
cisco Live!
Let's go
#CiscoLive

The bridge to possible

High Value Wavelength / Private Line Services

Understanding the Customer and Provider Perspective

Christian Schmutzer, Distinguished Engineer

BRKOPT-1005

Cisco Webex App

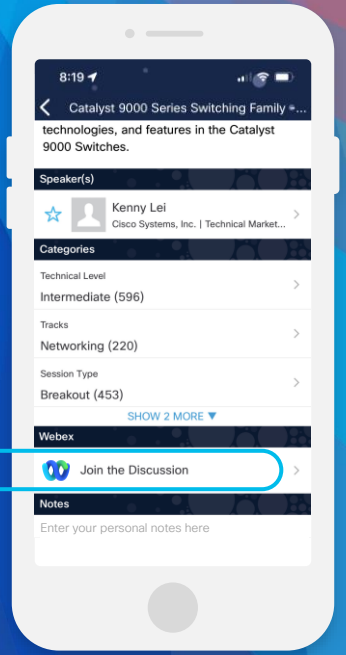
Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

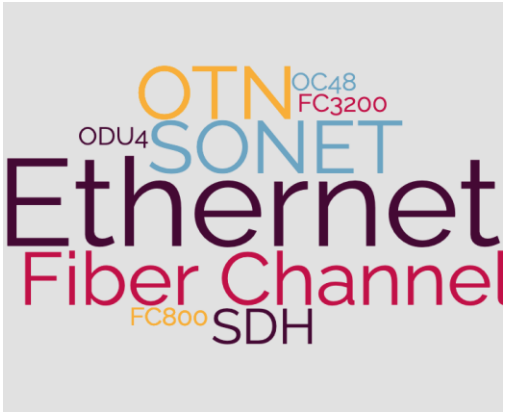
1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 9, 2023.



https://ciscolive.ciscoevents.com/ciscolivebot/#BRKOPT-1005

Wavelengths (Private Lines) are Premium Services

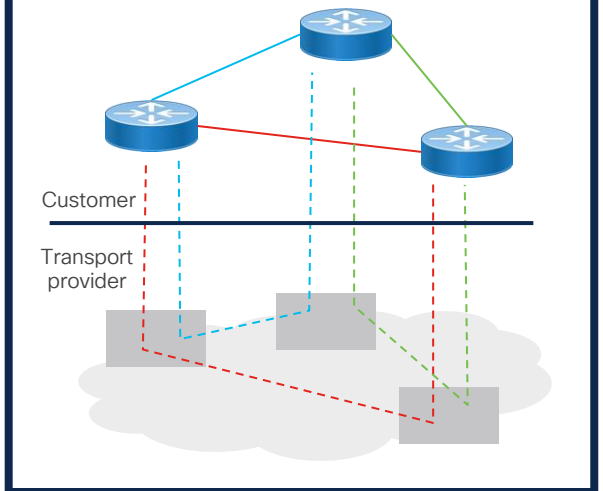


High bandwidth

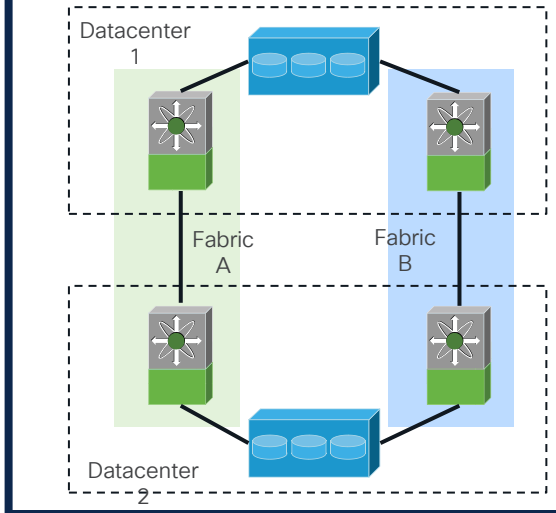
High availability

Diverse payloads

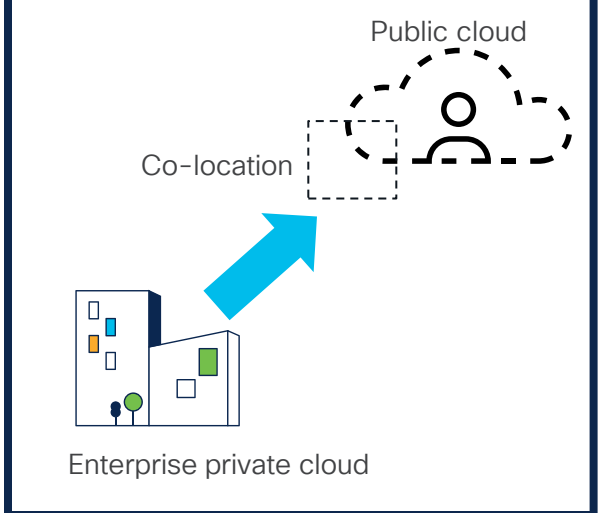
When Only the Very Best will Do...



Dedicated WAN pipes for your topology



Fibre Channel to build your fault tolerant SAN



Cloud connect

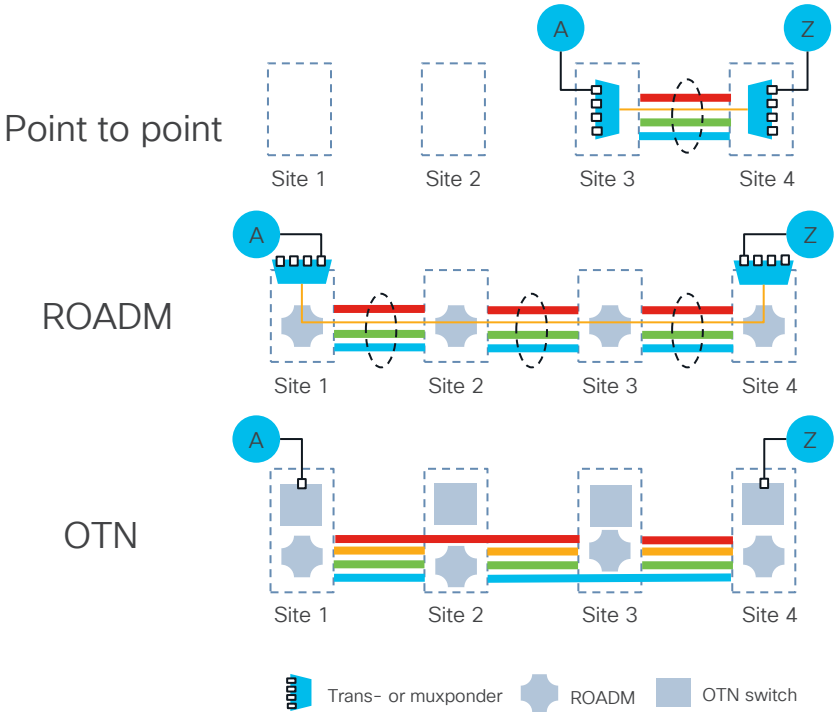
SAN ... Storage Area Network

Wavelengths, TDM have been the Gold Standard

Dedicated resources

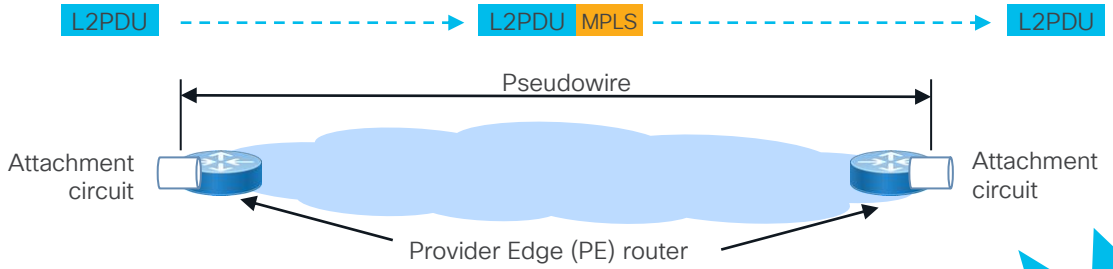
Embedded OAM

Payload agnostic



Meanwhile the Reality in Packet Networks

Brilliant Idea back in 2001 (RFC3985 and RFC3915) → Transport L2 payload over packet networks using a “pseudo” wire service



PROs

- One network for L2 and L3 services
- Efficient use of network bandwidth
- Simple and scalable

CONs & CONCERNS

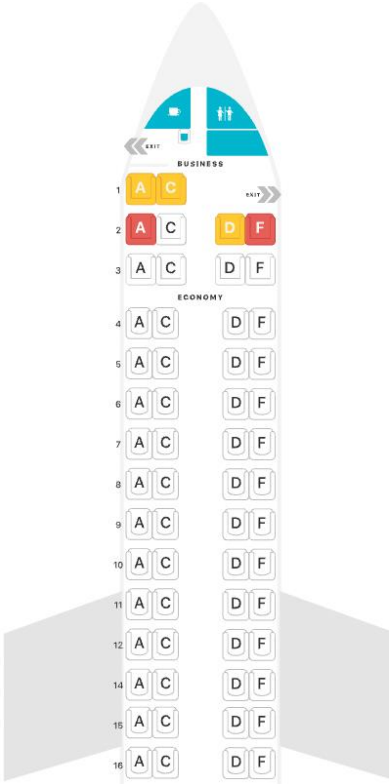
- Ethernet payloads only
- OAM
- Bandwidth commitment
- Load-balancing

Conclusion:

Good for many services but NOT for premium private line use cases

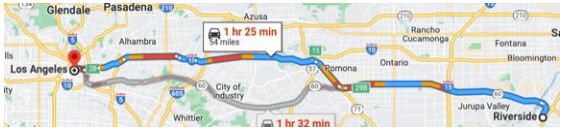
Packet is Like Driving in LA

- Always reserved for your team
- Runs on a schedule
- Put whoever you want in those seats

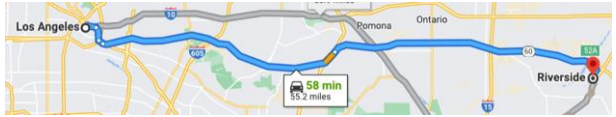


Source: Seatguru

Shared Bandwidth (rush hour)



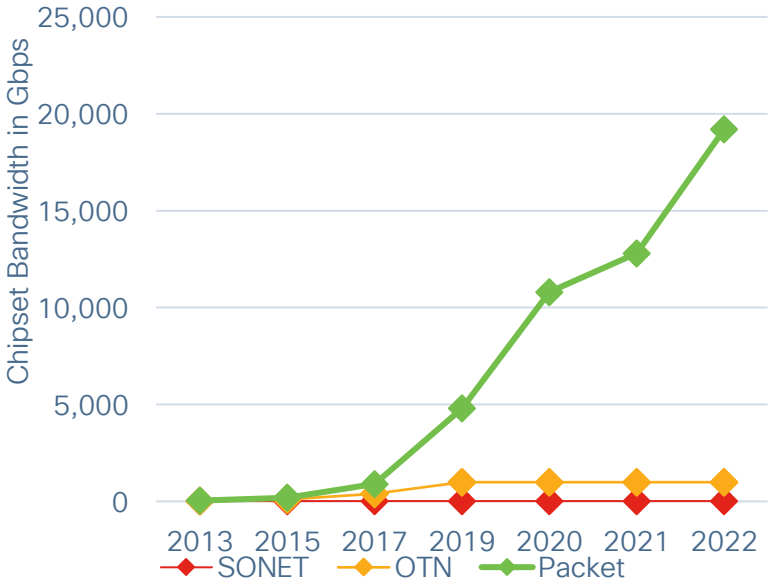
Asymmetric routes & latency



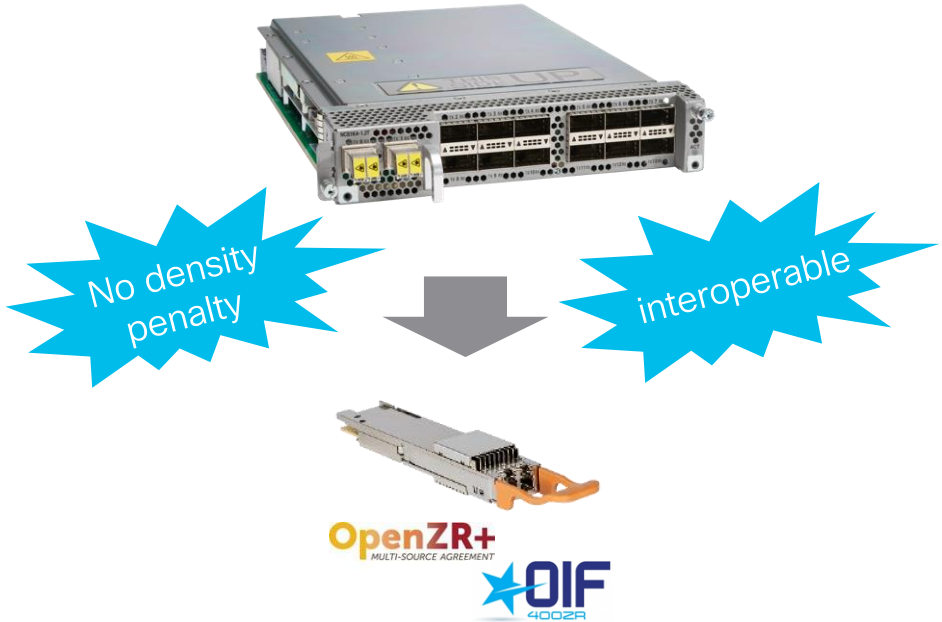
Source: google maps

Massive Shifts in Economics of Routing

Advances in silicon

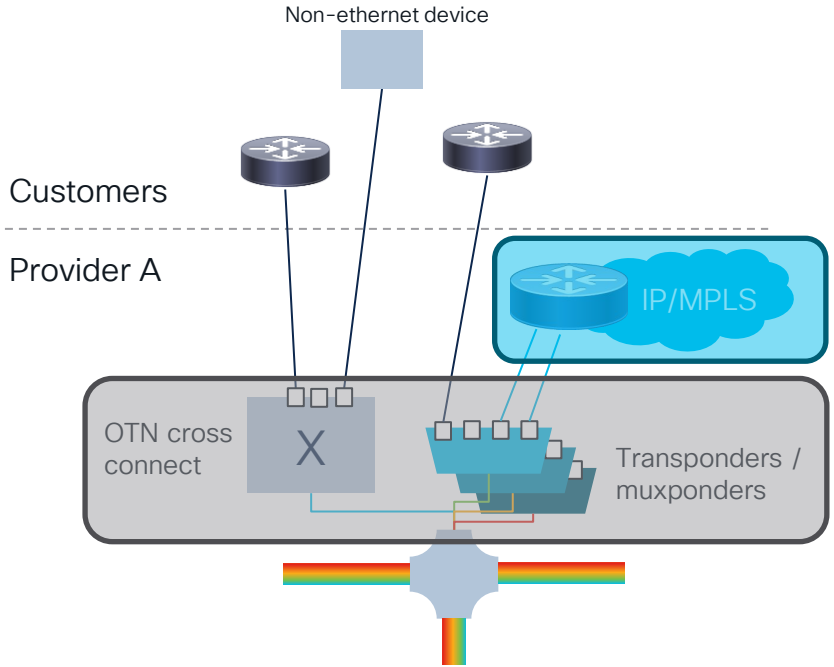


Coherent transceivers eliminate the need for DWDM transponders

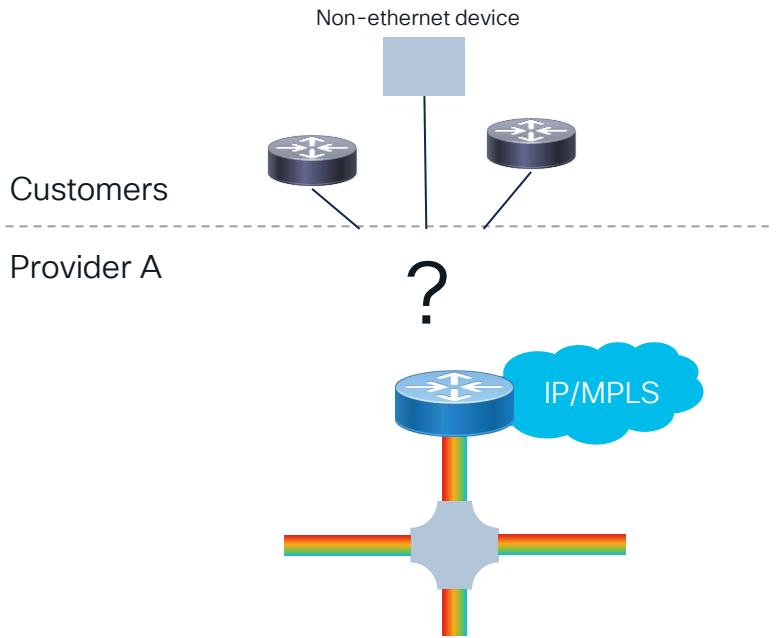


And there are challenges today

Operating 2 networks is hard & expensive



What if I only have an IP network?



Agenda

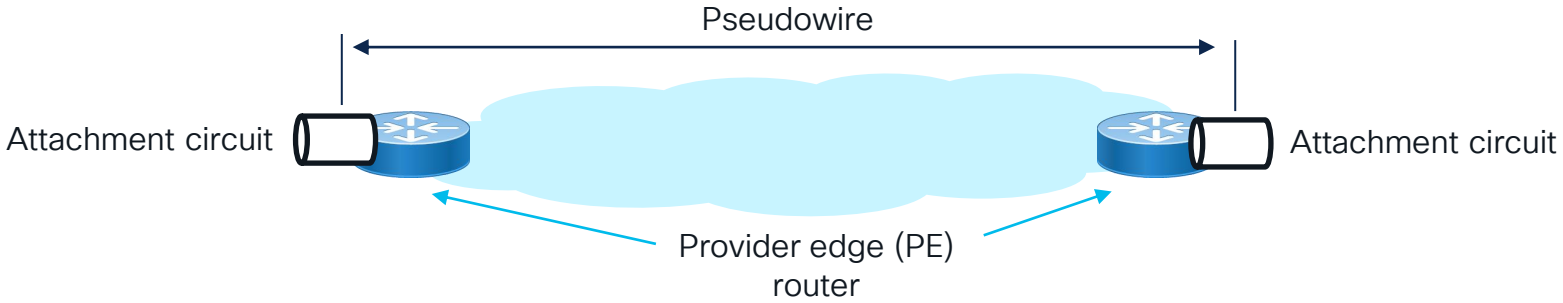
It is time for a Change, but..

- Bit-transparency & Non-etherent payloads
- Clocking
- Inband OAM
- MTU and overhead
- Co-routed, bidirectional paths
- Dedicated bandwidth

+ some extra thoughts

Non-Ethernet Payloads

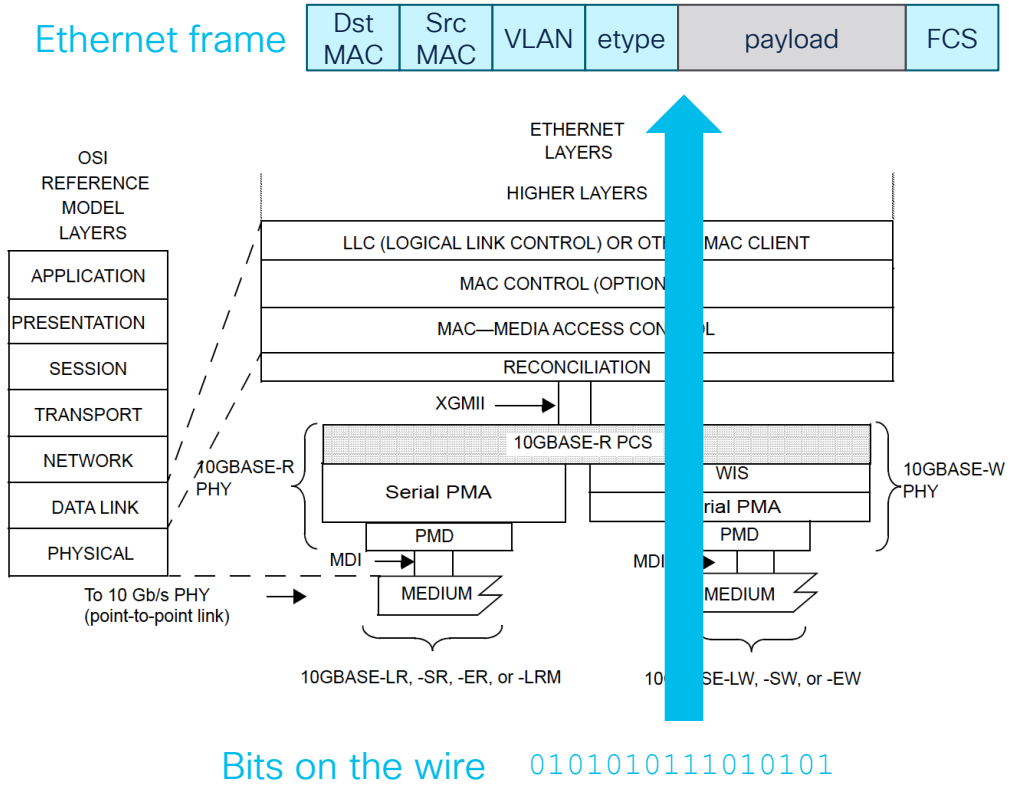
Virtual Private Wire Service (Pseudowire)



- Emulation of a “Telecommunications service”
- Initially defined by IETF PWE3 working group via RFC 3985
  - T-LDP for pseudowire signaling
  - Frame Relay, ATM, Ethernet and TDM
- Modernized by IETF BESS working group via RFC 8214
  - MP-BGP for pseudowire signaling → EVPN-VPWS

Routers do look for Frames on an Ethernet Port

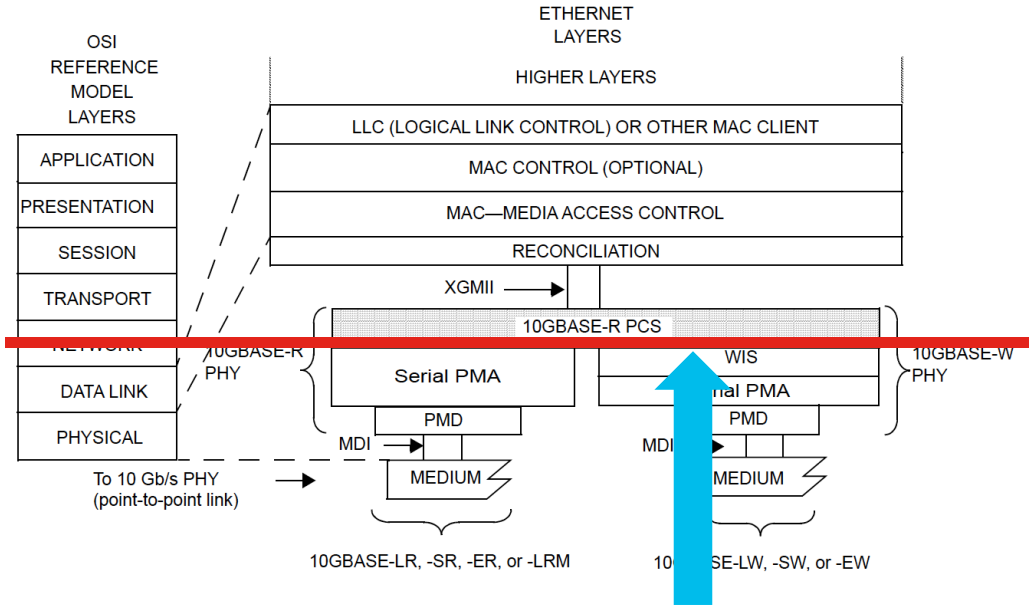
- Incoming bit stream is de-serialized
- Clock is recovered
- Align to PCS stream
  - 8B10B, 64B/66B or 256B/257B
- Identify MAC frames



Bits on the wire `0101010111010101`

What if we (want to) stop much earlier?

- We just process bits
- We are unaware of ethernet frames and their boundaries
- We can carry what we want
- (However) we need special hardware to do that



Bits on the wire `0101010111010101`

Solving the Control Protocol Challenge … finally

- Most products can tunnel/process spanning tree, LACP and OAM properly
- But there is always something new like
  - MACSEC uses EAPOL (IEEE 802.1X-2010) for key exchange
  - EtherType 0x888e and destination MAC 01-80-C2-00-00-03
  - MEF 45.1 is only “recommending” to pass those frames for ethernet private lines (EPLs)

Protocol Type	Protocol Identifier	L2CP Destination Address	L2CP Action
STP[3]/RSTP[4]/MSTP[4]	LLC Address: 0x42	01-80-C2-00-00-00	Pass
E-LMI[15]	EtherType: 0x88EE	01-80-C2-00-00-07	Pass[6]
LLDP[1]	EtherType: 0x88CC	01-80-C2-00-00-0E	Pass
PTP Peer Delay[8]	EtherType: 0x88F7	01-80-C2-00-00-0E	Pass
GARP[4]/MRP[4] Reserved Address	any	01-80-C2-00-00-20 through 01-80-C2-00-00-2F	Pass

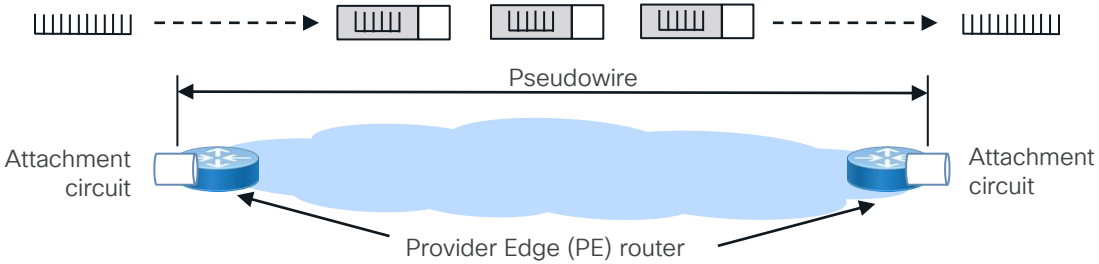
Table 9 – EPL Option 2 L2CP Processing Requirements

Protocol Type	Protocol Identifier	L2CP Destination Address	L2CP Action
PAUSE[7]	Etherype: 0x8808 Subtype: 0x0001	01-80-C2-00-00-01	Discard
LACP/LAMP[2]	EtherType: 0x8809 Subtypes: 0x01, 0x02	01-80-C2-00-00-02	Pass
Link OAM[7]	EtherType: 0x8809 Subtype: 0x03	01-80-C2-00-00-02	Pass
Port Authentication[6]	EtherType: 0x888E	01-80-C2-00-00-03	Pass
ESMC[11]	EtherType: 0x8809 Subtype: 0x0A	01-80-C2-00-00-02	Pass[7]

Table 10 – EPL Option 2 L2CP Processing Recommendations

Private Line Emulation (PLE) = Bits over Packets

- Encapsulation of bit-stream into packets
  - Extend idea from RFC 4553 (SATOP) beyond T1/E1 to Ethernet, Fibre Channel, OCn/STMn, anything
- Standardization at IETF underway
  - draft-schmutzer-pals-ple




```
0         1         2         3
01234567890123456789012345678901
+-------------------------------+
|     PSN and VPWS Demux        |
|          (MPLS/SRv6)          |
|                               |
+-------------------------------+
|0000|L|R|RSV|LEN| sequence nr |
+-------------------------------+
|           RTP Header          |
+-------------------------------+
|          Bit Stream           |
|           Payload             |
|                               |
+-------------------------------+
```

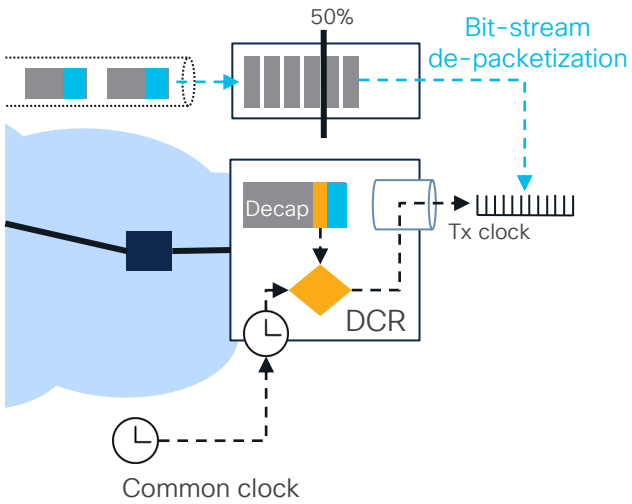
Figure 3: PLE Encapsulation Layer

Source: draft-schmutzer-pals-ple

PSN ... Packet Switched Network

Converting Packets back to Bits

- Store packets in a de-jitter buffer
  - Packet Delay Variation (PDV)
  - Packet reordering
- Bit from the packet payloads are transmitted using the recovered clock



PLE payload types

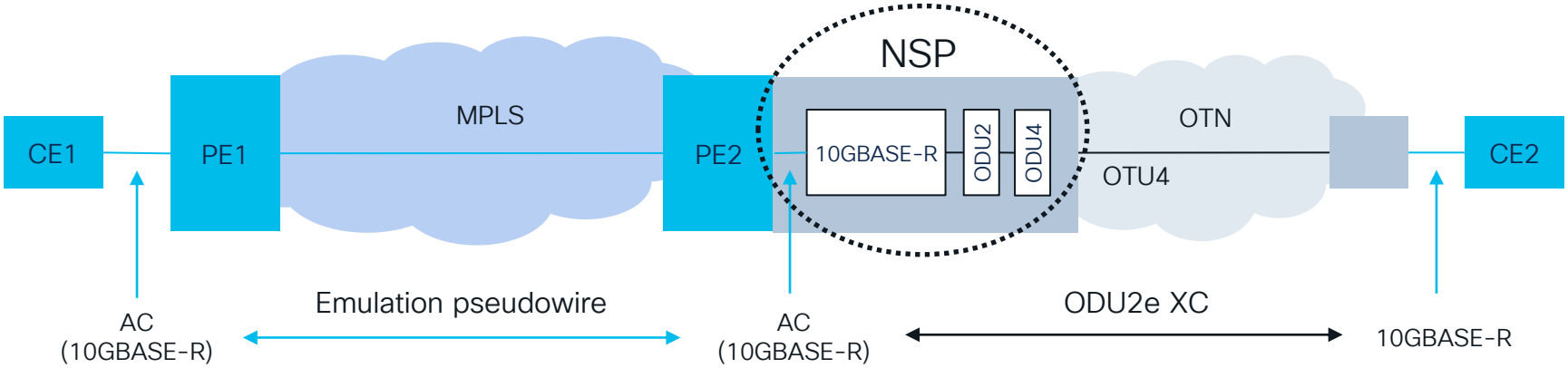
1Gbps

+100Gbps

SONET/SDH	OTN	Ethernet	Fibre Channel
	ODU0	1GE	1GFC
OC48/STM16	ODU1		2GFC
			4GFC
			8GFC
OC192/STM64	ODU2/ODU2e	10GE	
			16GFC
			32GFC
			64GFC
	ODU4	100GE	
			128GFC
		200GE	
		400GE	

Supported today

Emulation is independent of the physical Interface

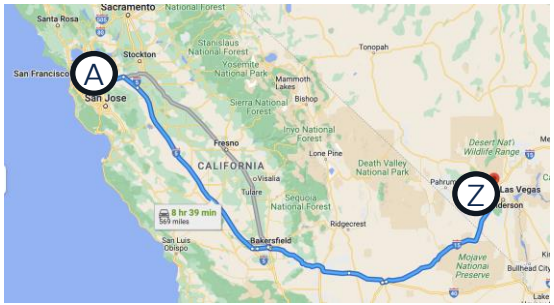


- Similar to SAToP (RFC 4553) the structure agnostic attachment circuit is independent from the physical port type
- It can either be a physical ethernet port or ODU2e mapped “logical” 10GE port inside a 100Gbps OTU4 interface extracted by the native service processing (NSP) function

It is Not about Latency, but about Jitter !

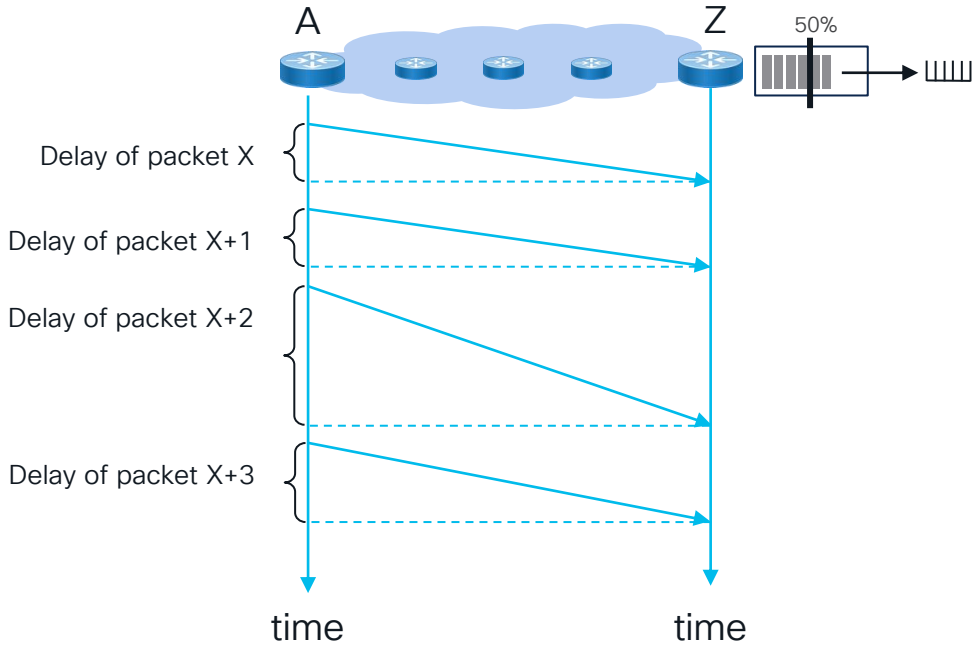
Router transfer latency

Optical propagation delay



Source: google maps

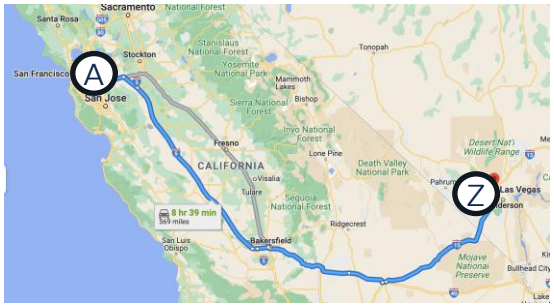
Network Packet Delay Variation (PDV)



Lets quantify this a bit

Router transfer latency

Optical propagation delay



Source: google maps

~5 µs / km

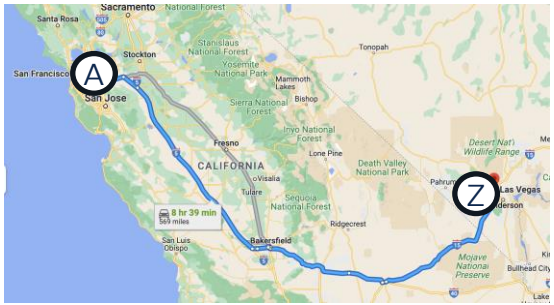
Lets quantify this a bit

Router transfer latency

Depening on NPU type and router architecture (number of NPUs passed)

~3-10 µs

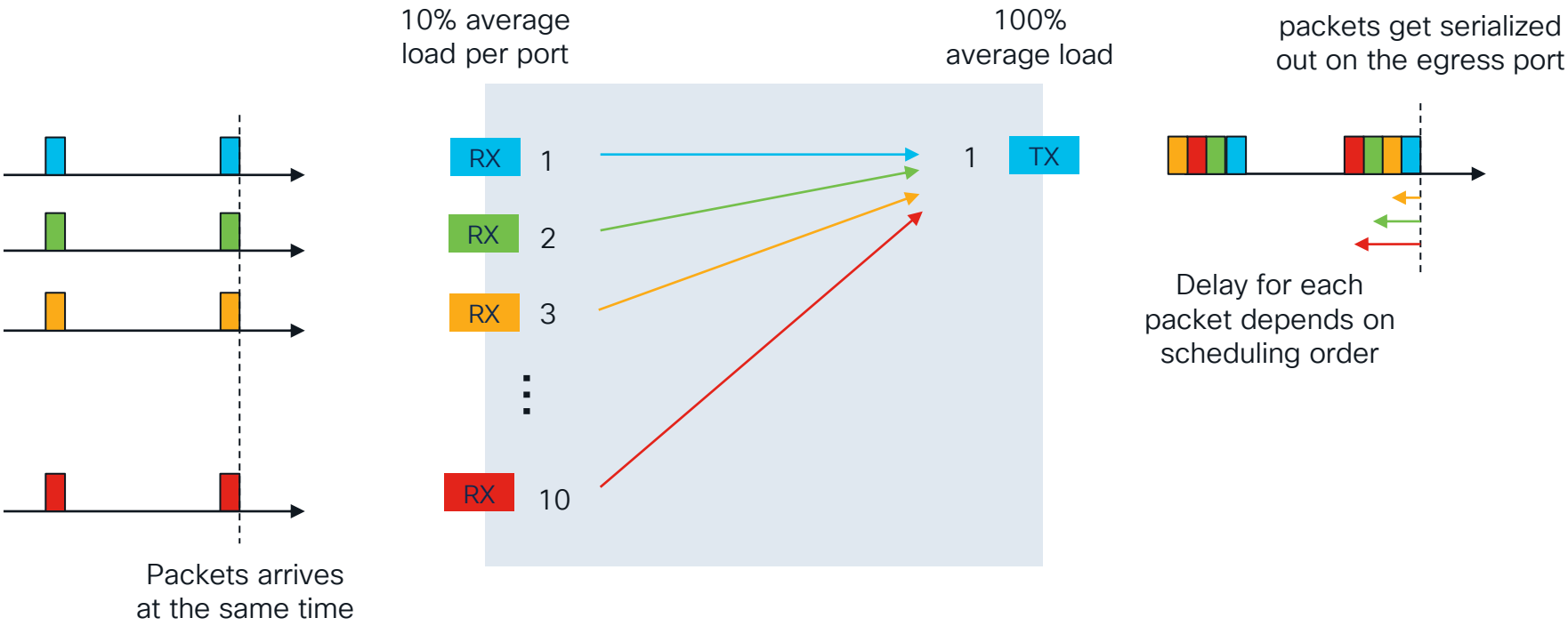
Optical propagation delay



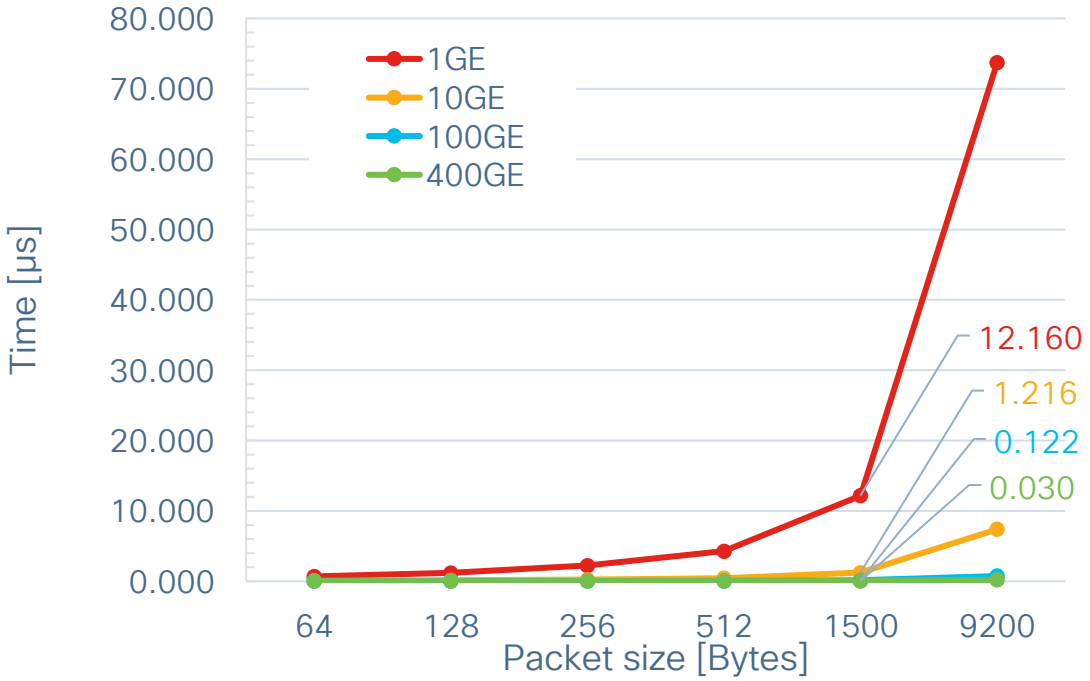
Source: google maps

~5 µs / km

No Oversubscription, buffering can still happen



Growing Interface Speed does help us



Sub-microseconds for 100GE and 400GE!

Lets quantify this a bit

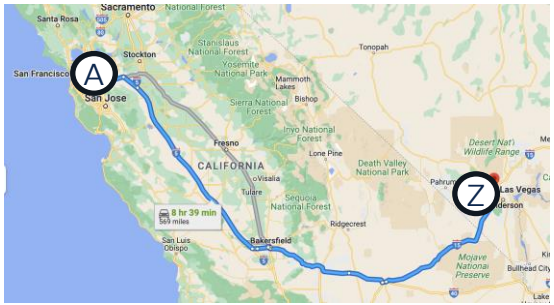
Router transfer latency

Depening on NPU type and router architecture (number of NPUs passed)

~3-10 µs

Max ±10 µs variance

Optical propagation delay



Source: google maps

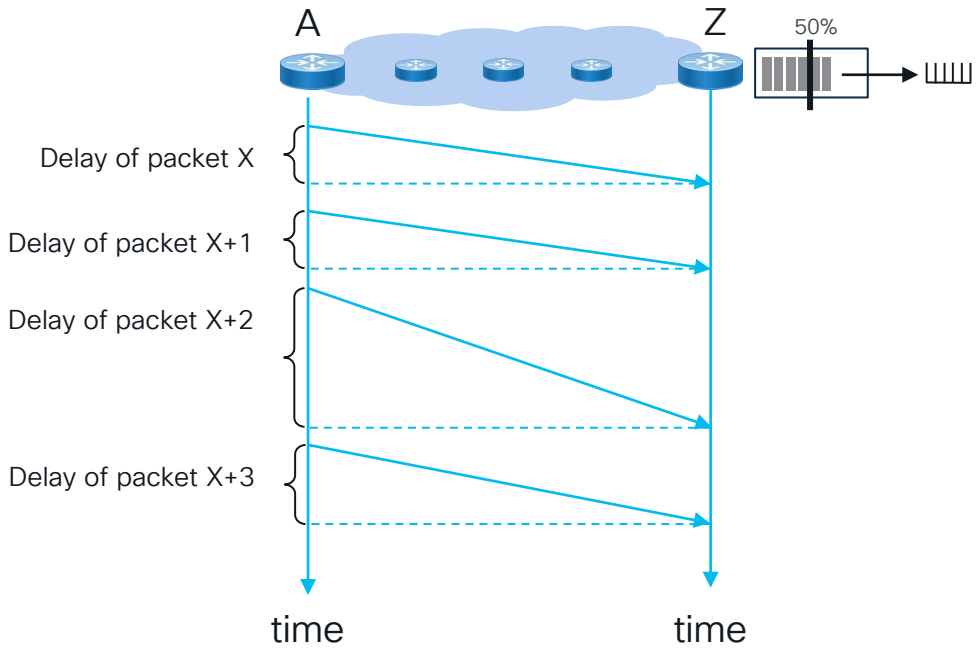
~5 µs / km

How much Buffer for our Network?

- Per router hop jitter ±10 µs
- Maximum 30 hops
- 30 * ±10 = ±300 µs PDV

Requires 100s of µs of dejitter buffer !

Network Packet Delay Variation (PDV)



For your Reference

Configuring a PLE Pseudowire

- Define client interface

```
controller Optics0/0/2/2
 port-mode FC framing cem-packetize rate FC8
```

- Enable L2VPN on the emulation interface

```
interface CEM0/0/2/2
 l2transport
```

- Configure EVPN-VPWS

```
l2vpn
 xconnect group PLE
  p2p ple_fc8
   interface CEM0/0/2/2
   neighbor evpn evi 103 target 16022 source 11022
    pw-class unprotected_circuit
```

```
pw-class unprotected_circuit
  encapsulation mpls
   preferred-path sr-te policy srte_c_1_ep_1.0.0.6
```

Agenda

cisco Live!

It is time for a Change

- Bit-transparency & Non-etherent payloads

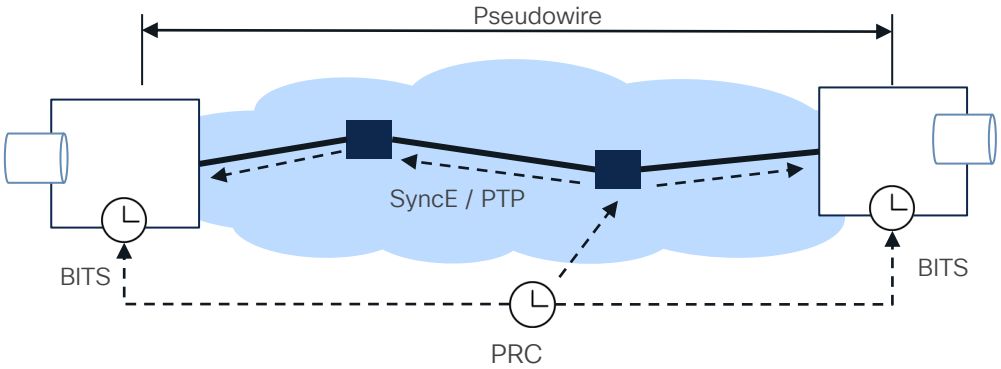
- Clocking
- Inband OAM
- MTU and overhead
- Co-routed, bidirectional paths
- Dedicated bandwidth

+ some extra thoughts

Clocking

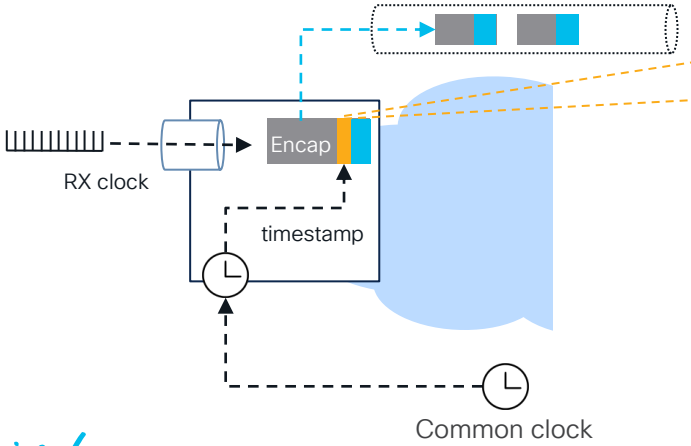
The Foundation is a Common Clock

- The two PLE endpoint routers must have synchronous system clocks
- Can be achieved by SyncE or BITS (or PTP)



Indicating Client Clock to the Egress

- Insert RTP timestamps (counter)
- Counter frequency = common clock
- Counter values = how long it takes to fill a packet (depends on client clock)




```
0          1          2          3
01234567890123456789012345678901
+--------------------------------+
|      PSN and VPWS Demux        |
|           (MPLS/SRv6)          |
|                                |
+--------------------------------+
|0000|L|R|RSV|LEN|  sequence nr  |
+--------------------------------+
|           RTP Header           |
+--------------------------------+
|           Bit Stream           |
|            Payload             |
|                                |
+--------------------------------+
```

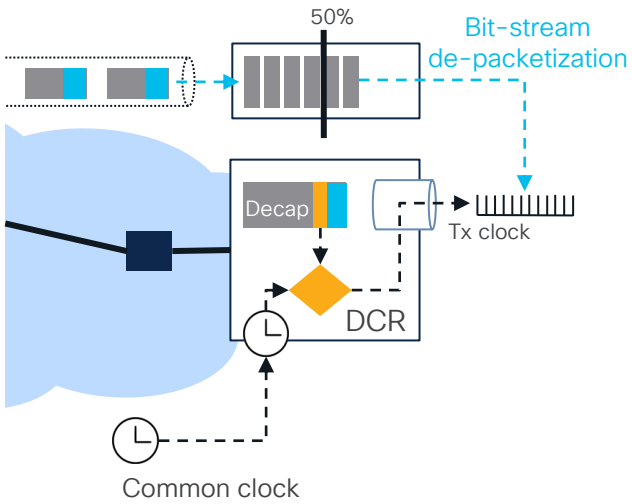
Figure 3: PLE Encapsulation Layer

Source: draft-schmutzer-pals-ple

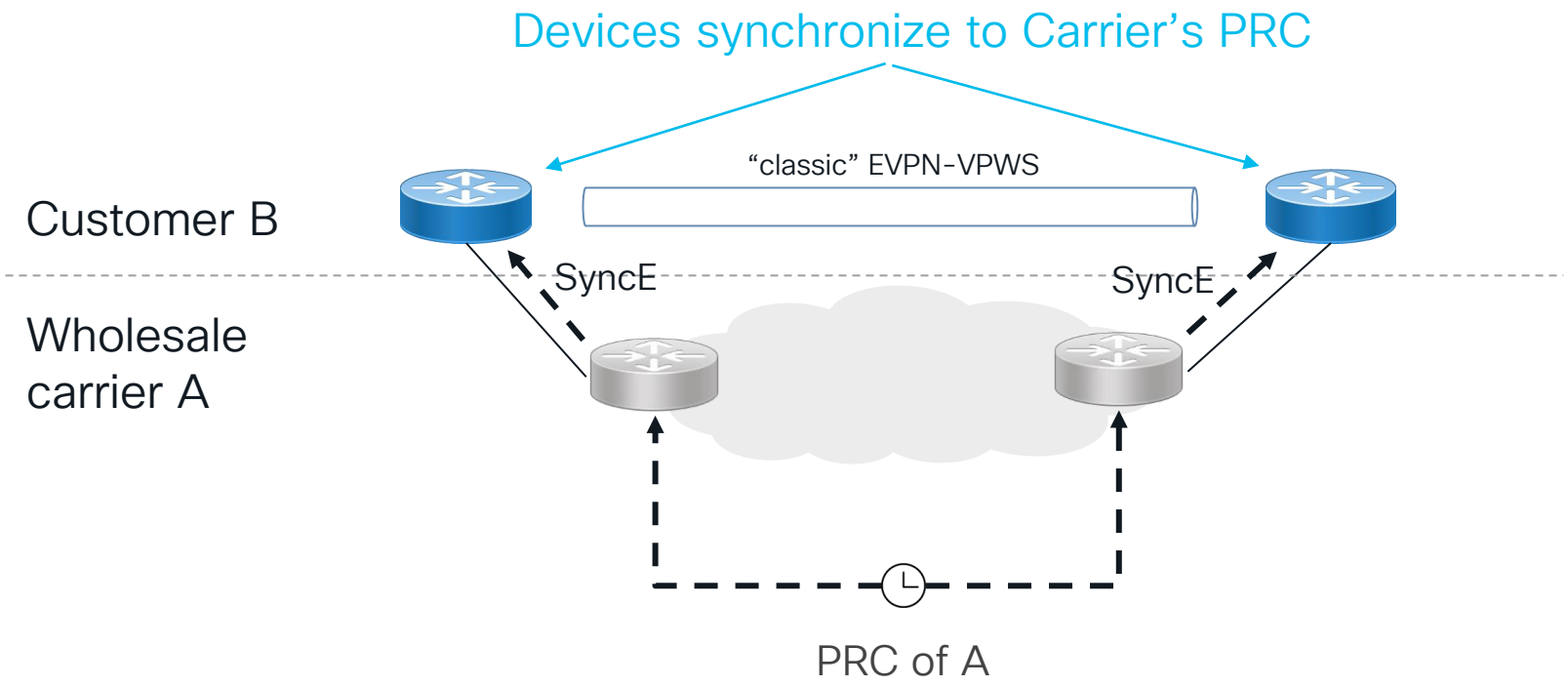
PSN ... Packet Switched Network

Recover the Client Clock on Egress

- Differential clock recovery (DCR)
  - The RTP timestamp indicates the ”difference” between the client and the common clock

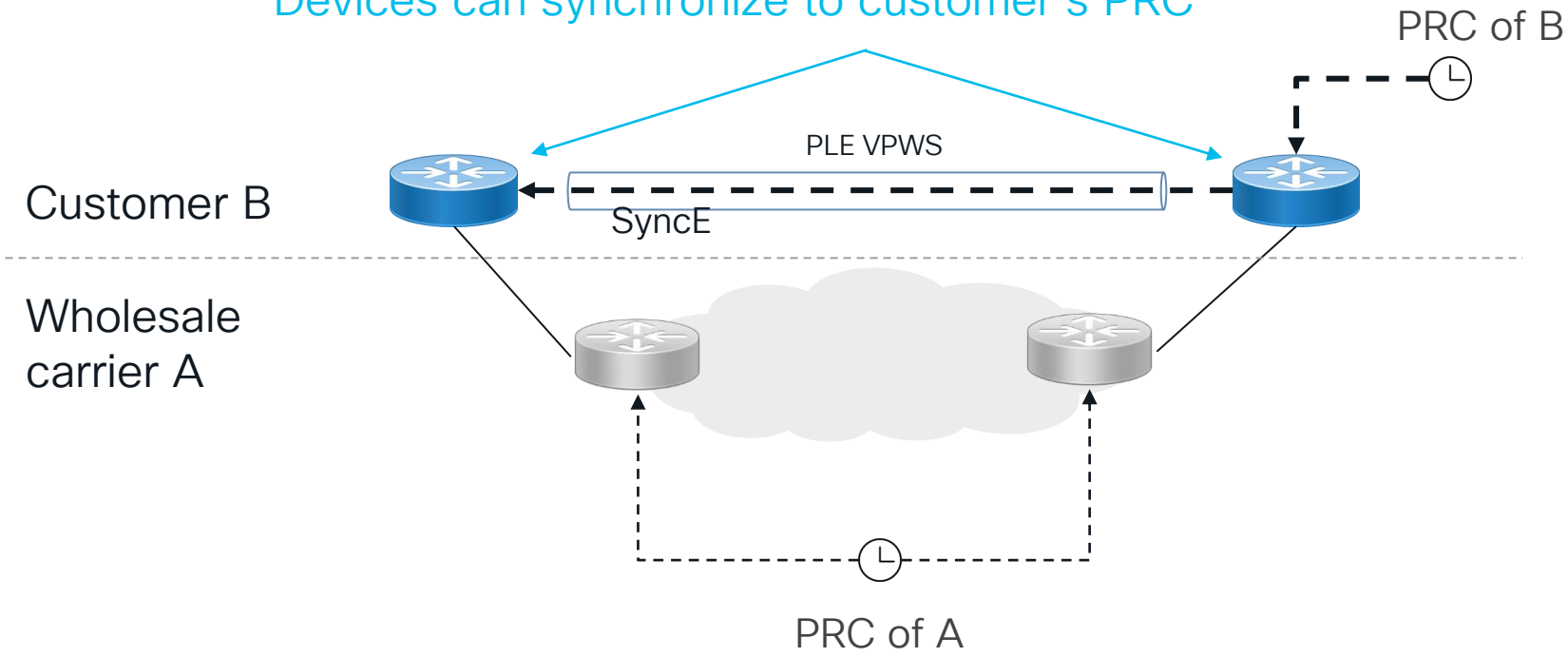


Carrier Ethernet merges SyncE Clock Domains



Carrier Ethernet merges SyncE Clock Domains

Devices can synchronize to customer's PRC



It is time for a Change, but …

- Bit-transparency & Non-etherent payloads

- Clocking 

- Inband OAM
- MTU and overhead
- Co-routed, bidirectional paths
- Dedicated bandwidth

Agenda

It is time for a Change, but..

- Bit-transparency & Non-etherent payloads 

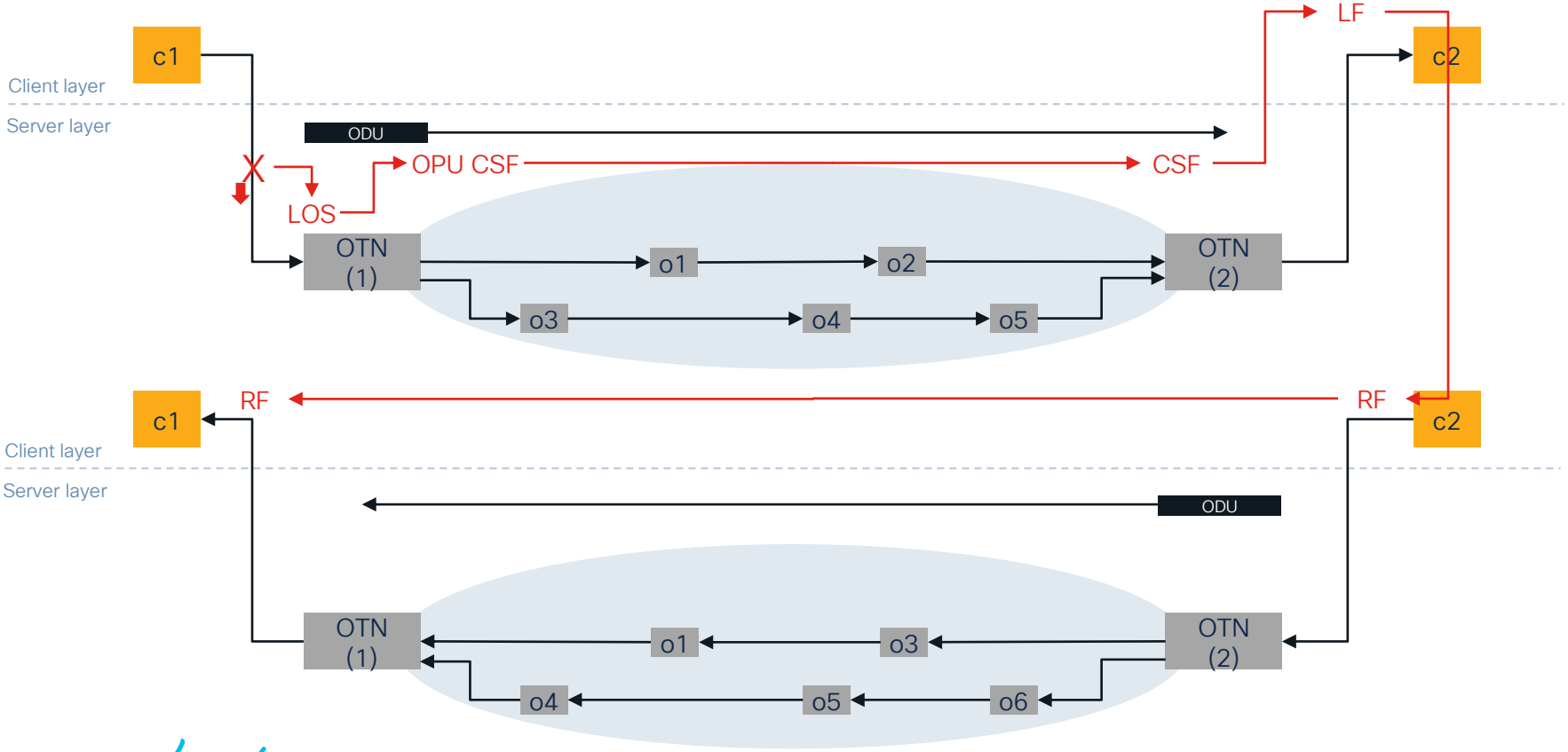
- Clocking 

- Inband OAM
- MTU and overhead
- Co-routed, bidirectional paths
- Dedicated bandwidth

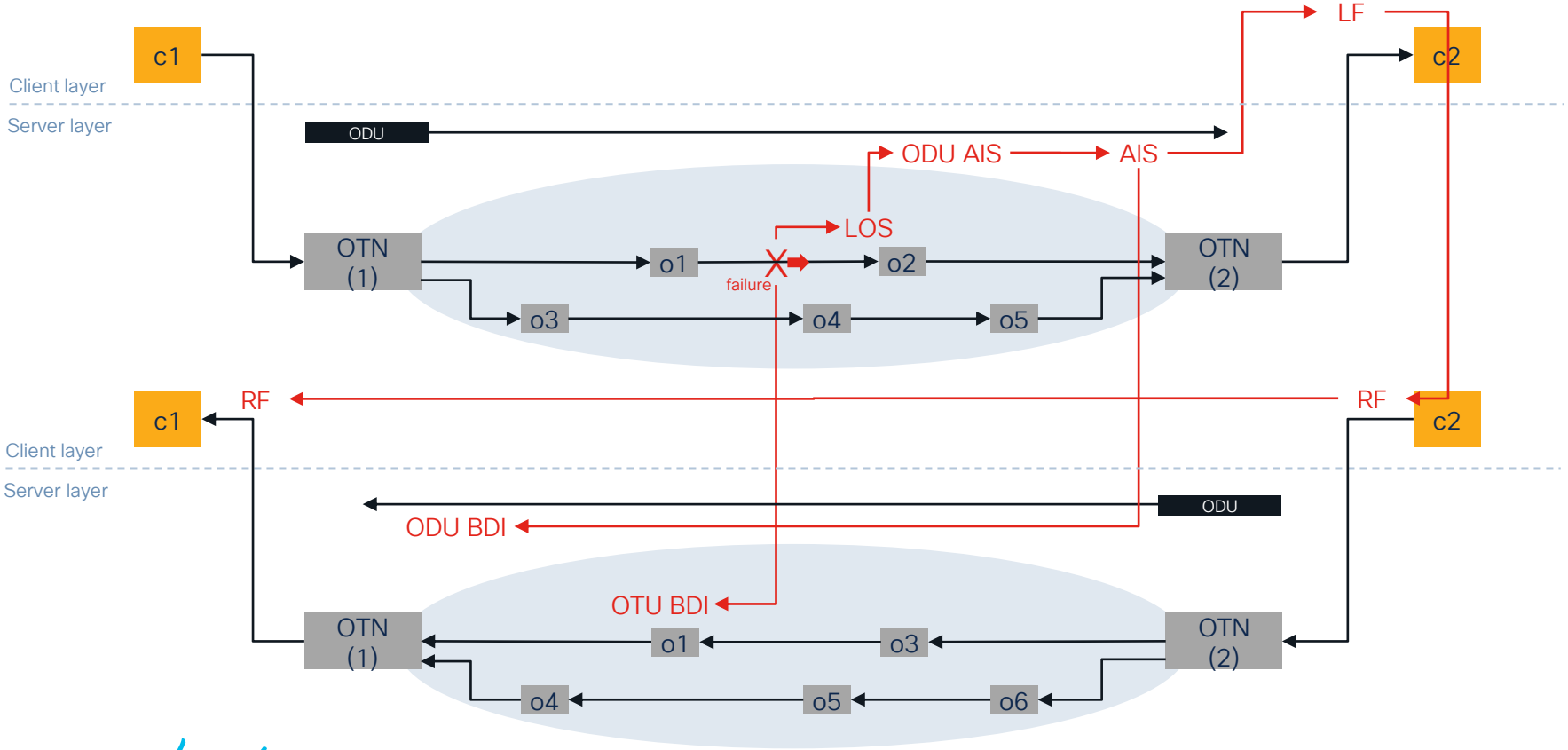
+ some extra thoughts

Inband OAM

What does OTN do in Case of a Client Failure?

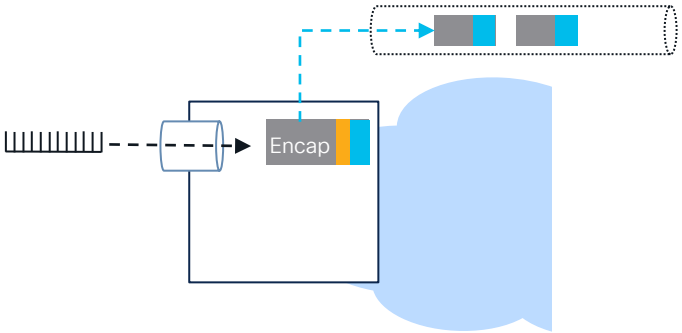


… what in case of a Network Failure?



OAM during Encapsulation (Ingress)

- Embedded OAM via PW control word
  - Client faults → L bit set
  - Rx pseudowire network fault → R bit set




```
0         1         2         3
01234567890123456789012345678901
+-------------------------------+
|     PSN and VPWS Demux        |
|          (MPLS/SRv6)          |
|                               |
+-------------------------------+
|0000|L|R|RSV|LEN| sequence nr  |
+-------------------------------+
|           RTP Header          |
+-------------------------------+
|          Bit Stream           |
|           Payload             |
|                               |
+-------------------------------+
```

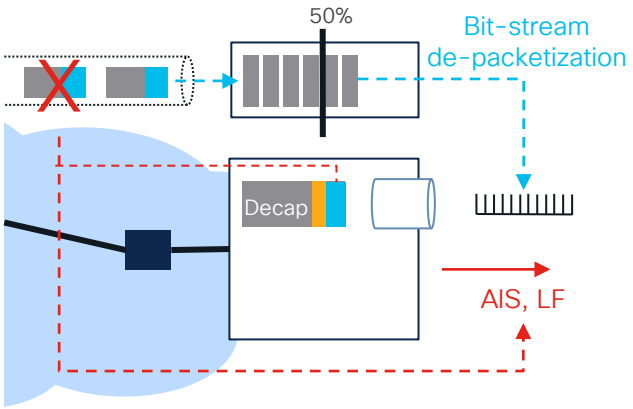
Figure 3: PLE Encapsulation Layer

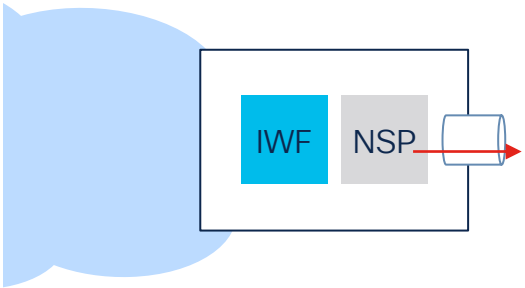
Source: draft-schmutzer-pals-ple

PSN … Packet Switched Network

OAM during Decapsulation (Egress)

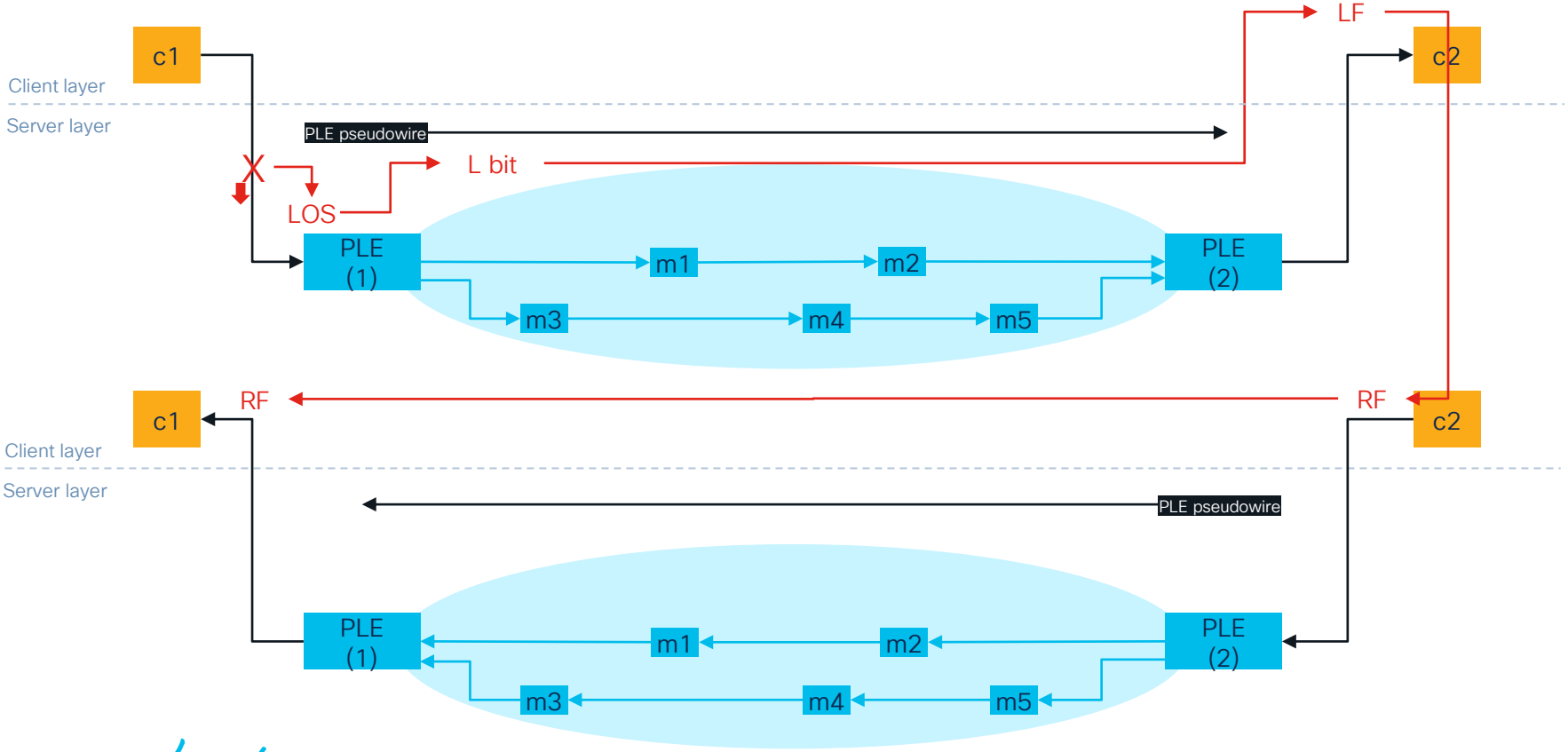
- Client fault indication
  - Too many packets lost
  - If L bit is set in PW Control word
  - Service specific by structure aware NSP
    - see section 4 of draft-schmutzer-pals-ple



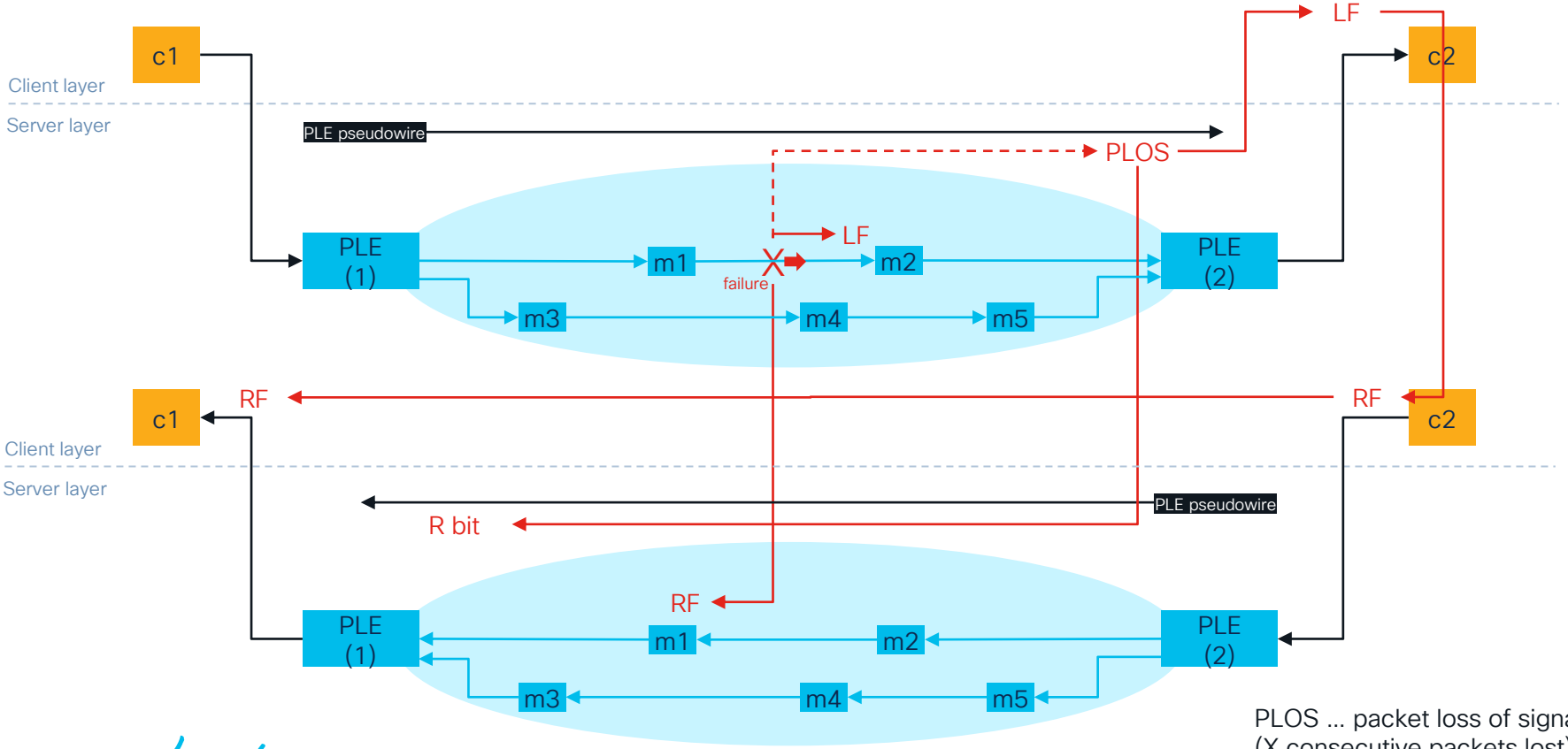


AIS ... Alarm Indication Signal
LF ... Local Fault
NSP ... Native Service Processing

PLE OAM for a Client Failure



PLE OAM for a MPLS Network Failure



PLOS ... packet loss of signal state
(X consecutive packets lost)

For your Reference

Verifying PLE OAM Statistics

```
RP/0/RP0/CPU0:peyto#show controllers cem 0/0/2/2
      .
      .
      .
Detected Alarms                    : None

Statistics Info
---------------
Ingress packets          : 569295050, Ingress packets drop     : 0
Egress packets           : 409658267, Egress packets drop      : 0
Total error              : 0
        Missing packets          : 0, Malformed packets        : 0
        Jitter buffer underrun   : 0, Jitter buffer overrun    : 0
        Misorder drops           : 0
Reordered packets        : 0, Frames fragmented        : 0
Error seconds            : 0, Severely error seconds   : 0
Unavailable seconds      : 0, Failure counts           : 0

Generated L bits         : 409658183, Received  L bits         : 409658267
Generated R bits         : 0,         Received  R bits         : 2541
```

We also have “ITU-style” Performance Monitoring

Agenda

It is time for a Change, but..

- Bit-transparency & Non-etherent payloads 

- Clocking 

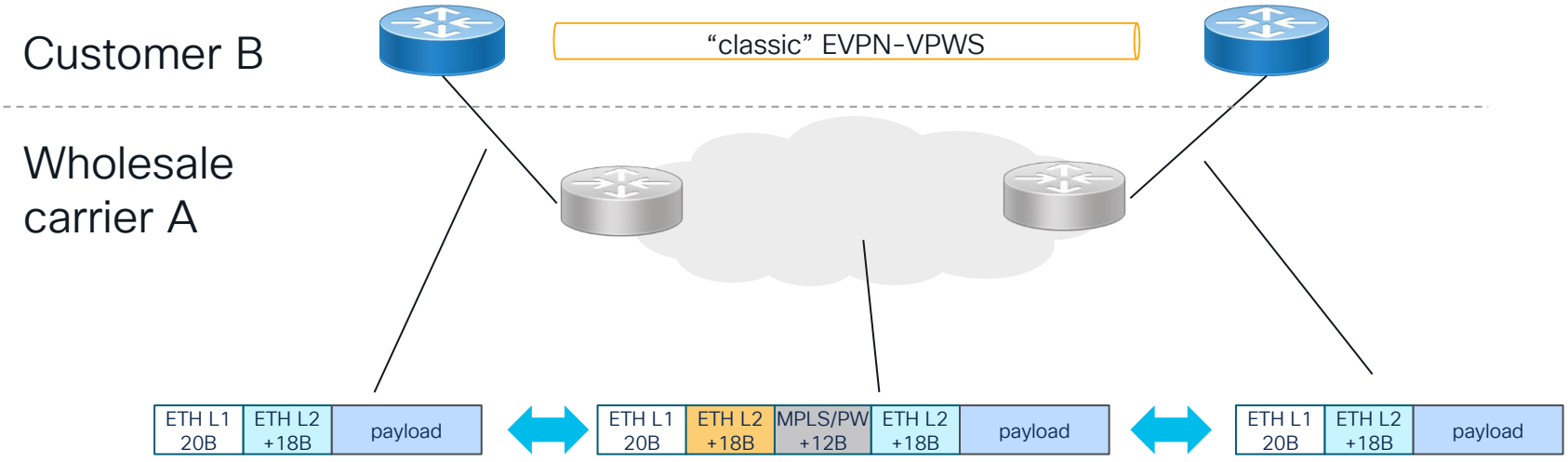
- Inband OAM 

- MTU and overhead
- Co-routed, bidirectional paths
- Dedicated bandwidth

+ some extra thoughts

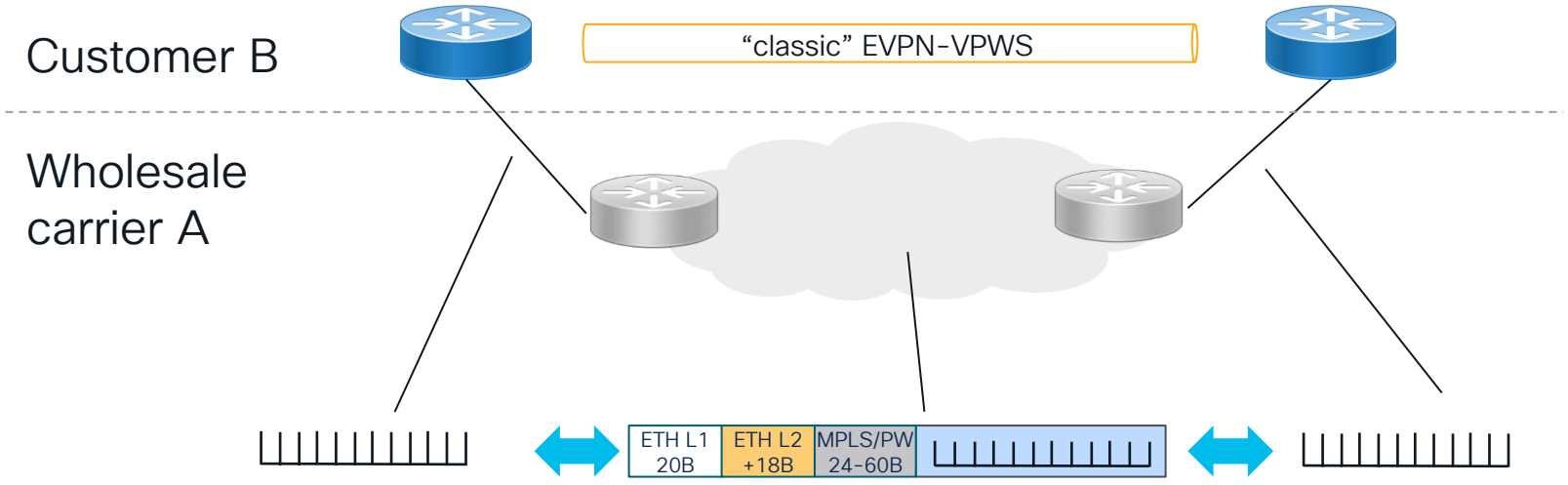
MTU and Overhead

Ethernet PW Overhead reduces Customer MTU



- For a carrier MTU of 9216, customer MTU must be 9176 or less
- Overhead % is frame size dependent

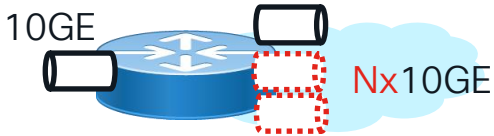
PLE VPWS Overhead is just Overhead



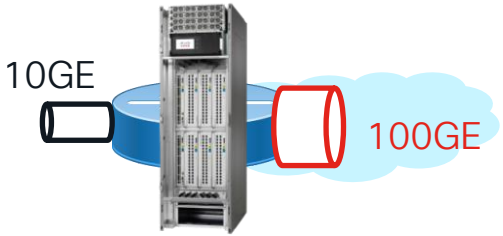
- Customer MTU can still be 9216 (actually anything)
- 6-10% overhead (frame independent)

Dealing with Overhead is Easy Nowadays

Past

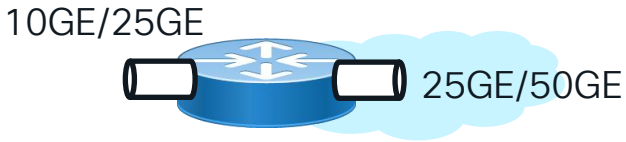


- Link bundle / ECMP
- 5-tuple hashing ?

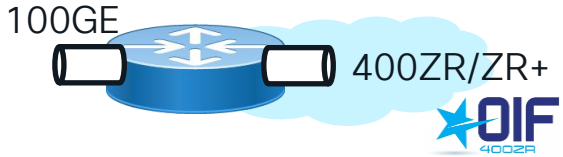


- 100GE expensive
- limited router choice

Present



- more than enough bandwidth
- Good reach variety



- Pluggable coherent optics
- Open ecosystem

Agenda

It is time for a Change, but..

- Bit-transparency & Non-etherent payloads 

- Clocking 

- Inband OAM 

- MTU and overhead 

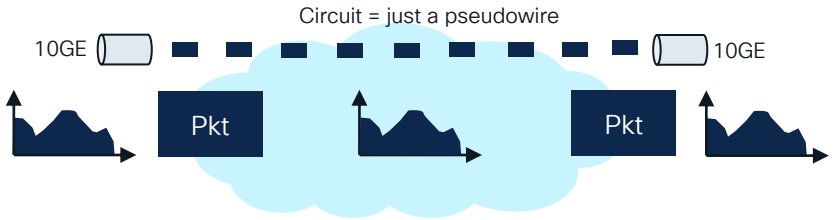
- Co-routed, bidirectional paths
- Dedicated bandwidth

+ some extra thoughts

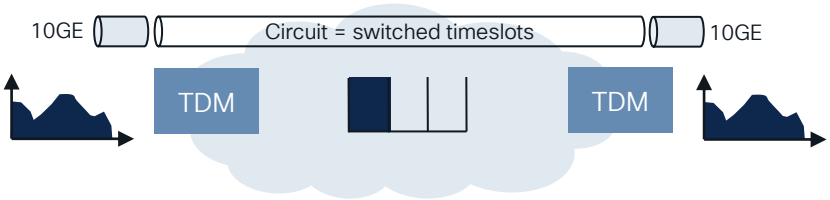
Co-routed, bidirectional Paths & Guaranteed Bandwidth

Classic View : Two very different Models

- Native packet transport
  - Bandwidth only consumed when customer is sending data
  - Allows for multiple traffic classes and forwarding behaviors

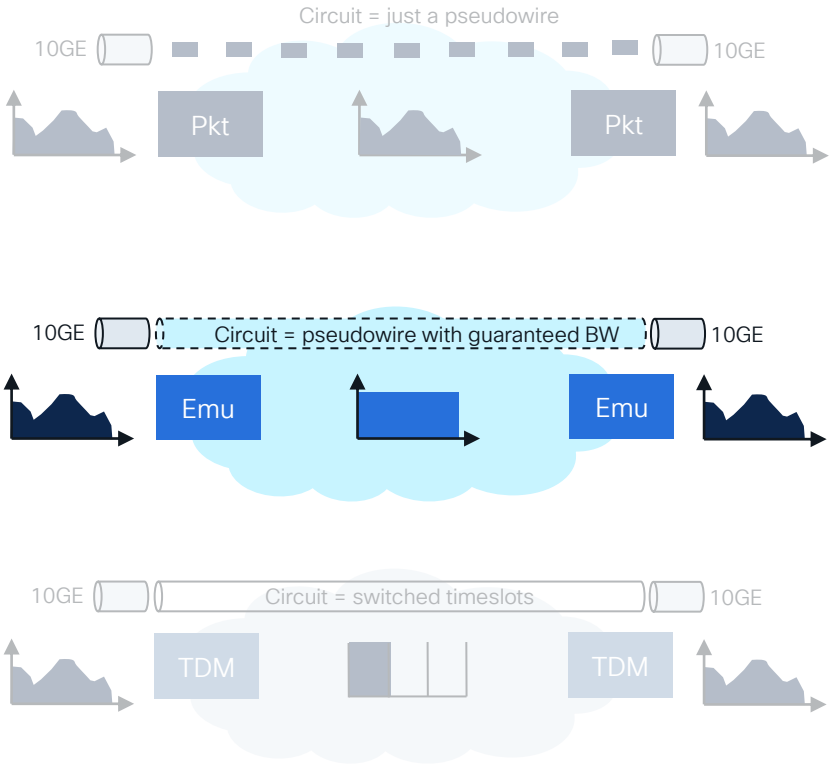


- TDM transport
  - Static timeslot allocation

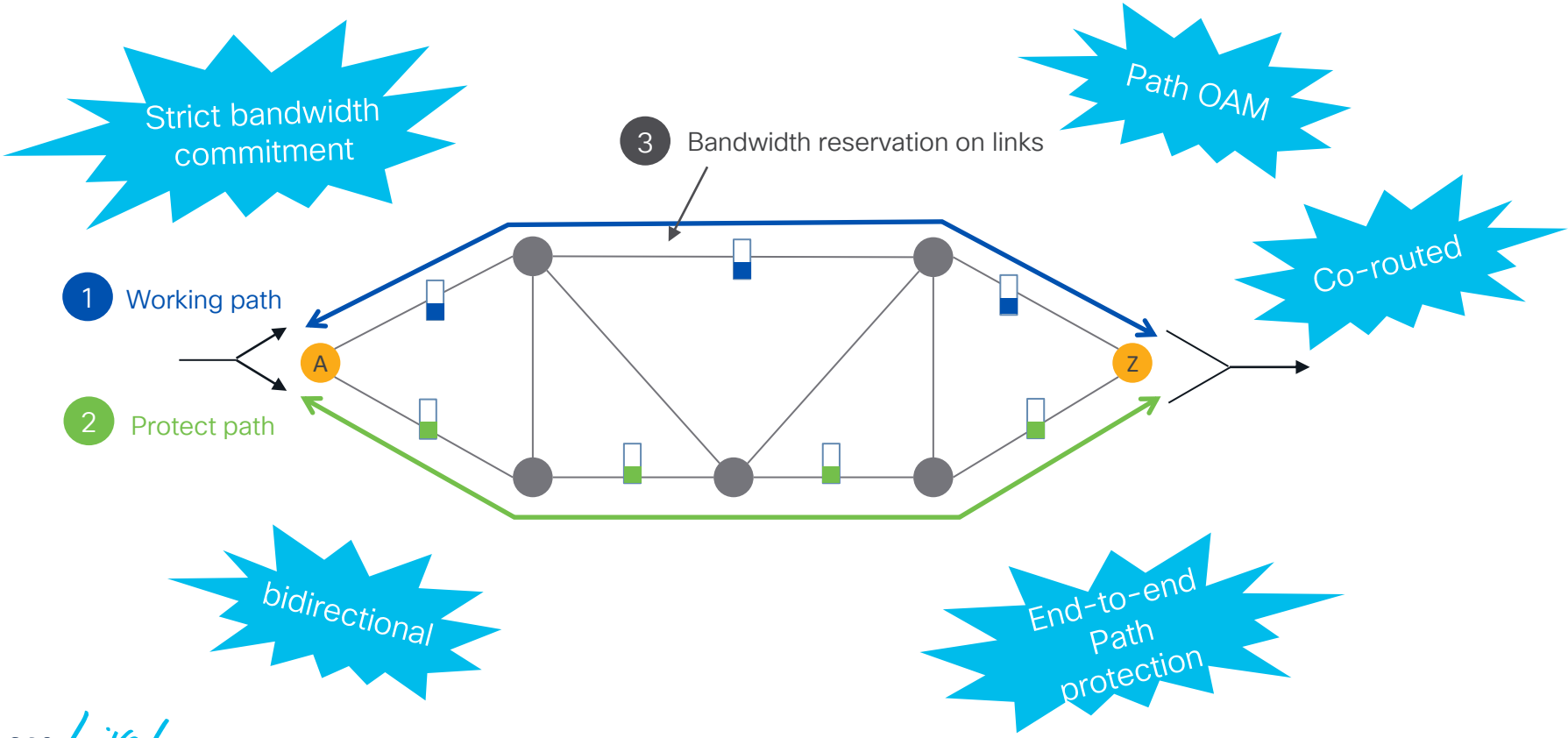


Circuit Emulation = Constant Traffic Load !

- Native packet transport
  - Bandwidth only consumed when customer is sending data
  - Allows for multiple traffic classes and forwarding behaviors
- Emulation
  - Bit transparency
  - Constant network load
- TDM transport
  - Static timeslot allocation



Premium Transport for Premium Services



Why do protection schemes matter?

MPLS-TE FRR

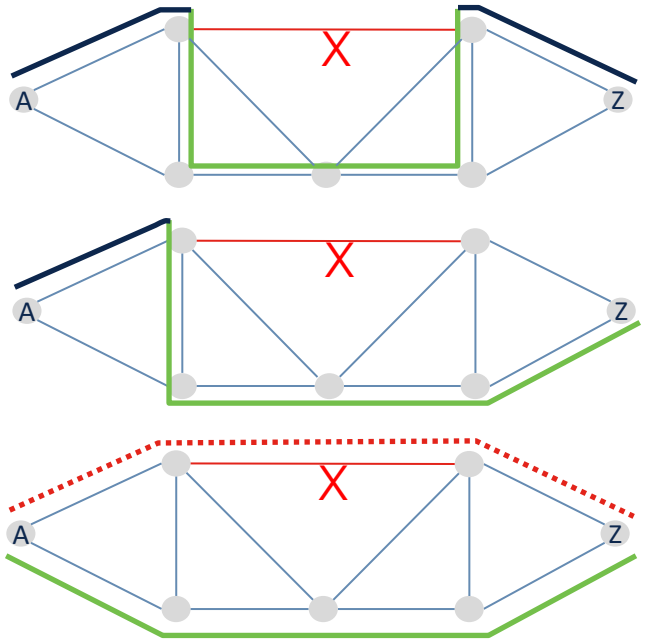
Local bypass protection, without bandwidth allocated

Loop Free Alternate (LFA)

Post convergence path, without bandwidth allocated

Path Protection

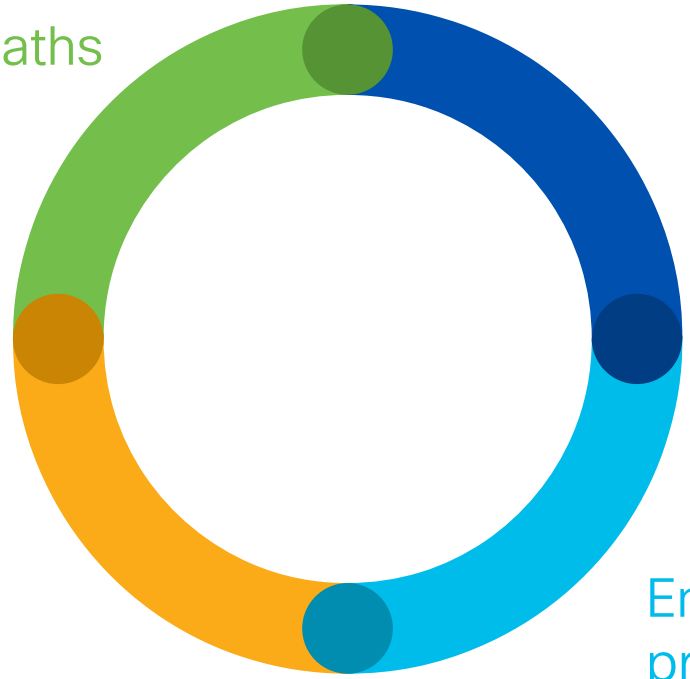
pre-allocated bandwidth end2end



Only with path protection you know “a priori” that there won’t be congestion

Circuit-Style Segment Routing (CS-SR)

Traffic engineered paths
- bidirectional
- co-routed
- persistent

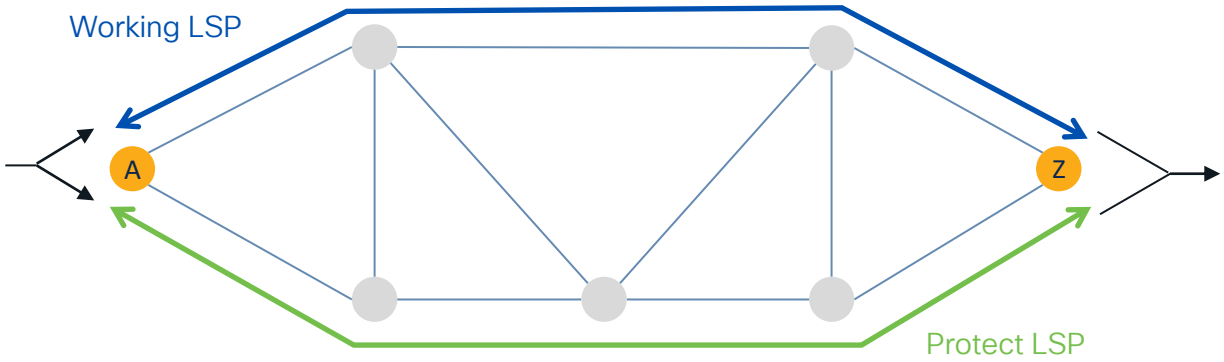


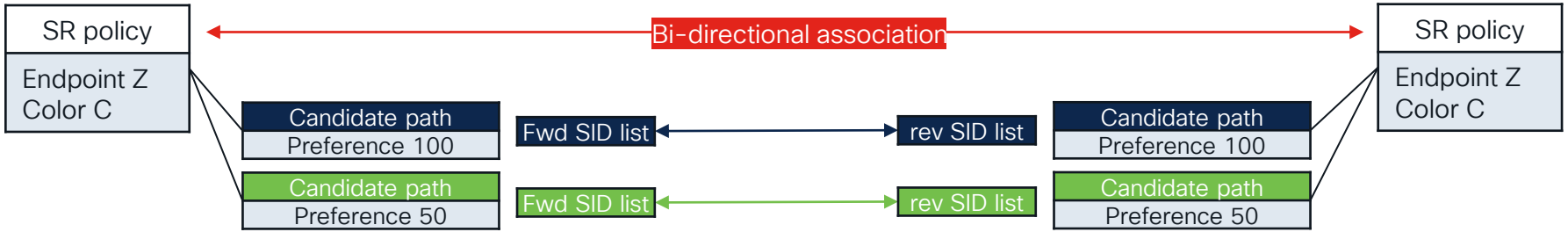
Strict bandwidth commitment

Path OAM

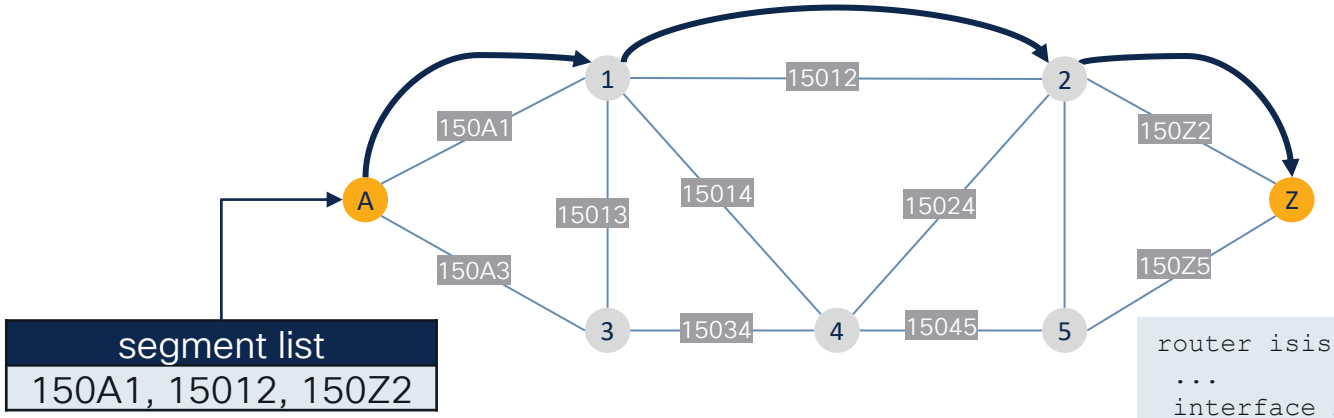
End-to-end path protection & restoration

Path protected, co-routed, bi-directional SR policy





Deterministic and persistent SR paths




```
router isis core
 ...
 interface HundredGigE0/0/2/0
  ...
  address-family ipv4 unicast
  ...
   adjacency-sid absolute 15012
```

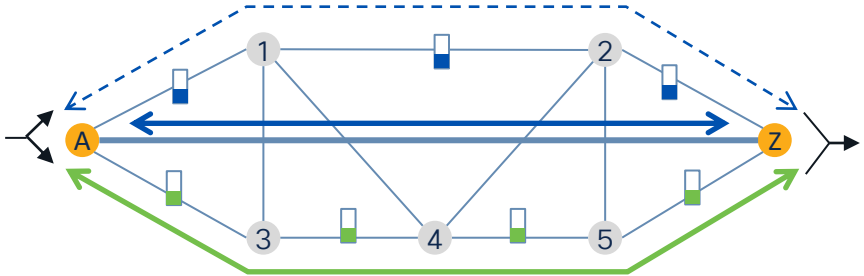
- Strict list of adjacency SIDs → no ECMP
- Manual adjacency SIDs → persistent across node reloads
- Unprotected adjacency SIDs → no traffic rerouting due to TI-LFA

Persistency – ignore newly added Resources

Classic TE behavior

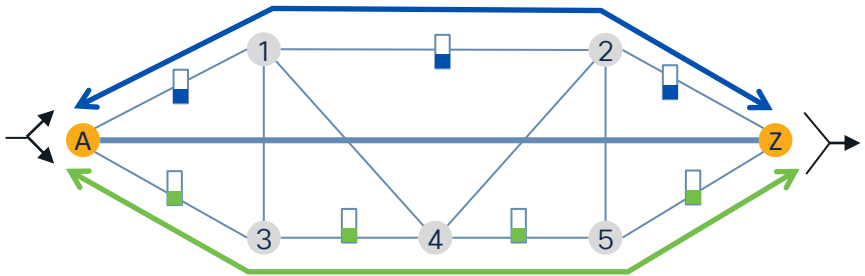
Transport expectation

1' NEW Working LSP



2 Protect LSP

1 Working LSP (stays where it was)

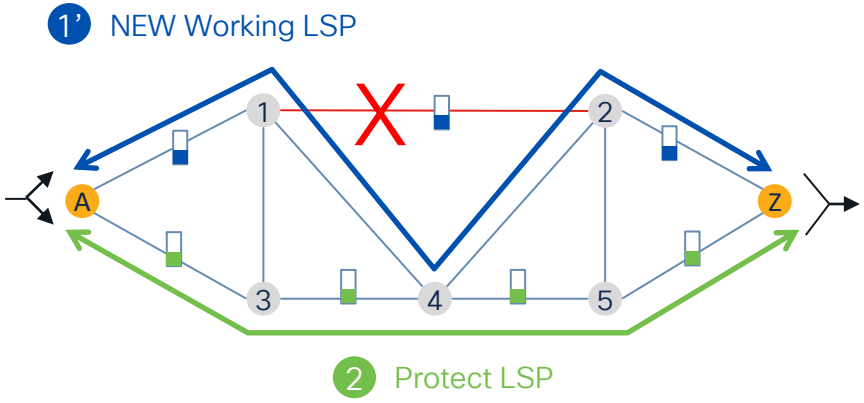


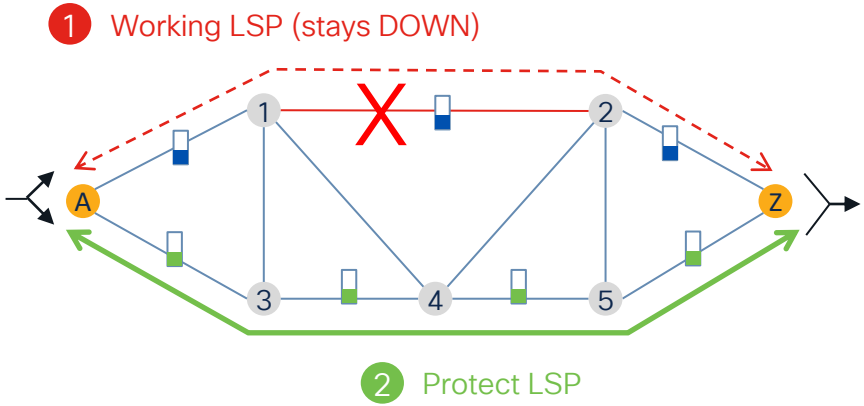
2 Protect LSP

Persistency – don't adjust to Network Changes

Classic TE behavior

Transport expectation



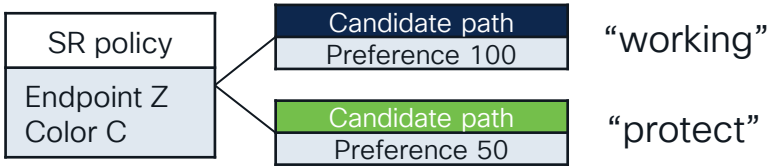


For your Reference

Configuring a “static” CS-SR Policy

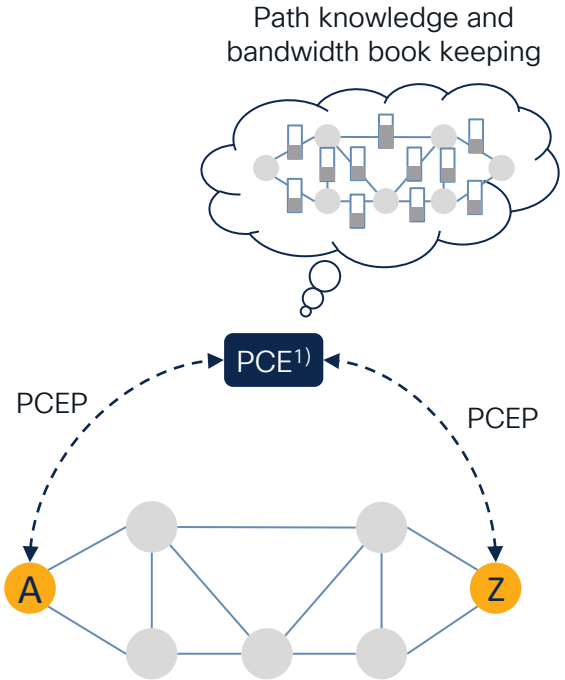
```
segment-routing
 traffic-eng
  policy protected_manual
   color 2 end-point ipv4 1.0.0.6
   path-protection
   !
   candidate-paths
    preference 50
     explicit segment-list via_three_eight
      reverse-path segment-list via_eight_three
     !
    !
    preference 100
     explicit segment-list via_two_four
      reverse-path segment-list via_four_two
```

```
segment-list via_four_two
  index 10 mpls label 15004
  index 20 mpls label 15002
  index 30 mpls label 15001
 !
 segment-list via_two_four
  index 10 mpls label 15002
  index 20 mpls label 15004
  index 30 mpls label 15006
 !
 segment-list via_eight_three
  index 10 mpls label 15008
  index 20 mpls label 15003
  index 30 mpls label 15001
 !
 segment-list via_three_eight
  index 10 mpls label 15003
  index 20 mpls label 15008
  index 30 mpls label 15006
```



PCC-initated CS-SR policy creation

- A SR policy is configured on both endpoints
- Each endpoint requests a path via PCEP from a central PCE
  - Common bi-directional association
  - Required bandwidth
  - Path constraints
- The central PCE maintains a real time view of
  - The network topology (BGP-LS)
  - All path/bandwidth requests (PCEP)



1) Cisco Crosswork Optimization Engine (COE)

For your Reference

Configuring a “dynamic” CS-SR Policy

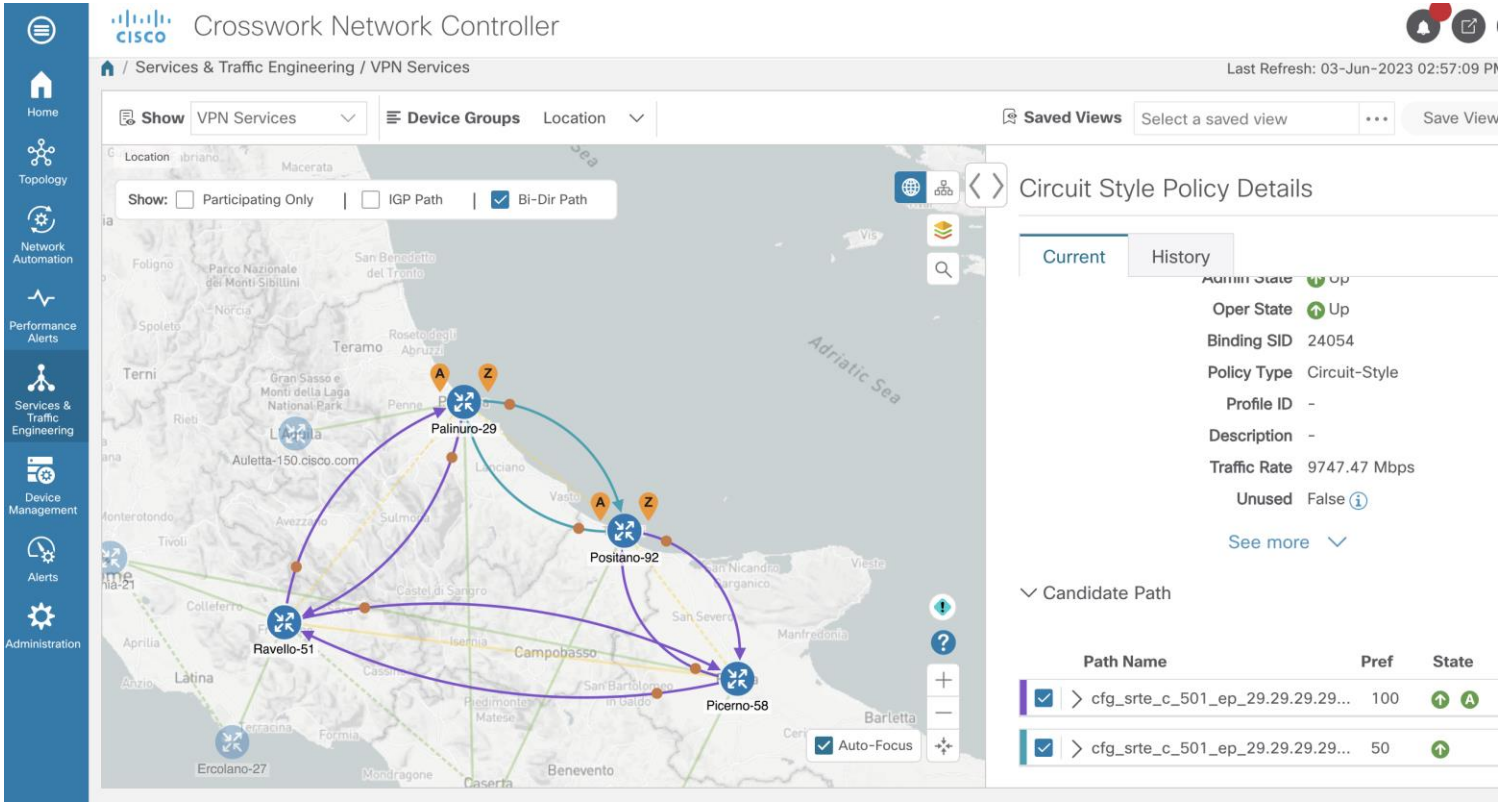
```
segment-routing
 traffic-eng
  policy protected_dynamic
   bandwidth 8000000
   color 3 end-point ipv4 1.0.0.6
   path-protection
   !
   candidate-paths
    preference 100
     dynamic
      pcep
      !
      metric
       type latency
      !
     !
     constraints
      segments
       protection unprotected-only
       adjacency-sid-only
      !
      disjoint-path group-id 16 type link
     !
     bidirectional
      co-routed
      association-id 10
```

```
    preference 50
     dynamic
      pcep
      !
      metric
       type latency
      !
     !
     lock
      duration 60
     !
     constraints
      segments
       protection unprotected-only
       adjacency-sid-only
      !
      disjoint-path group-id 16 type link
     !
     bidirectional
      co-routed
      association-id 11
     !
    !
```

```
    preference 10
     dynamic
      pcep
      !
      metric
       type igp
      !
     !
     lock
      duration 60
     !
     backup-ineligible
     !
     constraints
      segments
       protection unprotected-only
       adjacency-sid-only
      !
     !
     bidirectional
      co-routed
      association-id 12
```

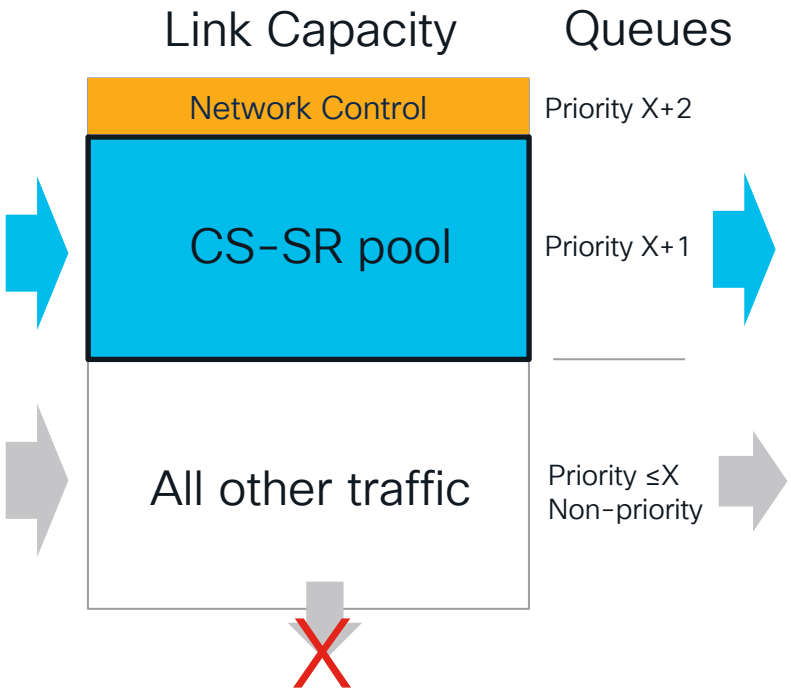
CS-SR PCE is in CNC5.0

Tip: Check out our demo @ World of Solutions



How to Guarantee Bandwidth?

- Two simple rules
  1. Traffic of CS-SR is limited (policed) to requested bandwidth
  2. During congestion only other traffic is dropped, not CS-SR
- To adhere to 2)
  - Don't commit what you don't have
  - Treat CS-SR at highest data priority



Agenda

It is time for a Change, but..

- Bit-transparency & Non-etherent payloads 

- Clocking 

- Inband OAM 

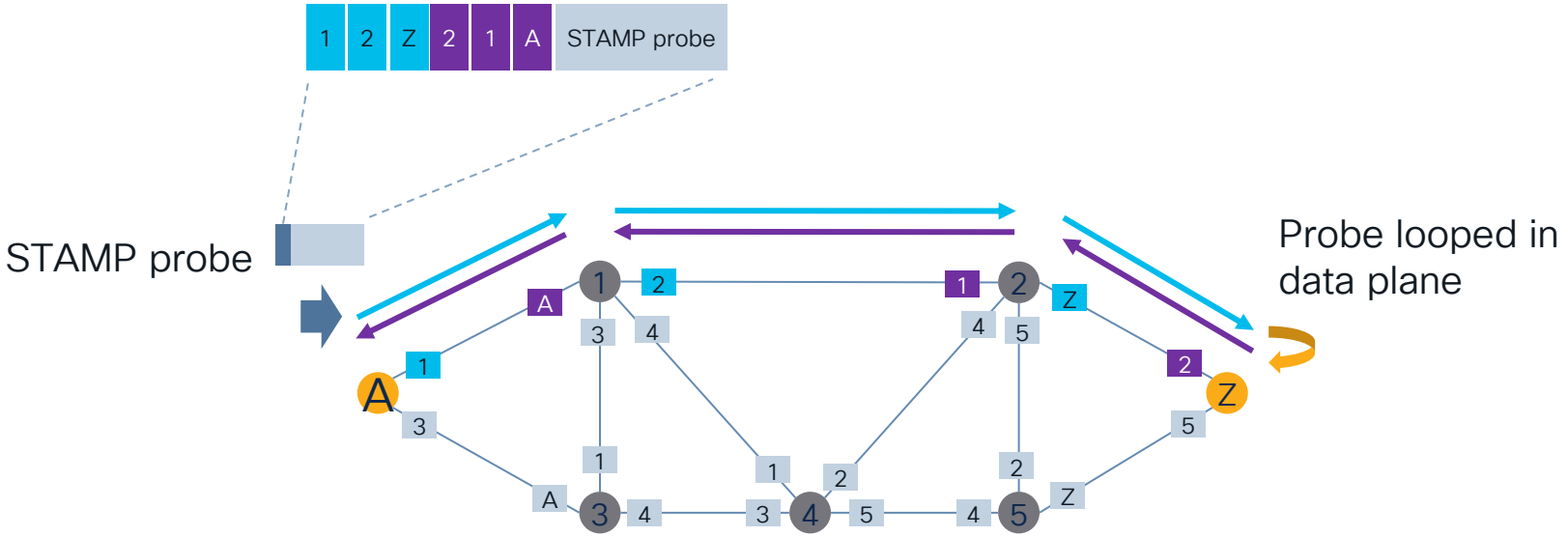
- MTU and overhead 

- Co-routed, bidirectional paths 

- Dedicated bandwidth 

+ some extra thoughts

CS-SR with STAMP adds OAM to Transport



- Bidirectional Liveness and performance measurement via a single protocol (STAMP)
- Loopback probes with exact reverse path encoded providing bidirectional detection

For more see draft-ietf-spring-stamp-srpm

For your Reference

Enabling Candidate Path Liveness

```
performance-measurement
 liveness-profile name cs_protect
  liveness-detection
   multiplier 3
  !
  probe
   tx-interval 100000
  !
 !
 liveness-profile name cs_working
  liveness-detection
   multiplier 3
  !
  probe
   tx-interval 3300
  !
  npu-offload
   enable
```

```
segment-routing
 traffic-eng
  policy protected_manual
   ...
   performance-measurement
    liveness-detection
     liveness-profile backup name cs_protect
     liveness-profile name cs_working
```

Agenda

It is time for a Change, but..

- Bit-transparency & Non-etherent payloads 

- Clocking 

- Inband OAM 

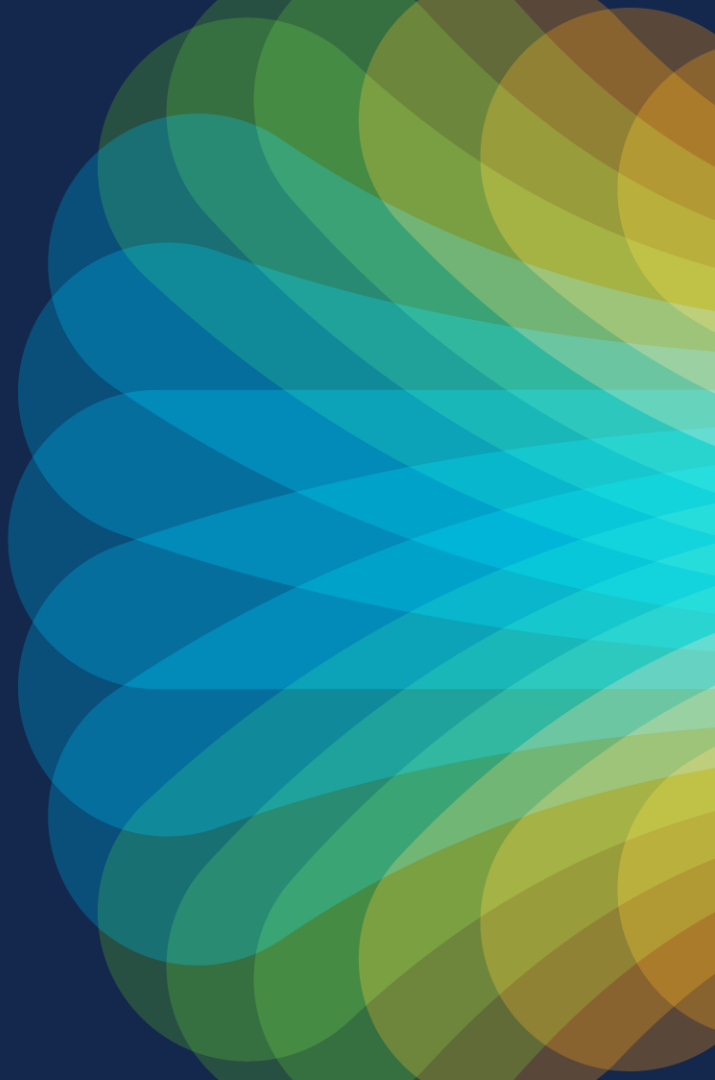
- MTU and overhead 

- Co-routed, bidirectional paths 

- Dedicated bandwidth 

+ some extra thoughts

So far so good ...
lets try to deploy this



PLE is real since XR 7.7.1 (July 2022) !

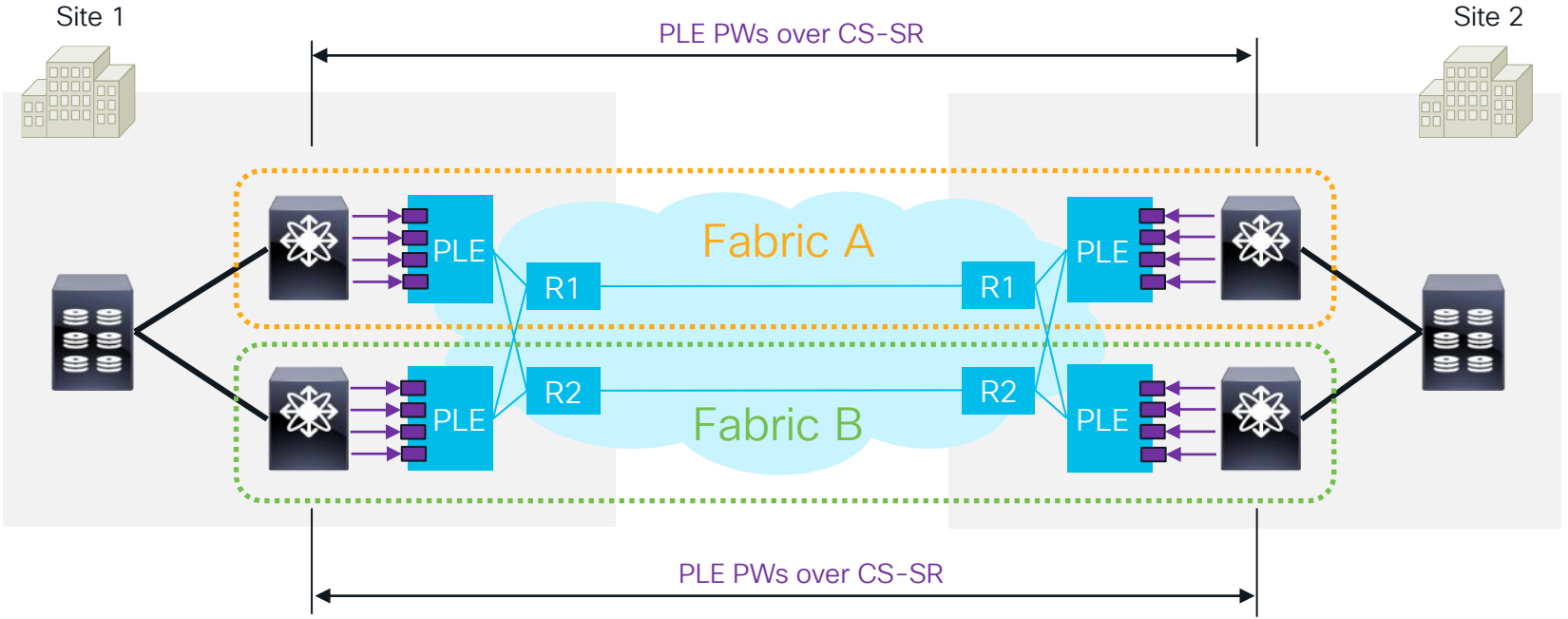
- Supported client types
  - 1GE, 10GE
  - OC48/STM16, OC192/STM64
  - Fibre channel (1, 2, 4, 8 ,10 ,16 and 32G)
  - OTU2, OTU2e
- Any mix of client types supported
- Supported in NCS-55A2 (peyto) and NCS-57C3 (Eryie)

NC55-OIP-02

NCS-55A2

NCS-57C3

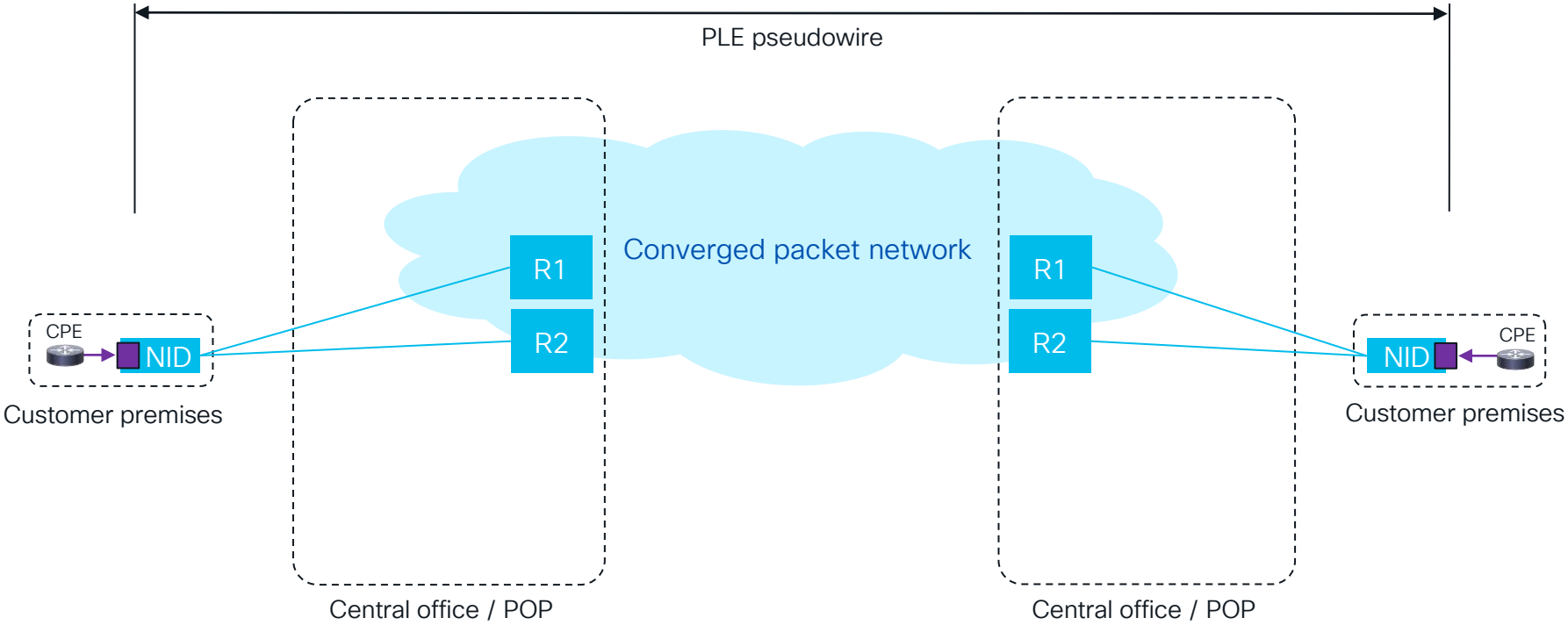
Applying PLE to Storage Area Networks



Tip: Check out our demo @ World of Solutions

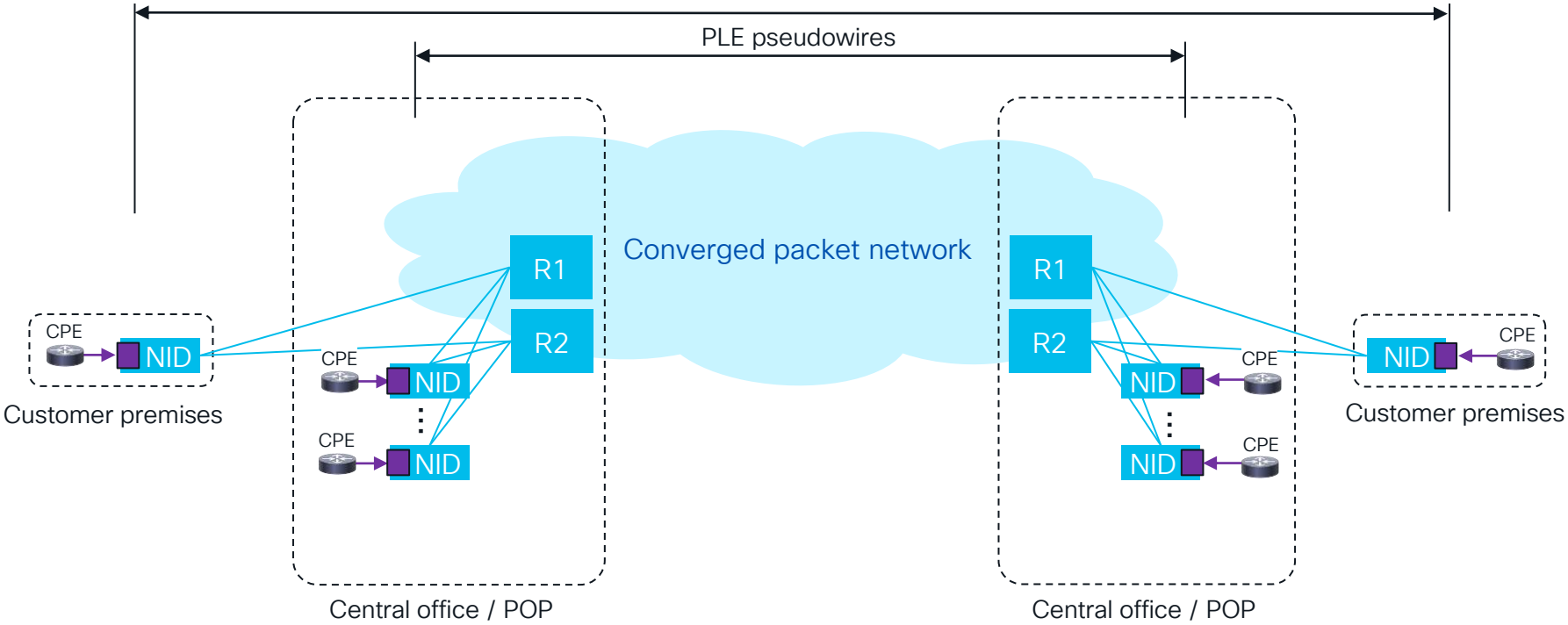
PLE hardware

Easy ”Pay-as-you-grow” Insertion for Carriers



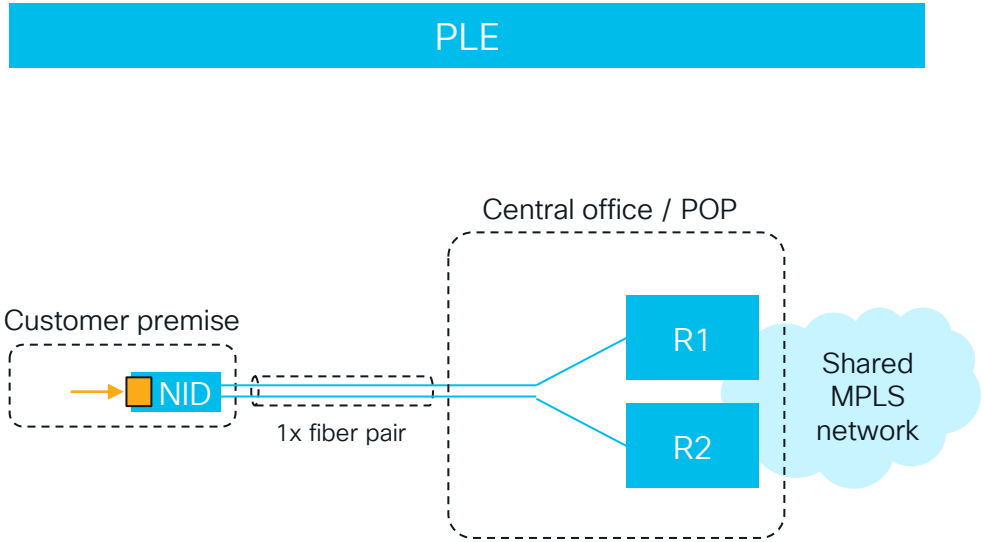
PLE hardware

At Customer Premise or in Central Office



Connecting a Customer – more Details

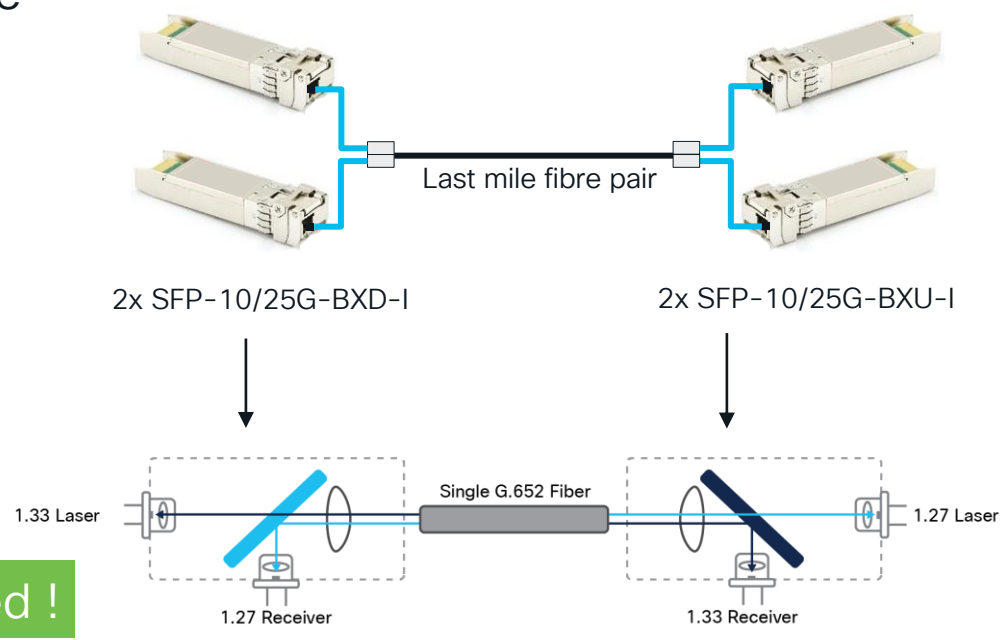
- One fiber pair to customer premise
- Generally there is a pair of routers in the CO/POP
- The challenge:
  - How to connect the NID to two routers via a single fiber pair?



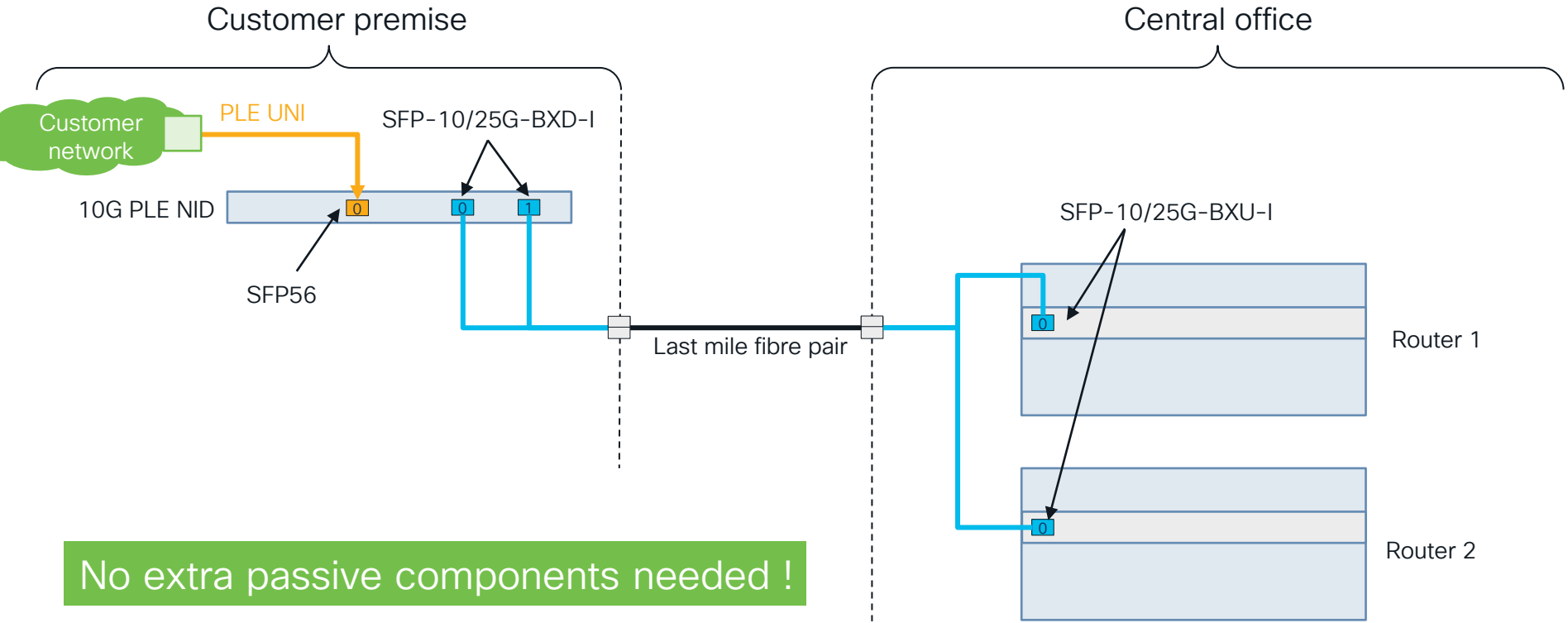
25GE BiDi Transceivers

- Allows for a 25GE link over a single strand of single-mode fiber
- Two SFP28 variants
  - SFP-10/25G-BX40D-I (1330nm TX)
  - SFP-10/25G-BX40U-I (1270nm TX)
- Achievable distance
  - ~18dB power budget
  - 40km reach (CD limited)

No extra passive components needed !

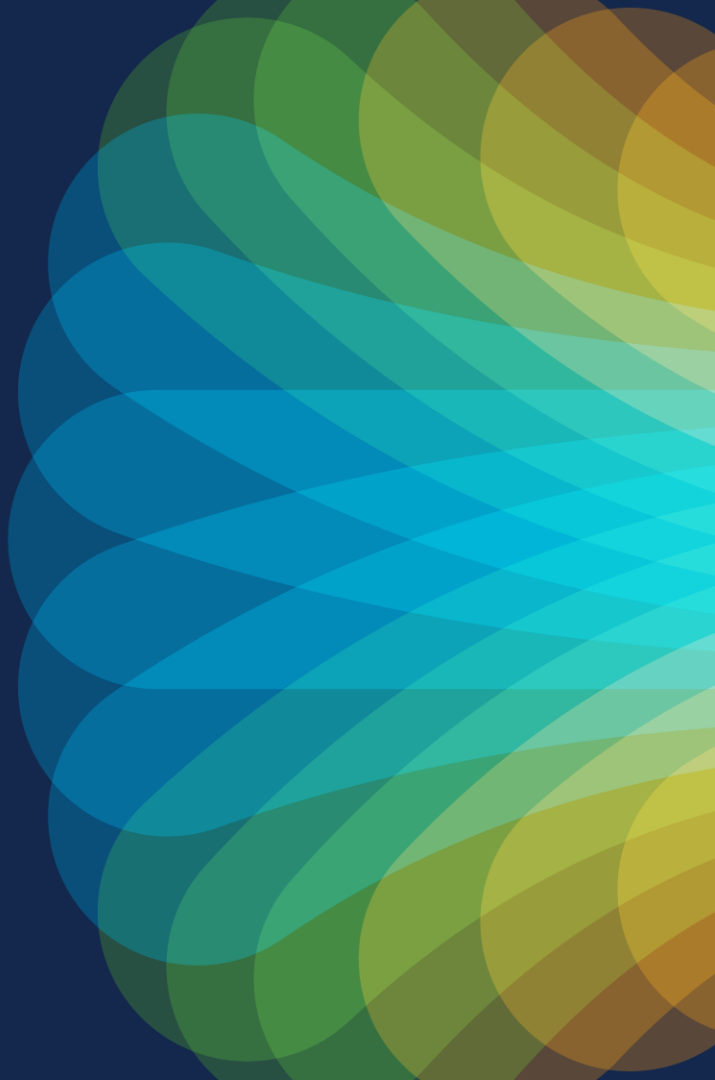


10G PLE at Customer Premise

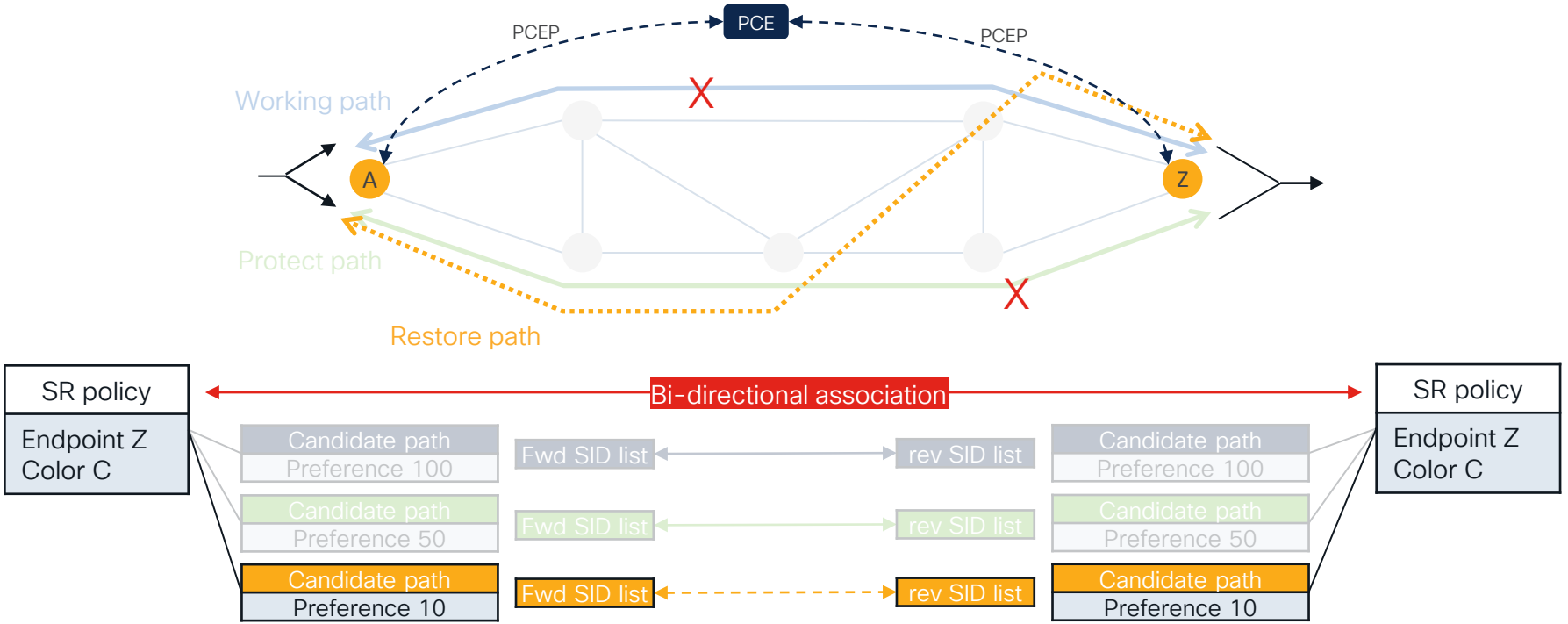


No extra passive components needed !

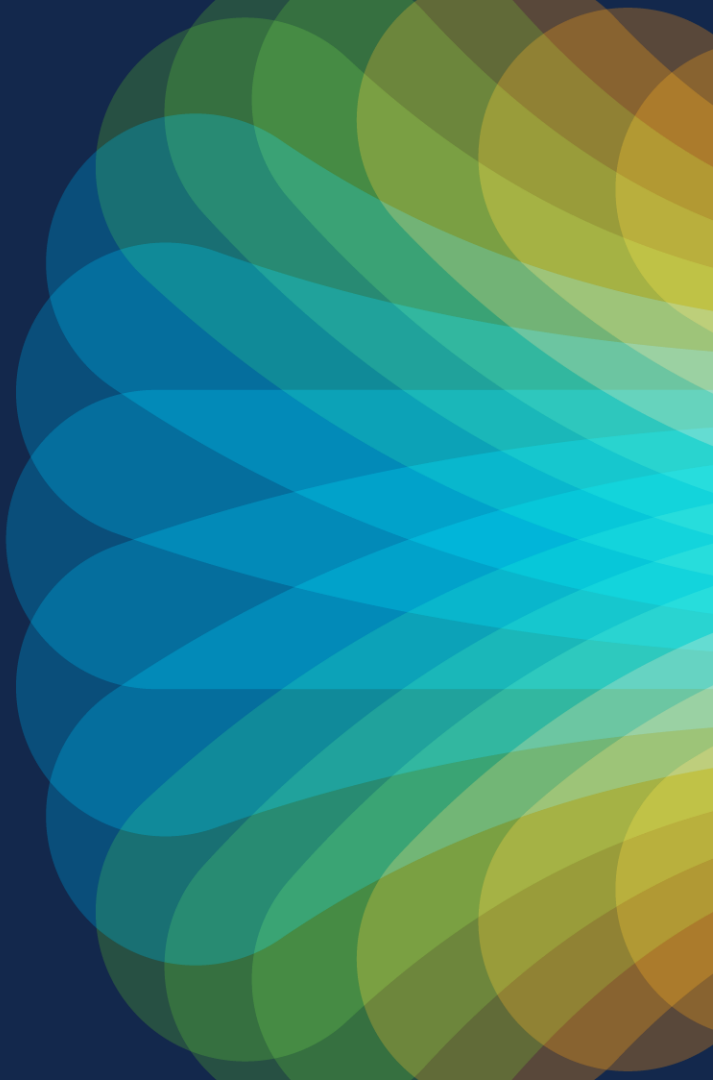
... ah there is
one more thing



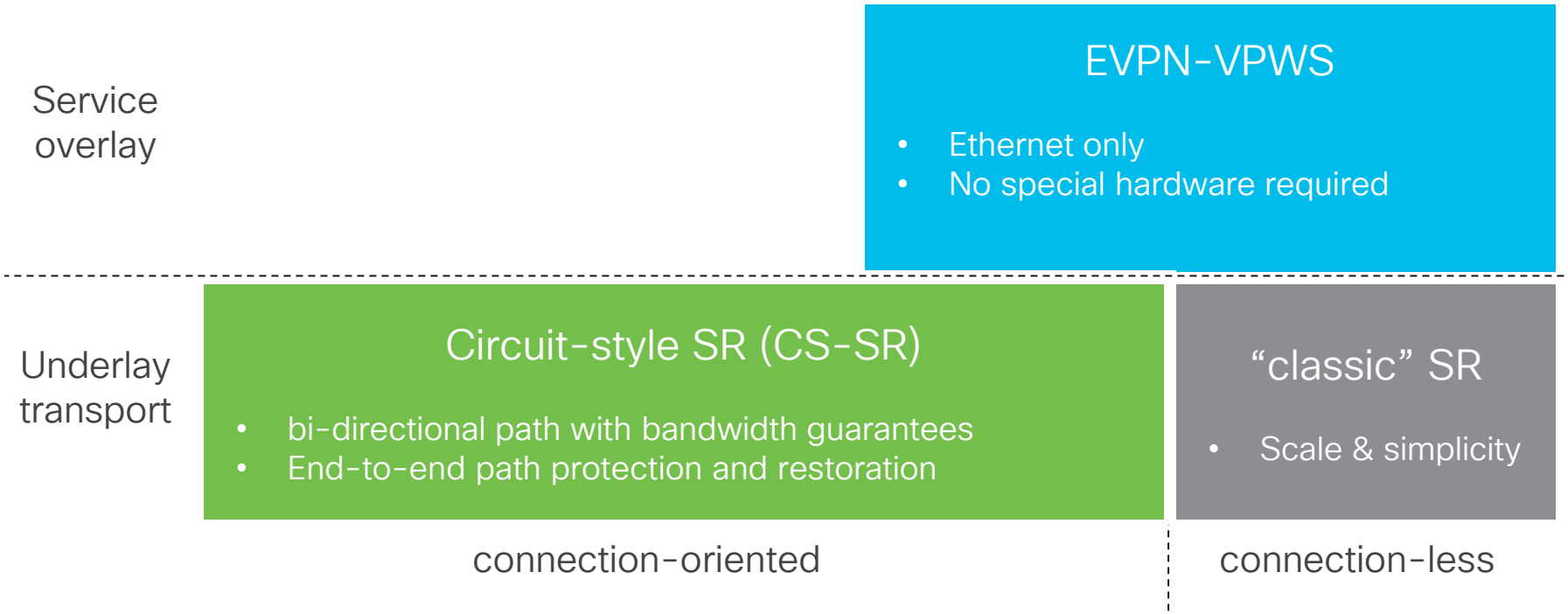
CS-SR, Dynamic Recovery from Double Failures



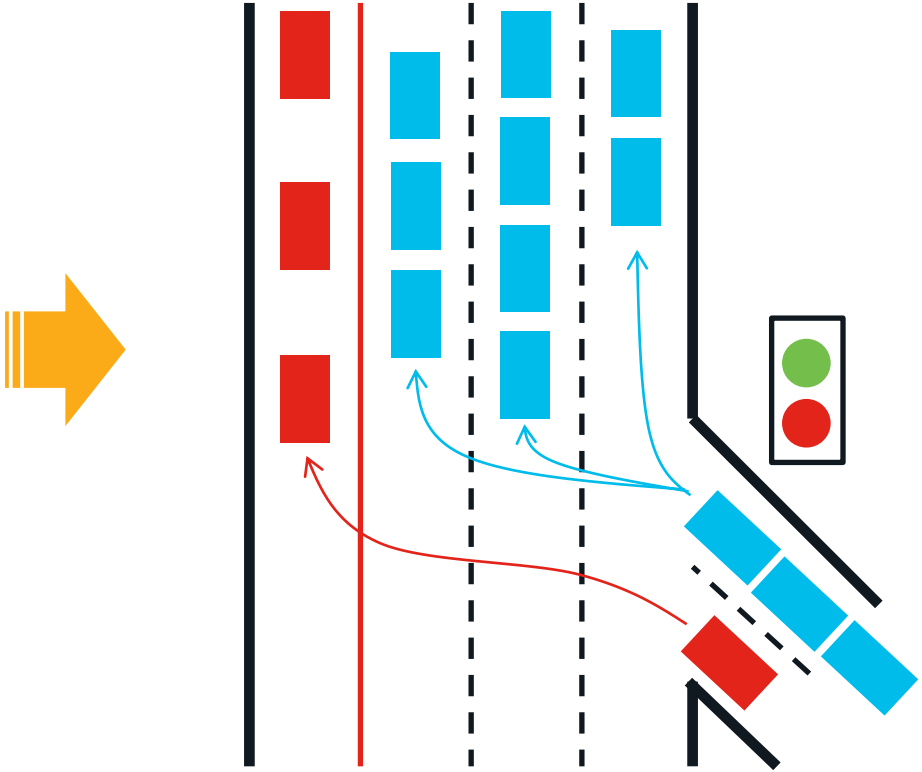
Key take aways



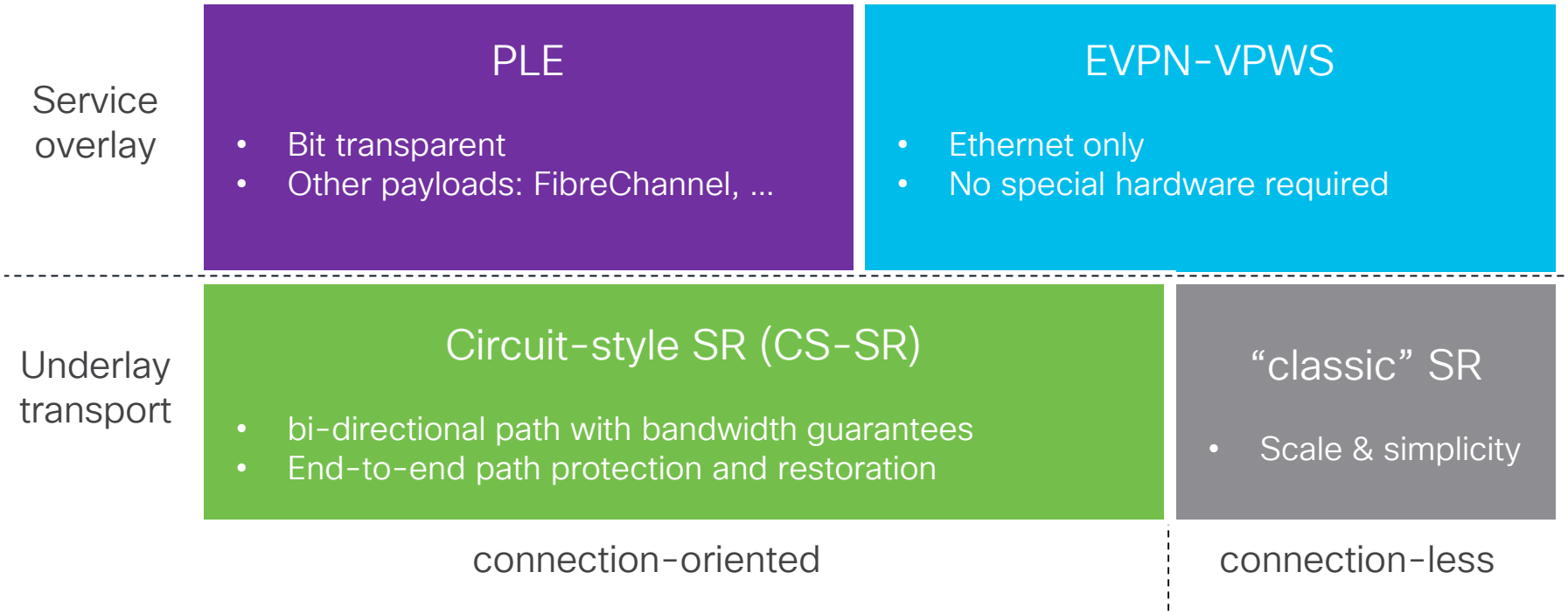
CS-SR … a new Transport Behavior



CS-SR is bringing Order onto “Planet Packet”



PLE completes the Service Portfolio



From Closed Optical to Open Packet Private Lines

Big Routers with standard interfaces

OpenZR+ MULTI-SOURCE AGREEMENT

OIF 400ZR

IETF®

draft-schmutzer-pals-ple
draft-schmutzer-spring-cs-sr-policy
draft-sidor-pce-circuit-style-pcep-extensions

Cost-efficient, standards-based solution

Fill out your session surveys!

Attendees who fill out a minimum of four session surveys and the overall event survey will get **Cisco Live-branded socks** (while supplies last)!

Attendees will also earn 100 points in the **Cisco Live Challenge** for every survey completed.

These points help you get on the leaderboard and increase your chances of winning daily and grand prizes

Tip: Check out related sessions by searching for “BRKOPT-” in the session catalog

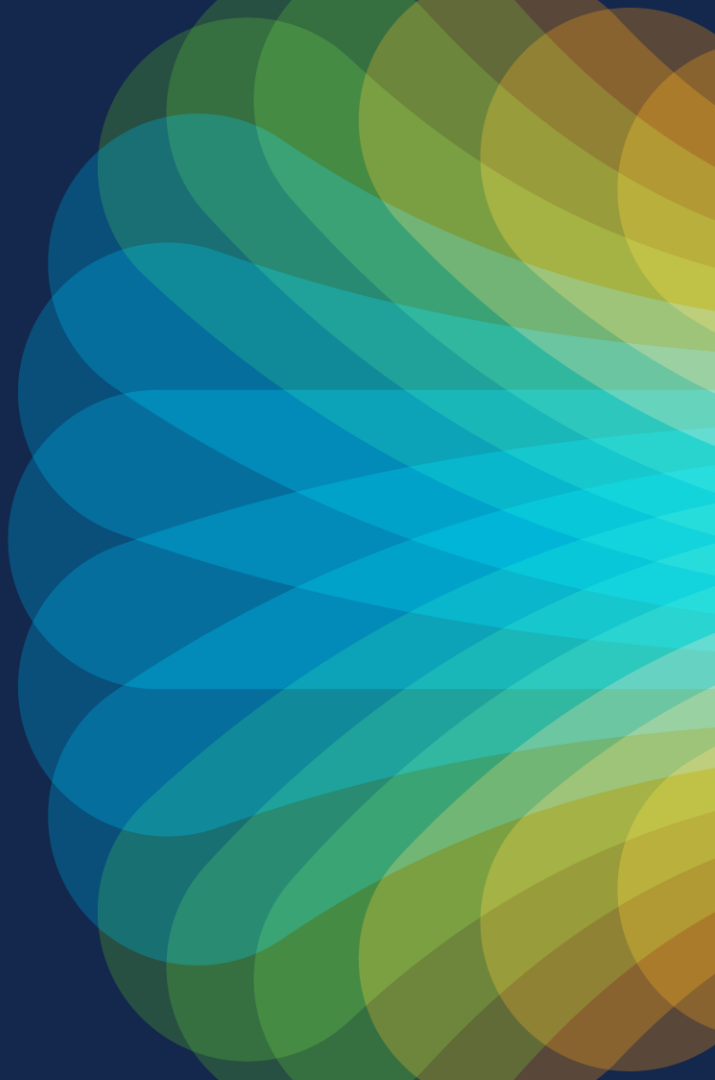
Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

cisco Live!

The bridge to possible

Thank you



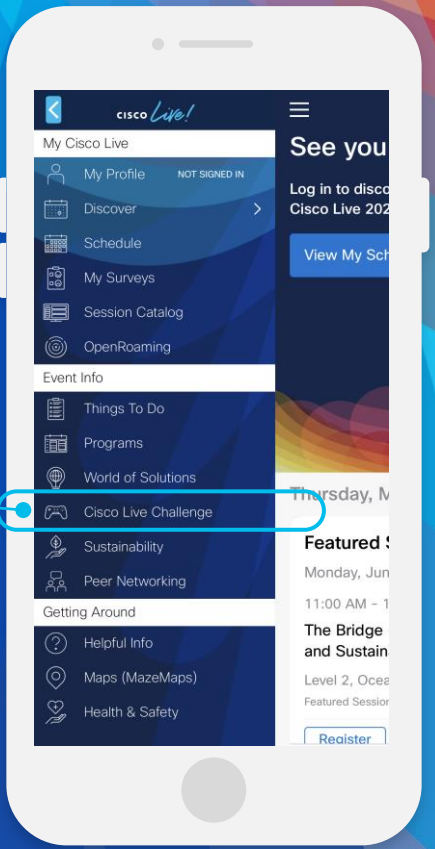
Cisco Live
Challenge

Gamify your Cisco Live experience!

Get points for attending this session!

How:

1. Open the Cisco Events App.
2. Click on 'Cisco Live Challenge' in the side menu.
3. Click on View Your Badges at the top.
4. Click the + at the bottom of the screen and scan the QR code:



cisco Live!


cisco Live!
Let's go
#CiscoLive