Bridge-Based Ethernet Service Provision

Norman Finn

Current Standards Activities

- IETF Provider Provisioned Virtual Private Networks Working Group (PPVPN) is defining “L2-VPNs”.
- Metro Ethernet Forum (MEF) is now defining requirements, and may develop standards, as well.
- ITU SG15 has a Question for Ethernet Service.
- IEEE 802.1 is developing a Project Authorization Request to start work on Metro Ethernet Services.

IETF PPVPN L2-VPNs (1)

- **Very much a work in progress: Choices for L3 tunneling mechanism not settled. Martini draft[1] over MPLS a front runner, L2TP[2] not far behind.**
- **Current and planned documents are based on the assumption that an L3-VPN is essentially the same thing as an L2-VPN.**
  - **This assumption is valid if (and only if!) the L3 tunnel runs end-to-end and no decisions are based on MAC addresses.**
  - **Some currently popular architecture drafts[3] violate this assumption.**

[1] draft-martini-ethernet-encap-mpls-01.txt

[2] draft-heinanen-dns-l2tp-vpls-01.txt

[3] draft-lasserre-vkompella-ppvpn-vpls-02.txt

IETF PPVPN L2-VPNs (2)

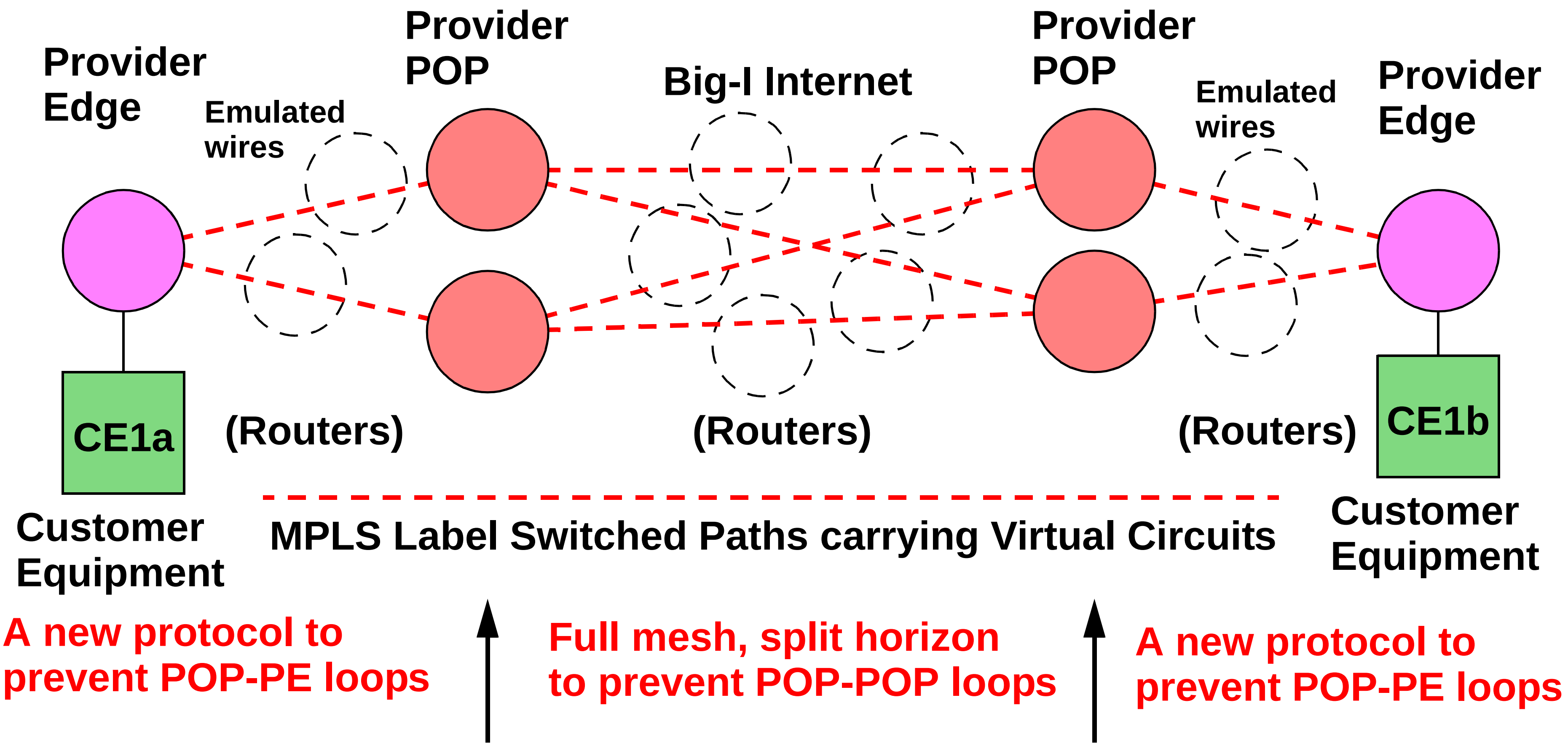
- **A draft[1] has been accepted by the PPVPN Working Group as a baseline for the requirements for L2-VPNs.**
- **Another draft[2] has been accepted as a framework for further work.**
- **See Slide 56 for an update.**

[1] [draft-augustyn-ppvpn-l2vpn-requirements-00.txt]

[2] [draft-ietf-ppvpn-l2-framework-01.txt]

Lasserre-VKompella

- **Typical of IETF PPVPN drafts.**



Provider Edge

Emulated wires

Provider POP

Big-I Internet

Provider POP

Emulated wires

Provider Edge

CE1a

(Routers)

(Routers)

(Routers)

CE1b

Customer Equipment

MPLS Label Switched Paths carrying Virtual Circuits

Customer Equipment

A new protocol to prevent POP-PE loops

Full mesh, split horizon to prevent POP-POP loops

A new protocol to prevent POP-PE loops

Provider Points Of Presence (POPs) switch packets by MAC address, learn MAC address / virtual wire associations, and run protocols to prevent loops, *but they’re not bridges!* ☺

Ethernet Emulation over MPLS

- **The IETF solutions include features which are, in effect, “Ethernet Emulation over MPLS” (EEoMPLS).**
  - **The elements constituting this feature are never grouped together and given a single name. They are buried in the description of the overall solution.**
- **“EEoMPLS” is characterized by:**
  - **A full mesh of point-to-point tunnels is built among the ports.**
  - **No frame is ever relayed directly from one tunnel to another tunnel on the same EEoMPLS instance (“split horizon”).**
  - **Devices learn and forget MAC address / tunnel associations.**
- **The requirements and framework drafts describe a function that quacks, flies, and swims like EEoMPLS. It’s a duck!**

EEoMPLS is simpler than ATM LANE

— If additional requirements are added, such as avoiding a full mesh in order to scale to a larger number of ports, it will grow in complexity, just as did LANE.

— The Layer 3 world, as a substrate, is much richer than ATM.

— EEoMPLS is not afraid to do MAC learning at each node.

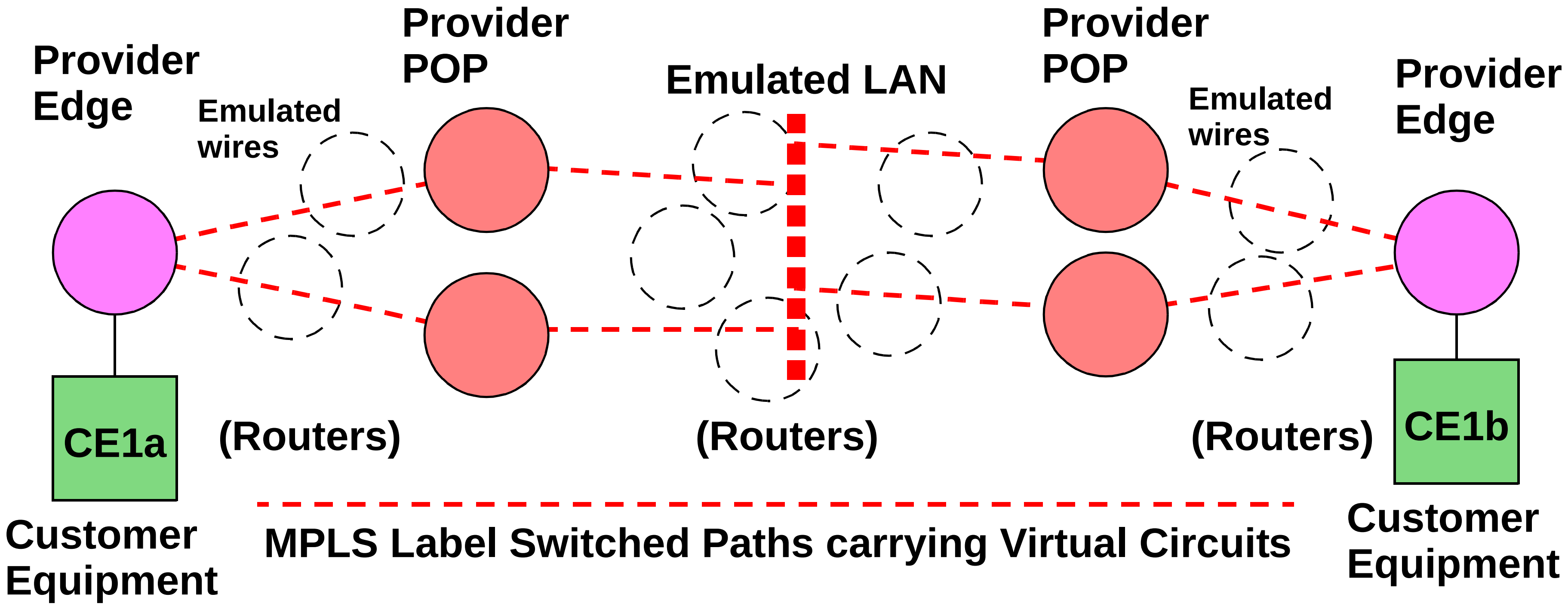
— EEoMPLS is not trying to scale to numbers of nodes greater than a full mesh of connections can support. Hence, no Broadcast and Unknown Server (BUS).

— Participating ports use a common method of discovering each other’s existence. BGP and DNS have been proposed as methods.

— EEoMPLS uses MPLS’s point-to-multipoint capabilities for broadcast/multicast destination MAC addresses.

— It accepts the inefficiency of point-to-point links for unknown unicast addresses, to avoid the BUS and out-of-order delivery.

Hence, Lasserre-VKompella is really:



Provider Edge

Emulated wires

Provider POP

Emulated LAN

Provider POP

Emulated wires

Provider Edge

CE1a

(Routers)

(Routers)

(Routers)

CE1b

Customer Equipment

MPLS Label Switched Paths carrying Virtual Circuits

Customer Equipment

A new protocol to prevent POP-PE loops

Emulated LAN

A new protocol to prevent POP-PE loops

Provider POP? Provider Edge? No!

They’re bridges, running something besides Spanning Tree!

Avoiding Loops in L2-VPN

- **Clearly understood by IETF:**
  - — **MAC frames have no TTL to make up for transient loops.**
  - — **Full mesh split horizon prevents loops across the big-I Internet.**
  - — **Limiting the topologies at the edge and running a simple backup link protocol prevents loops in edge triangles.**
  - — **Some POP-to-POP protocol or extension of triangle edge protocol is required to prevent multiple POPs from connecting one Service Instance to the full mesh twice, causing Grand Loops.**
  - — **A semi-automatic means for the configuration is required.**
- **Perhaps not yet fully understood by IETF:**
  - — **The required simplified topologies are enforced by configuration and signaling protocols which have transients.**
  - — **During transients, the actual topology may violate the simplified protocols' assumptions.**

MAC Learning and MPLS

- **MPLS is an optimization (: many claim :) of routing.**
  - — **First, you build a routed infrastructure. Packets flow like a river over a flat region — they meander a bit, and sometimes flood.**
  - — **MPLS then puts up levees: a Label-Switched Path (LSP).**
  - — **You may have alternate LSPs for fast recovery from failure.**
  - — **Any LSP can be broken, and packets routed from the broken ends, because MPLS paths are based on routing algorithm(s).**
- **MPLS is not (yet) an optimization of bridging.**
  - — **There is, as yet, no routing algorithm for MAC addresses on which to build LSPs. (Spanning tree provides it own levees!)**
  - — **Since LSPs are not built on a MAC address infrastructure, MAC address decisions must take the LSPs for granted, as “wires”.**
  - — **This is logically equivalent to bridging among Ethernets!**

Of Course, IETF *Can* Re-Invent Bridging

- **Perhaps the result would be better than IEEE 802.1D! However, there are issues:**
  - **Certain layer 2 protocols (e.g. NetBEUI) cannot tolerate out-of-order or duplicated frame delivery. MAC “routing” must be able to guarantee in-order delivery, even in the face of transient events.**
  - **It is difficult to see how new bridges based on routing algorithms could interoperate with existing spanning-tree based bridges.**
  - **Features that are affected and/or must be reinvented: L2 QoS, VLAN tagging details, MIBs, optimal multicast distribution, techniques used for configuring preferred topologies, VLAN pruning, etc.**

A Few Terminology Issues

- **There is no such formal entity as a “switch”. IEEE 802.1D defines a bridge. RFCs define routers.**
  - **— *Your* company builds boring old routers and bridges.**
  - **— *My* company builds sexy new switches!**
- **The bridge port on the Provider’s bridge facing the customer is a User-Network Interface (UNI).**
  - **— The UNI is at different places for different layers — the Layer 1 UNI, for example, is typically at the customer’s site.**
  - **— The Layer 2 UNI is on the PEB. That’s the focus of these slides.**
  - **— Also called the “demarc”.**

Starting Points: Drivers to a Solution (1)

- **The service seen from the Customer's side of the UNI must be standardized to one or a few choices.**
- **The Customer Equipment needs no new software or hardware to take advantage of the Providers' services.**
  - **Long term, this may change.**
  - **As far as the first round of standards are concerned, violating this goal would seriously hurt the standard's chances of success.**
- **An existing bridge should be able to serve in a Provider's network with a minimum of changes.**
  - **One assumes that carrying an extra MAC tag is a small change.**
  - **One can imagine a feature being so valuable that it is worth violating this goal, but there should be no other practical way to accomplish the feature.**

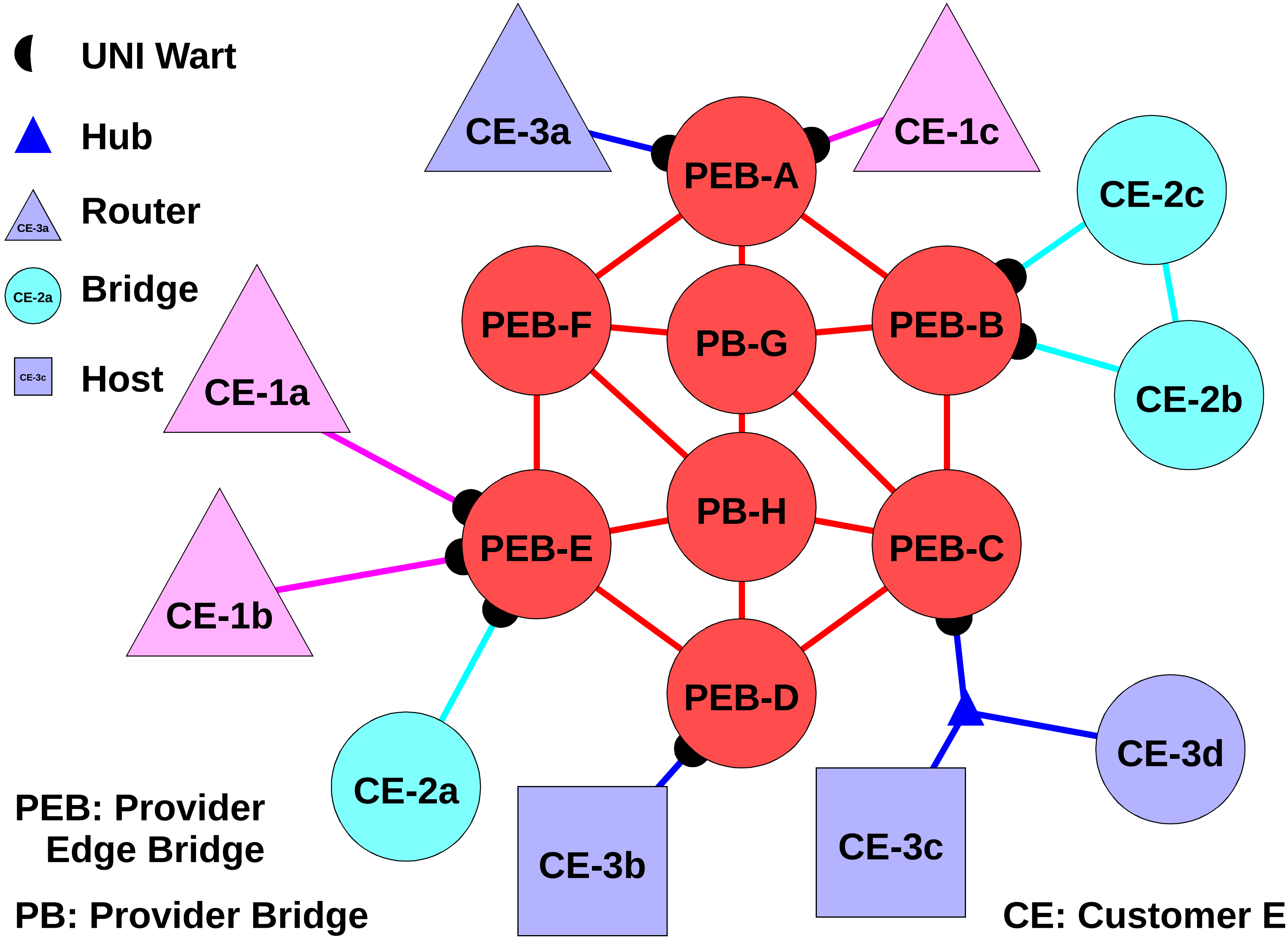
Starting Points: Drivers to a Solution (2)

- **The UNI is clearly *not* a bridge port within the current standards.**
  - **The Customer's spanning tree is *not* concatenated to the Provider's spanning tree!**
  - **An extra IEEE 802.1Q tag must be added in some cases.**
- **Therefore, changes to IEEE 802.1D, Q, etc., should be restricted to that which can be accomplished by adding a "UNI wart" between a bridge port and the Customer.**
  - **To the extent that the differences between a standard bridge port and a UNI port can be isolated, the documents defining bridges need not be re-written.**
  - **For management purposes, if nothing else, this "wart" should be considered part of the Provider Bridge, not the CE or the wire.**

Starting Points: Drivers to a Solution (3)

- **The Provider's network must be scalable to a larger number of bridges than is practical with the current standard (and about-to-be-standard) spanning tree protocols.**
  - — **Providers need to support more than 4k customers.**
  - — **Providers' networks may have more than 7-8 hops.**
  - — **Some customers' connections may span multiple Providers' networks.**
- **The IEEE 802.1 solution needs to interoperate with the IETF solution(s).**

All-Bridged Provider Model

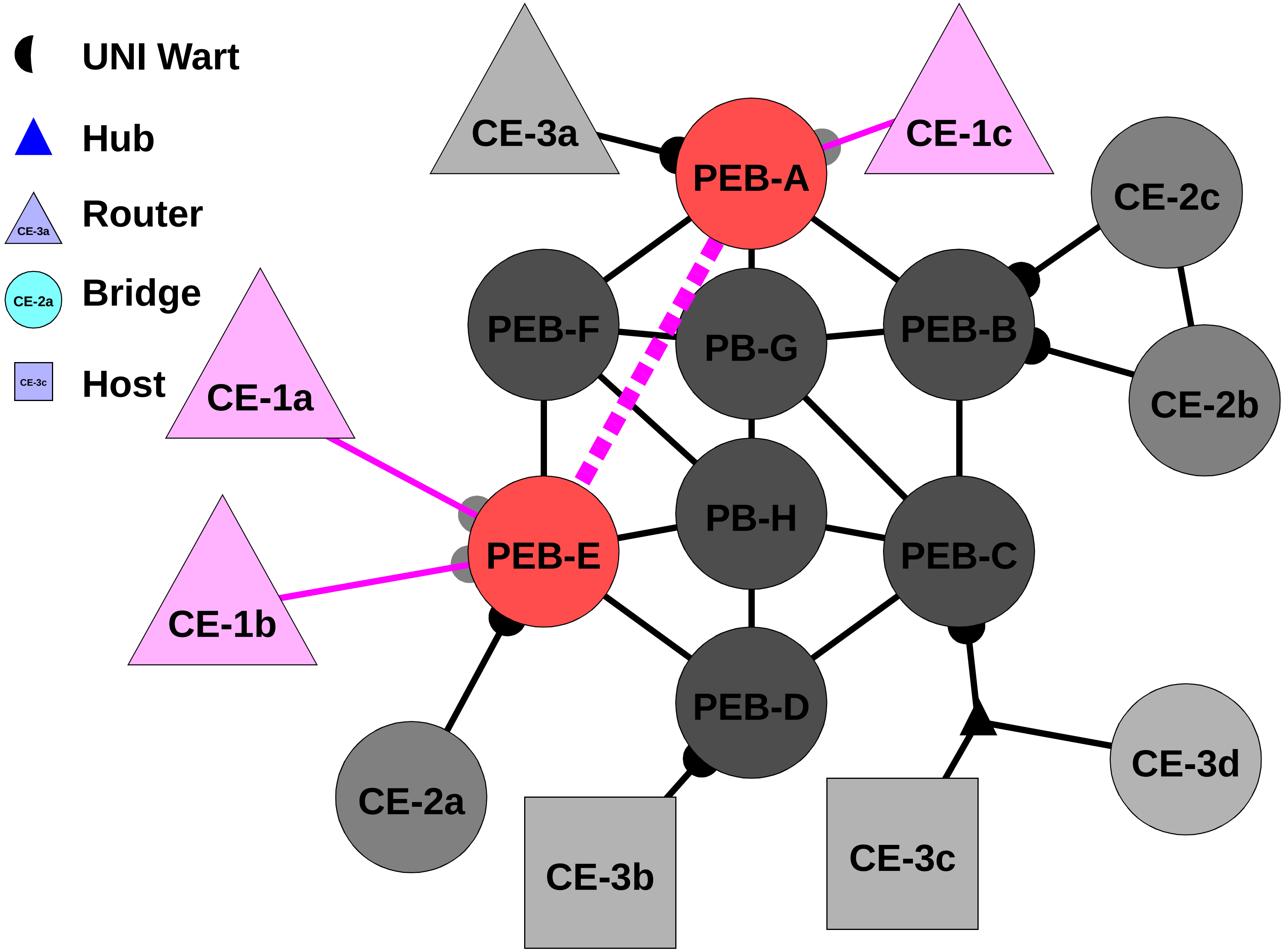


PEB: Provider Edge Bridge

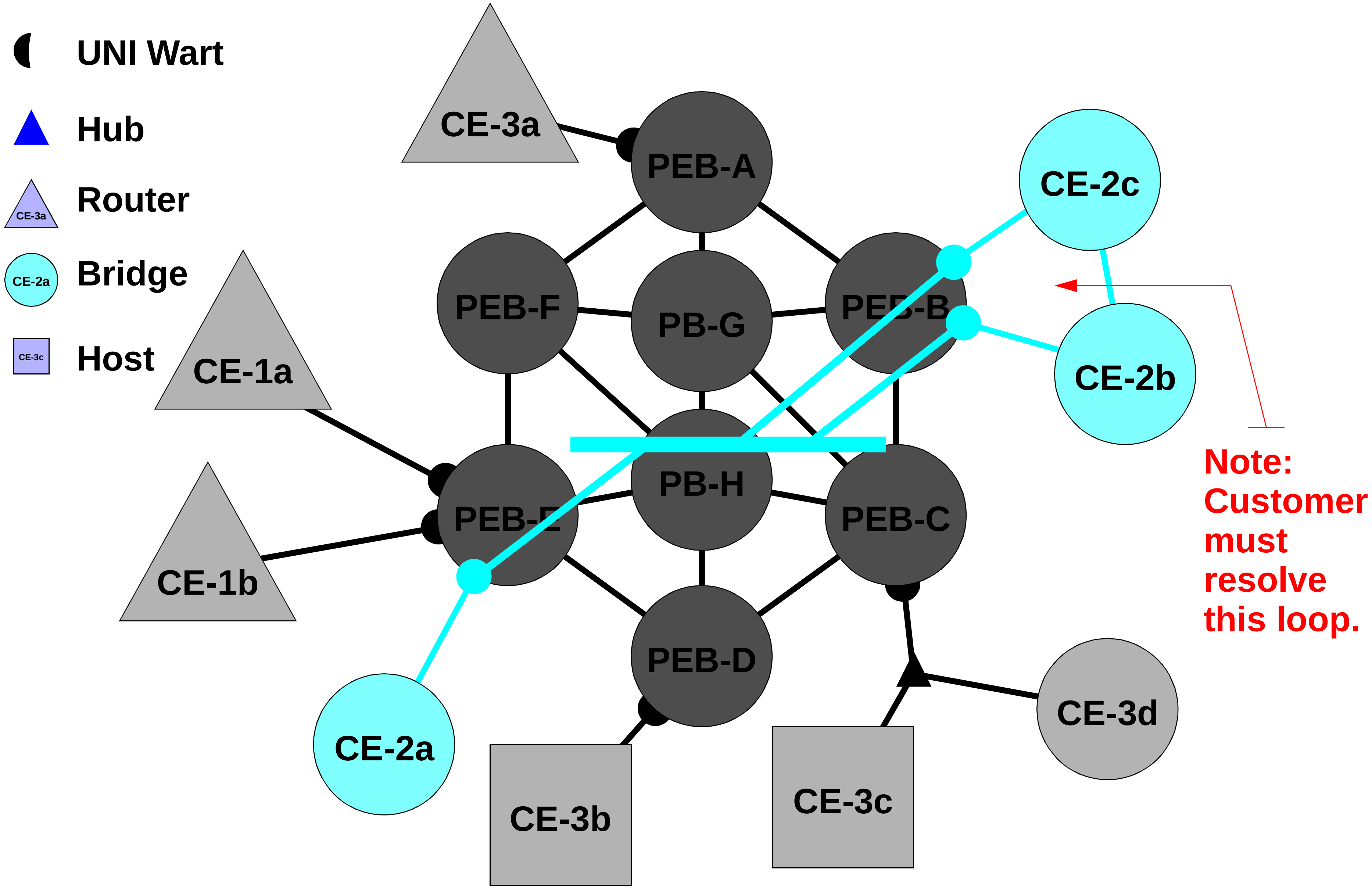
PB: Provider Bridge

CE: Customer Equipment

Customer 1 Sees Bridges:

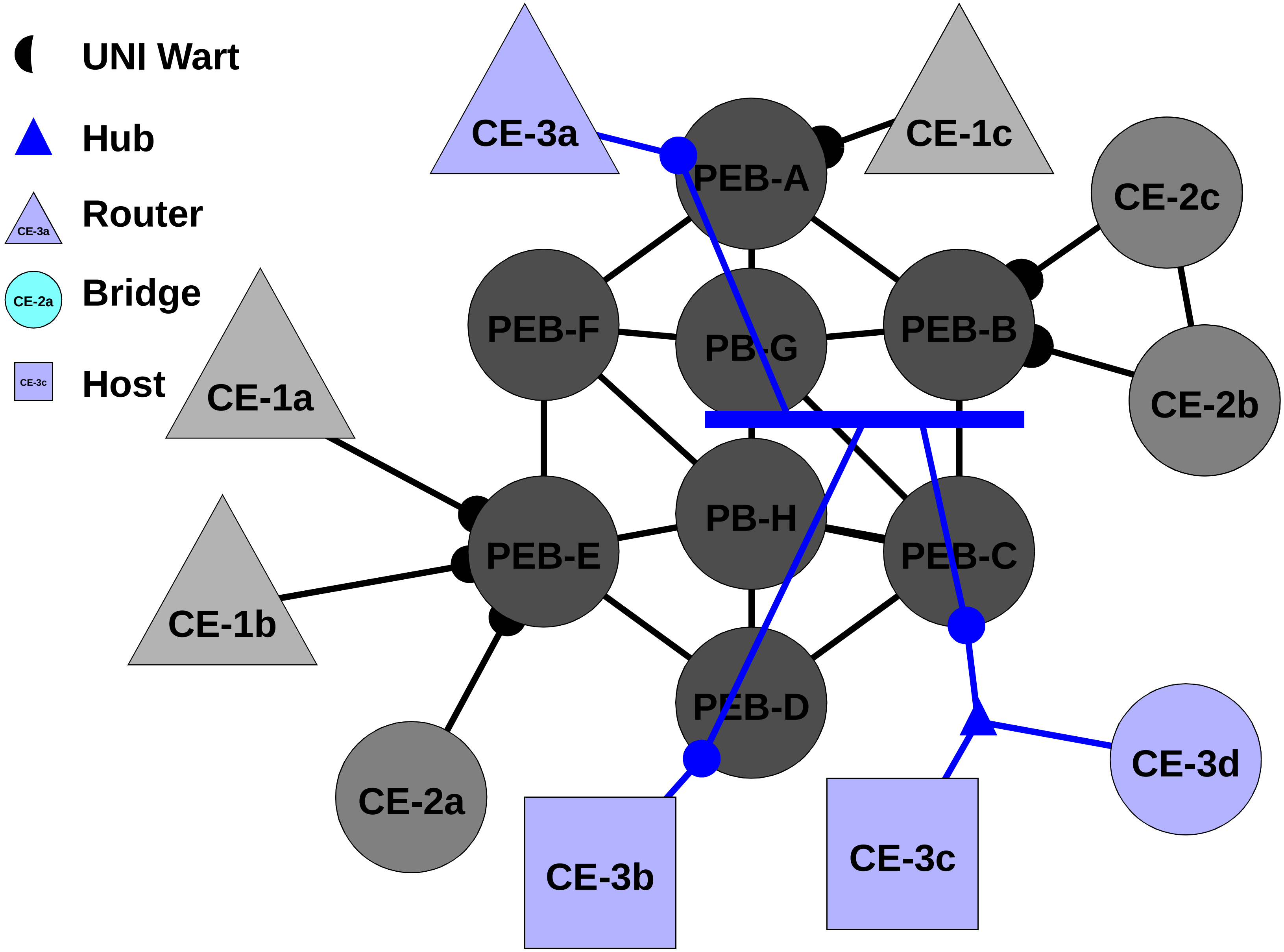


Customer 2 Sees a Shared LAN

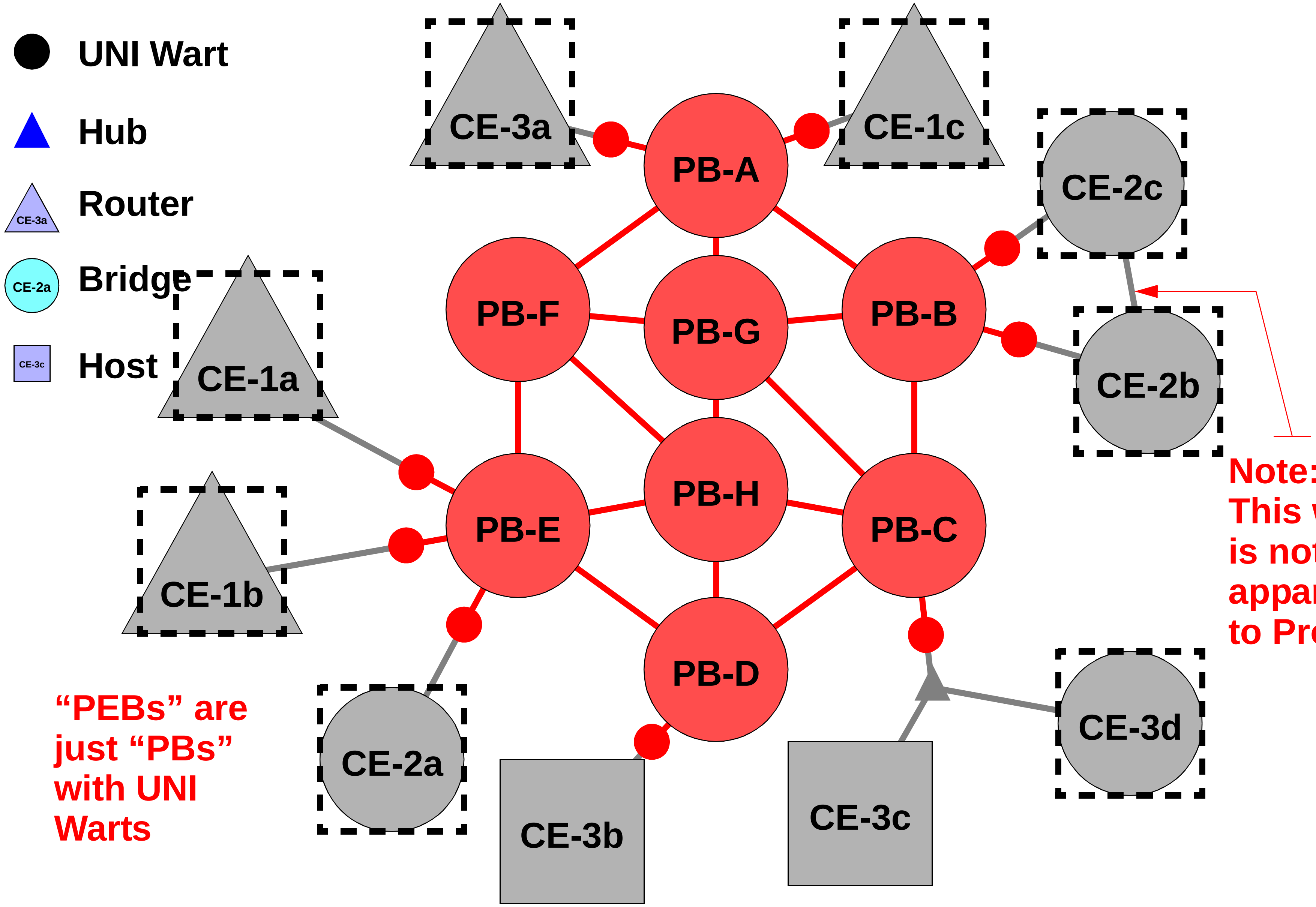


Note: Customer must resolve this loop.

Customer 3 Sees Another LAN



Provider Sees Access Ports + UNI Warts



Note:
This wire
is not
apparent
to Provider

“PEBs” are
just “PBs”
with UNI
Warts

UNI Wart Controls L2 Protocols

- Provider does *not* transmit BPDUs towards Customer.
- Customer's BPDUs may be discarded (bridge-like service) or transported transparently across Provider's network (wire-like service).
- IEEE 802.3x link pause packets must be handled normally, by hardware, even on a wire-like service.
- IEEE 802.1X link authentication might well be used between Customer and Provider to authenticate the use of a bridge-like service.
- Other protocols vary between discard, transparently transmit, or Provider/Customer handshake.

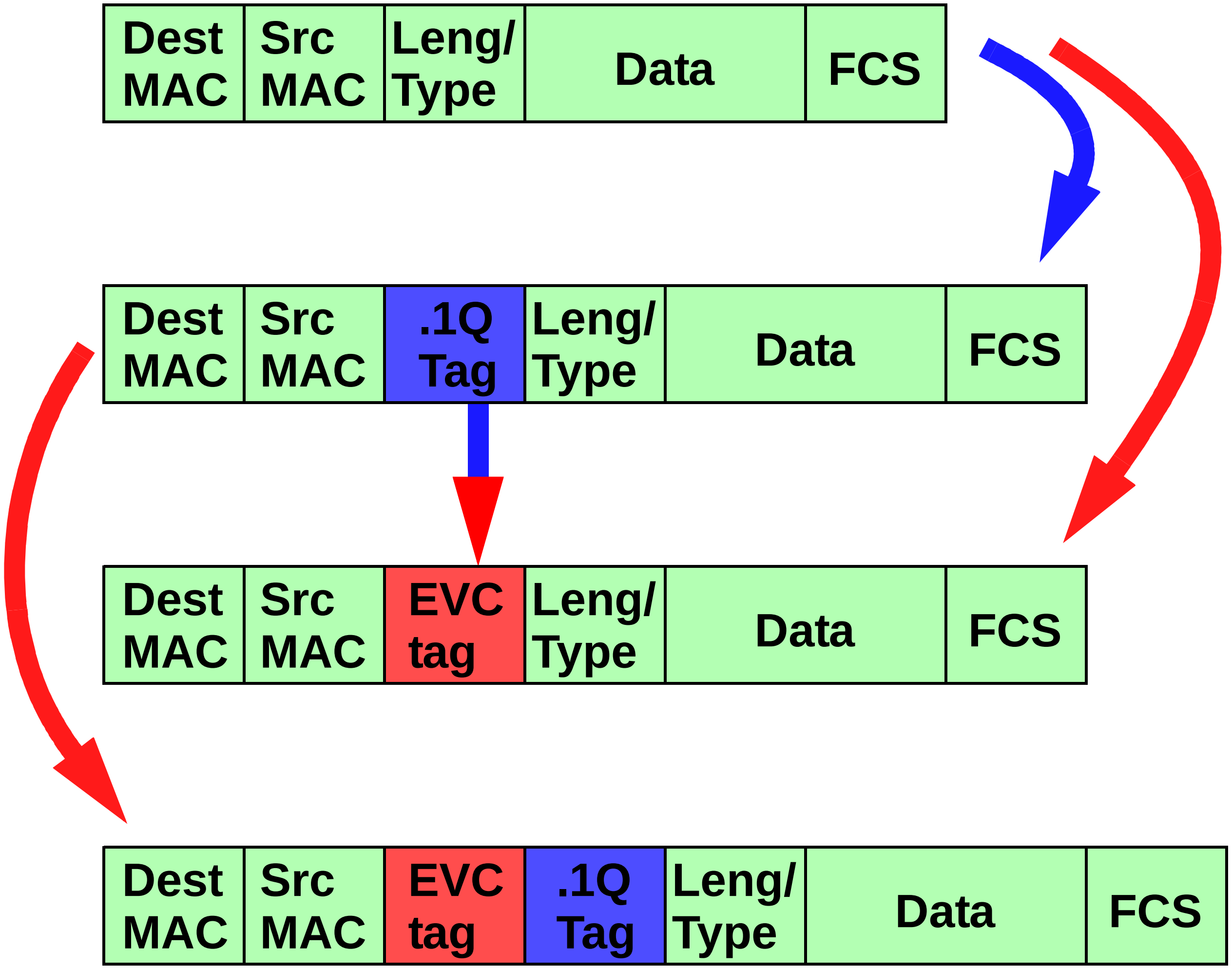
Layer-2 Protocols across UNI

Protocol Dest. MAC Addr.	Bridge-like Service	Wire-like service
GMRP	?	Transmit
GVRP	?	Transmit
Other GARP	Discard	Transmit
BPDUs: STP, RSTP, MSTP	Discard	Transmit
“All Bridges” MAC address	Discard? Handshake?	Transmit
802.3x Pause	Handshake?	Discard or Handshake
Link Aggregation	Discard? Handshake?	Transmit
802.1X Port Authentication	Discard? Handshake?	Transmit
Other “slow protocols”	Discard?	Transmit
LLDP	Discard? Handshake?	Transmit
Various proprietary protocols	?	?

Customers' and Provider's Spanning Trees do *not* Interoperate

- The complete spanning tree could be arbitrarily large.
- Customers mustn't purposefully nor accidentally hijack Provider's spanning tree.
- To a first approximation, if Customer has a loop which passes through the Provider's network, Customer uses (and pays for) the maximum bandwidth. (Of course, Provider will assist Customer to do better than that.)
- Customer spanning tree BPDUs traversing a Provider network, which necessarily is slower and less reliable than a wire, raises issues that must be identified, avoided, and/or resolved.

Frame Formats



Original frame from PC

Dest MAC	Src MAC	Leng/ Type	Data	FCS

Customer’s bridge may add 802.1Q tag

Dest MAC	Src MAC	.1Q Tag	Leng/ Type	Data	FCS

Dest MAC	Src MAC	EVC tag	Leng/ Type	Data	FCS

PEB may add EVC tag or may translate .1Q tag to EVC tag

Dest MAC	Src MAC	EVC tag	.1Q Tag	Leng/ Type	Data	FCS

“EVC tag” may be additional .1Q tag, or may be something new

[Aside] Standard 802.1Q VLAN Tagging

- **INGRESS: Static-configured default VLAN-IDs:**
  - **Untagged Internet Protocol packets default (802.1v): 6**
  - **Default Port VLAN (802.1Q PVID): 7**

802.1Q tag	Protocol	Assigned to VLAN:
untagged	IP	6
untagged	other (e.g. NetBEUI)	7
VLAN-ID == 0 (.1p)	any	6 or 7, as for untagged
VLAN-ID == 5	any (even IP!)	5
VLAN-ID == 6	any (even NetBEUI!)	6
VLAN-ID == 12	any	12

- **EGRESS: Static-configured which VLANs strip tags.**

[Aside] Standard 802.1Q VLAN Tagging

- **What about BPDUs?**
  - — **Untagged BPDUs are assigned to no VLAN!**
  - — **There is no such thing (in the standards) as a tagged BPDU.**
  - — **GARP frames are treated like data frames w.r.t. tagging.**
- **Transformations which can be configured:**
  - — **Tagged with VLAN-ID *X* on ingress, same on egress.**
  - — **Tagged with VLAN-ID *X* on ingress, no tag on egress.**
  - — **No tag, or tagged VLAN-ID = 0 on ingress, no tag on egress.**
  - — **No tag, or tagged VLAN-ID = 0 on ingress, tagged *X* on egress.**
- **Transformations not allowed in IEEE 802.1Q:**
  - — **Tagged with VLAN-ID *X* on ingress, tagged with *Y* != *X* on egress.**
  - — **Anything tagged with VLAN-ID = 0 on egress.**

Simplified View: The UNI Wart

Customer side

UNI Wart

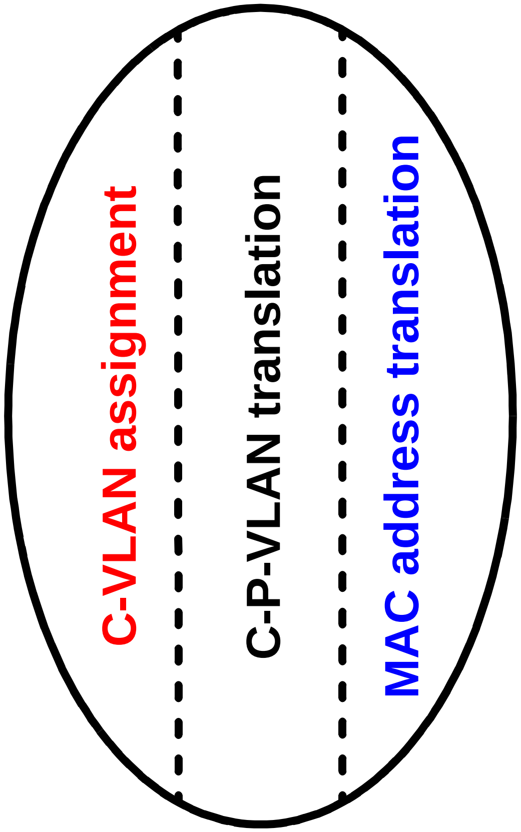
Provider side

Ingress direction ->

<- Egress direction

C-VLAN IDs

P-VLAN IDs



C-VLAN tagged control
0-tagged control
Untagged control
C-VLAN tagged data
0-tagged data
Untagged data

Untagged control
P-VLAN tagged control
P-VLAN tagged data

(MAC address translation could be in any order relative to the other two functions.)

UNI Wart: Ingress (1)

- **C-VLAN ID Assignment (standard 802.1Q behavior)**
  - — **Every data frame is assigned to a C-VLAN, perhaps using 802.1v.**
  - — **Every control frame is assigned either to a C-VLAN or assigned to no VLAN.**
  - — **If the frame is untagged and assigned to a C-VLAN, an 802.1Q tag for that C-VLAN may be added.**
- **C-P-VLAN Translation**
  - — **Depending on C-VLAN ID (or “none”), the UNI Wart either:**
    1. **Translates the C-VLAN ID to a P-VLAN ID; (This may be the “null” translation, which changes nothing.)**
    2. **Adds a P-VLAN ID just after the Source MAC address; or**
    3. **Discards the frame.**
  - — **Priority field handled in a similar manner.**

UNI Wart: Ingress (2)

- **MAC Address Translation**
  - **(This is needed for some methods for carrying Customer BPDUs across the Provider network. One method discussed, here.)**
  - **If the frame belongs to a protocol configured for “handshake” with the Provider bridge, the frame is unchanged.**
  - **If the frame belongs to a protocol configured to be “discarded”, the frame is discarded.**
  - **If the frame belongs to a protocol configured to be “transmitted”, the destination MAC address is translated to another value in the “Provider Only” range.**
  - **If the frame’s destination MAC address is in the “Provider Only” range, it is discarded.**

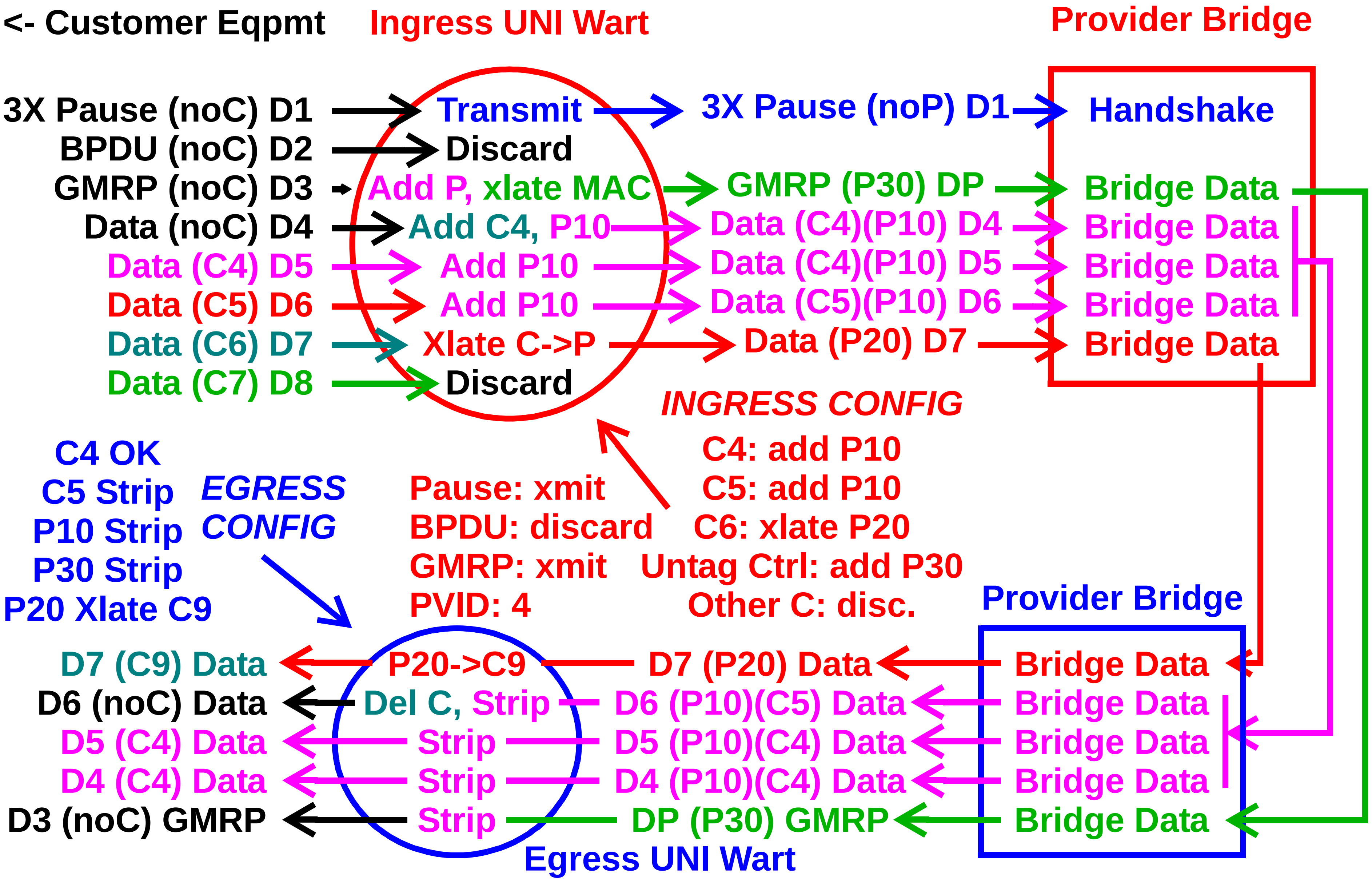
UNI Wart: Egress (1)

- **MAC Address Translation (Again, one Possibility)**
  - — **If the frame belongs to a protocol configured for “handshake” with the Provider bridge, the frame is unchanged.**
  - — **If the frame belongs to a protocol configured to be “discarded”, the frame is discarded.**
  - — **If the frame’s destination MAC address is in the “Provider Only” range, it is to be “transmitted”. The destination MAC address is translated to the appropriate value for the protocol.**
- **C-P-VLAN Translation**
  - — **Depending on P-VLAN ID (or none), the UNI Wart either:**
    1. **Strips the P-VLAN tag off; or**
    2. **Translates P-VLAN ID and priority to C-VLAN values; or**
    3. **Passes the frame through (for untagged handshake packets).**

UNI Wart: Egress (2)

- C-VLAN ID Assignment (standard 802.1Q behavior)
  - Depending on the C-VLAN ID, the 802.1Q tag may be removed before transmission, perhaps using 802.1v.

A Complex UNI Example



<- Customer Eqpmt

Ingress UNI Wart

Provider Bridge

3X Pause (noC) D1 → Transmit → 3X Pause (noP) D1 → Handshake

BPDU (noC) D2 → Discard

GMRP (noC) D3 → Add P, xlate MAC → GMRP (P30) DP → Bridge Data

Data (noC) D4 → Add C4, P10 → Data (C4)(P10) D4 → Bridge Data

Data (C4) D5 → Add P10 → Data (C4)(P10) D5 → Bridge Data

Data (C5) D6 → Add P10 → Data (C5)(P10) D6 → Bridge Data

Data (C6) D7 → Xlate C->P → Data (P20) D7 → Bridge Data

Data (C7) D8 → Discard

INGRESS CONFIG

Pause: xmit
BPDU: discard
GMRP: xmit
PVID: 4

C4: add P10
C5: add P10
C6: xlate P20
Untag Ctrl: add P30
Other C: disc.

C4 OK
C5 Strip
P10 Strip
P30 Strip
P20 Xlate C9

EGRESS CONFIG

Provider Bridge

D7 (C9) Data ← P20->C9 ← D7 (P20) Data ← Bridge Data

D6 (noC) Data ← Del C, Strip ← D6 (P10)(C5) Data ← Bridge Data

D5 (C4) Data ← Strip ← D5 (P10)(C4) Data ← Bridge Data

D4 (C4) Data ← Strip ← D4 (P10)(C4) Data ← Bridge Data

D3 (noC) GMRP ← Strip ← DP (P30) GMRP ← Bridge Data

Egress UNI Wart

Notable Points in Slide 32 Example (1)

[D*x* = dest. MAC address *x*, (C*y*) = C-VLAN tag, (Pz) = P-VLAN tag.]

- **Ingress UNI Wart had to distinguish among untagged frames 3X Pause (noC) D1, BPDU (noC) D2, GMRP (noC) D3, and Data (noC) D4, based on protocol, in order to decide what to do with them.**
- **Data (noC) D4 entered with no 802.1Q tag, but left with a (C4) tag. Data (C5) D6 entered with an 802.1Q tag, but left with no tag. This is normal bridging.**
- **Ingress C-VLANs 4 and 5 are "bundled" together in P-VLAN 10.**
- **Data (C7) D8 discarded: no service exists for C7. BPDU (noC) D1 discarded: BPDUs are not allowed.**

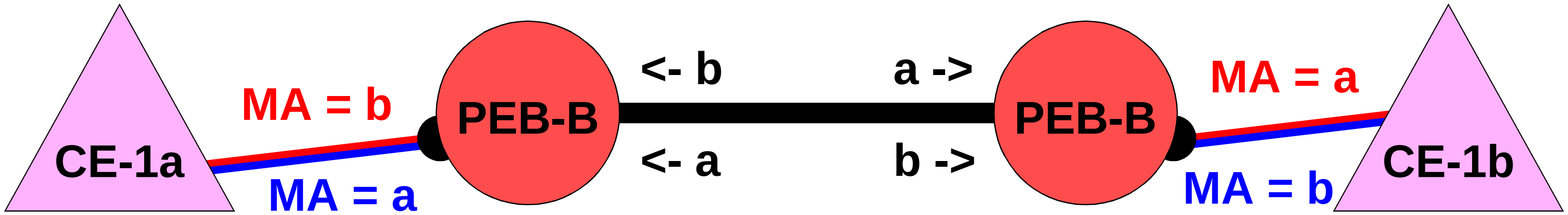
Notable Points in Slide 32 Example (2)

- **Data (C6) D7 entered with one tag, but left with another tag (C9). This is not normal IEEE 802.1Q bridging, but it is:**
  - **often desirable, e.g. connecting an ISP to a subscriber; and**
  - **hard to prevent, given that the C-P-translation is configured in two different UNI Warts.**
- **Why so much flexibility?**
  - **Translation is easy to describe: Current bridge port rules, plus two 4k x (12+1) tables of VLANs and add/xlate bits, plus items necessary for transporting Customer BPDUs.**
  - **Separating the three parts of the UNI Wart allows “bundling” to be used for directing streams to appropriate Provider Edge Bridges, independently of wire-like/bridge-like service selection.**

Notable Points in Slide 32 Example (3)

- **Full service is described. Useful service can be performed by UNI Wart without C-P-translation:**
  - **Ability to add/strip extra P-VLAN tag to all frames allows a single “virtual wire” service over one UNI.**
  - **Lack of translation capability between C-VLANs and P-VLANs means that Provider must specify C-VLAN (= P-VLAN), which is acceptable to many Customers for many uses.**
  - **The lack of a destination MAC address translation capability would make transparent BPDU transmission difficult.**

Well-Known Problems and Solutions (1)



- MAC addresses are not unique. VRRP causes routers on opposite sides of the network to have the same locally-administered MAC address. *Solutions*:
  - — Don’t do that, i.e. reconfigure the routers.
  - — Assign any pair of C-VLANs for which this problem could occur to different P-VLAN bundles.
  - — Configure all of the Provider bridges with all of the Customers’ Independent/Shared Learning Constraints and have them learn both the P-VLAN tag and the C-VLAN tag.
  - — Be very efficient at VLAN pruning, and don’t learn MAC addresses at all; flood all frames within the P-VLAN.

Well-Known Problems and Solutions (2)

- **Customers may use any number of standard or proprietary multiple spanning tree protocols, resulting in one MAC address being behind two different UNIs on different C-VLANs.**
  - — **Disallow multi-homed Customers. (That is, disallow any given Customer Layer 2 network from having two UNIs.)**
  - — **Assign to each P-VLAN bundle only C-VLANs sharing the same spanning tree, or at least the same STP parameters.**
  - — **Configure all of the Provider bridges with all of the Customers' Independent/Shared Learning Constraints, and have them learn both the P-VLAN tag and the C-VLAN tag.**
  - — **Be very efficient at VLAN pruning, and don't learn MAC addresses at all; flood all frames within the P-VLAN.**

Well-Known Problems and Solutions (3)

- A spanning-tree topology change in the Customer’s network may mean that the Provider needs to forget MAC addresses on just one or a few P-VLANs.
  - — Disallow multi-homed Customers.
  - — Be very efficient at VLAN pruning, and don’t learn MAC addresses at all; flood all frames within the P-VLAN.
  - — Monitor the Customers’ traffic for the various flavors of standard and proprietary spanning tree topology change notifications, and trigger MAC address forgetting in the Provider’s network.
- Flooding every Customer frame to all that Customer’s UNIs is unacceptable in Wide-Area Networks.
  - — *DO NOT:* Be very efficient at VLAN pruning, and don’t learn MAC addresses at all; flood all frames within the P-VLAN.

Using 802.1Q Tags to Separate Customers

- Using 802.1Q tags on standard bridges within the Provider’s network has certain limitations:
  - No more than 4k - 2 Customers per 802.1Q-bridged network.
  - All of the C-VLANs bundled together in one P-VLAN must be able to share the same Filtering Database. That requires unique MAC addresses, and a single Customer spanning tree instance.
- But .1Q has great advantages over any new tag:
  - Changes to bridges (double tags, UNI Warts) are minimal.
  - Some P-VLANs can be “normal” VLANs used to carry normal traffic, e.g. for management traffic or Customer Layer 3 services.
  - Solution for one problem, e.g. larger networks, is applicable to all bridged networks, and not just to Ethernet Providers.
  - Existing solutions, e.g. 802.1p priority, are usable by Providers.

Enlarging Provider Networks

- ***Many*** **techniques have been implemented, and even more discussed, to expand the size of a bridged network.**
  - — **Separate spanning trees at edges which run on top of central spanning tree.**
  - — **Running two disconnected spanning trees in one bridge.**
  - — **Substituting hop count for Max Age in Rapid Spanning Tree. (Standardized by 802.1y: 64k hops allowed across network!)**
  - — **Enforce topology restrictions by some non-spanning tree means.**
- **We will look at just one: Topology restrictions.**
  - — **This also promises to interact well with the IETF solution(s).**

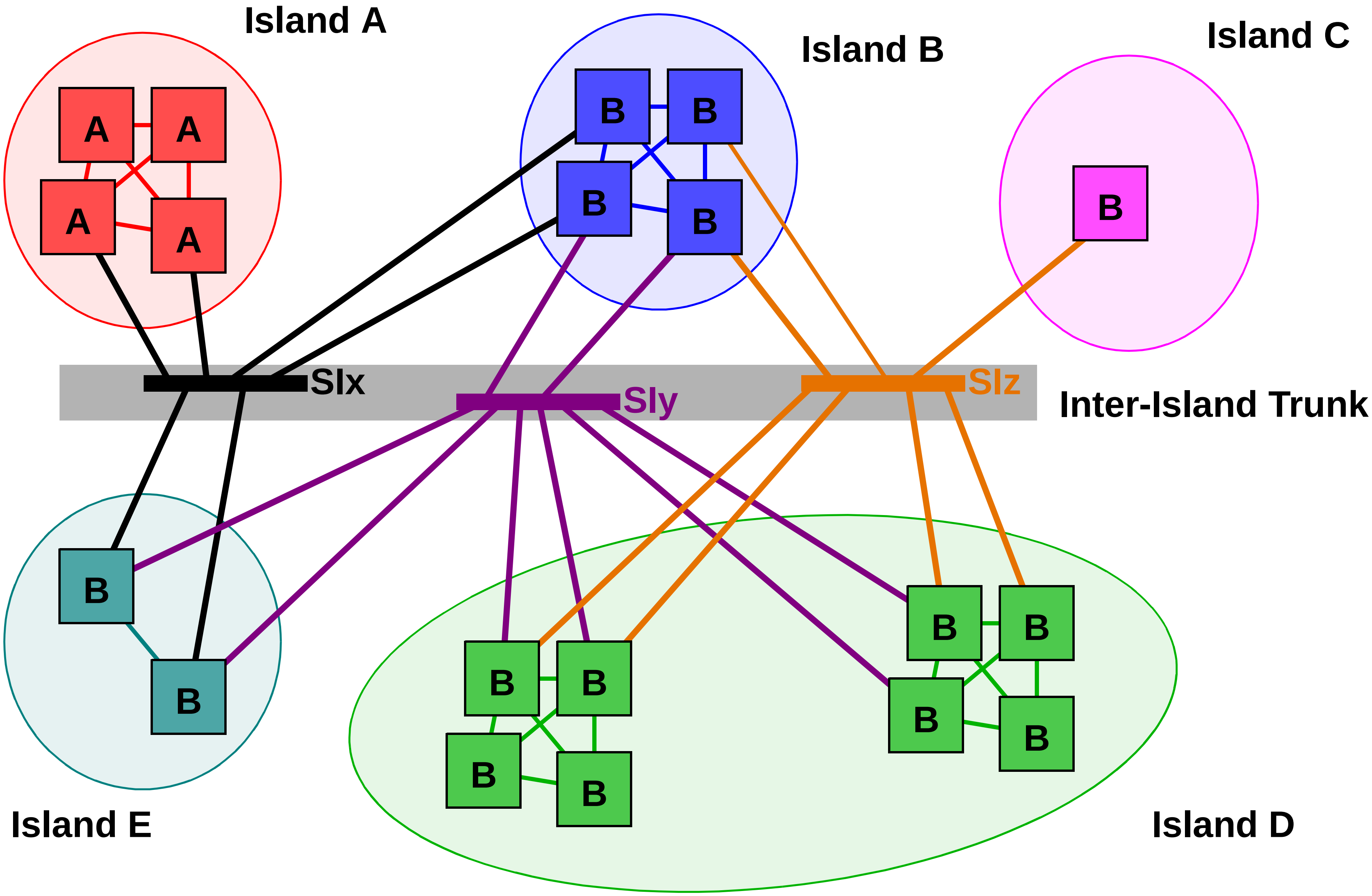
SIs, Islands, and Inter-Island Trunks (1)

- **A “Service Instance” is the analog, in a very large network, of a VLAN in an 802.1Q network.**
- **The Provider Network carries Service Instances via “Islands” connected by “Inter-Island Trunks”.**
- **Islands**
  - **An Island consists of one or more bridges connected by normal LAN segments and/or Inter-Island Trunks.**
  - **Different Islands must be connected *only* via Inter-Island Trunks.**
  - **Within any one Island, a given SI is identified by a unique P-VLAN ID.**
  - **Two different Islands may have completely independent associations between P-VLANs and SIs.**

SIs, Islands, and Inter-Island Trunks (2)

- **Inter-Island Trunks**
  - — **An Inter-Island Trunk behaves, logically, like a Shared Medium LAN segment.**
  - — **Among different Islands, a given Service Instance is carried over, at most, one Inter-Island Trunk.**
  - — **The identification of a SI on an Inter-Island Trunk is dependent on the medium, e.g. Ethernet, or emulated Ethernet over layer 3 tunnels.**
  - — **Any bridge port connected to an Inter-Island Trunk must make 1:1 translations between the P-VLAN IDs used within the bridge’s Island and the SI identifiers used on the IIT.**
- **These rules prevent SI loops among Islands.**
- **Spanning trees prevent SI loops within an Island.**

SIs, Islands, and Inter-Island Trunks (3)



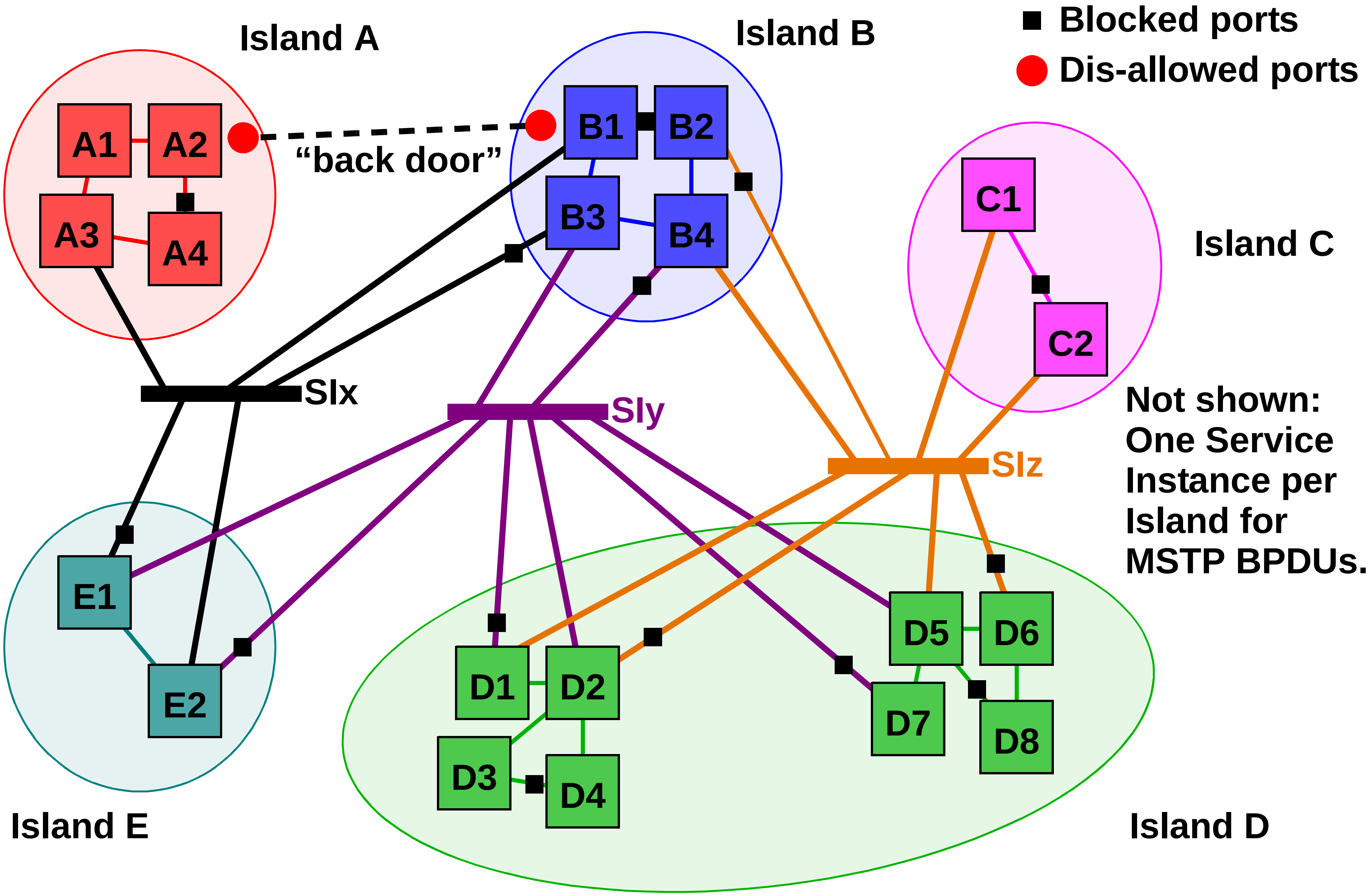
Interconnecting Islands(1)

- A separate MSTP Service Instance per Island carries MSTP BPDUs.
  - Bridges in the same Island interchange BPDUs and guarantee that, for any given frame on a P-VLAN, only one copy will be transmitted to the Inter-Island Trunk.
  - Similarly, by interchanging BPDUs, the bridges can guarantee that a frame on an Inter-Island Trunk will be delivered, at most, once to any LAN segment within the Island.
  - Since we are guaranteed (See Slide 51) that no SI exists on more than one Inter-Island Trunk, and that there are no back doors, failing to receive the BPDUs from bridges in other Islands cannot cause a loop, but does manage to limit the size of any one Spanning Tree Instance.

Interconnecting Islands(2)

- **A new Topology Change Notice is required.**
  - — **Islands may have completely different MSTP configurations.**
  - — **Forgetting all addresses is too much, but we must forget some.**
  - — **The other “Island” may, in fact, be an IETF implementation.**
  - — **This new Topology Change Notice must signal topology changes based on Service Instance (Inter-Island Trunk).**

Interconnecting Islands (3)

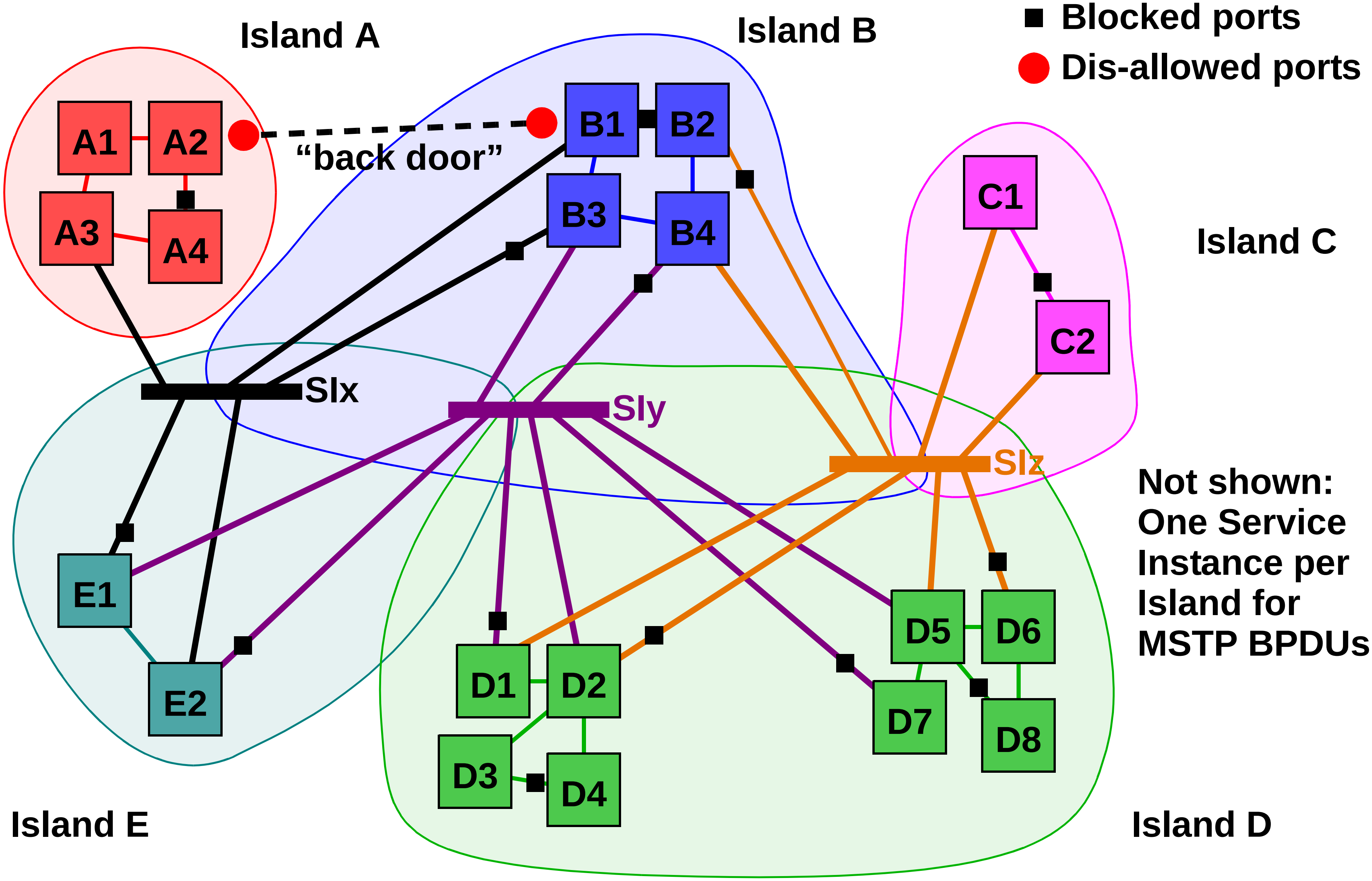


Not shown: One Service Instance per Island for MSTP BPDUs.

Interconnecting Islands (4)

- (See Slide 46) Island A has only one connection to SIx. If it is lost, the Island loses data connectivity with Islands B and E. It must not use the “back door” link to Island B; the two Islands may have different mapping between SIs and P-VLAN IDs.
- Island B has arranged to “load share” its connections to other islands; each bridge B1, B3, and B4 has one.
- Island C has elected to connect C1 and C2 via SIz, rather than using its internal LAN connection.
- Island D is forced to use SIy and SIz to connect the two parts of the Island. D5-D8 are not load-sharing; only D5 is active on the Inter-Island Trunk.

Why Do Bridges Need EEoMPLS? (1)



Not shown:
One Service Instance per Island for MSTP BPDUs.

Why Do Bridges Need EEoMPLS? (2)

- **Slide 48 is a copy of Slide 46, but gives a more accurate picture of the Islands as viewed by bridges.**
- **One MSTP SI per Island extends that Island’s spanning tree to include the EEoMPLS shared media, but not the bridges in other Islands.**
- **This permits the bridges with each Island to see each other, and thus unblock just the right connections to the SIs to enable connectivity, yet prevent loops.**
- **If the emulated shared media were replaced by some combination of point-to-point links, we would have to run spanning trees throughout the Provider’s network in order to prevent loops through multiple Islands.**

SIs and EEoMPLS Instances

- **Does one EEoMPLS instance carry one SI or many?**
  - **In some drafts, MPLS is used to establish a full mesh of Label Switched Paths (LSPs) among all devices in the union of a set of EEoMPLS instances (Service Instances).**
  - **Inside each point-to-point MPLS path, a number of point-to-point Virtual Circuits (VCs) can be carried, up to one per SI.**
  - **Not every port in the MPLS mesh needs a VC for every SI.**
  - **For each SI, the devices construct a full mesh of VCs among the participating devices, which is a subset (or is equal to) the full MPLS mesh.**
  - **In this plan, the MPLS mesh is the shared medium, and the SIs are the VLANs.**
  - **802.1 adds an additional SI for each Island for untagged control traffic (MSTP) among all of the Island's bridges.**

Provisioning

- How do we guarantee that any given Service Instance is carried over at most one Inter-Island Trunk?
  - Short answer: We rely on IETF.
  - Long answer: Current IETF proposals envision a control plane such that, if each Provider bridge participates in only one “universe” of control, all L3 connected bridges in a Provider’s network will, eventually, have the same SI/IIT configuration.
- How are Islands created?
  - By configuring individual bridges with SI / P-VLAN maps and MSTP Service Instances.

IETF - 802.1 Interoperability

- **In some scenarios, each EEoMPLS port is tied to a single specific UNI. (Not their terminology!)**
  - **In this scenario, the IETF is *not* reinventing bridging; the EEoMPLS port to UNI connection requires no MAC address learning or forwarding.**
  - **This provides a perfect point of interoperability between IETF and IEEE 802.1. IETF ports on one side of the EEoMPLS instance would interoperate with bridges on the other side.**
- **EEoMPLS is a good fit for the Inter-Island Trunk.**
  - **It supports a large number of SIs, one EEoMPLS per SI.**
  - **If an SI can *only* be on an Inter-Island Trunk, then it is easy to distinguish between an IIT port and an interior port.**
  - **The ubiquity of the Layer 3 world makes it easy to scale one Provider network or to interconnect Providers’ networks.**

Plenty of Room for Discussion

- **For “bridge-like” UNI, which protocols are “handshake”? Which are “discard”? Optional?**
- **Does “UNI Wart” model include MAC/PHYs, or does it act instantaneously to transform frames on the wire?**
- **Is 802.1Q tag sufficient for all P-VLAN purposes? (Certainly it must be allowed as a P-VLAN tag.)**
- **How do we transport Customer BPDUs? MAC address mapping? (Corollary: *Should* we transport them?)**
- **What other features (e.g. ping) should Wart support?**
- **Will 802.3 let us add another tag?**

A Few Stray Puzzle Pieces

- **Providing services to Customers with service agreements highlights some other components of the solution, as well as features long missing in 802.1D:**
  - — **There is no practical bridging equivalent to the information obtainable from routers by the Layer 3 “Traceroute” function.**
  - — **There is no practical equivalent to the Layer 3 “Ping” function.**
  - — **There is no equivalent to the “Link Management Interface” of ATM or Frame Relay.**
- **If one were well-versed in telephony techniques, one could perceive many more missing features.**
  - — **Many, however, would insist that this lack *is* a feature.**

Summary

- Islands of independent bridges carry Customer’s VCs using Q-over-Q tags.
- Islands are interconnected using Emulated Ethernet over MPLS.
- Bridges within one Island run MSTP to prevent loops without requiring a universal spanning tree.
- The IETF defines Emulated Ethernet over MPLS, whether they admit it or not.
- Bridging solution interoperates perfectly with PPVPN solution because bridges can connect to an Emulated Ethernet over MPLS function.

Recent Communication

- **From: Loa Andersson <loa.andersson@utfors.se>, editor of L2VPN requirements document:**

```
+-----+      From L2VPN Requirements:    +-----+
+ CE1 +--+                           +---| CE2 |
+-----+  |    ...................    |   +-----+
 VPLS A  |  +----+           +----+  |    VPLS A
         |  |VPLS|           |VPLS|  |
         +--| PE |--Service--| PE |-+
            +----+  Provider +----+
           /  .     Backbone    .  \     -   /\-_
+-----+   /   .      |          .   \   / \ /   \     +-----+
+ CE  +--+    .      |          .    +--\ Access \----| CE  |
+-----+       .    +----+       .       | Network |   +-----+
 VPLS B       .....|VPLS|........        \       /     VPLS B
                   | PE |     ^           -------
                   +----+     |
                     |        |
                     |        |
                  +-----+     |
                  | CE3 |     +-- Logical bridge
                  +-----+
                   VPLS A
```

Andersson’s version of 802.1’s needs:

```
            +--------------------------------------------+
            |               provider network             |
            | +-------------+                +---------+ |
            | |  vppls-pe   |                | vpls-pe | |
  vpls1    +-----+          |                |      +-----+    vpls2
+-----+    | vsi |.......   |----------------|      | vsi |   +-----+
| ce1 |....|     |   ,,,,,,,|,,,, tunnel ,,,,|,,,,,,|     |,,,| ce2 |
+-----+    |     |.. ,  .   |----------------|   ,,,|     |   +-----+
           +-----+ . ,  .   |                |   ,  +-----+
            | |    . ,  .   |                |   ,     | |
  vpls2    +-----+ . ,  .   |                |   ,     | |
+-----+    | vsi | . ,  .   |                |   ,     | |
| ce3 +,,,,|     |,,,   .   |-------         |   ,     | |
+-----+    |     |,.,,,,.,.,|.,.,.,.\        |   ,     | |
           +-----+ .        |------ t\       |   ,     | |
            | |    ....     |       \u\      |   ,     | |
            | +-------------+        \n\     +---------+ |
            |        |.|              \n\      |,|      |
            |        |.|               \e\     |,|      |
            | +-------------+            \l\  +---------+ |
            | |   vppls-pe  |             \,\-| vpls-pe | |
            | |      .      |              \.,|., ,     | |     vpls2
  vpls1    +-----+    .     |              --| . ,  +-----+    vpls1
+-----+    | vsi |.....     |----------------| ,.,.,| vsi |   +-----+
| ce4 |....|     |..........|.... tunnel.....|......|     |.,.| ce5 |
+-----+    |     |          |----------------|      |     |   +-----+
           +-----+          |                |      +-----+
            | |             |                |         | |
            | +-------------+                +---------+ |
            |                                            |
            +-------------------------------------loa----+
```

Is Andersson right?

- Is that our model?
  - — Can we respond to him with a model as 802.1?
  - — As individuals?