Foundry Networks
Principal member of the MPLS Forum

Challenges Facing L2 VPN

MPLS Japan 11/8/2001

Gary Blankenship
Foundry Networks
Systems Engineer - Japan

FOUNDRY
NETWORKS

Agenda

➡️ L2 Technologies in MAN (Metro)

- MPLS VLL Draft-Martini

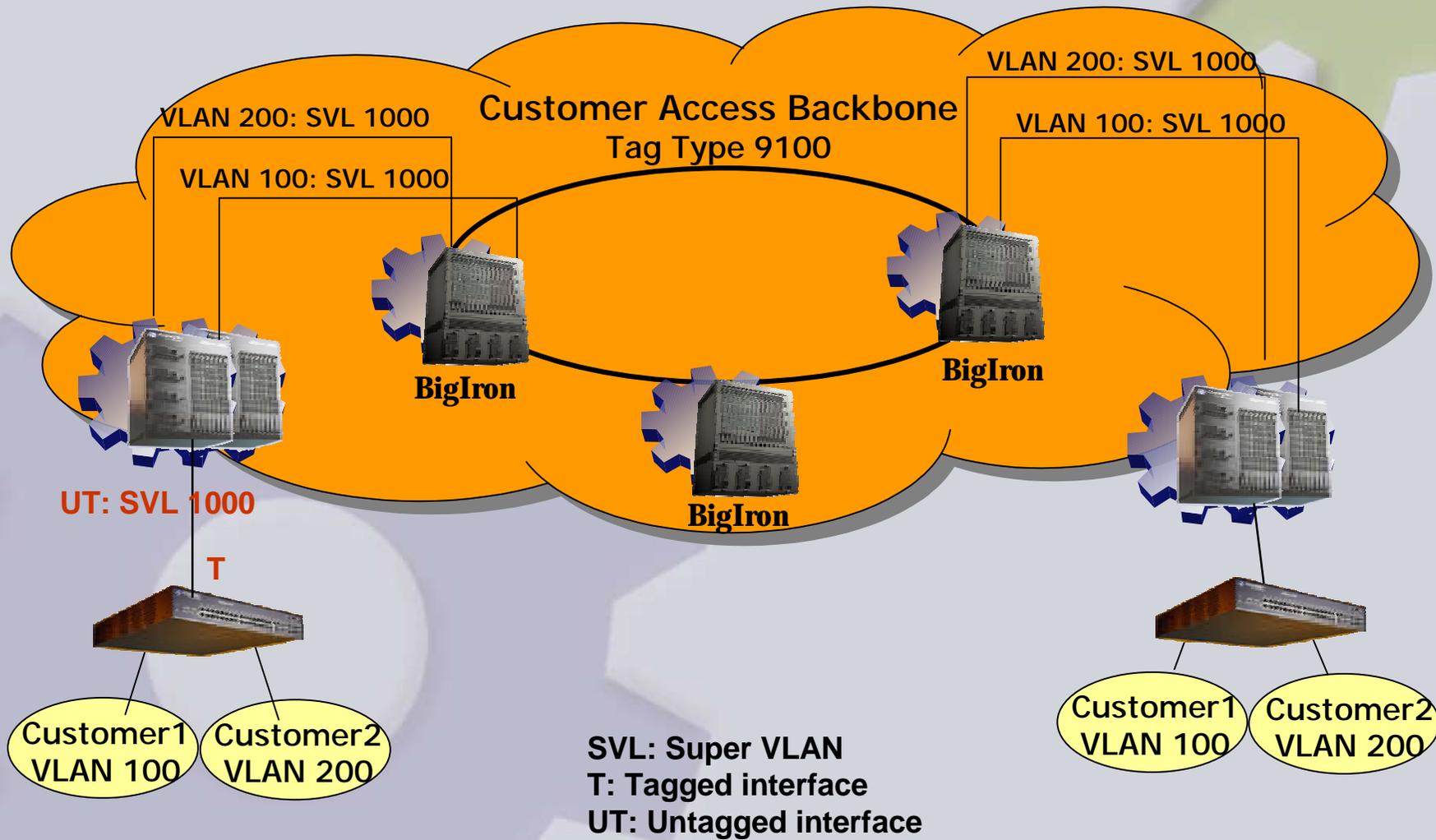- MPLS VPLS

- Interoperability

L2 Technologies in Metro Area Networks (MAN)

- 802.1Q allows a maximum of 4,096 unique VLANs
- SuperVLAN. A stack of two 802.1Q headers to be carried in an Ethernet frame. 16 million unique VLANs
  - Potential solution through VLAN tag stacking – 4096x4096 VLANs
  - Can work with Spanning Tree
  - Customers can define their own VLAN ID
  - Core of the network remains limited to 4,096 VLANs
  - Large number of MAC addresses in the core
  - Management of the VLAN hierarchy can become rather complex

L2 Technologies in Metro Area Networks (Cont.)

- SuperVLAN
  - Lacks visibility of individual customer traffic within the network core
    - Customers VLANs within the same Super VLAN must share its resources (Aggregation = Loss of 802.1p information)
    - Broadcast/unknown unicast traffic may be forwarded to more edge switches than necessary
  - Limited means for resources allocation and QoS treatment from 802.1Q
  - STP provides only crude traffic placement ability and is not straightforward to manage
- 802.1w – **Rapid Spanning Tree Protocol** for rapid convergence time of 1 to 5 seconds
- 802.1s – **STP per VLAN Group** provides VLAN and STP scalability, and utilizes dark fiber efficiently
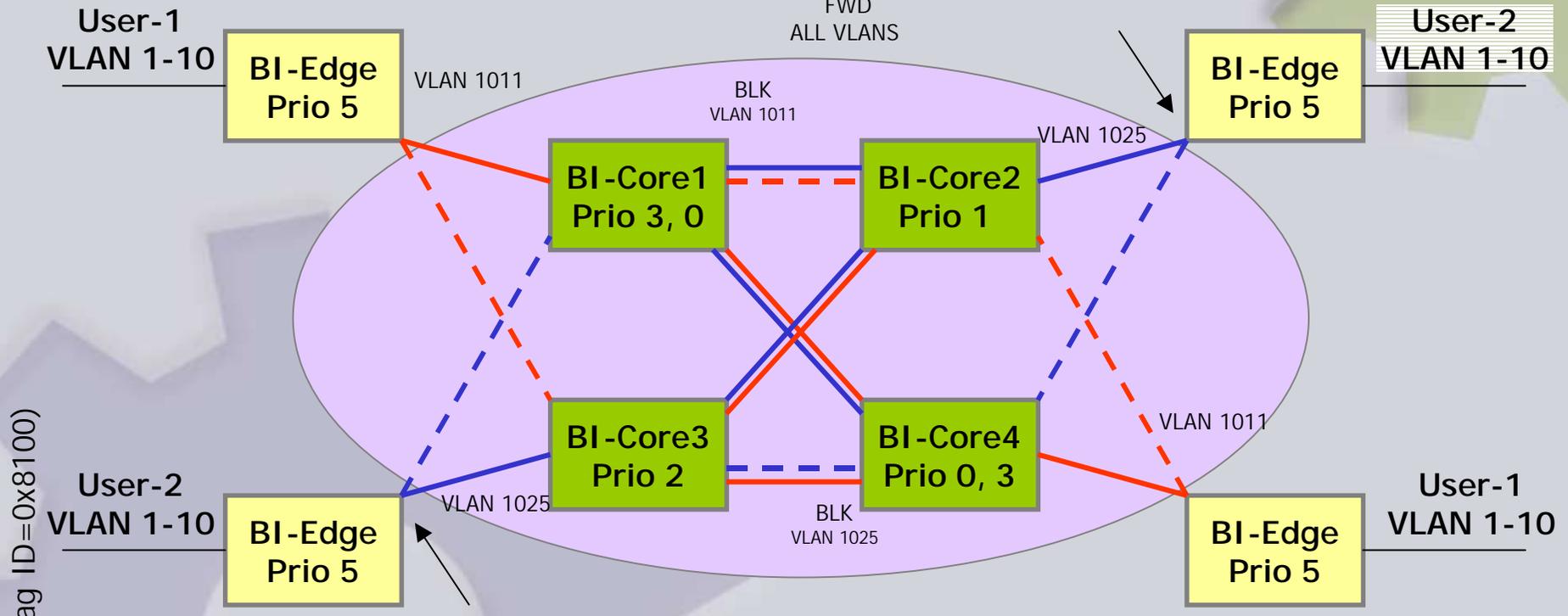
Super VLAN Aggregation Solution to VLAN Growth

VLAN 200: SVL 1000

VLAN 100: SVL 1000

VLAN 200: SVL 1000

Customer Access Backbone
Tag Type 9100

VLAN 100: SVL 1000

BigIron

BigIron

BigIron

UT: SVL 1000

T

Customer1
VLAN 100

Customer2
VLAN 200

Customer1
VLAN 100

Customer2
VLAN 200

SVL: Super VLAN
T: Tagged interface
UT: Untagged interface

How 802.1s Works
STP with VLAN Grouping

Using 802.1 S
To Support 4 to 8 groups of Spanning Tree

Spanning Tree in Service
Provider's network
(Tag ID=0x9100)

FWD
ALL VLANS

User-1
VLAN 1-10

BI-Edge
Prio 5

VLAN 1011

BLK
VLAN 1011

User-2
VLAN 1-10

BI-Edge
Prio 5

BI-Core1
Prio 3, 0

BI-Core2
Prio 1

VLAN 1025

BI-Core3
Prio 2

BI-Core4
Prio 0, 3

VLAN 1011

(Tag ID=0x8100)

User-2
VLAN 1-10

BI-Edge
Prio 5

VLAN 1025

BLK
VLAN 1025

User-1
VLAN 1-10

BI-Edge
Prio 5

Spanning Tree in Service
Provider's network
(Tag ID=0x9100)

VLAN 1-1024   -> VLAN Group 1 =
VLAN 1025-2048 -> VLAN Group 2 =
VLAN 2049-3072 -> VLAN Group 3 =
VLAN 3073-4096 -> VLAN Group 4 =

All Core Switches forwarding unless
otherwise noted

Dash line indicates blocked state
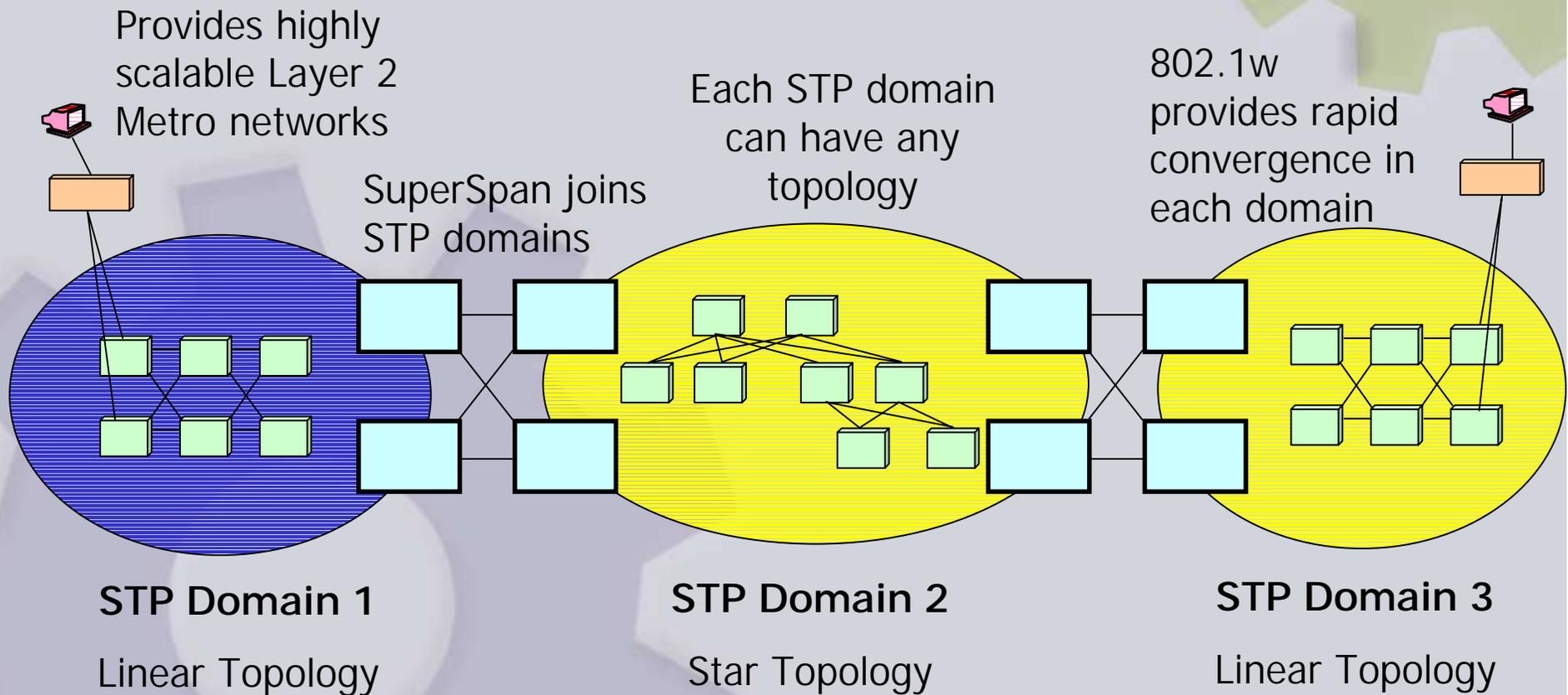
FOUNDRY
NETWORKS

Foundry's SuperSpan™
A Breakthrough For Layer 2 Network Scalability and High Availability

- Organizes Layer 2 Metro Core into smaller STP domains that are easy-to-manage and simple to configure

- Localizes STP to the domain – Provides rapid convergence within each domain with 802.1w Rapid STP

- Dramatic improvements in scalability – By joining simple, faster-converging STP domains, SuperSpan allows highly scalable large Layer 2 Metro networks

- Improves High Availability – A link failure is confined to the STP domain as other STP domains will be unaffected

- Allows any flexible topology within each STP domain – star, ring or mesh networks

- Interoperates with Super VLAN aggregation to allow end-to-end VLAN based services

FOUNDRY
NETWORKS

How Foundry's SuperSpan Works

Easy-to-manage, scalable, highly available Layer 2 Metro Core Networks!

Provides highly scalable Layer 2 Metro networks

Each STP domain can have any topology

802.1w provides rapid convergence in each domain

SuperSpan joins STP domains

STP Domain 1

Linear Topology

STP Domain 2

Star Topology

STP Domain 3

Linear Topology

FOUNDRY
NETWORKS

MPLS Layer-3 vs. Layer-2 VPNs

Layer-3

- Provider has to do some routing on the subscriber's behalf on the PE routers – extra burden for the provider

- Each PE router must have multiple routing and forwarding tables, one per customer site

- Introduces operational and management complexity

- Scalability challenges

- IP-only solution as it is a routed service

Layer-2

- Provider doesn't handle any routing on the subscriber's behalf, instead only simple mapping of Layer-2 frames to VCs is done

- No multiple routing and forwarding tables, only the global routing table

- Much simpler to operate and manage

- No scalability challenges in point to point mode
  Scalability challenges in VPLS mode

- Subscriber IP addresses are opaque to the provider, that way overlapping address spaces are handling without any extra work on the provider side

- Solution allows for carrying any protocol contained within the Layer-2 frames

Agenda

- L2 Technologies in MAN
- ➡ MPLS VLL Draft-Martini
- MPLS VPLS
- Interoperability

What Is Virtual Leased Line/Draft-Martini???

- "Draft-Martini" refers to two IETF Internet Drafts co-authored by Luca Martini of Level3

- draft-martini-l2circuit-encap-mpls-03.txt
  - Defines encapsulation formats for ATM, FR, Ethernet, Ethernet VLAN, HDLC, and PPP
  - We focus on Ethernet and Ethernet VLAN for now

- draft-martini-l2circuit-trans-mpls-07.txt
  - Proposes an extension to Label Distribution Protocol (LDP)
  - Purpose is to coordinate encapsulations between two routers
  - Introduces a new VC FEC TLV, to be used within Label Mapping messages

VLL Feature Highlights

- Capable of selecting LSP tunnels based on COS criterion
- Load balancing among multiple LSPs, tunnel failover
- The two endpoints of a VLL can have
  - Mixed tagged and untagged
  - Different tagged VLAN IDs
- VLAN ID can be reused
  - Subject to same DMA restriction
- Label withdrawal when local port goes down

The Martini Drafts

- The Martini drafts introduce the concept of Virtual Circuits (VCs)

- LSP are used as tunnels, an LSP might carry multiple VCs

- To accomplish this, the drafts leverage the MPLS label stacking ability by adding an extra MPLS label that distinguishes the VC used

- An MPLS frame traversing a service provider network has two labels:
  - Tunnel Label   ←   *For the LSP*
  - VC Label   ←   *For the VC*

The Martini Drafts

- draft-martini-l2circuit-trans-mpls-07.txt
  - Distribution of VC labels via LDP
  - Introduces a new VC FEC TLV, to be used within Label Mapping messages
- draft-martini-l2circuit-encap-mpls-03.txt
  - Defines encapsulations for ATM, FR, Ethernet, Ethernet VLAN, HDLC, and PPP
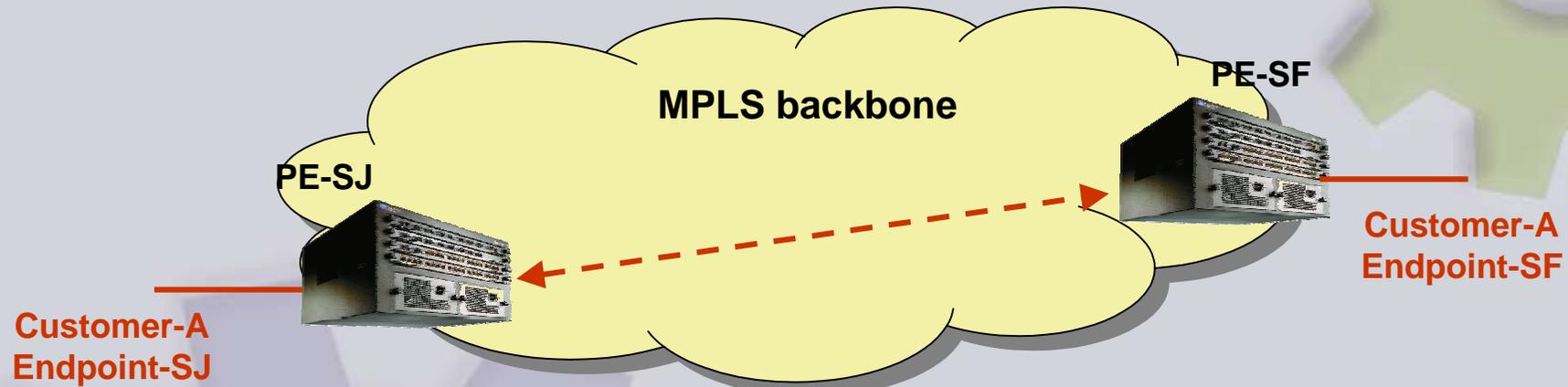
Format for Ethernet/VLAN Encapsulation:

PPP-MPLS or Enet Hdr.	Tunnel Label	VC Label	Optional Control Word	Tagged or untagged Ethernet payload

The Virtual Lease Line (VLL) Service

- The mechanisms specified in the Martini drafts could be used to offer a Virtual Leased Line (VLL) service for point to point connectivity

- The Virtual Leased Line is just an abstraction of the Virtual Circuits involved in realizing the point to point connectivity, making the service more manageable

- A VLL is composed of two VCs one in each direction, thus giving it bidirectionality

VLL

LSP
VC

PE-A

PE-B

What Draft-Martini Is and Is NOT

MPLS backbone

PE-SF

PE-SJ
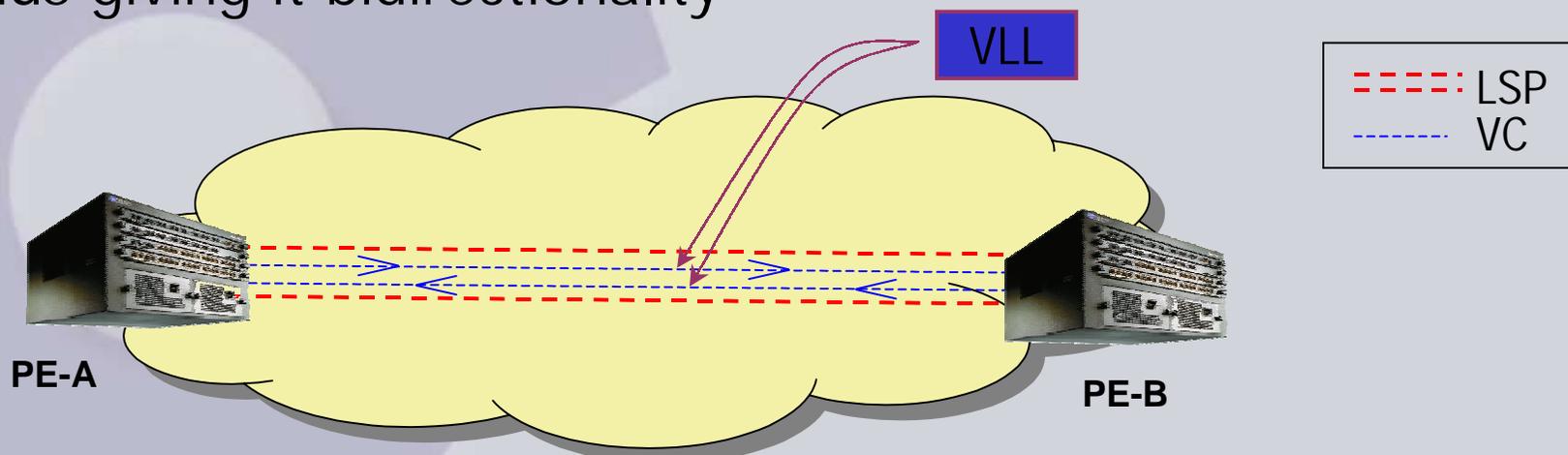
Customer-A Endpoint-SF

Customer-A Endpoint-SJ

- Layer-2 packet relay between two customer endpoints using MPLS encapsulation
- The VLL is only known to the two Provider Edge (PE) routers involved
- Point-to-Point ONLY
- Draft-Martini does not include Point-to-MultiPoint support – it's not transparent LAN switching (TLS)

What Draft-Martini Is and Is NOT (cont.)

- No switching means
  - PE routers do not have to care and learn about MAC addresses
  - Forwarding simplicity
  - No MAC table scalability concern

Anatomy of a VLL
What Needs to Be Configured?

- VLL name and ID
  - VLL ID must be unique on the two PE routers – it allows the two PEs to know which VLL they are referring to
  - Name is just for user reference and is not signaled
- VLL Peer
  - IP address of the other PE router
- Local Endpoint
  - Customer-facing interface
  - Can be either an untagged port or an tagged port with a specified VLAN-ID
  - Benefit of tagged endpoint: can have multiple endpoints on one physical port

Sample Configurations

- Untagged endpoint

VLL ID

```
router mpls
    vll foundry-sj-to-sf 40000
        vll-peer 192.168.2.100
        untagged e 2/1
```

- Tagged endpoint

```
router mpls
    vll foundry-sj-to-la 50000
        vll-peer 200.100.1.100
        vlan 20
            tagged e 3/1
```

- **Options: statically configured label values, vll cos (class of service)**

Anatomy of a VLL
How Does a VLL Transport Packets?

- Peers forming a single virtual pipe
- Peers exchange VC-label to VLL-ID binding through LDP session with Draft-Martini extension

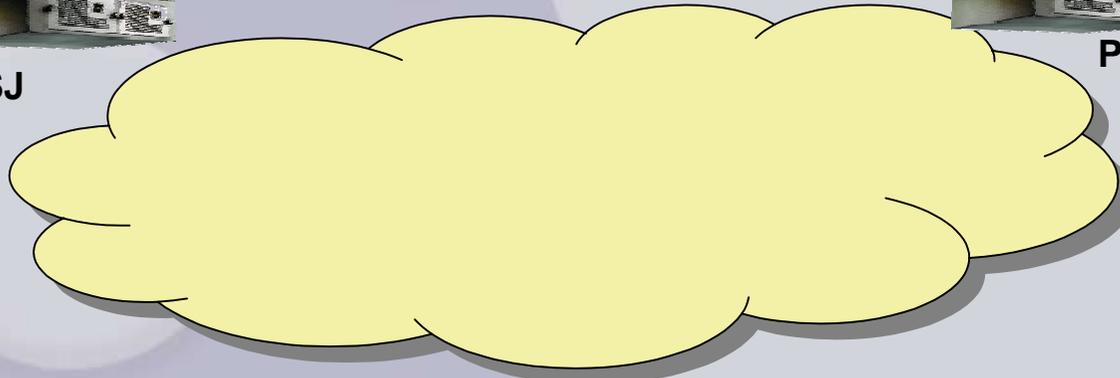*"To send packets for VLL-ID 1000 to me, use VC label X"*

"To send packets for VLL-ID 1000 to me, use VC label Y"

PE-SF

PE-SJ

Anatomy of a VLL
How Does a VLL Transport Packets?

- LSP tunnels, established by RSVP, provide the transport through MPLS backbone

PE-SJ can be reached by LSP TL-20
Outgoing tunnel label is L

PE-SF can be reached by LSP TL-10
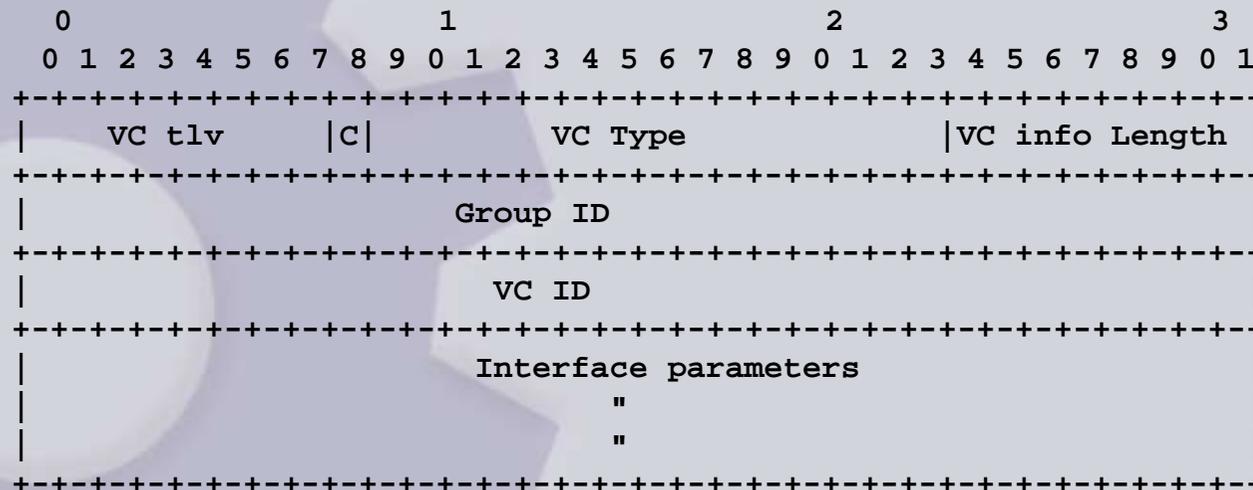*Outgoing tunnel label is M*

PE-SJ

PE-SF

Outgoing MPLS packet generated by PE-SJ:

PPP Hdr Type-MPLS	Tunnel Label M	VC Label Y	Tagged or untagged Ethernet payload

Anatomy of a VLL
VC Label Signaling

- Once all configurations are complete, a PE router sends an LDP Label-Mapping message to the other end (in downstream unsolicited mode)

- The message indicates the binding of a local VC label (in Label TLV) to the VC FEC TLV (introduced by draft-martini)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    VC tlv     |C|      VC Type        |VC info Length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Group ID                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       VC ID                                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Interface parameters                        |
|                         "                                    |
|                         "                                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Anatomy of a VLL
VC Label Signaling

Local label X
Remote label Y

"I bind label X to VLL-ID 1000"

"I bind label Y to VLL-ID 1000"

Local label Y
Remote label X

PE-A

PE-B
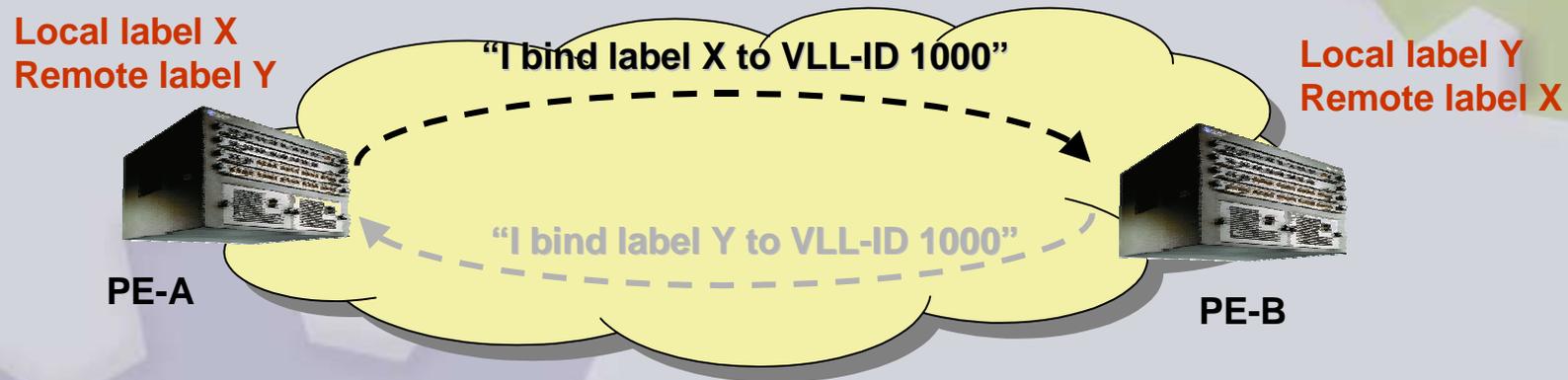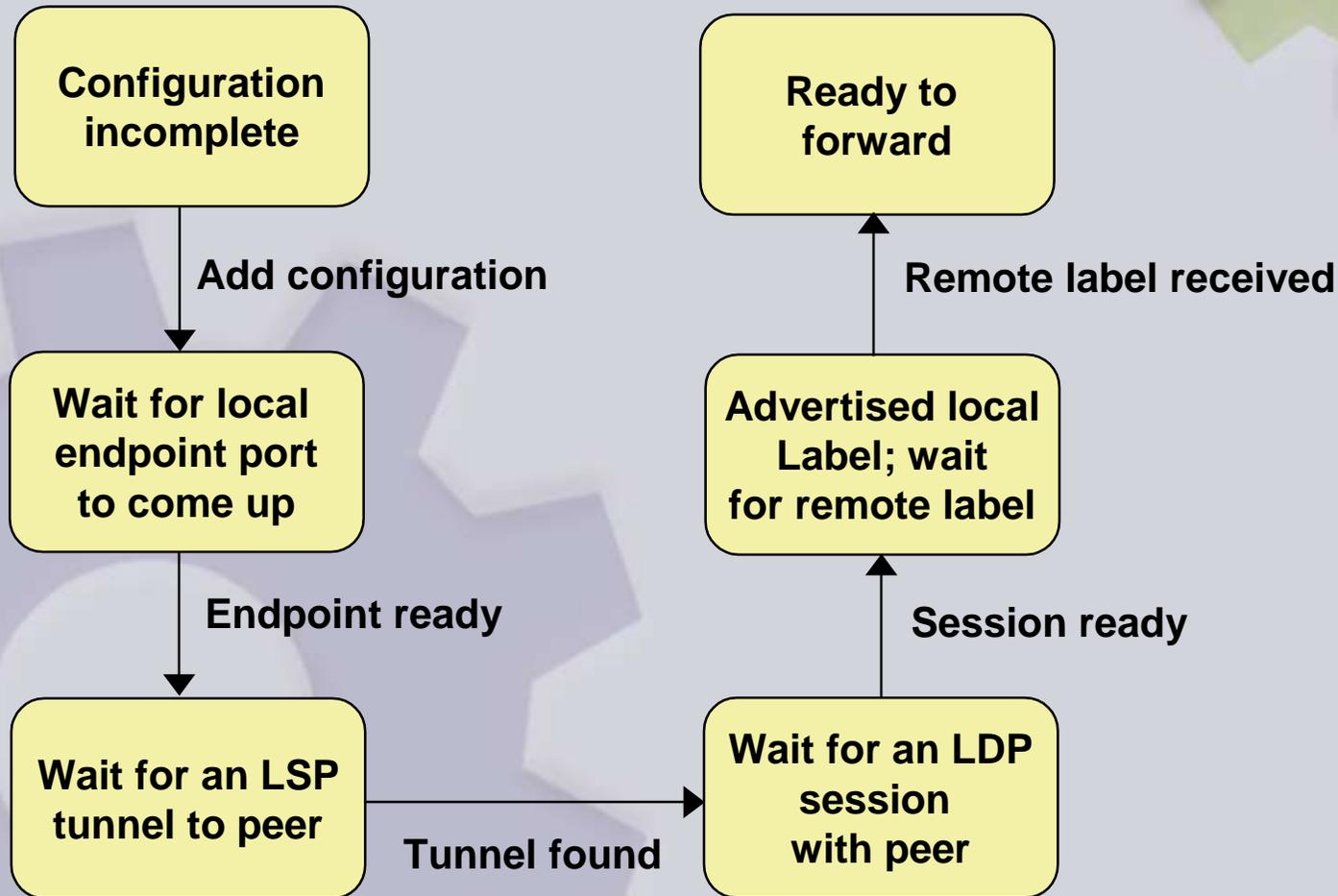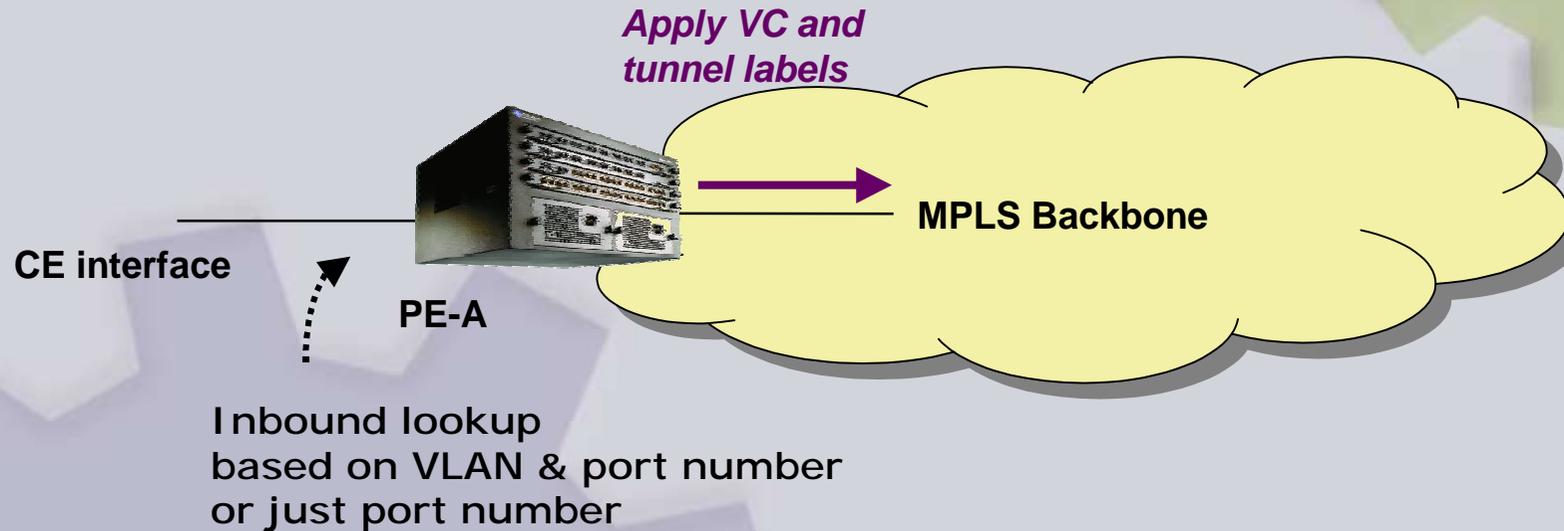
- PE router sends an LDP Label-Mapping message to the other end

- May send Label-Withdraw message when local CE interface goes down

Anatomy of a VLL
# How Does a VLL Get Up and Running?

```
┌─────────────────┐                          ┌─────────────────┐
│  Configuration  │                          │    Ready to      │
│   incomplete    │                          │    forward       │
└─────────────────┘                          └─────────────────┘
         │                                             ▲
         │ Add configuration                           │ Remote label received
         ▼                                             │
┌─────────────────┐                          ┌─────────────────┐
│  Wait for local │                          │ Advertised local │
│  endpoint port  │                          │   Label; wait    │
│  to come up     │                          │  for remote label│
└─────────────────┘                          └─────────────────┘
         │                                             ▲
         │ Endpoint ready                              │ Session ready
         ▼                                             │
┌─────────────────┐     Tunnel found         ┌─────────────────┐
│  Wait for an LSP │ ───────────────────────▶│  Wait for an LDP │
│  tunnel to peer  │                         │    session       │
│                  │                         │   with peer      │
└─────────────────┘                          └─────────────────┘
```

Fowarding Frames:
CE to Backbone

Apply VC and tunnel labels

MPLS Backbone

CE interface

PE-A

Inbound lookup
based on VLAN & port number
or just port number

Outgoing MPLS Packet Format:

PPP-MPLS or Enet Hdr.	Tunnel Label M	VC Label Y	Optional Control Word	Tagged or untagged Ethernet payload

Forwarding Frames: Backbone to CE

CE interface

PE-A

MPLS Backbone
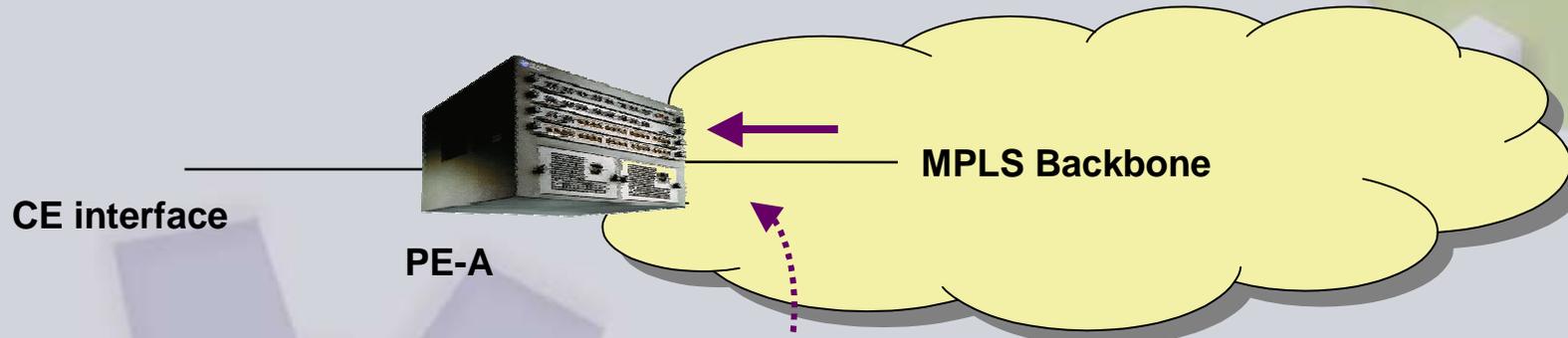
Incoming label X indicates VLL
payload and which CE interface to send to
and VLAN-ID to use

Incoming MPLS Packet Format:

PPP-MPLS or Enet Hdr.	Tunnel Label M	VC Label Y	Optional Control Word	Tagged or untagged Ethernet payload

Incoming MPLS Packet Format in Case of Pen-Ultimate Hop Popping:

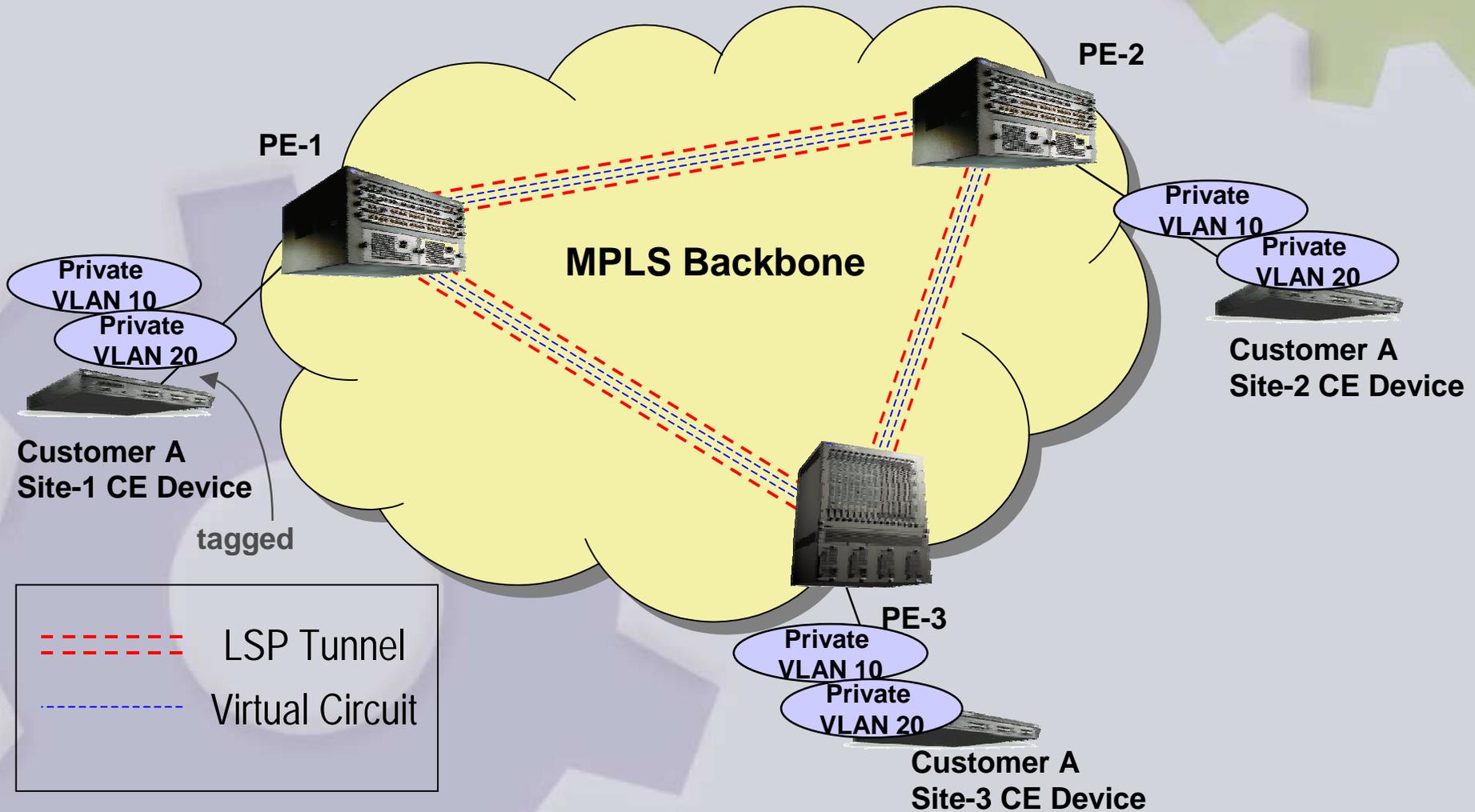| PPP-MPLS or Enet Hdr.	VC Label X	Optional Control Word	Tagged or untagged Ethernet payload

VLL Example Application:
CE Device Handles Switching

- Router at central location responsible for data forwarding among all remotes; switches or routers at remotes

Remote

Central Site Router

VLL for subnet 1

UT

The VLAN-ID's on these PEs are not required to match with the central site PE

T

Multinetting w/ one subnet to each remote

VLL for subnet 2

UT

Remote

T: tagged
UT: untagged
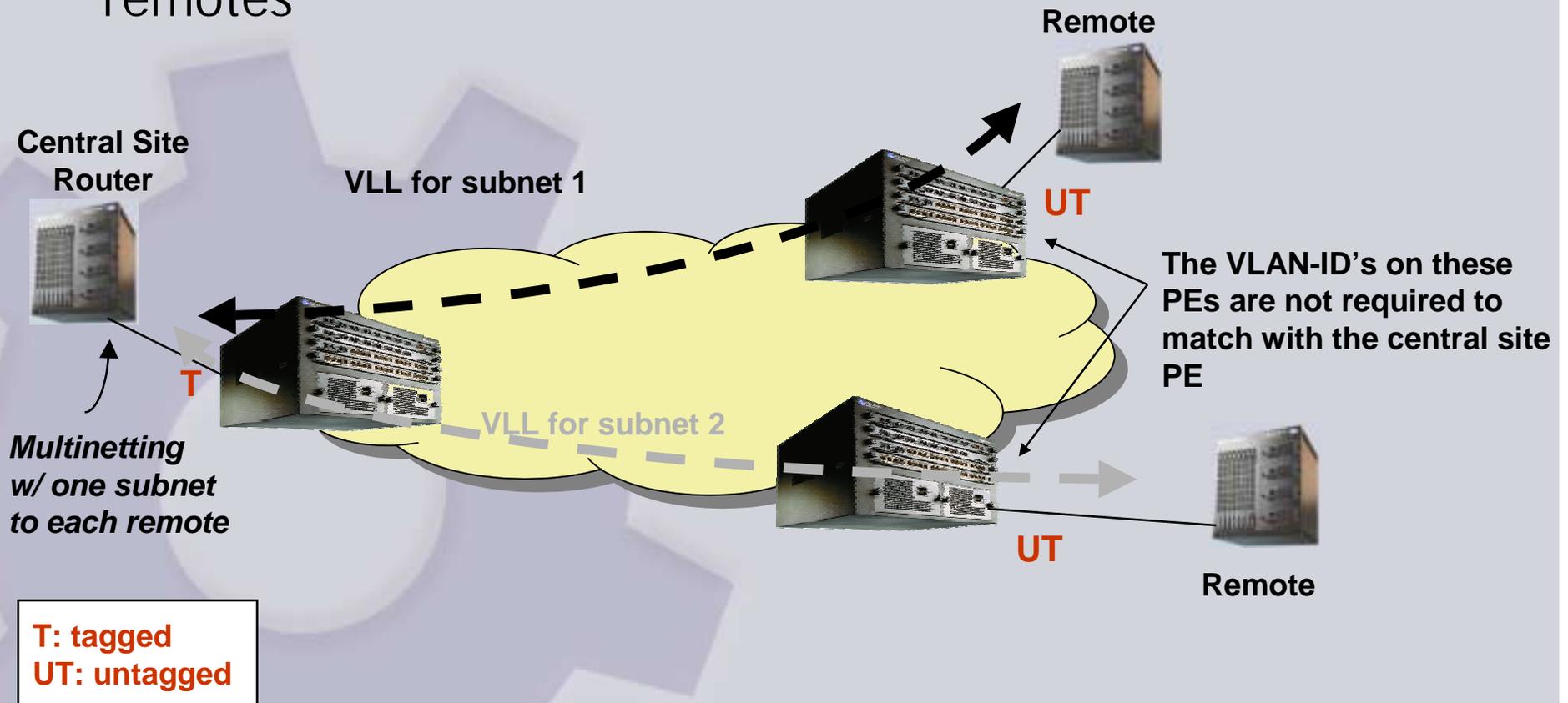
VLL Example Application: SP Handles Switching

Tagged interface shared among VLANs

One VLAN per customer

Customer 1 Site C

Customer 2 Site A

SP network

MPLS Backbone

Customer 1 Site A

Customer 2 Site C

Edge switches perform SA learning, broadcasting/ unknown unicasting, transit forwarding, and STP

Customer 1 Site B

Customer 2 Site B
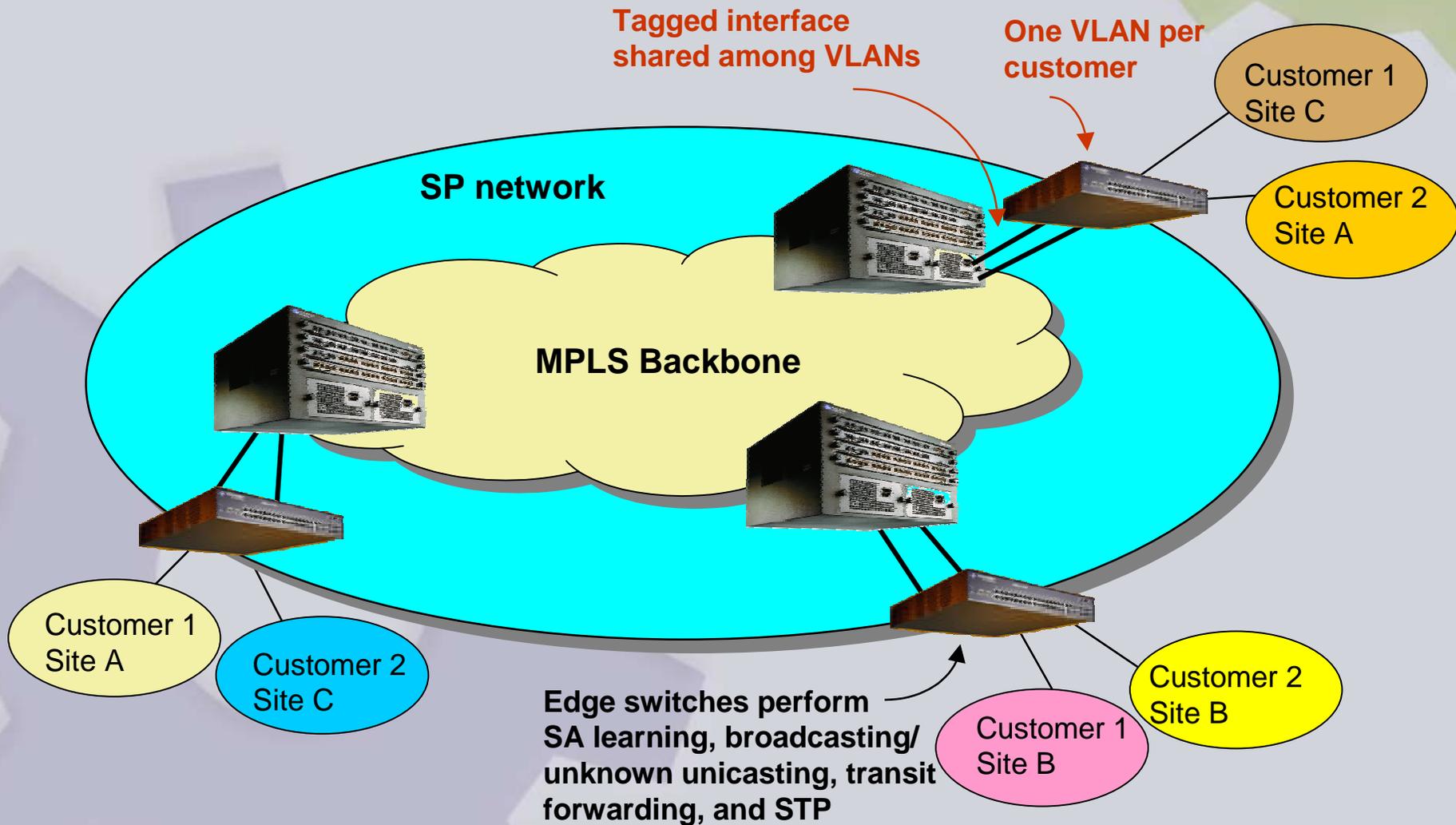
Service Provider View of the Service

PE-2

PE-1

MPLS Backbone

Private
VLAN 10

Private
VLAN 20

Private
VLAN 10

Private
VLAN 20

**Customer A
Site-2 CE Device**

**Customer A
Site-1 CE Device**

tagged

PE-3

Private
VLAN 10

Private
VLAN 20

**Customer A
Site-3 CE Device**

- - - - - - - LSP Tunnel

- - - - - - - Virtual Circuit

Customer's View of the Service

Customer's CE devices think that they are all in the same subnet and that they are connected through a switch

Private
VLAN 10

Private
VLAN 20

Private
VLAN 10

Private
VLAN 20

Customer A
Site-2 CE Device

Customer A
Site-1 CE Device

tagged

PE-3

Private
VLAN 10

Private
VLAN 20

Customer A
Site-3 CE Device
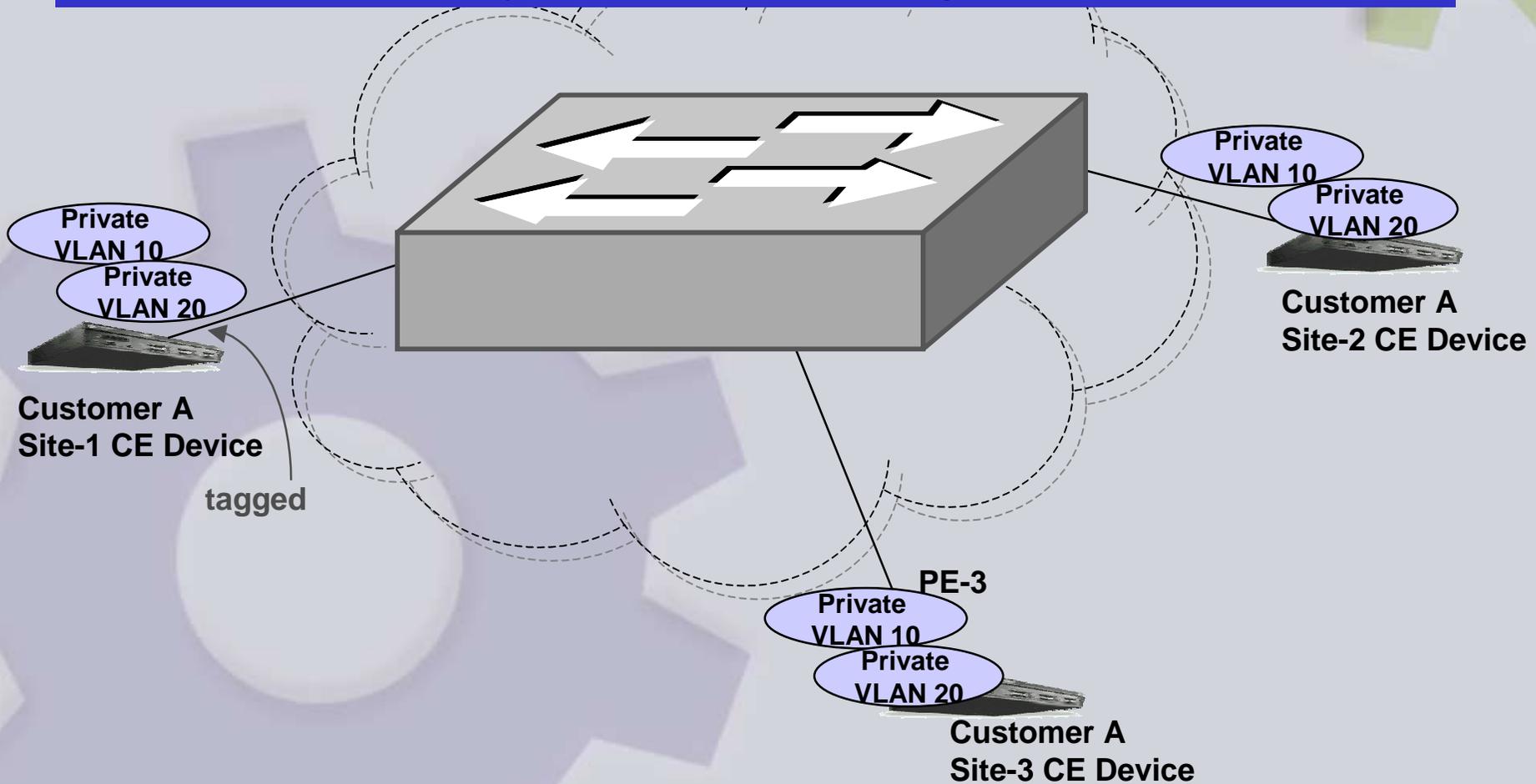
Agenda

- L2 Technologies in MAN
- MPLS VLL Draft-Martini
- → MPLS VPLS
- Interoperability

Service Terminology

- In this scenario, the provider is said to be offering:

  A Virtual Private LAN Segment (VPLS)

  or (alternative name):

  A Transparent LAN Service

Not to be confused with "Virtual Private LAN Services" which is an IP only solution described in "draft-tsenevir-vpl-ip-00.txt"

VPLS Defining Documents

- RFC 2764 – "A Framework for IP Based Virtual Private Networks":
  - Provides the general framework for IP Based VPNs – MPLS included
- draft-lasserre-tls-mpls-00.txt:
  - Describes the forwarding of Ethernet/802.3 frames between multiple customer sites, as if they were in the same layer-2 broadcast domain
  - Defines a new parameter "VPN ID" to identify PE routers participating in the same customer VPN
- draft-vkompella-ppvpn-vpsn-mpls-00.txt:
  - Describes MAC address learning and aging
  - Describes MAC address signaling
- A single WG document will be formulated

Functional Requirements

- As frames are transported over a provider backbone, efficient use of resources is crucial

- Frames should be forwarded only over the needed VCs

- PE routers should be capable of learning Ethernet source MAC addresses, just like a regular switch

- A PE router keeps a VPLS Forwarding Table (quite similar to a bridge's MAC table) that distiguishes the entries belonging to the different VPNs

- Before forwarding a frame, the PE router checks its VPLS forwarding table to determine which VCs the frame should go over

Functional Requirements (Cont.)

- Broadcast or unknown unicast frames are flooded over all outgoing VCs belonging to the VPN

- The PE routers have to be connected via VCs forming a full mesh topology

- For loop prevention, running an instance of STP per VPN in the provider network would not scale:

  – Solution:
    When forwarding frames a PE follows a split-horizon rule: it never forwards a frame from one VC to another in the same VPN

Handling Multicast Frames

- Multicast frames are currently handled like broadcasts
- Extensions for performing smarter mutlicasts are to be developed in the future
- Possible scenarios:
    - Interaction with 802.1 GMRP
    - IGMP snooping
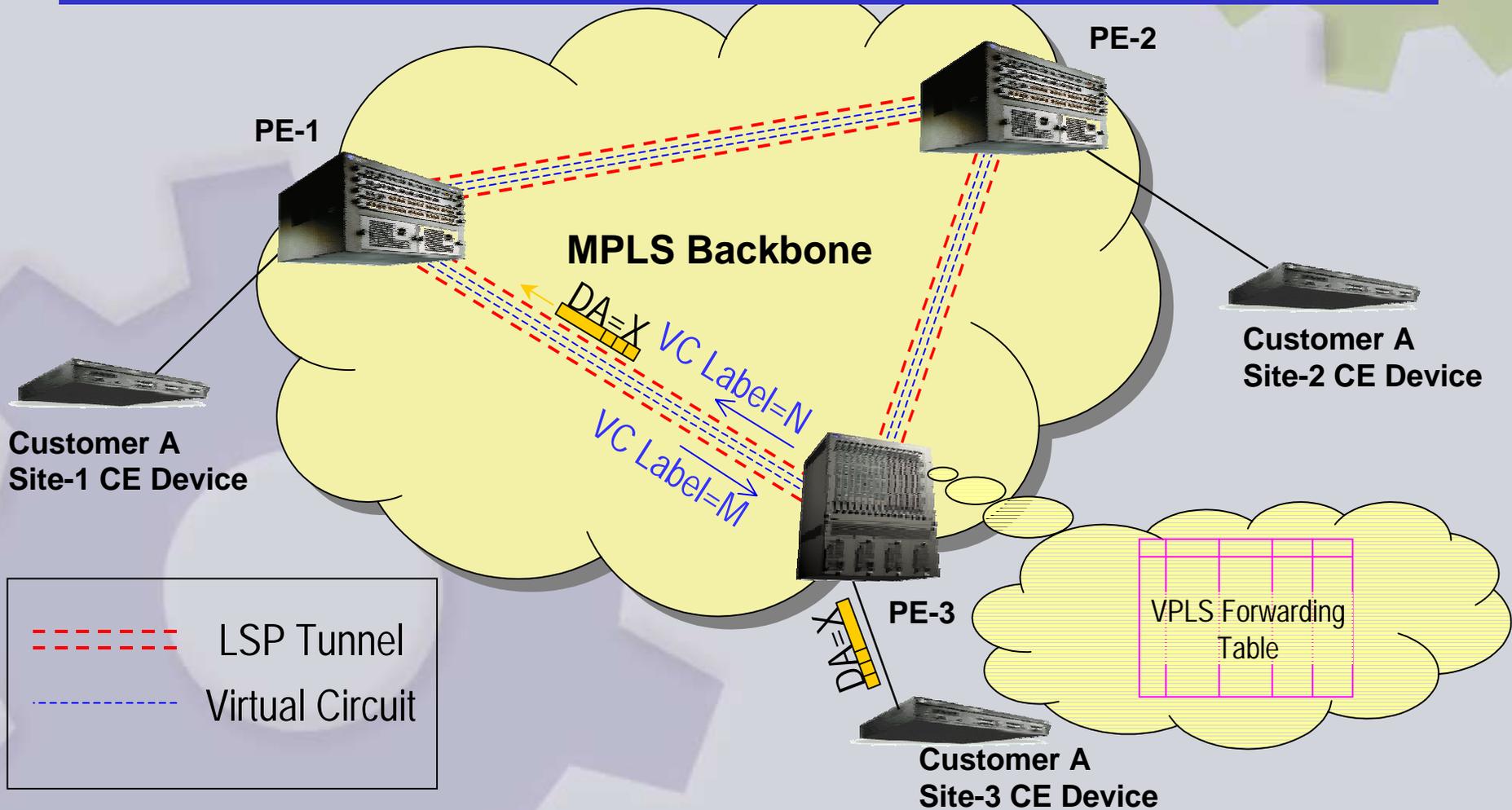   - Static MAC multicast filters

PE Router Performs Source Address Learning
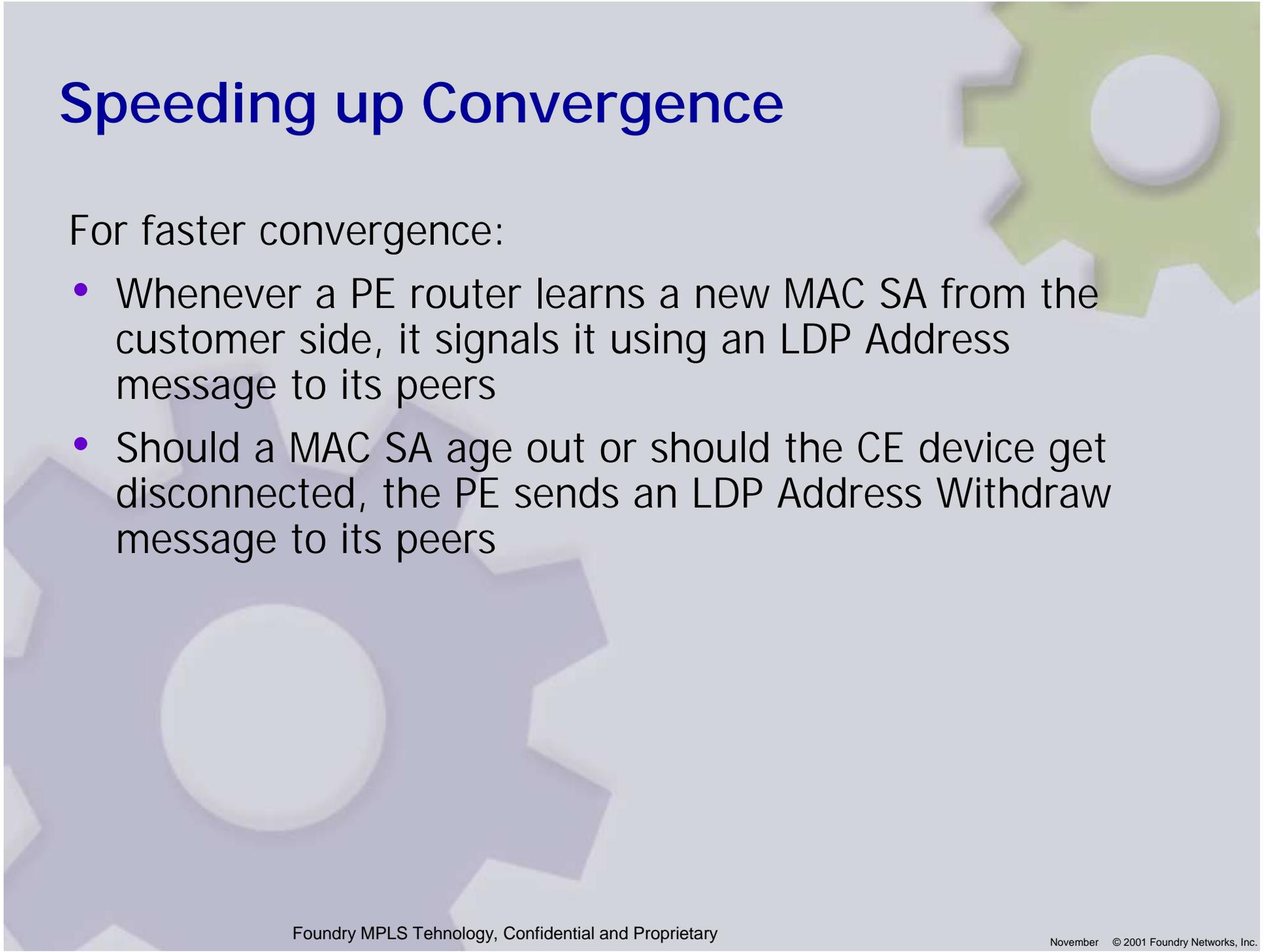
PE-3 associates MAC X with the outgoing VC (VC Label N)

PE-2

PE-1

MPLS Backbone

SA=X

VC Label=N

VC Label=M

**Customer A
Site-2 CE Device**

**Customer A
Site-1 CE Device**

PE-3

VPLS Forwarding Table

---- ---- LSP Tunnel

-------- Virtual Circuit

**Customer A
Site-3 CE Device**
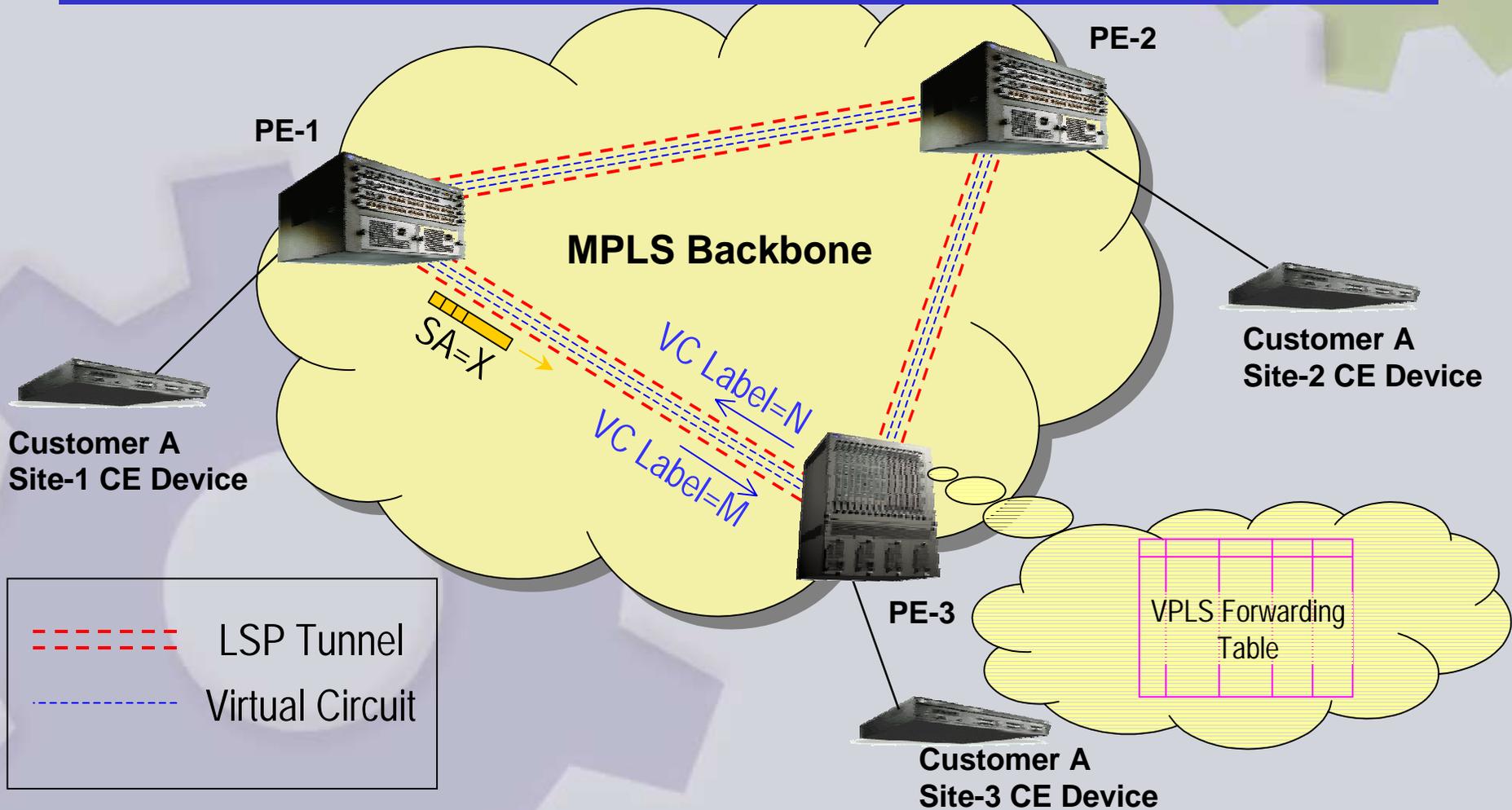
PE Router Performs Forwarding

PE-3 Checks its VPLS forwarding table and selects VC Label N

PE-2

PE-1

MPLS Backbone

DA=X

VC Label=N

VC Label=M

**Customer A
Site-2 CE Device**

**Customer A
Site-1 CE Device**

PE-3

DA=X

**VPLS Forwarding
Table**

- - - - - - - LSP Tunnel

- - - - - - - Virtual Circuit

**Customer A
Site-3 CE Device**

Speeding up Convergence
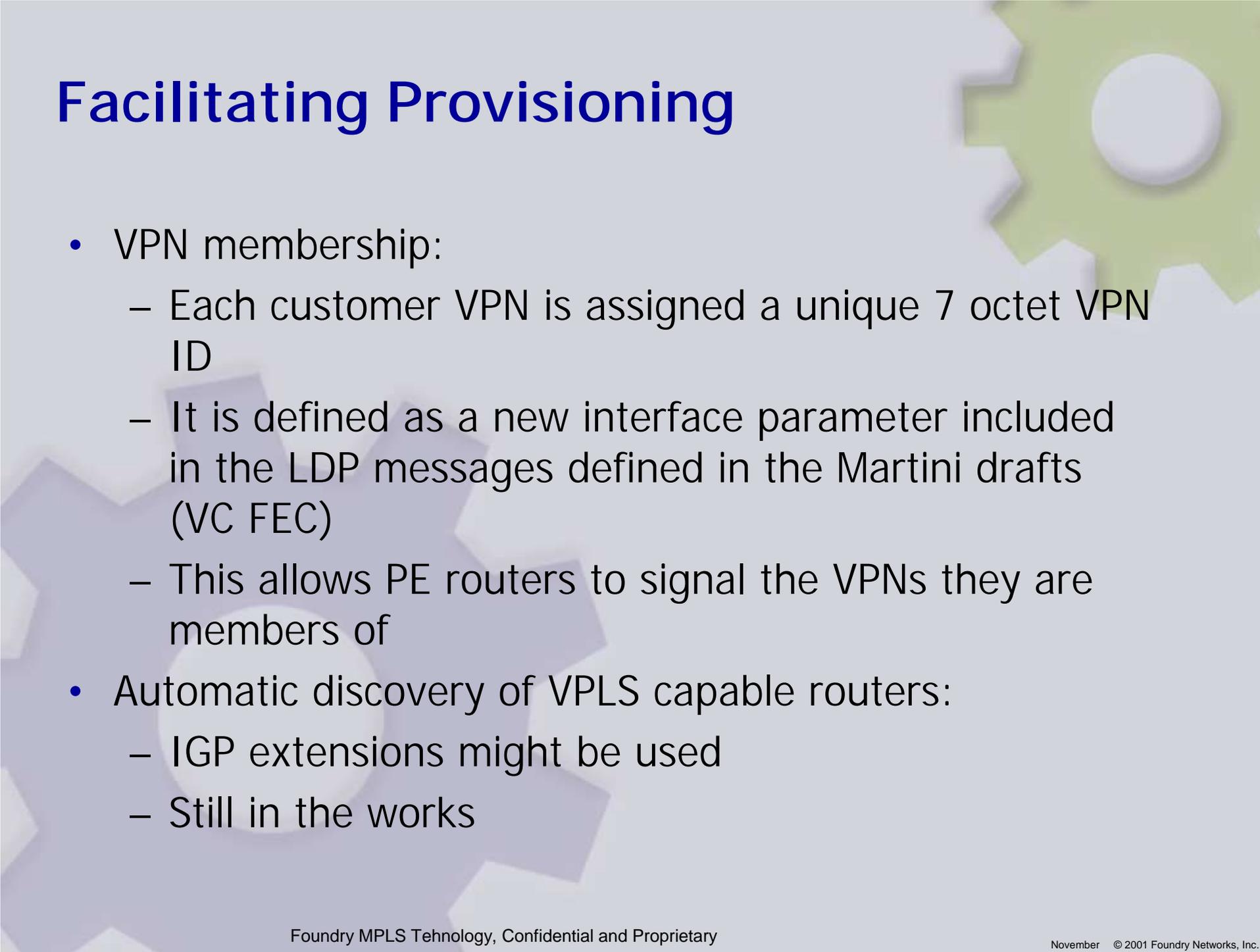
For faster convergence:

- Whenever a PE router learns a new MAC SA from the customer side, it signals it using an LDP Address message to its peers

- Should a MAC SA age out or should the CE device get disconnected, the PE sends an LDP Address Withdraw message to its peers

MAC Signaling

PE-3 signals the new MAC to speed up convergence

PE-2

PE-1

MPLS Backbone

LDP Msg

LDP Msg

Customer A
Site-2 CE Device

Customer A
Site-1 CE Device

New MAC.
Notify peers.

PE-3

SA=Y

- - - - - - - LSP Tunnel
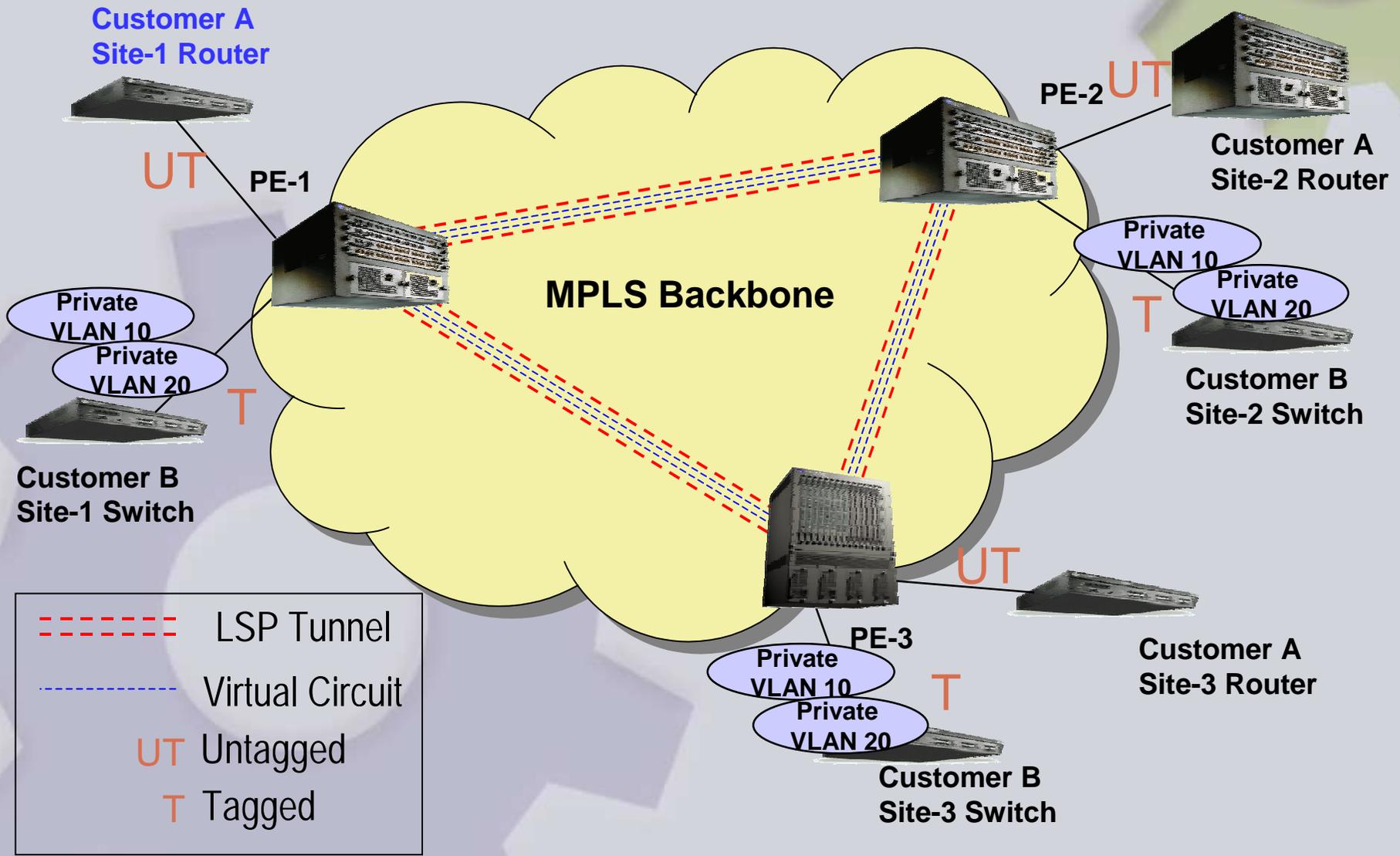- - - - - - - Virtual Circuit

Customer A
Site-3 CE Device
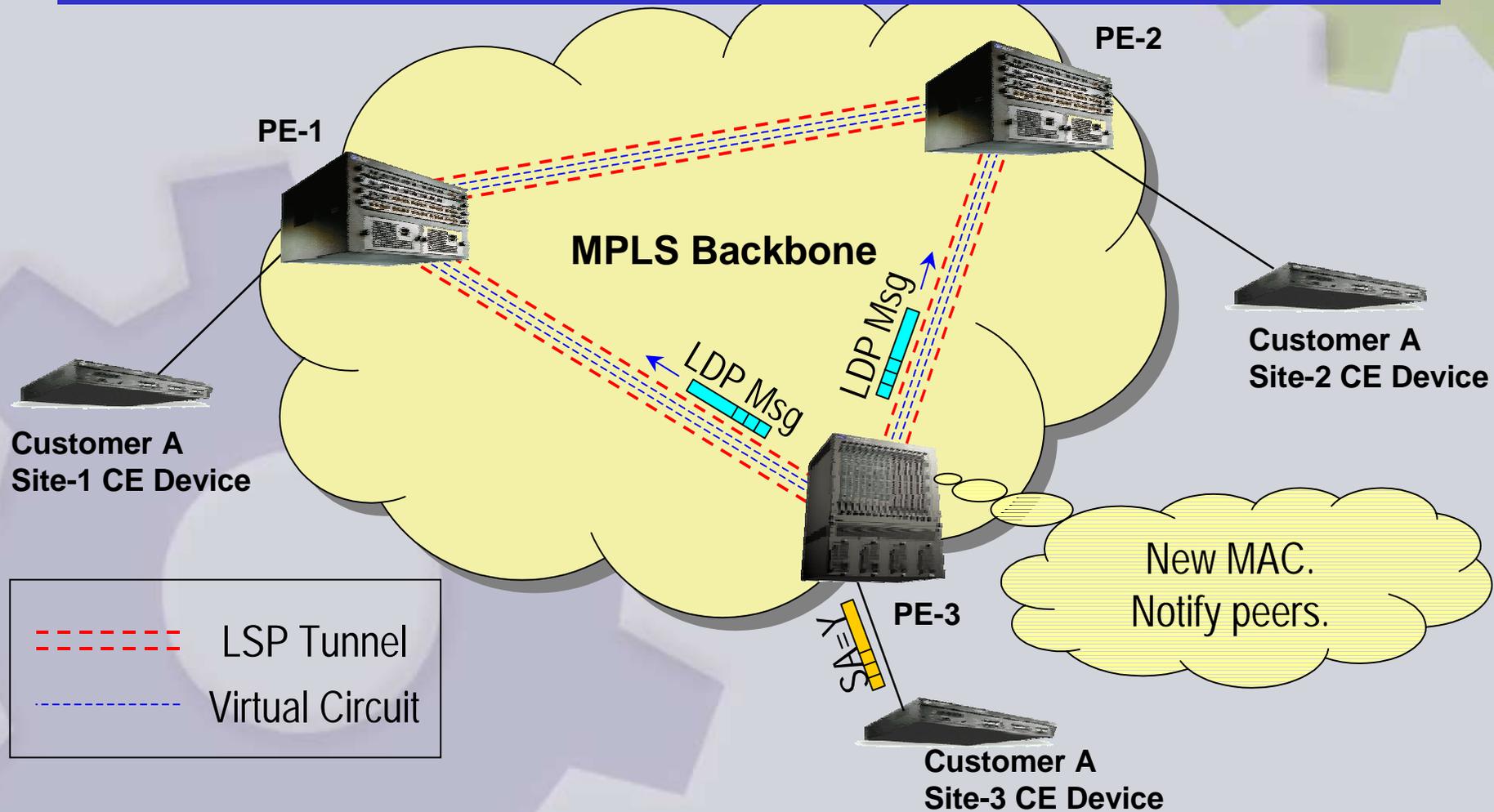
Facilitating Provisioning

- VPN membership:
  - Each customer VPN is assigned a unique 7 octet VPN ID
  - It is defined as a new interface parameter included in the LDP messages defined in the Martini drafts (VC FEC)
  - This allows PE routers to signal the VPNs they are members of
- Automatic discovery of VPLS capable routers:
  - IGP extensions might be used
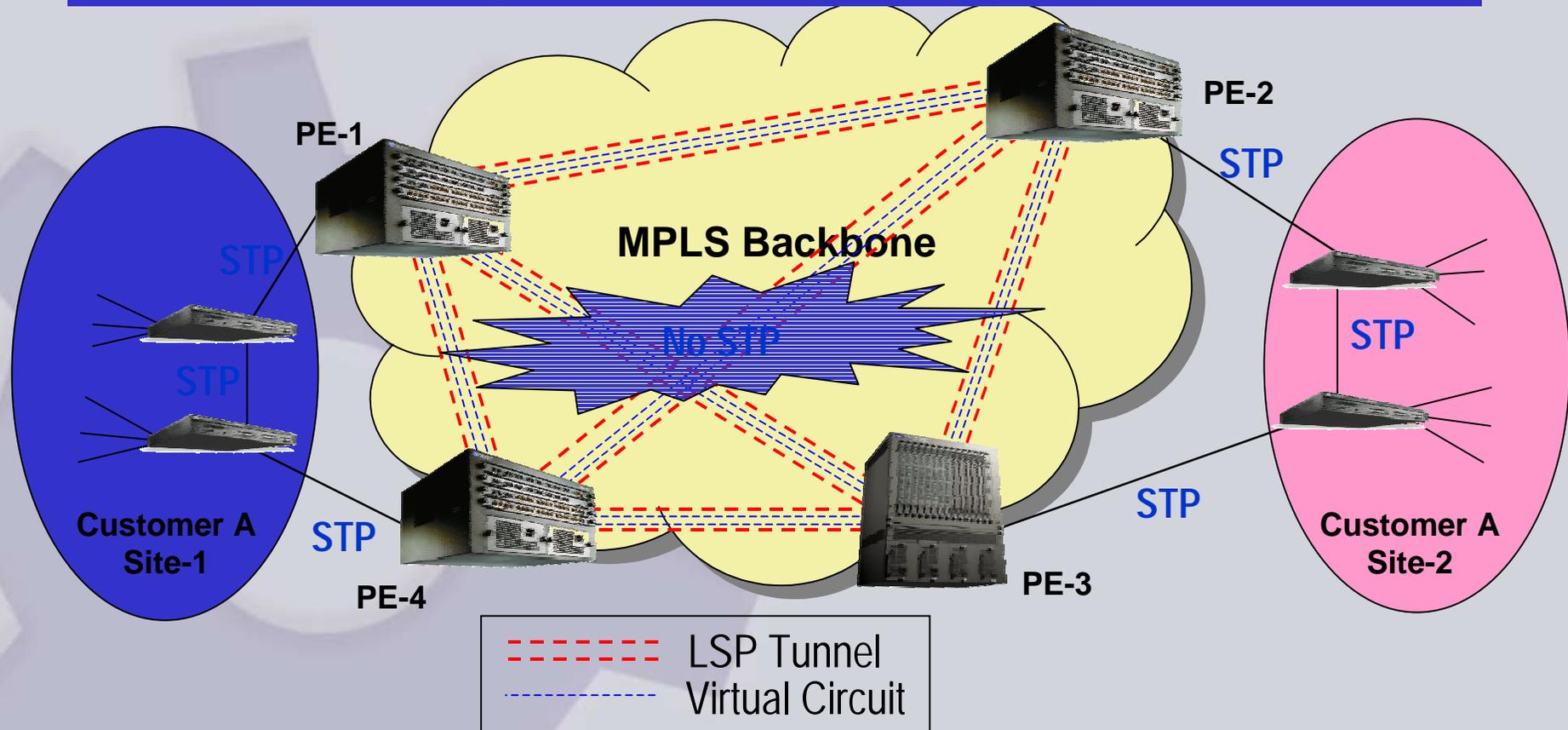 - Still in the works

VPLS Example Application: Providing TLS

**Customer A
Site-1 Router**

UT

PE-1

UT **PE-2**

**Customer A
Site-2 Router**

MPLS Backbone

**Private
VLAN 10**

**Private
VLAN 20**

T

**Private
VLAN 10**

**Private
VLAN 20**

T

**Customer B
Site-1 Switch**

**Customer B
Site-2 Switch**

UT

PE-3

**Private
VLAN 10**

**Private
VLAN 20**

T

**Customer A
Site-3 Router**

**Customer B
Site-3 Switch**

- - - - - - - LSP Tunnel

- - - - - - - Virtual Circuit
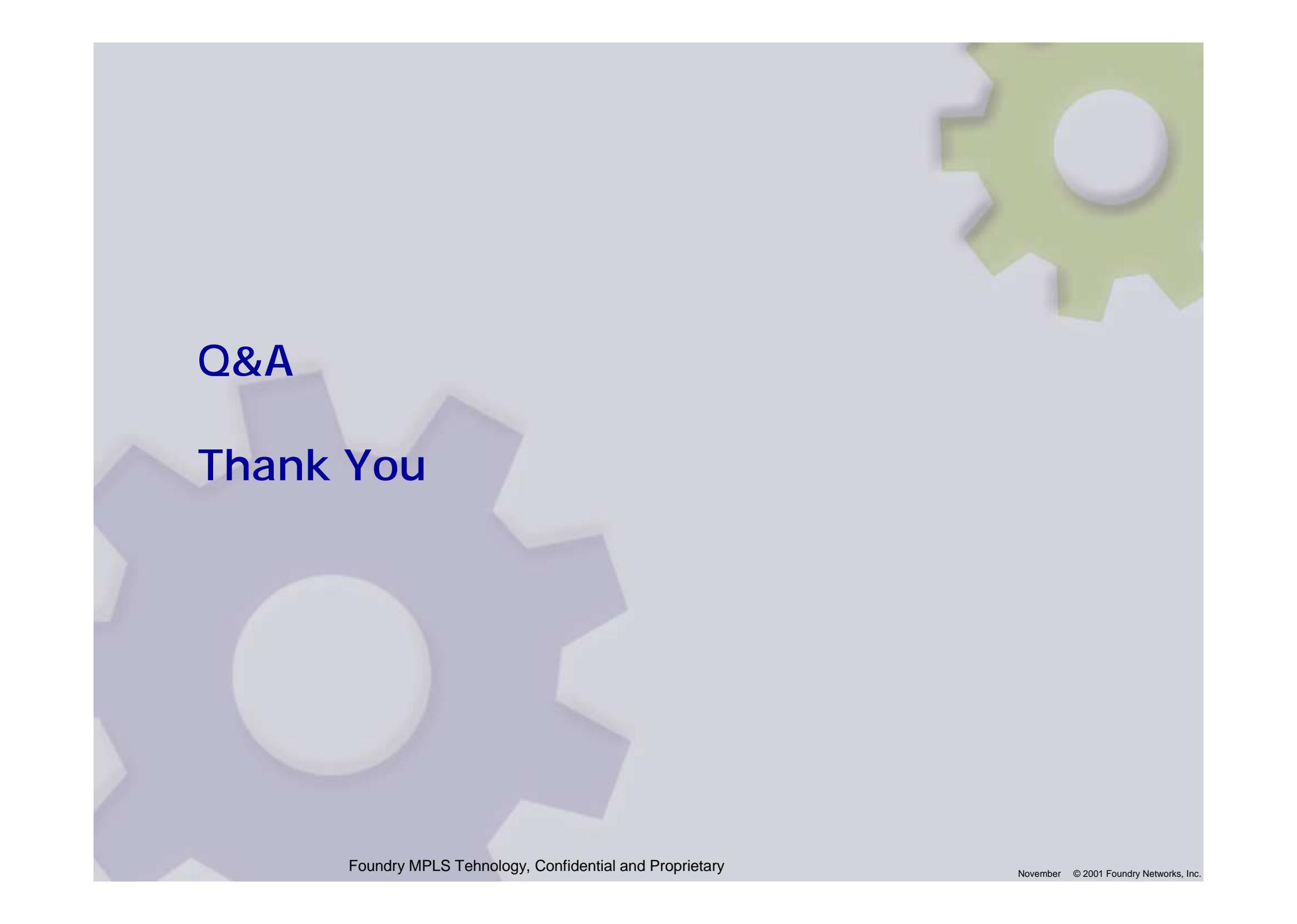
UT Untagged

T Tagged

VPLS Example Application: Dual Homing the Customer

- The customer might run and manage their own instance of STP
- The provider does not run STP on their backbone, they just carry the customer's BPDUs

STP used by the customer but not by the provider

PE-1
PE-2
PE-3
PE-4

MPLS Backbone

No STP

STP

Customer A Site-1

Customer A Site-2

- - - - - - - LSP Tunnel
- - - - - - - Virtual Circuit

Q&A

Thank You