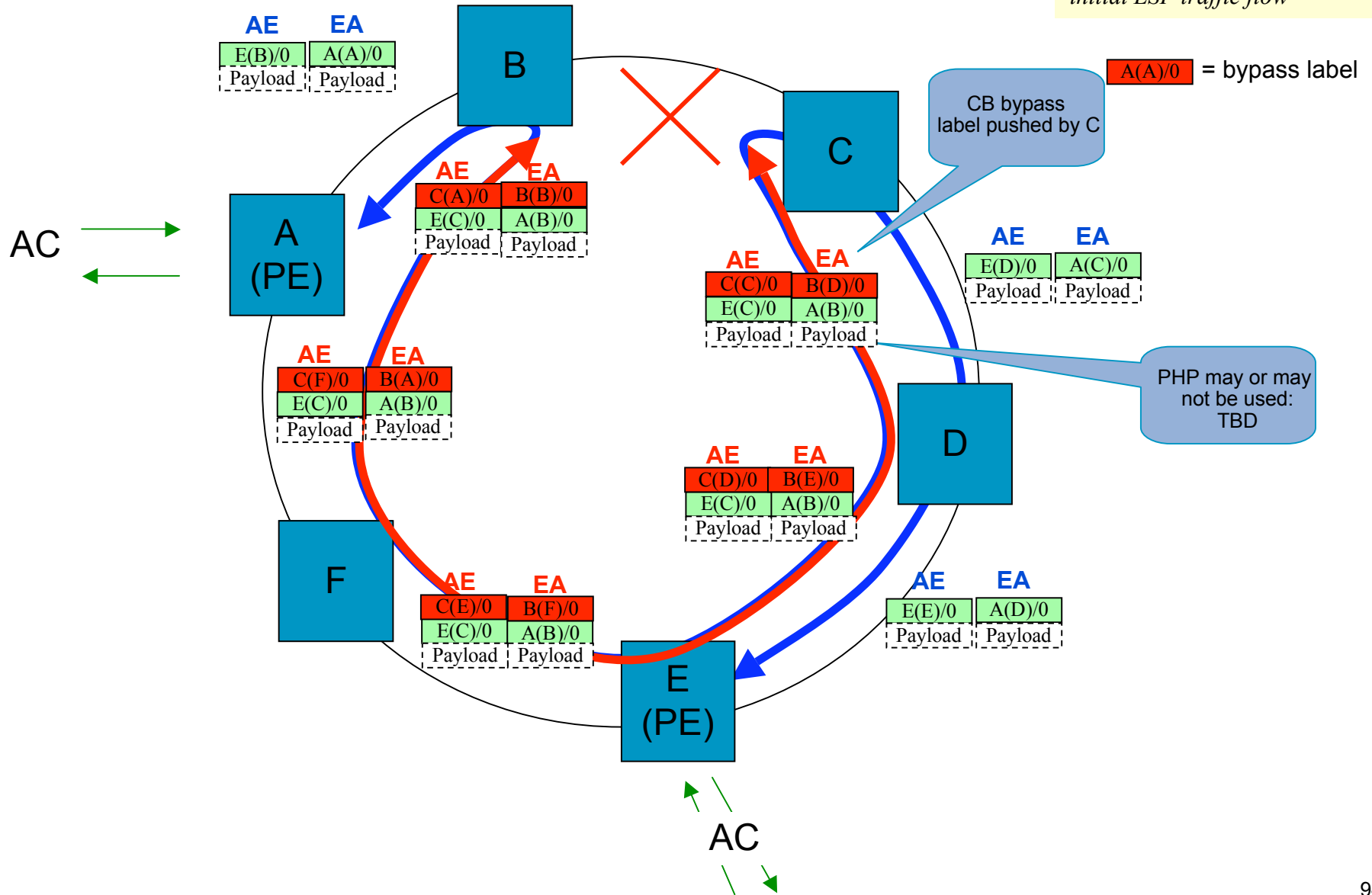
**AE** = Initial clockwise ring  
**AE** = bypass for **AE**  
*Spin is relative to initial LSP traffic flow*



# MPLS Facility Bypass Label Stack

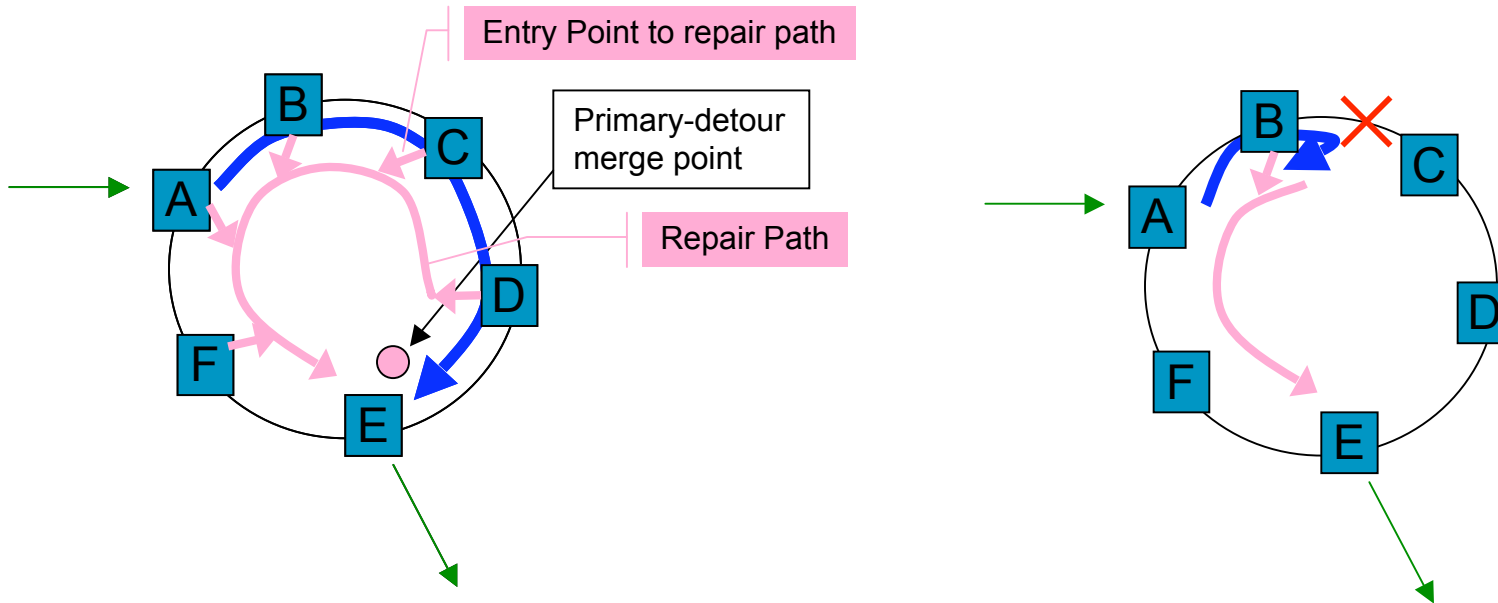
## Failure state, Bidirectional LSP

**AE** = Initial clockwise ring  
**EA** = Initial anticlockwise ring  
**AE** = bypass for **AE**  
**EA** = bypass for **EA**  
*Spin is relative to initial LSP traffic flow*



# MPLS 1:1 Detours

## - Optimized Restoration



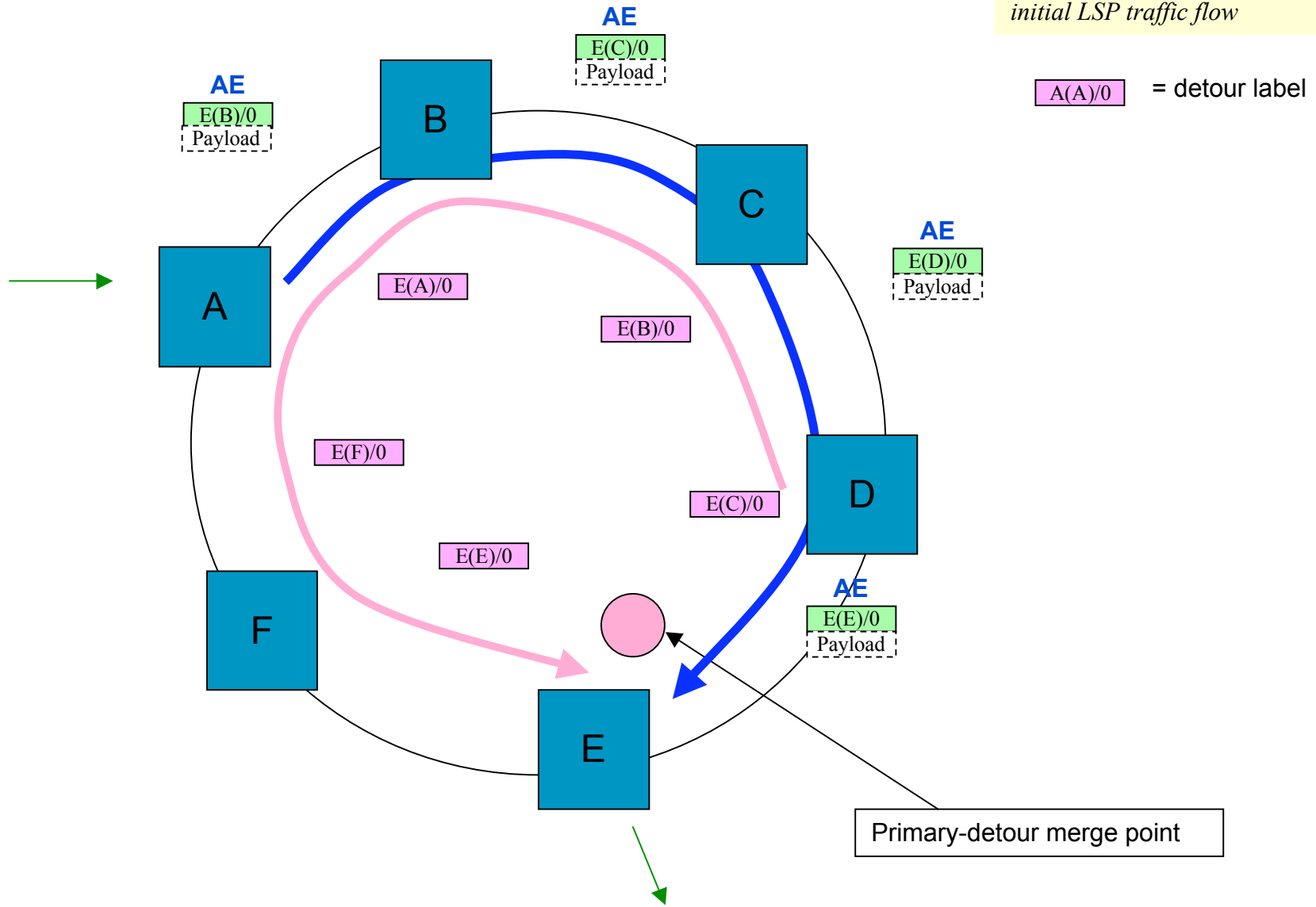
### Example

- Assume ingress to ring is at A and egress is at E
- Detour established to protect link B-C merges with primary path at E, resulting in protection through B-A-F-E
- Link B-C in the ring goes down
- Detour carries traffic to E
- Optimizes on conventional optical ring and facility bypass failure recovery
- Requires one-label stack to redirect the LSP around the failure
- Scale issue:
  - One detour per LSP is required for each working LSP
  - The detour LSP can be used to protect the failure of any link on the ring

# MPLS 1:1 Detours - Label Stacks .1

## Initial state, Unidirectional LSP

**AE** = Initial clockwise ring  
**EA** = Initial anticlockwise ring  
**AE** = detour for **AE**  
**EA** = bypass for **EA**  
*Spin is relative to initial LSP traffic flow*

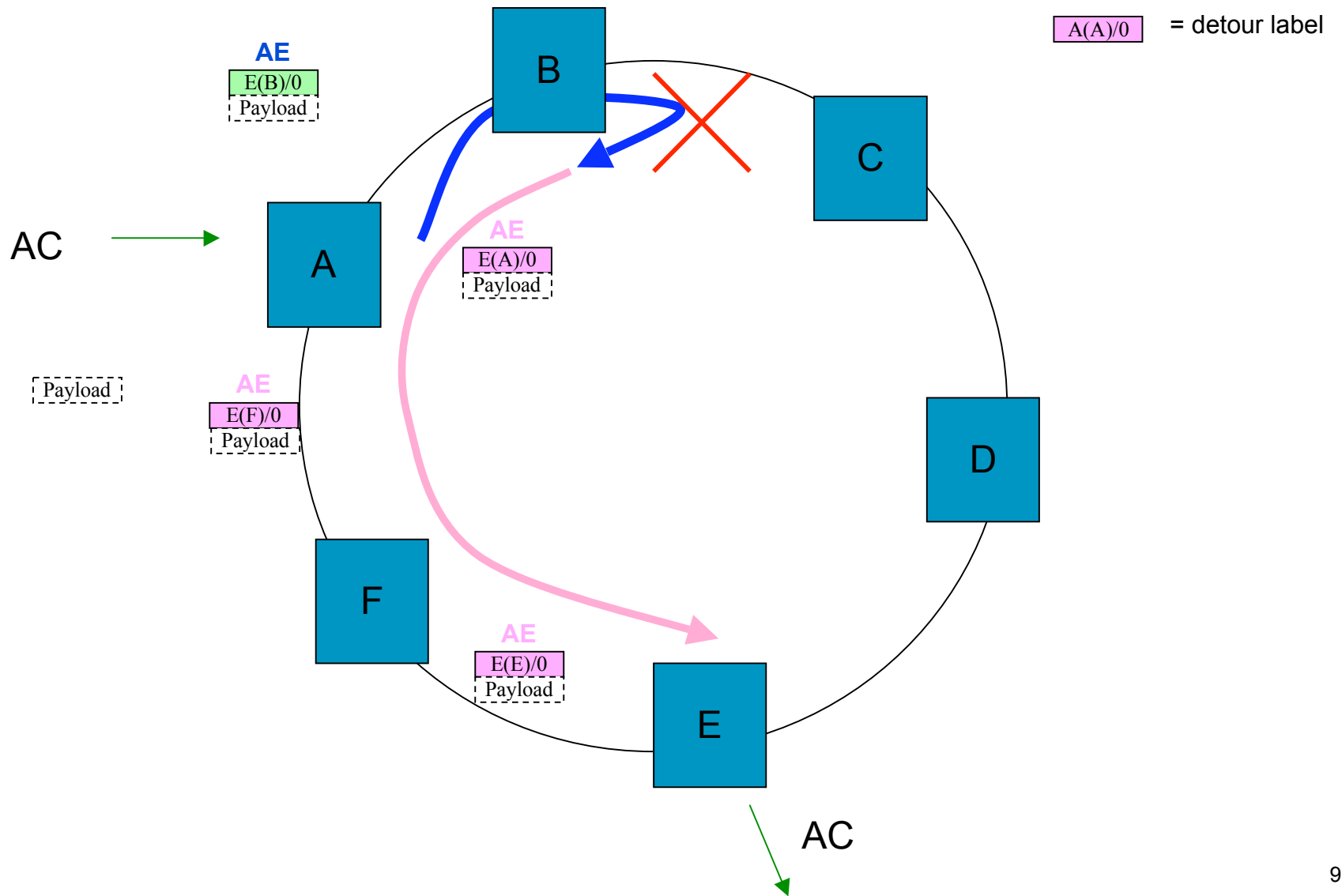




# MPLS 1:1 Detours - Label Stacks .2

## Failure state, Unidirectional LSP

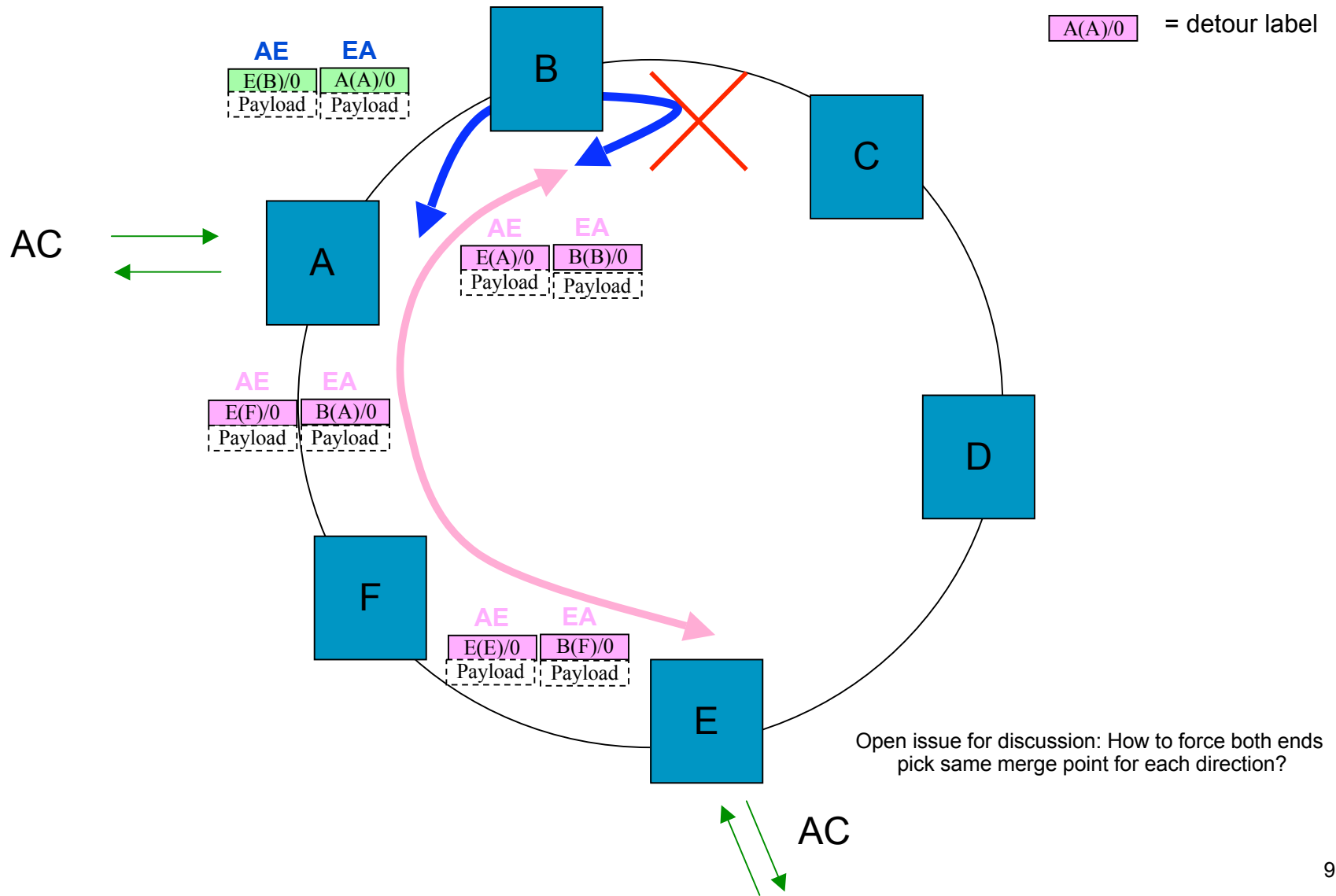
AE = Clockwise ring  
AE = bypass for AE  
*Spin is relative to initial LSP traffic flow*



# MPLS 1:1 Detours - Label Stacks

## Failure state, Bidirectional LSP

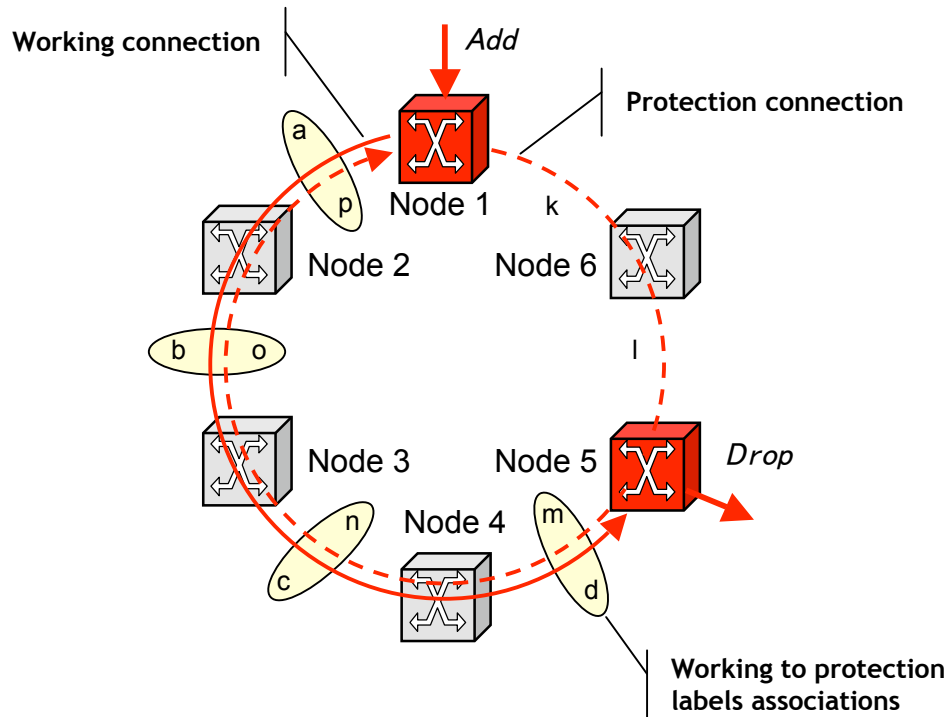
**AE** = Clockwise ring  
**EA** = Anticlockwise ring  
**EA** = bypass for **AE**  
**AE** = bypass for **EA**  
*Spin is relative to initial LSP traffic flow*



# Open questions on MPLS Facility bypass/detours

- **No showstoppers** but, to be solved in the design phase
  - *Loop avoidance*
  - *Implementation of bi-direction switching*
  - *Implementation of manual switching/operator requests*
  - *Implementation of switching priorities*
    - *Faults conditions, operator commands*
  - *Node configuration so that it is aware of the ring*
  - *Multiple failures*
    - *Ring segmentation*
  - *p2mp LSPs*

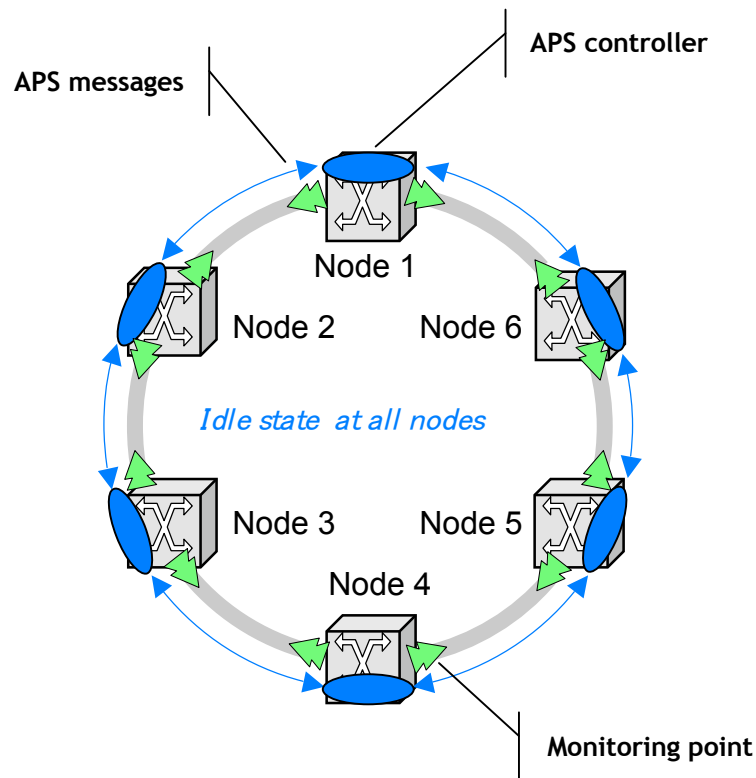
# Review: TM-SPRing labels allocation



Labels allocation and association

Wk Labels	Pr Labels
a	$p \leftrightarrow a$
b	$o \leftrightarrow b$
c	$n \leftrightarrow c$
d	$m \leftrightarrow d$
	$l \leftrightarrow \text{nil}$
	$k \leftrightarrow \text{nil}$

# Review: TM-SPRing OAM monitoring and APS messages



## ■ Monitoring:

- Each section (span) in the ring is monitored by sending CV OAM with periodicity of 3.3ms
- Span failures are detected as absence of 3 consecutive CV frames

## ■ APS:

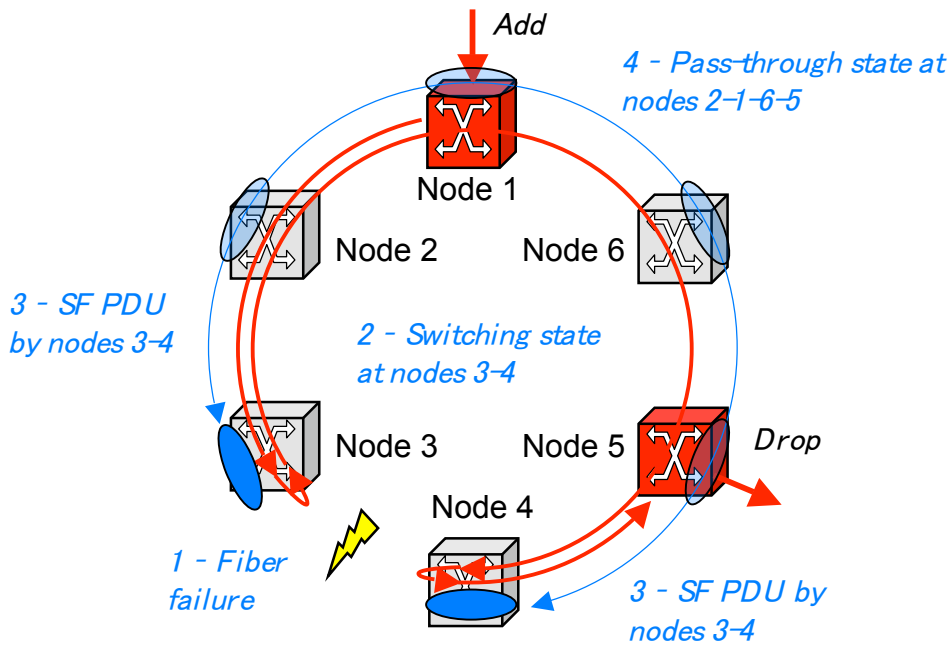
- Each node has an APS controller that sends and receives APS PDUs using an ACH
- In normal state APS controller generates NR (no request) PDUs to its neighbours in both directions
- When there is no failure each node in the ring is in the Idle state i.e. frames are not forwarded on the protection LSP

When there is no failure in the ring:

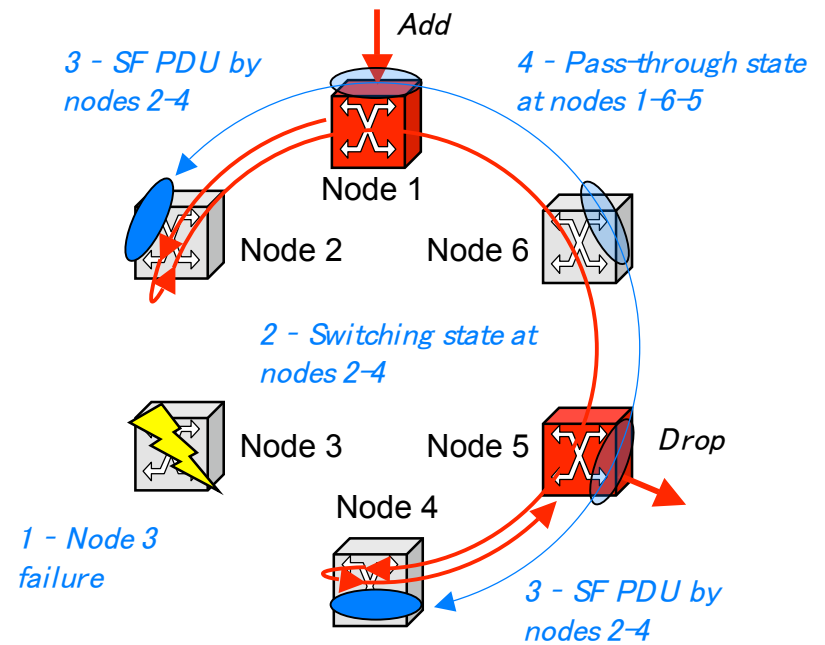
- All nodes are in the idle state
- All nodes generate and terminate APS NR PDUs to their neighbours

# TM-SPRing point-to-point example

Fiber failure



Node failure

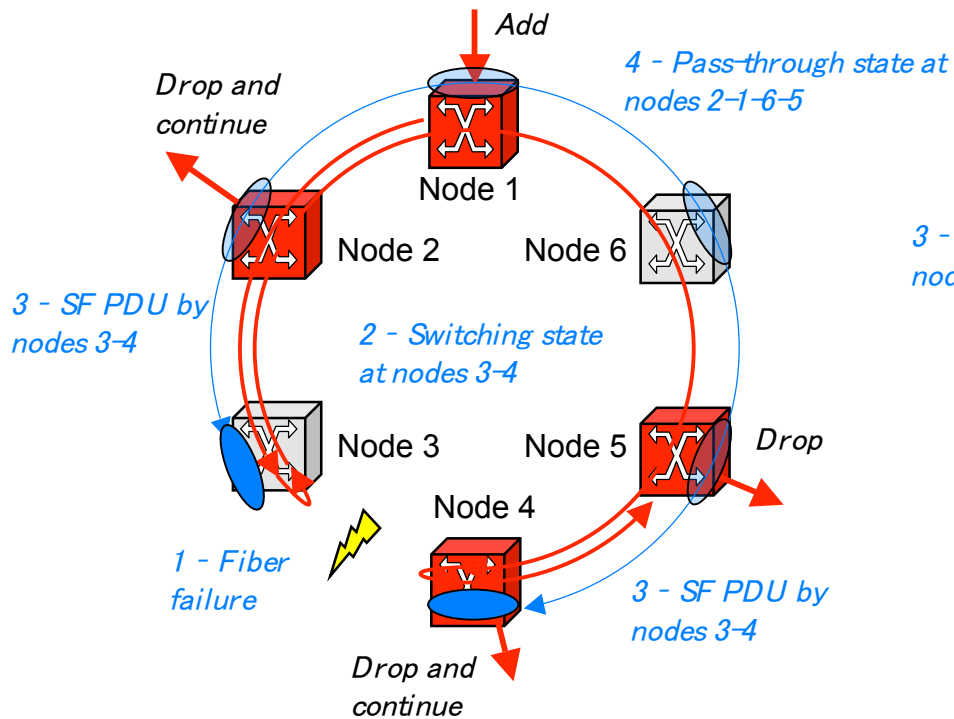


When failure occurs:

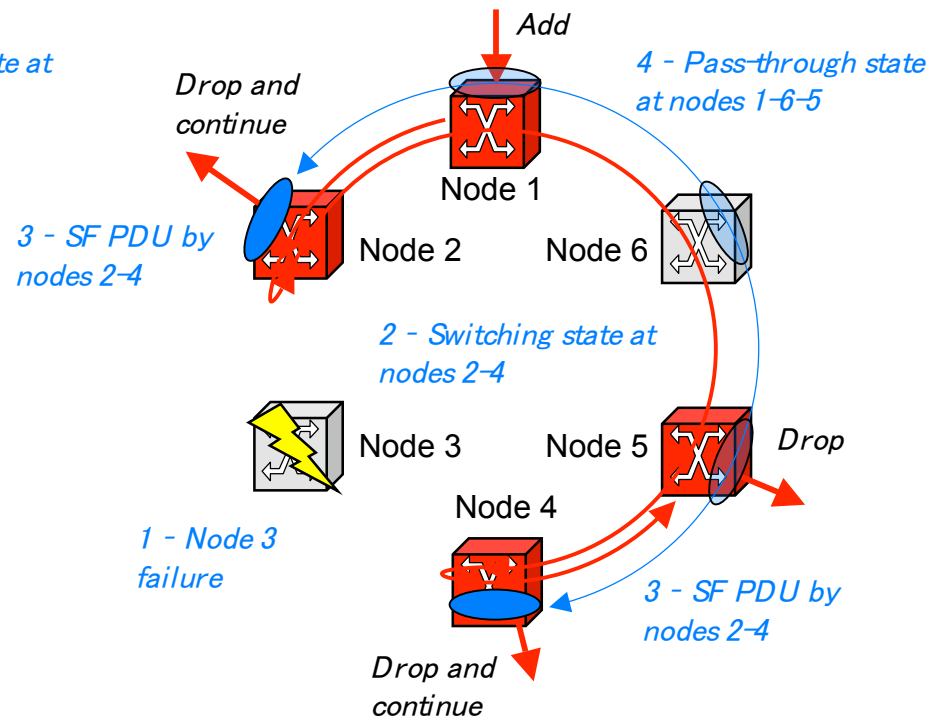
- The nodes adjacent to the failure enter the switching state and sends APS SF PDUs to neighbors
- When the other nodes in the ring receive the SF PDU they enter pass-through state (i.e. allow forwarding on the protection LSP) and forward the APS PDUs without modification

# TM-SPRing point-to-multipoint example

Fiber failure



Node failure

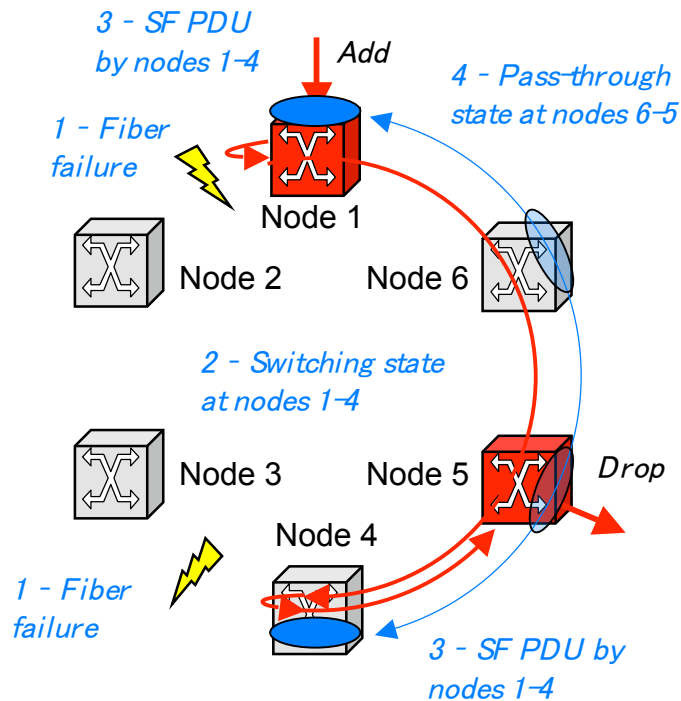


The same mechanism:

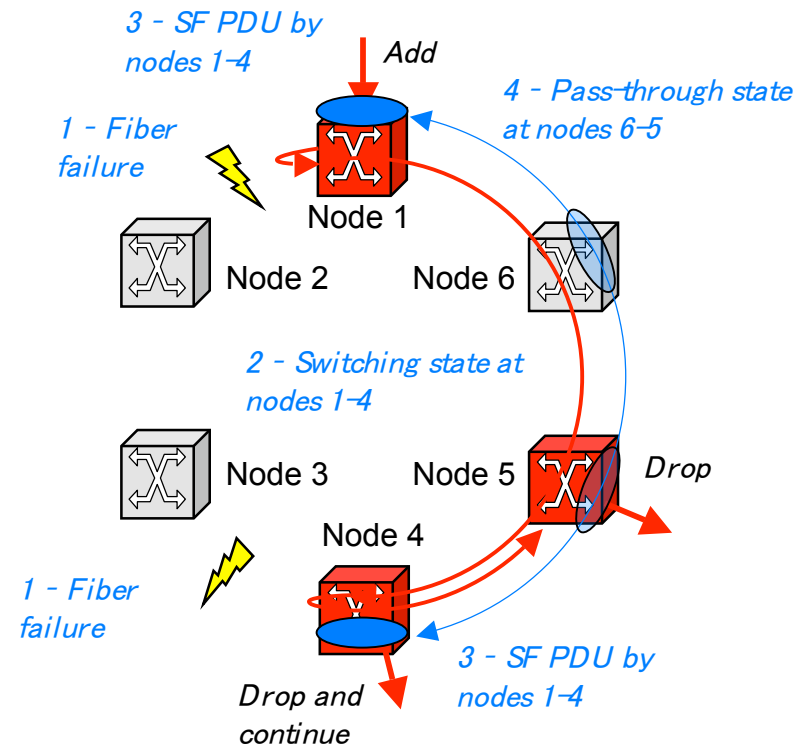
- For p-t-p and p-t-mp connections
- For fiber or node failure
- For single or multiple failures

# TM-SPRing multiple failures example

## Multiple failures, p-t-p



## Multiple failures, p-t-mp



The same mechanism with a single protection connection restores all traffic possible:

- For p-t-p and p-t-mp connections
- For fiber or node failure
- For single or multiple failures



# Network Management

# Advice

- Network Management sub team has not found any issues that prevent the creation of an MPLS transport profile
- Therefore option 1 can be selected

**No Showstoppers found**

# Conclusions - I

- Need to be able to provision and manage a LSP or PW across a network where some segments are managed by IETF (e.g. netconf) and other segments that are managed by ITU/TMF (XML/CORBA) interfaces.
  - LSP establishment
    - MPLS management in the IETF already supports the ability to independently setup LSP segments (using different tools) to create a concatenated (end to end) LSP
  - LSP maintenance
    - It is possible to run maintenance on an LSP independent of the mechanism used to establish the LSP
  - The ITU/TMF interface supports the management of multiple technologies
    - Management of MPLS-TP needs to be added to these multi technology interfaces
  
- No need to explicitly support the case of a single NE that offers both the IETF and ITU/TMF interface
  - This is a NE implementation issue

# Conclusions - 2

- Network Management (NM) requirements
  - Configuration
    - No issues
  - Fault, PM
    - If the OAM can provide the measurement primitives then no reason that NM cannot report them
    - Need to allow each operator to determine the performance of the segment (plus end to end).
  - Accounting
    - Limited functionality – e.g. reporting of unavailable time, providing PM data
  - Security (of the management interface)
    - Not specific to MPLS-TP networks
    - Dependent on:
      - Management protocol
      - Management application
      - Bearer for the management traffic
    - Security implementation is per network segment

# Management – Background IETF

- IETF architecture is layered and the functionality is allocated in separate processes, e.g.:
  - Performance management
    - Netflow/IPfix
      - Sample packets with a defined label – allows inspection of contents
    - SNMP MIBs (e.g. packet counts on LSPs, Octets on an LSP, Queue drops, CRC errors from lower layers – LSP not identified)
  - Fault management
    - SNMP traps, informs, BFD and syslog
  - Configuration management
    - Netconf, SNMP
  - Security
    - IPsec, tls, eap, Radius etc
  - Accounting
    - TACACS, netflow, ippm, ppmi
- IETF doesn't use TMF style CORBA/XML interfaces

# Management – Background ITU

- TMF/ITU approach
  - Provides both a NE and Network level interface to the OSS
  - Protocol neutral model (in UML), requirements and use cases
  - Protocol specific interface definitions

# ITU-T PM objectives

- PM Requirements for a MPLS-TP LSP/PW
- Same measurements and processing as Ethernet
  - Connectivity defects present in a 1-second period
  - number of lost (circuit/packet) frames in a 1-second period
  - near-end and far-end (severely) errored second
  - 10 seconds being severely errored/not severely errored to enter/exit unavailable time (UAT)
  - 15min and 24hr PM parameter reporting
- To define how LM (loss measurement) and DM (delay measurement) information, as defined in Y.1731 & draft G.8114, is registered in 15min/24hr bins (G.7710)

Dependent on OAM providing the primitives to make these measurements

# Summary



# Summary

To date we have found no showstoppers and everyone is in agreement that we have a viable solution

## *Recommend Option 1*

It is technically feasible that the existing MPLS architecture can be extended to meet the requirements of a Transport profile

The architecture allows for a single OAM technology for LSPs, PWE and a deeply nested network

From probing various SGs, WGs it appears that label 14 has had wide enough implementation and deployment that the solution may have to use a different reserved label (e.g. Label 13)

Extensions to Label 14 should cease

This architecture also appears to subsume Y.1711 since the requirements can be met by the mechanism proposed here

# Some open discussion points

1. One way delay measurement techniques need to be defined although not required for initial design
  - Decision: architecture can not preclude a solution for one-way delay measurement
  - No issues w/ 2-way delay
2. Measurement of packet loss to support PMs and detection of degraded performance need to be defined
  - One approach is to encapsulate the appropriate Y.1731 pdus in an ACH

**The End**