midas

Ethernet in Fiber Access Networks

Control Plane & Management

R.Thirumurthy

Banyan Networks

Chennai

E-mail: rtm@banyannetworks.com

midas

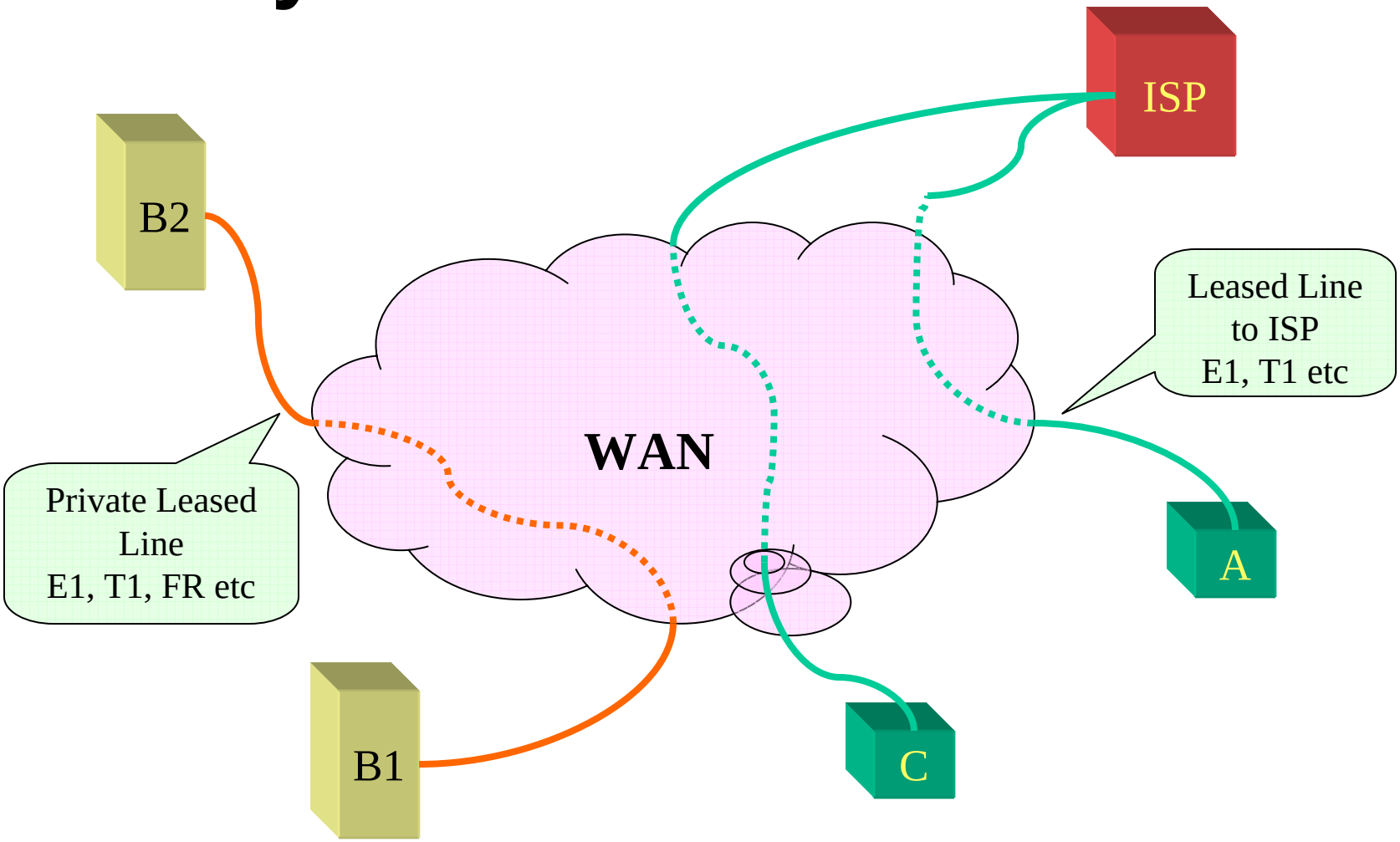
Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

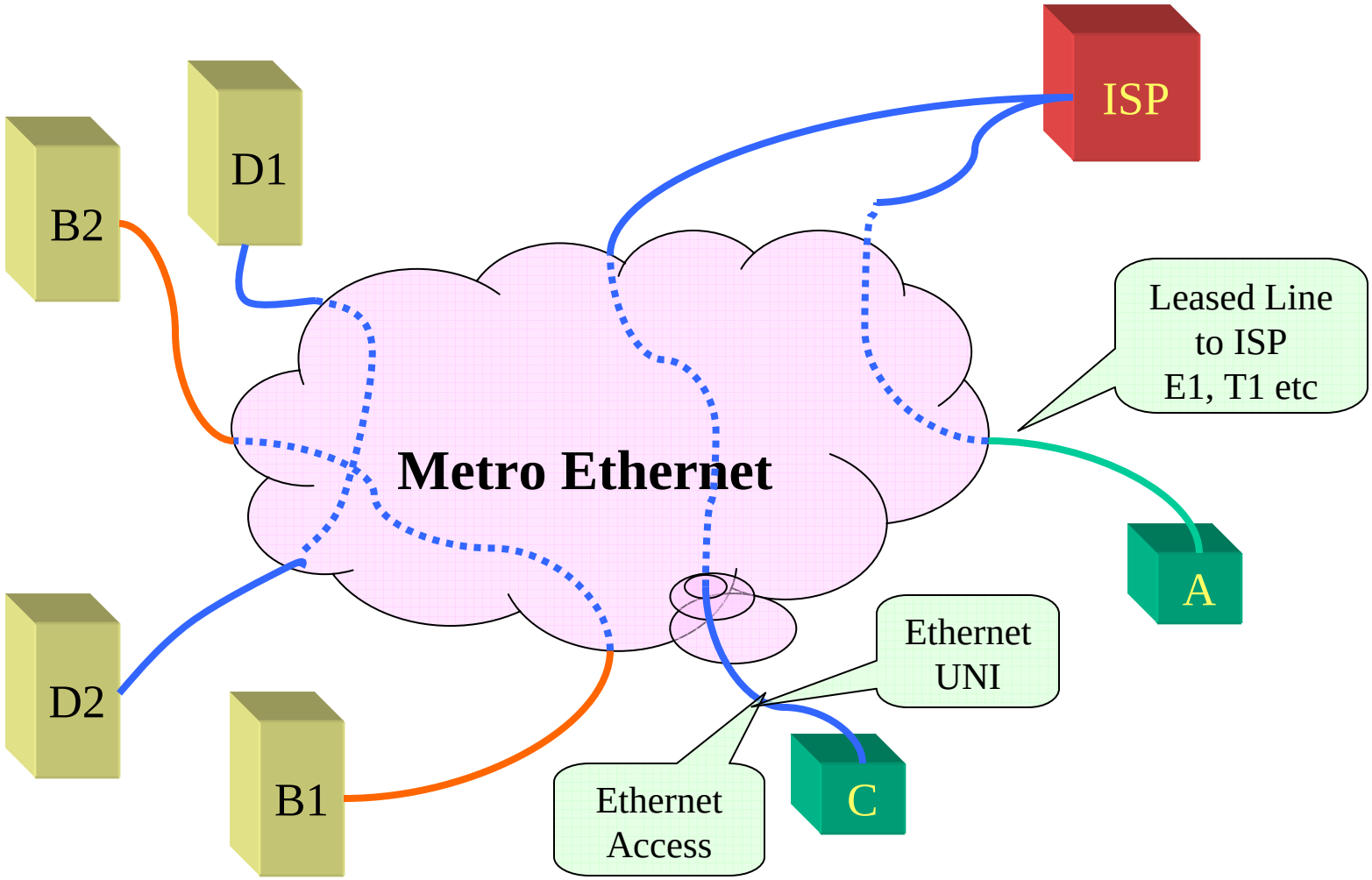
Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

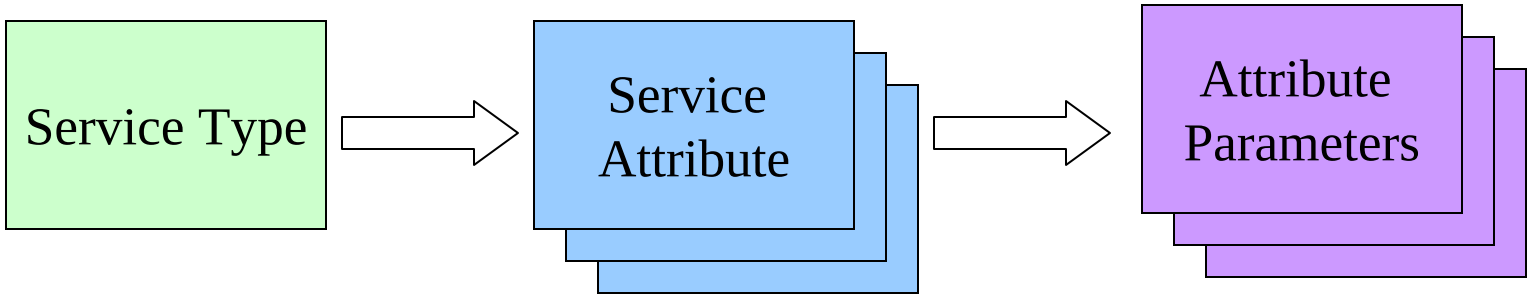
Today's Access Network



Metro Access Network

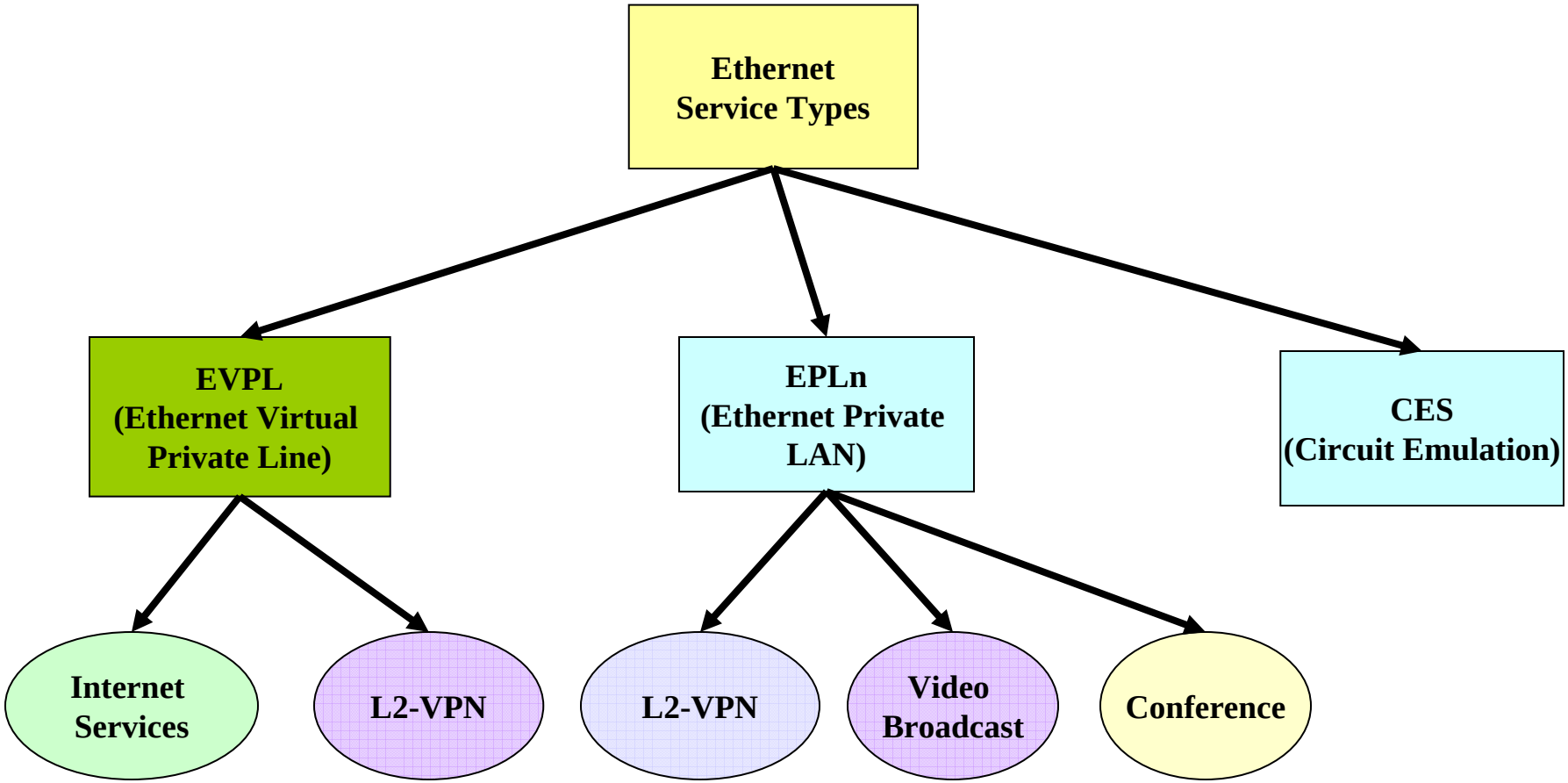


Ethernet Service Model

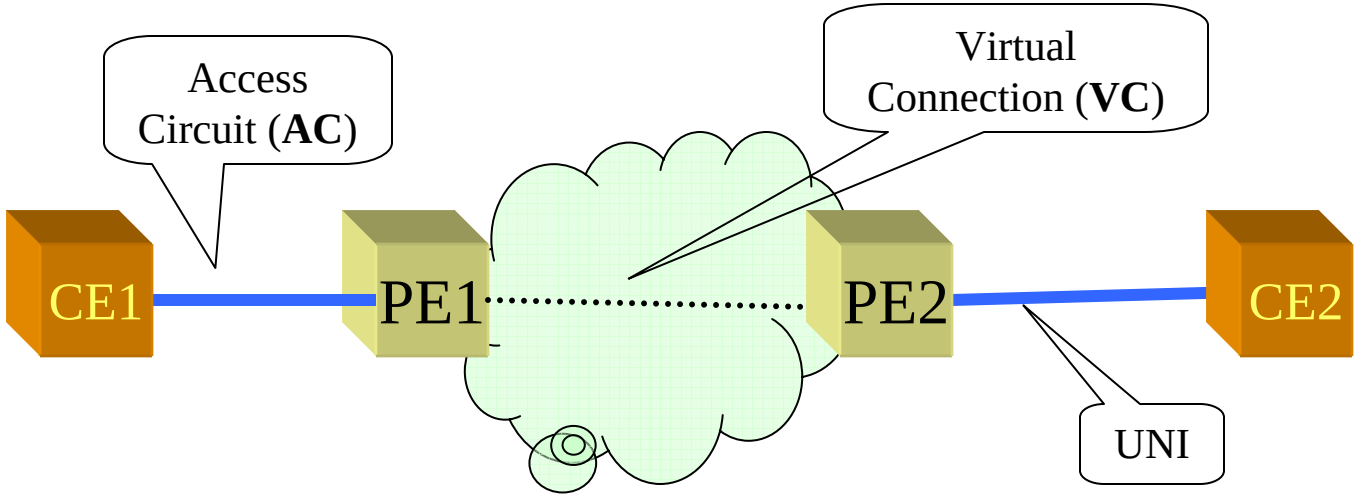


- Service Type, abstractly define the various access mechanism
  - E.g. Point to Point, Multipoint etc
- Service Attributes, define the capabilities of the Ethernet service type (e.g Bandwidth, Speed etc)
  - Some Apply to the EVC
  - Some apply for the UNI (e.g Speed, Mode)
- Parameters quantify the service attribute
  - E.g, for Bandwidth Attribute, the parameters are CIR, CBS etc.

Service Types

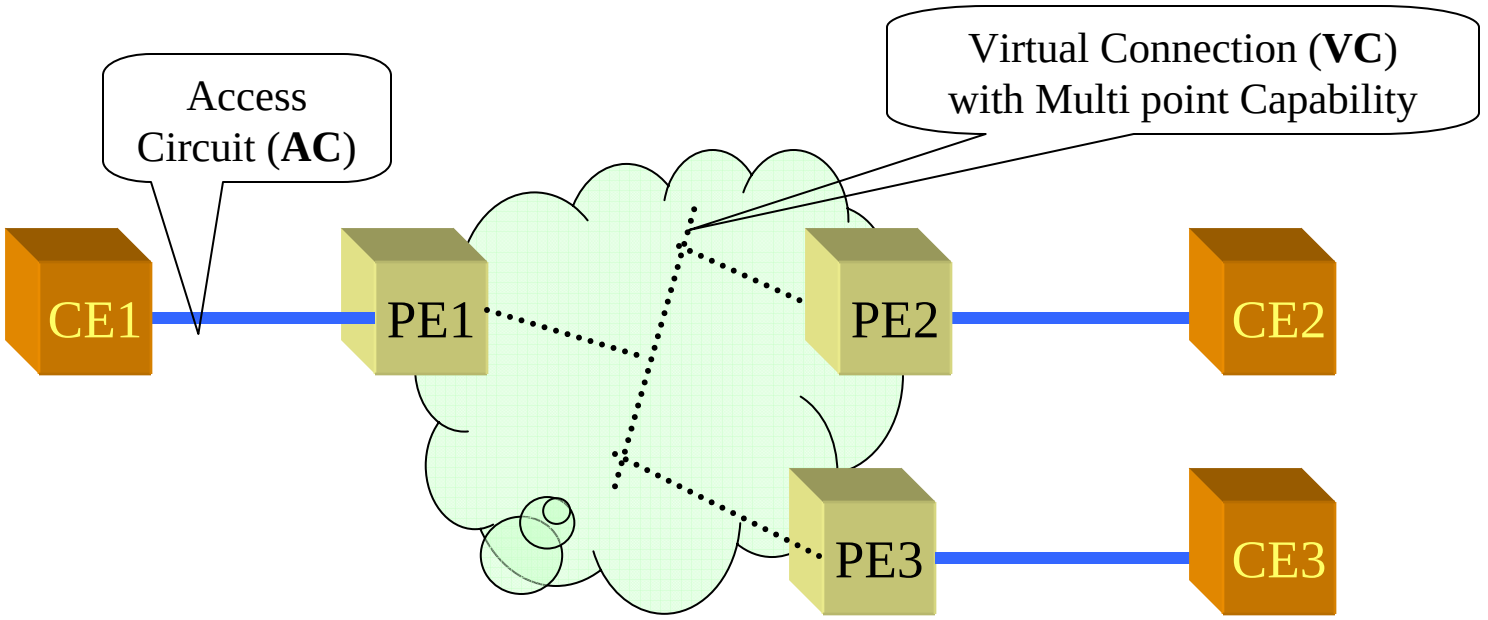


Ethernet Virtual Private Line



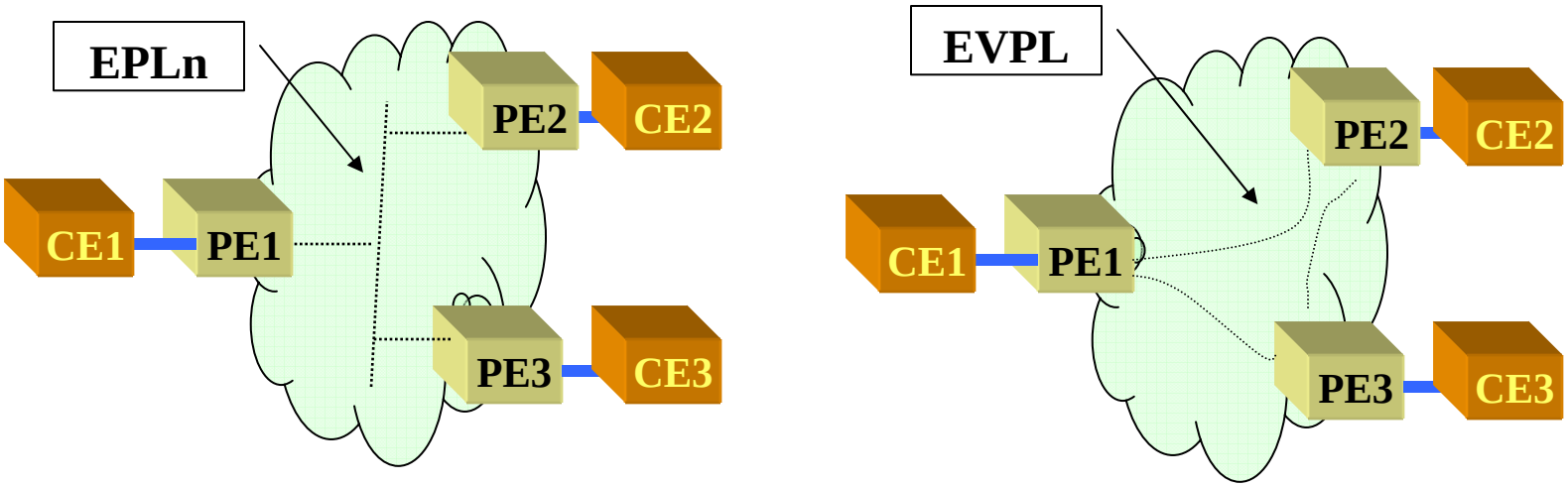
- Point-to-Point Virtual Connection
- Packets from CE1 is transparently sent to CE2,
- The packets from the CE1 is delivered to CE2 unaltered
- Mapping between the AC and VC is static.
- No Packet based switching is required in PEs

Ethernet Private LAN Service



- Provides Multipoint connectivity
- L2VPN Capability connecting multiple site over a metro-Ethernet
- The packets from the CE1 is delivered to CE2 unaltered
- Mapping between the AC and VC is static.
- PE should run a virtual switch instance for each VPLS.
- Switching based on MAC address

EVPL & EPLn



- Adding a new site through EVPL requires creating EVC for each peer and configuring in other locations
- Adding a site in EPLn, just requires configuration of the new site

midas

Other Service Attributes

- Service Multiplexing
  - Capability of the UNI to handle more EVCs
- Bundling
  - Two or more VLANs mapped to the single EVC
  - All-In-One Bundling is mapping all VLANs from a UNI to a single EVC

Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

Ethernet Switching : Recap

- Forwarding based on the Destination MAC Address
- FDB table stores destination MAC Address to Port mapping
- FDB Table is built by the source address learning mechanism
- Required loop free topology for Operation, if loops exist for redundancy then Spanning Tree Protocol (STP) is used to Prevent the loop.
- STP, Selects the port for forwarding; redundant ports are blocked

midas

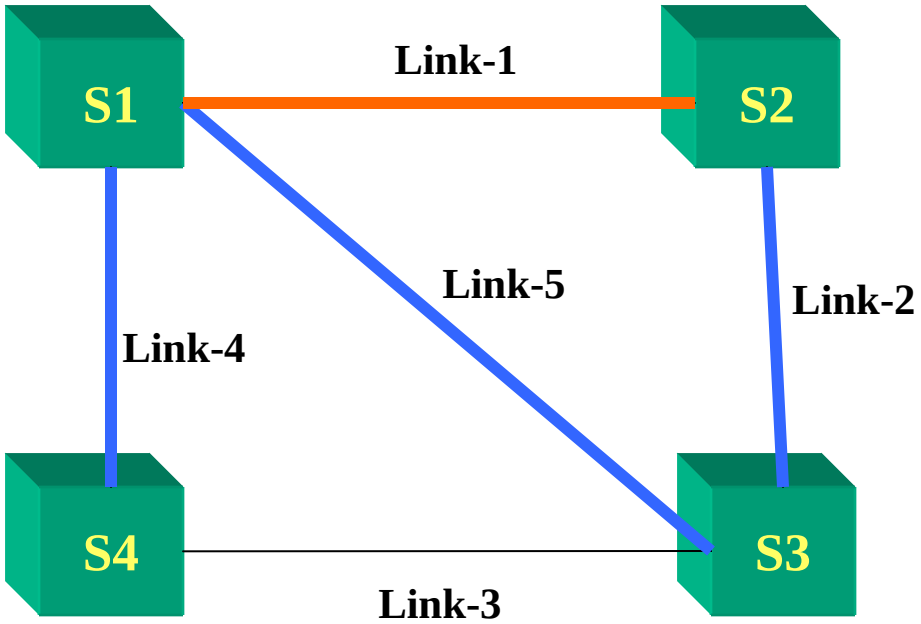
FDB Table

MAC Address	Port Map	Other Info
00:00:12:22:33:44	P1	
01:11:22:33:44:55	P1,P2	

- Bridges learn the FDB entries through source address learning
- Dynamic Entries are timedout if not refreshed
- Mgmt can create a static entry
- Port map can have more than one port as the destination (Multicast)

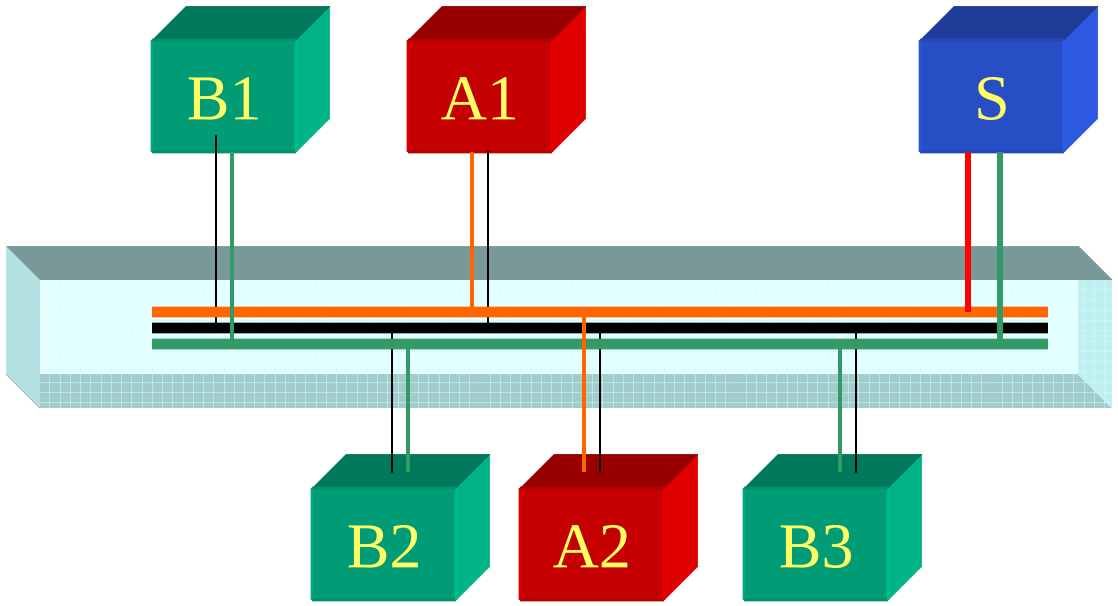
STP: Overview

- Distributed algorithm, involving all the switches in the network
- Prevents Loop in the Network
- Uses Redundant link, on active link failure
- Convergence time, in the order of tens of second
- Controls the port state of each switch



VLAN : An Introduction

- **Creates Multiple Virtual LANs within the same physical Network**
- **Restricts the Broadcast/Multicast packets within the VLAN**
- **Classification of the group is based on Port/MAC/IP Address etc**



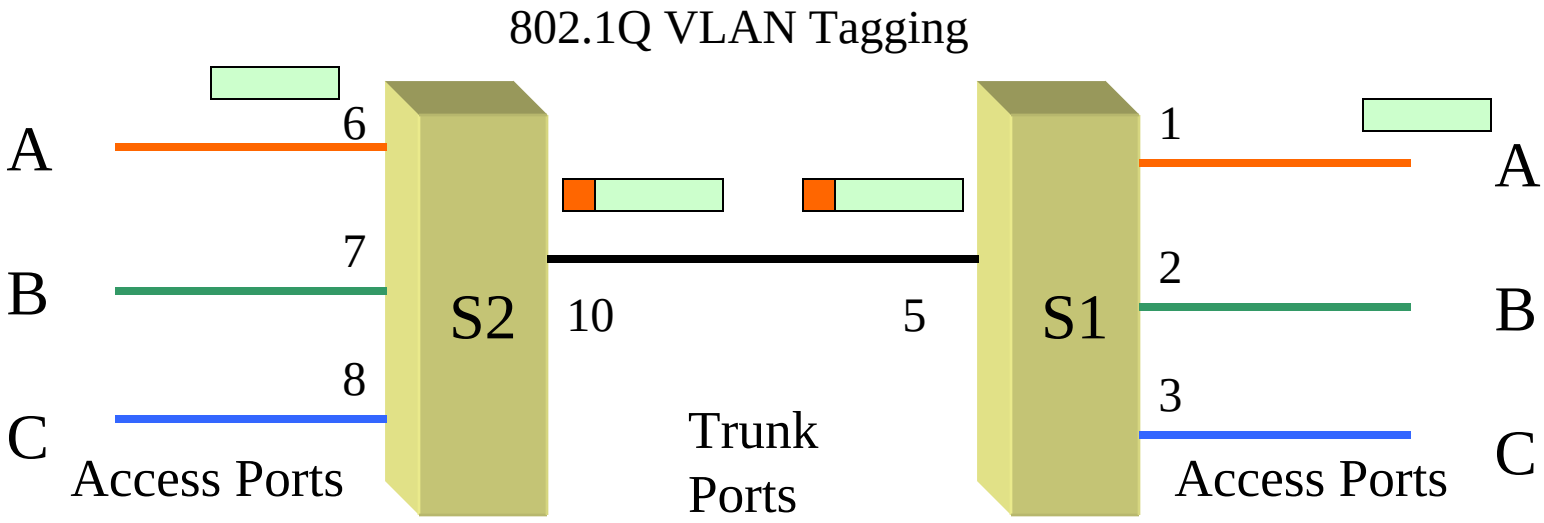
Port Based VLAN

VLAN	Portmap
A	6, 10
B	7, 10
C	8,10

VLAN Table

VLAN	Portmap
A	1, 5
B	2, 5
C	3, 5

Port	Default VLAN
1	A
2	B
3	C



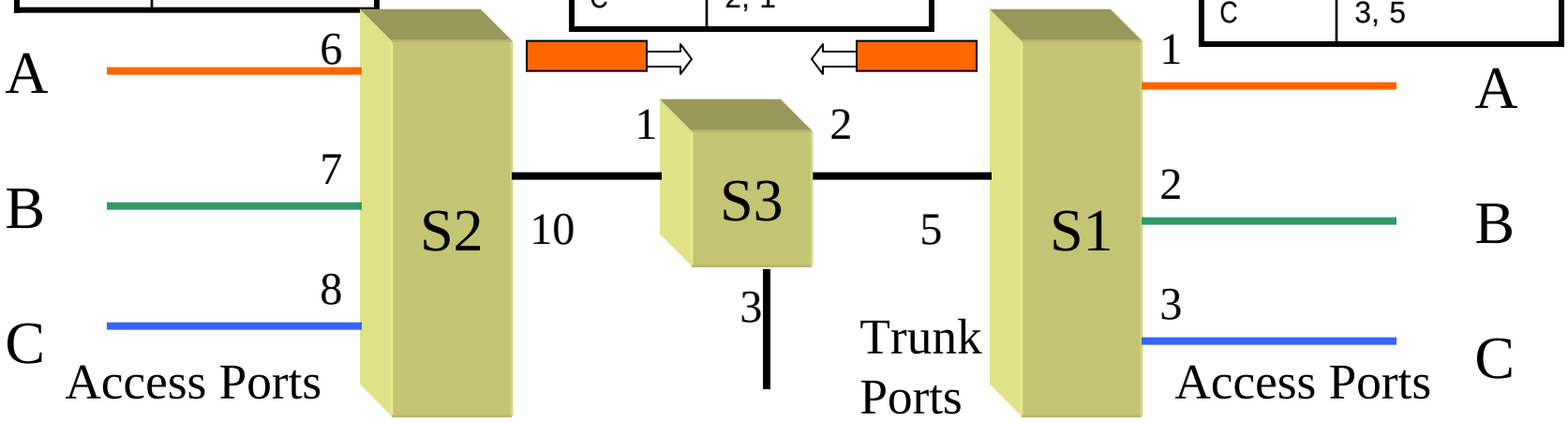
GVRP: Generic VLAN Registration Protocol

- Announces its VLAN membership to the other switches in the Network
- Bridges Update their **VLAN table** based on the GVRP Info.

VLAN	Portmap
A	6, 10
B	7, 10
C	8,10

VLAN	Portmap
A	2, 1
B	2, 1
C	2, 1

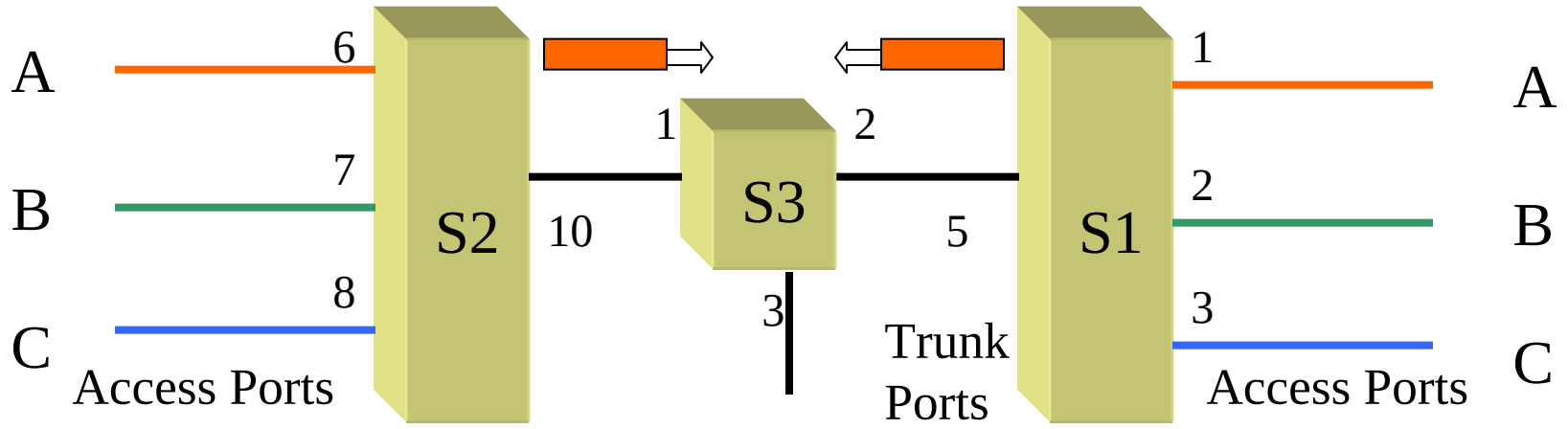
VLAN	Portmap
A	1, 5
B	2, 5
C	3, 5



GMRP: Generic Multicast Registration Protocol

- Bridge, announces its Multicast membership to other switches in the network
- Reduce the Multicast flooding in the network
- Bridges update the **FDB table** based on the GMRP information

MAC	Portmap
Multicast-1	1,2



VLAN Packet Format

Normal Ethernet Packet

7	1	6	6	2	48-1500	4
PA	SFD	DA	SA	Type/Len	Data	CRC

IEEE 802.1Q Tagged Frame

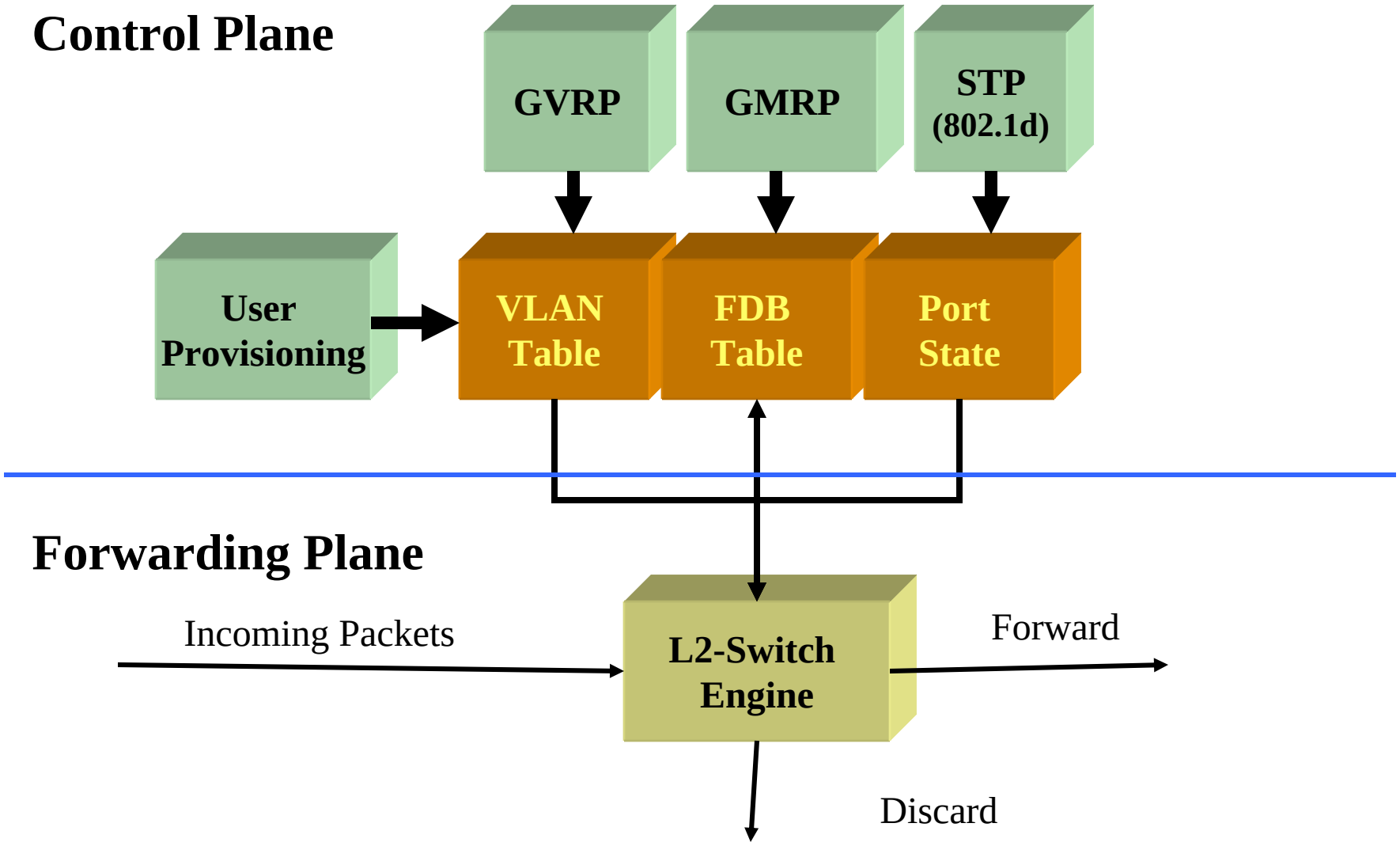
7	1	6	6	2	2	2	48-1500	4
PA	SFD	DA	SA	TPI	Tag	Type/Len	Data	CRC

802.1p CoS

Canonical Format Identifier

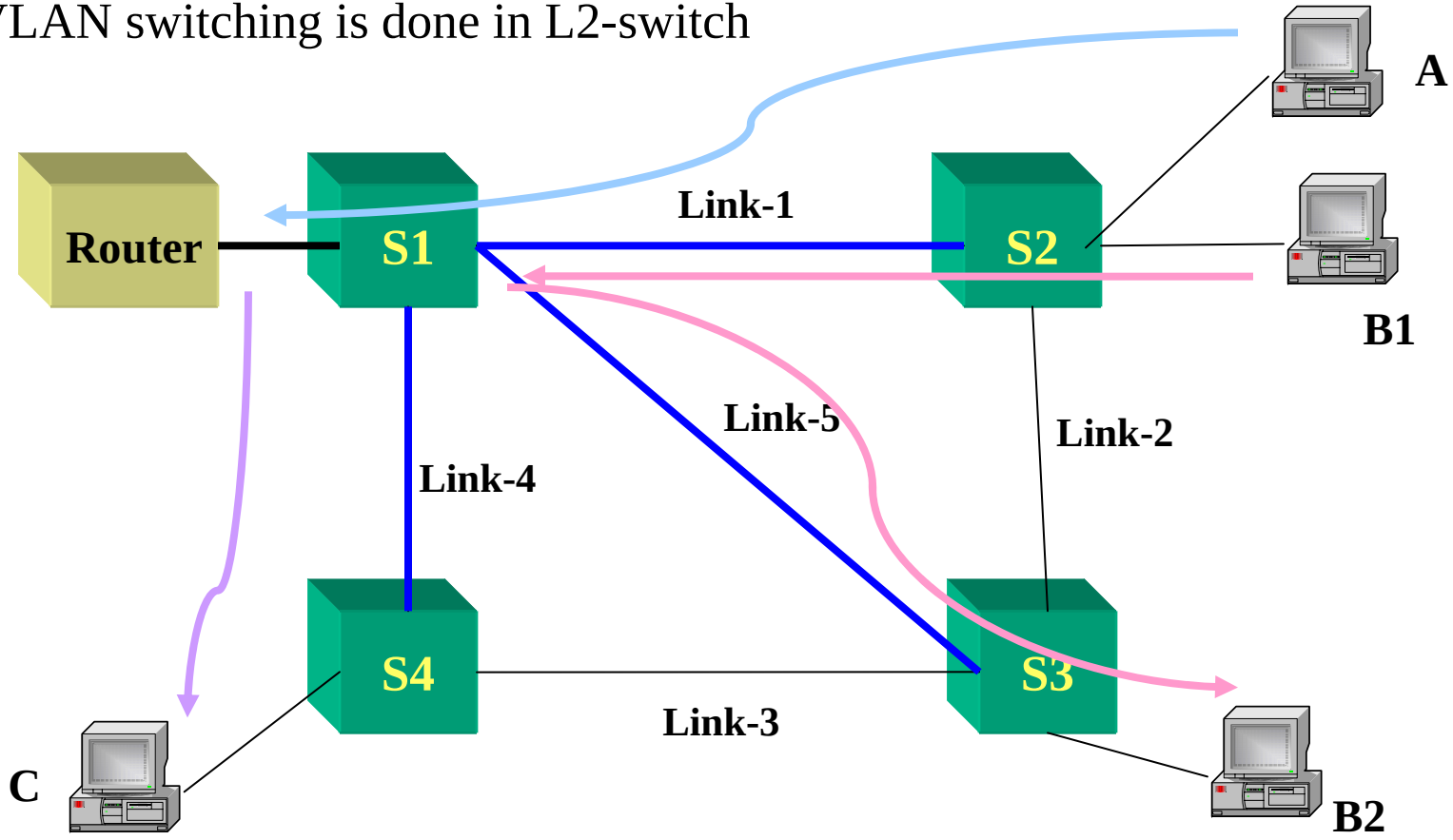
Pri	CFI	VID
3 bits	1 bits	12 bits

Ethernet Control & Forwarding



Ethernet in Metro

- Each Subscriber is assigned a Unique VLAN Id
- Inter-VLAN Switching is done via the L3-Switch/Router
- Redundant Links are blocked through STP
- Intra VLAN switching is done in L2-switch



midas

Enhancements Required

- Ethernet technology is not sufficient for the Access Network
  - STP Convergence time is in the order of few 10's of seconds
  - Redundant Links in STP are left unused
  - Broadcast traffic in the network has to be controlled
  - Transparent tunneling of Enterprise network traffic across the metro for EVPL service

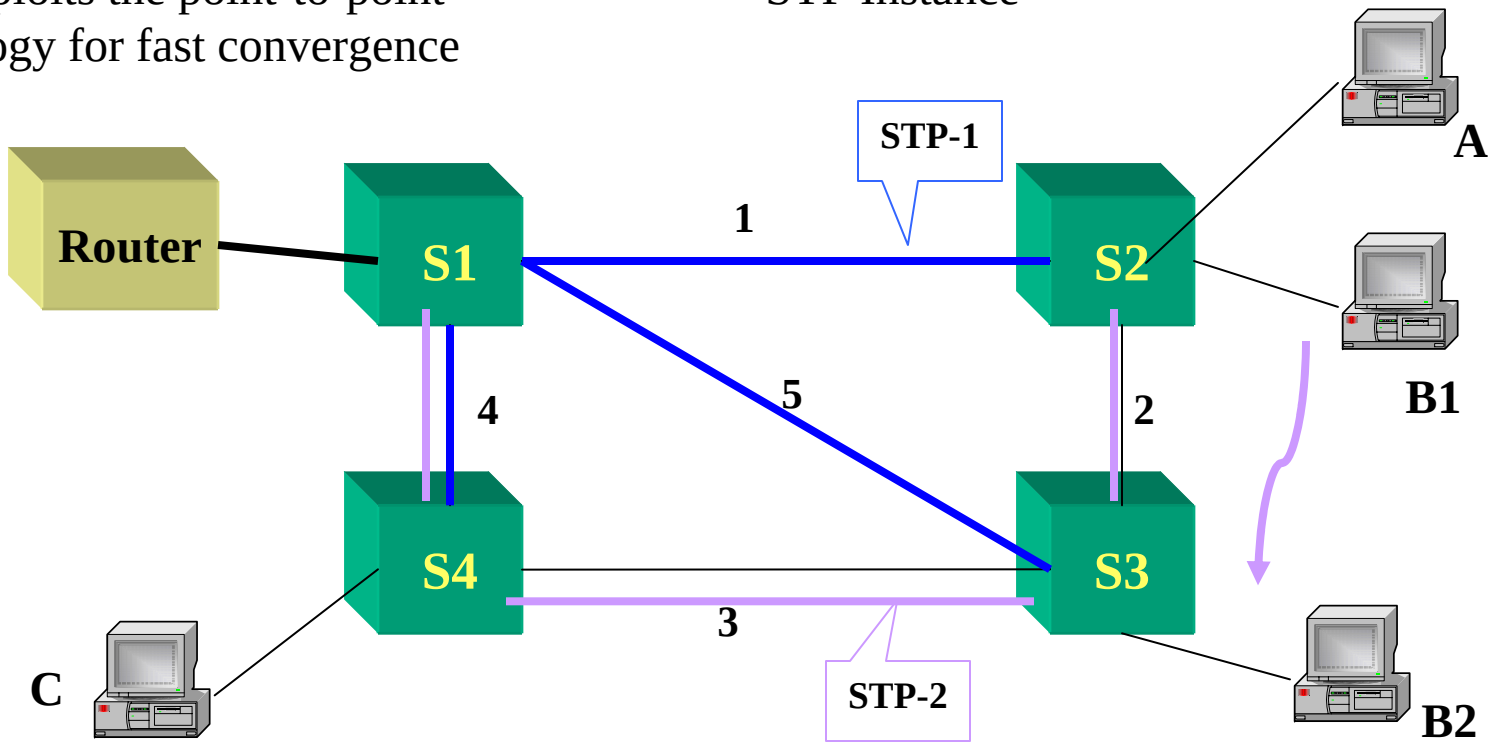
STP Drawbacks

RSTP (IEEE 802.1w)

- In STP, the convergence time for fail-over is in the order of few tens of seconds.
- RSTP (IEEE 802.1w), to reduce the convergence time.Inter-VLAN
  - Exploits the point-to-point topology for fast convergence

MSTP (IEEE 802.1s)

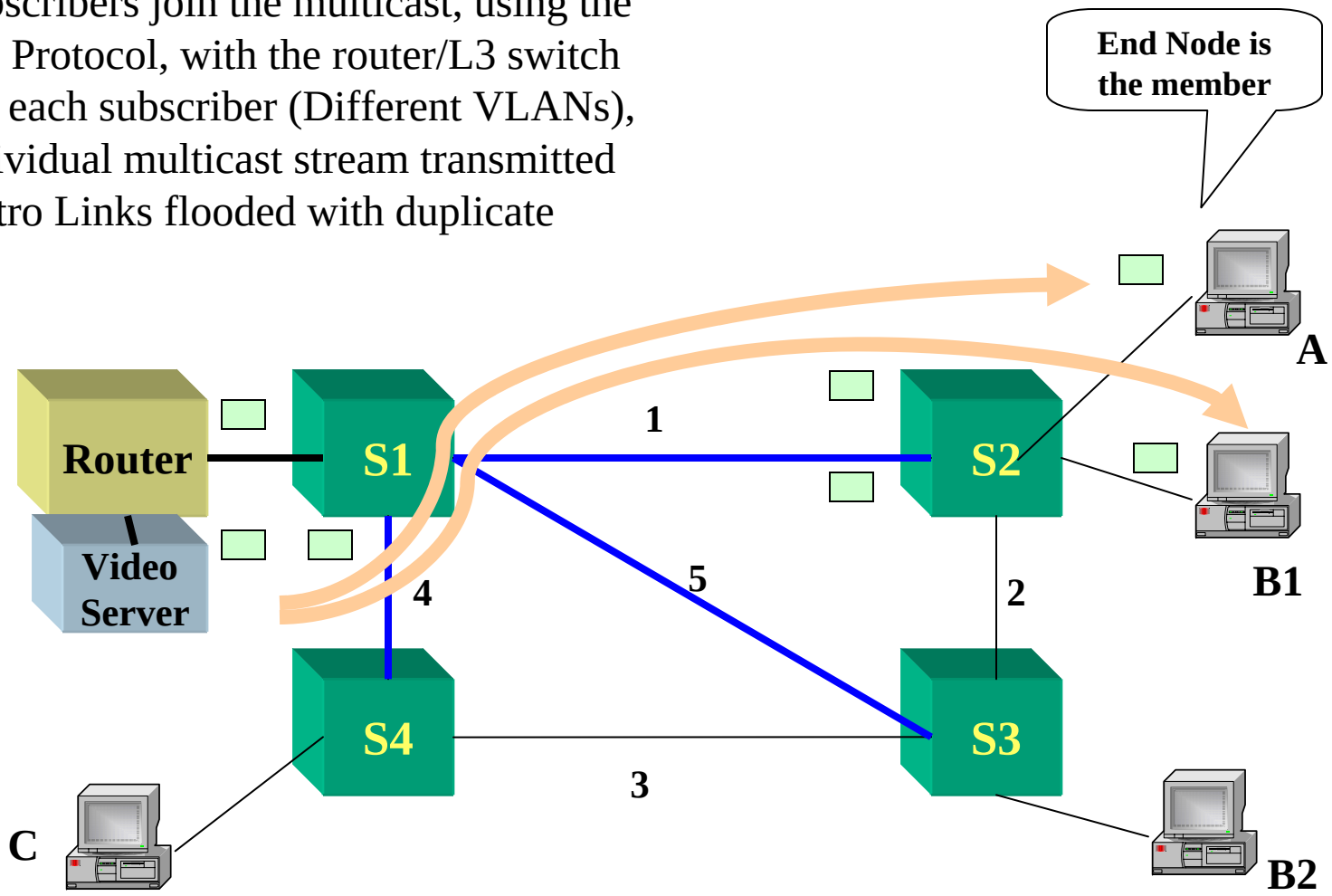
- Redundant Links are unused in the STP
  - MSTP (IEEE 802.1s) defines multiple instances of spanning trees
  - Set of VLANs can be mapped to a STP Instance



Multicast Streaming

IGMP

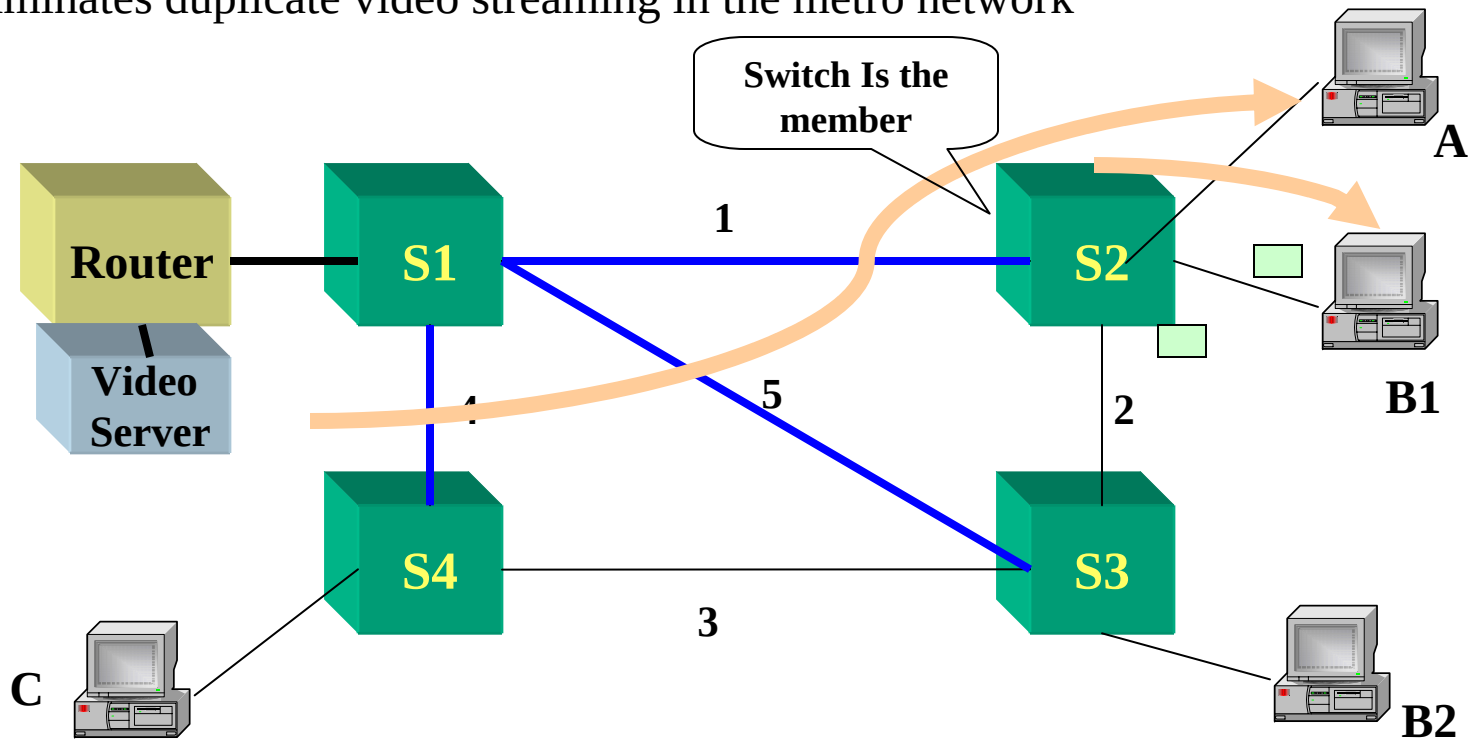
- Subscribers join the multicast, using the IGMP Protocol, with the router/L3 switch
- For each subscriber (Different VLANs), & individual multicast stream transmitted
- Metro Links flooded with duplicate traffic



Multicast Streaming

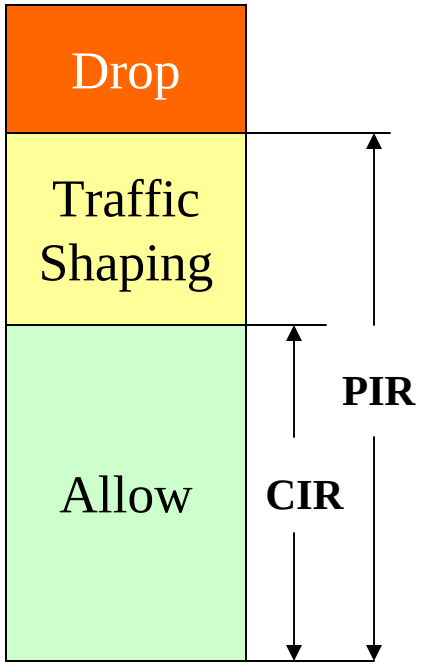
IGMP Snooping & Multicast Streaming

- Multicast Stream is Broadcasted in a separate VLAN
- L2-Switch snoops the IGMP requests for multicast streams
- Generates only one IGMP request for a particular stream
- Switches video streaming from the multicast VLAN to the subsciber ports
- Eliminates duplicate video streaming in the metro network



Bandwidth Provisioning

- Bandwidth profiling for each class
  - Classification based on Application, UNI etc
  - Either in the UNI level or in EVC level
  - On Ingress or Egress
  - CIR (Commited Information Rate)
  - PIR (Peak information rate)
  - CBS (Commited Burst Size)
  - MBS (Maximum Burst size)



Layer-2 Control Protocols

- Does the PE need to process L2 Control protocols from CE (e.g STP, GMRP, GVRP)?
  - **Peer**
    - PE treats CE as one of the nodes in the Networks and participates in the protocol.
  - **Discard**
    - Discard any control packets from the CE.
  - **Tunnel**
    - Tunnel the control packets, like any other data packets across the metro (requires VLAN transparency).

What is VLAN Transparency?

- In current 802.1Q Based Access Network, the packet from the subscriber is tagged with the Metro VLAN Id
- The VLAN-Id significant to the enterprise LAN is lost when the packet reaches to the other end
- Layer-2 Control Protocols cannot be used within the Enterprise network connected across the metro.

- Some method of preserving the subscriber VLAN across the access network is called VLAN transparency

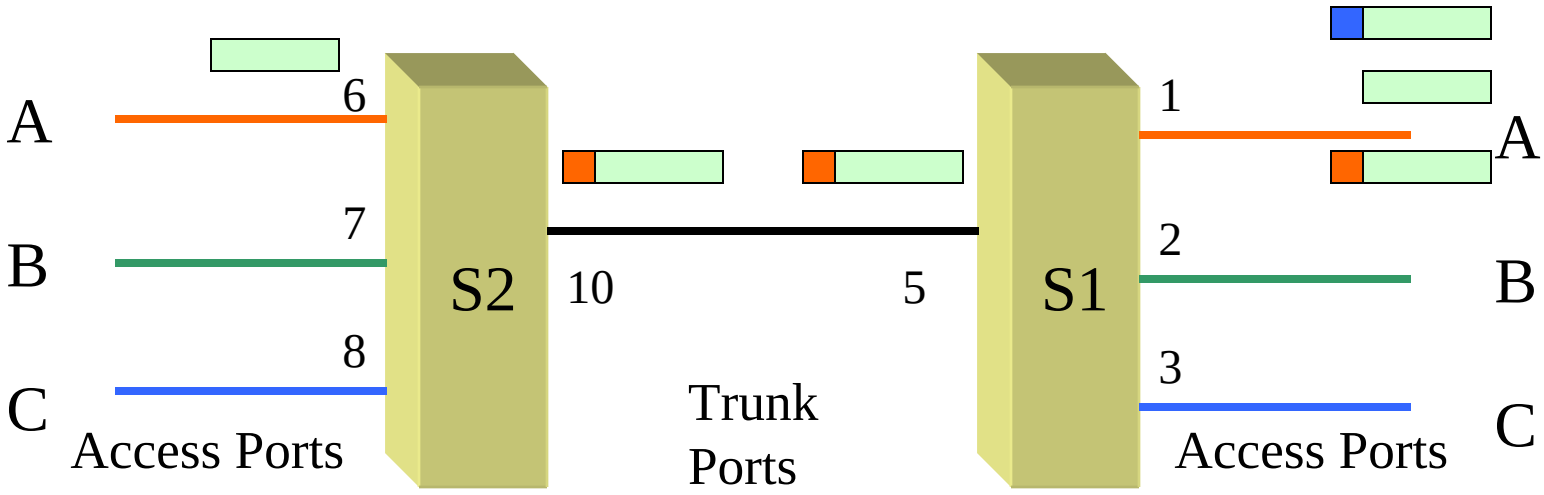
VLAN Transparency?

VLAN	Portmap
A	6, 10
B	7, 10
C	8,10

VLAN Table

VLAN	Portmap
A	1, 5
B	2, 5
C	3, 5

Port	Default VLAN
1	A
2	B
3	C



Is VLAN Transparency Essential?

- For Domestic subscribers, VLAN transparency is not required
- Ethernet service without VLAN transparency is sometimes referred as
  - ERS (Ethernet Relay Service)
  - ERMS (Ethernet Relay Multipoint Service)
- It is required for some enterprise LANs to interconnect.
- With VLAN transparency, switches can be used as CE device.
- Required for tunneling L2 control packets

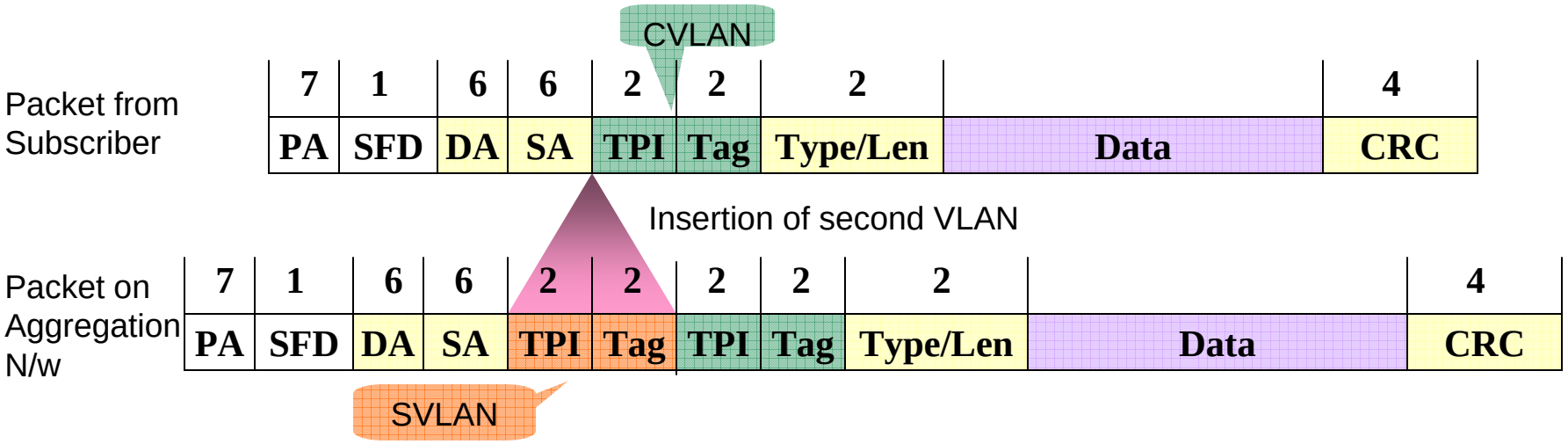
802.1Q VLAN for Metro : Pros & Cons

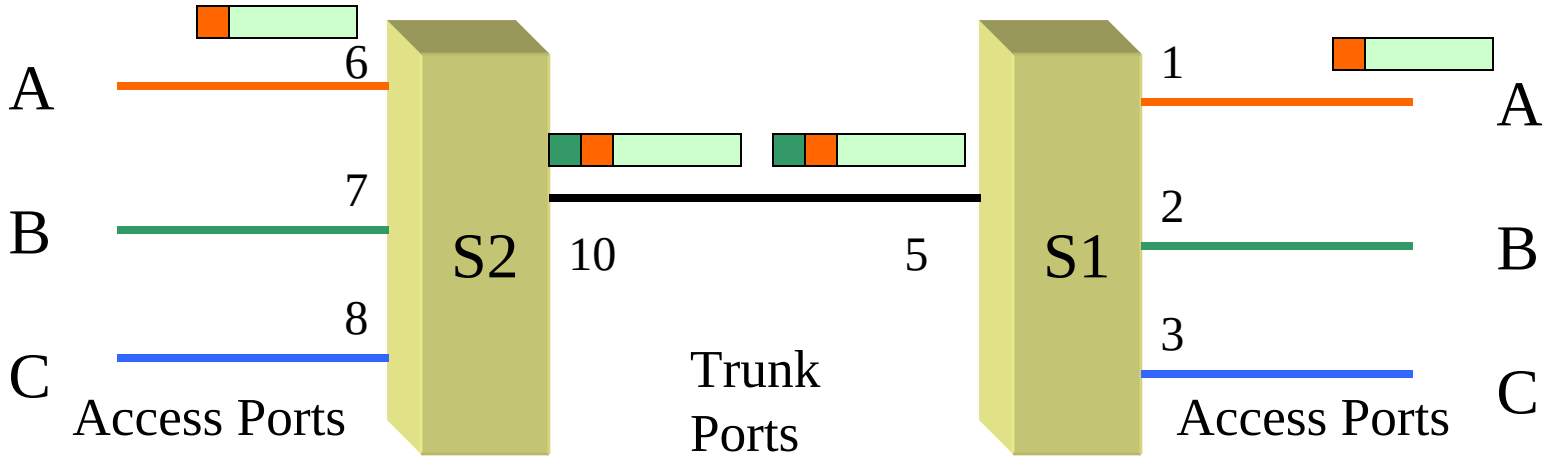
Parameter	*VLAN Solution*	*Drawbacks*
User Identification	Assigning Individual VLAN for Each User	Limited to 4096 per VLAN domain
Security	Assigning unique VLAN Id for each subscriber	-
Redundancy	STP	-
Resiliency	RSTP	Order of seconds
Load Sharing	MSTP	-
Multicasting	IGMP Snooping	-
End to End VLAN Transparency	**None**	**No VLAN Transparency**

VLAN Transparency in Ethernet

- In Ethernet Metro, VLAN transparency can be done, by **double 802.1Q** tagging also referred as **Q-In-Q**
- Above the subscriber VLAN tag, Metro VLAN tag is added by the PE
- Inter-operability Issues?

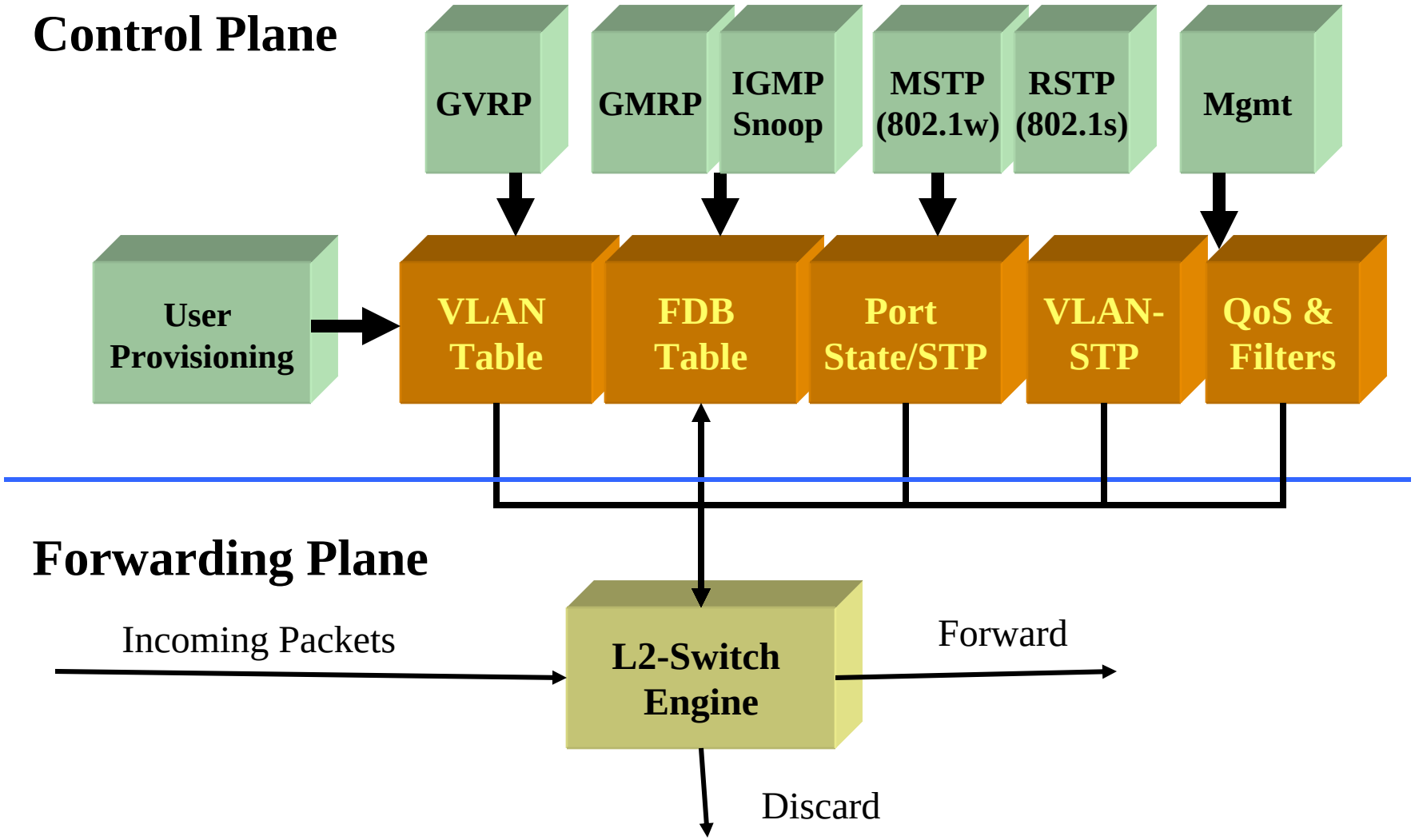
Q-In-Q


CVLAN
Packet from Subscriber
7 1 6 6 2 2 2 4
PA SFD DA SA TPI Tag Type/Len Data CRC
Insertion of second VLAN
Packet on Aggregation N/w
7 1 6 6 2 2 2 2 2 4
PA SFD DA SA TPI Tag TPI Tag Type/Len Data CRC
SVLAN


A
B
C
6
7
8
S2
10
5
S1
1
2
3
A
B
C
Access Ports
Trunk Ports
Access Ports

midas

Control & Forwarding for Metro Net



midas

Ethernet Summary

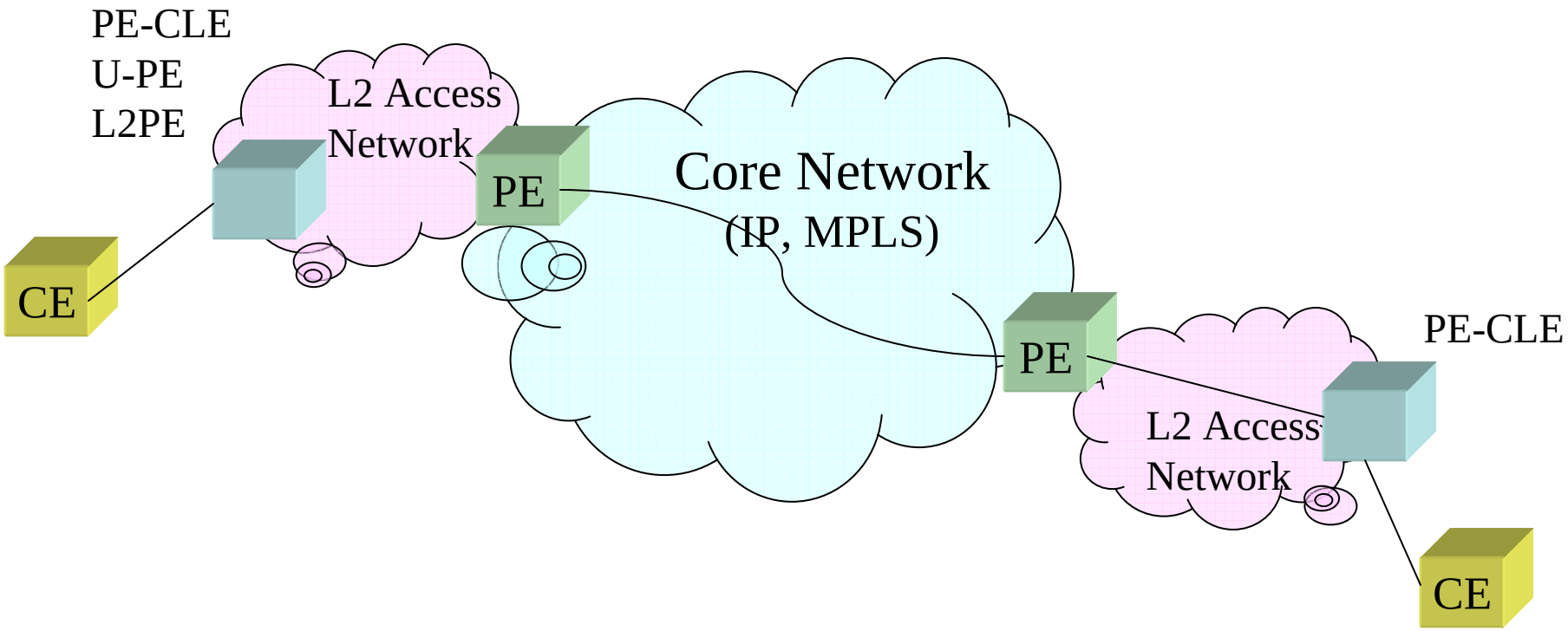
	802.1Q	Q-In-Q
VLAN Transparency	No	Yes
Handling L2 Control Protocols	Peer or Discard	Peer, Discard or Tunnel
Services	ERS, ERMS	ERS, ERMS, EVPL, EPLn
Maximum Subscriber	Maximum subscribers 4096	Maximum subscribers 4096
Scalability	Limited	Limited

midas

Why not Scalable?

- In large Networks, STP convergence is not proven to the satisfaction.
- Large Networks also requires huge MAC address table in each PE switches, making it more expensive
  - For 10,000 subscribers each with an average of 10 MAC address requires 100,000 MAC address size
- Ensuring QoS on the huge Ethernet metro is not easy

Hierarchical Network



PE-CLE : Provider Edge, Customer Located Equipment
U-PE : User PE
L2PE : Layer-2 PE

midas

Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

midas

Core Networks

- Mostly IP/MPLS based Network
  - Running above high speed links like POS, ATM, etc.
- Highly scalable
- Efficient Traffic Engineering
  - Traffic Re-routing
- Manageable
  - Proven management
- Stable control plane
  - Routing Algorithms deployed in the core are time-tested (e.g OSPF, BGP etc)

midas

Why not extend the core towards Subscriber?

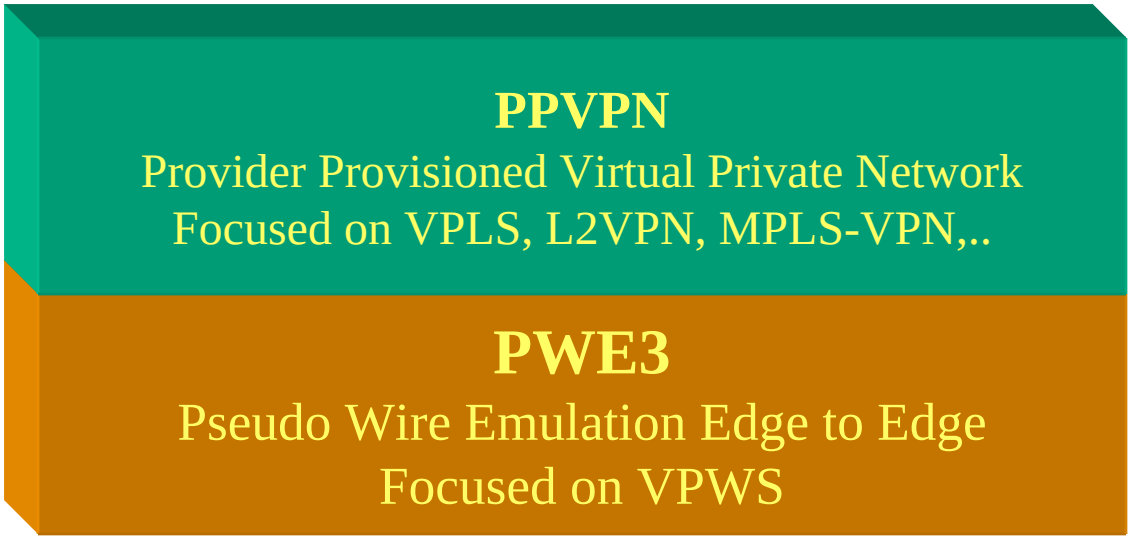
- Scalable
- POP requires a expensive router/MPLS switch
- With Fiber towards home, POP device (MTU-Multi-tenant Units) with IP/MPLS switching is expensive
- Manageability of IP/MPLS MTU is difficult compare to L2 device

midas

Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

IETF Workgroups



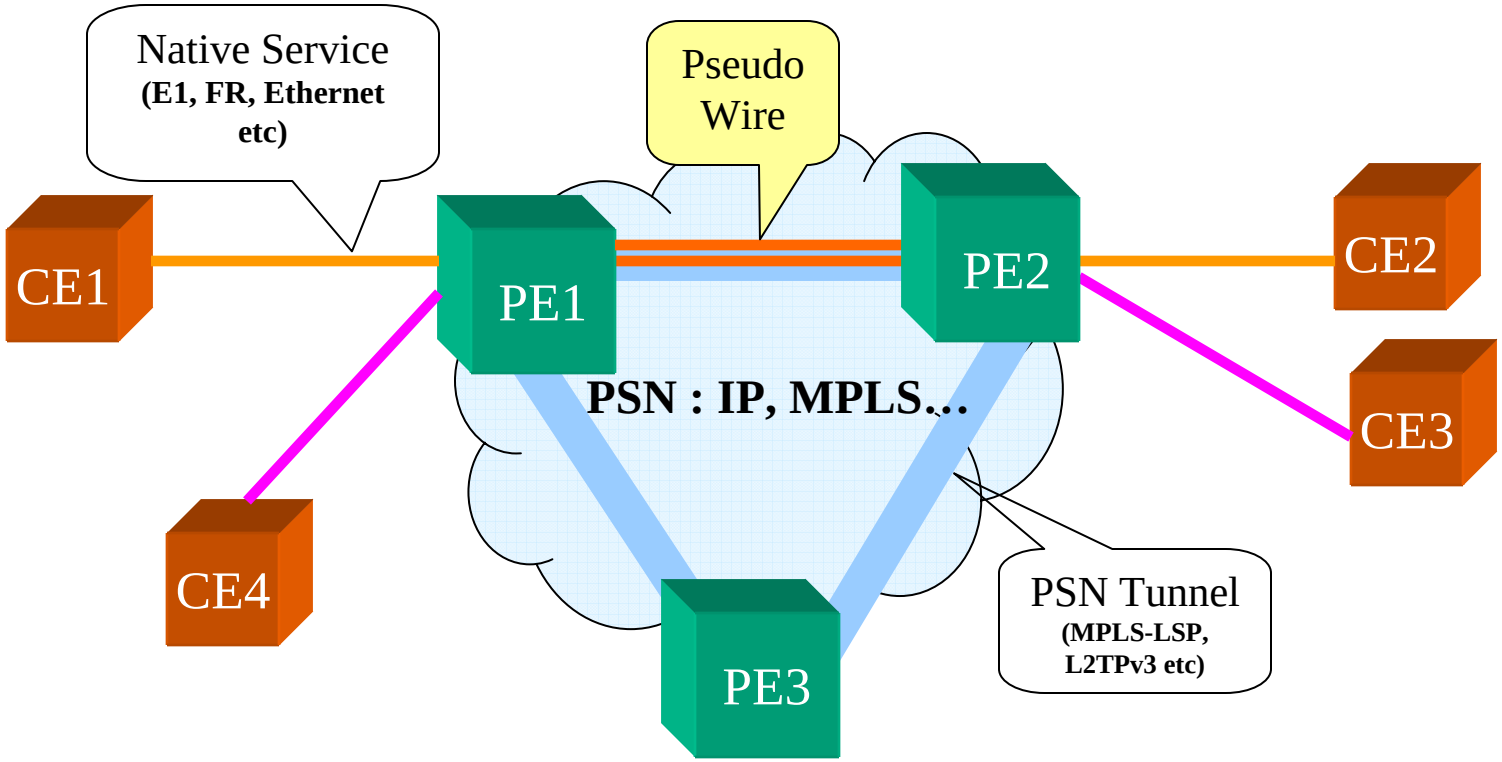
- Defines WAN & LAN services over Packet Switched Network (IP, MPLS ...)

midas

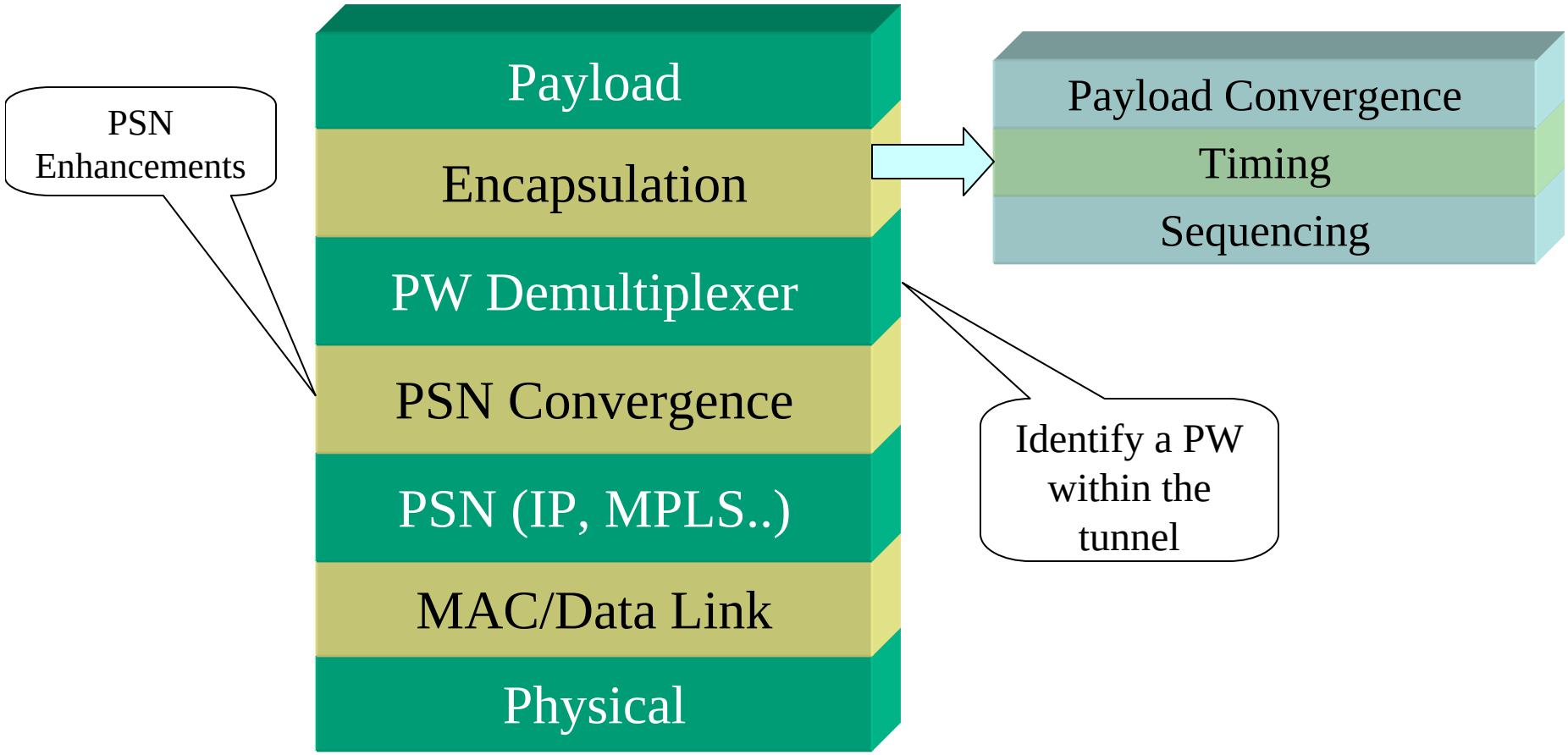
PWE3

- PWE3 is a mechanism that emulates the essential attribute of service over a PSN
  - PW will subject payloads to delay, jitter and re-ordering
  - Traffic Engineering required on PSN for certain services
- From Customer perspective, a PW is a unshared link or circuit of a service.

PWE3 Network Model



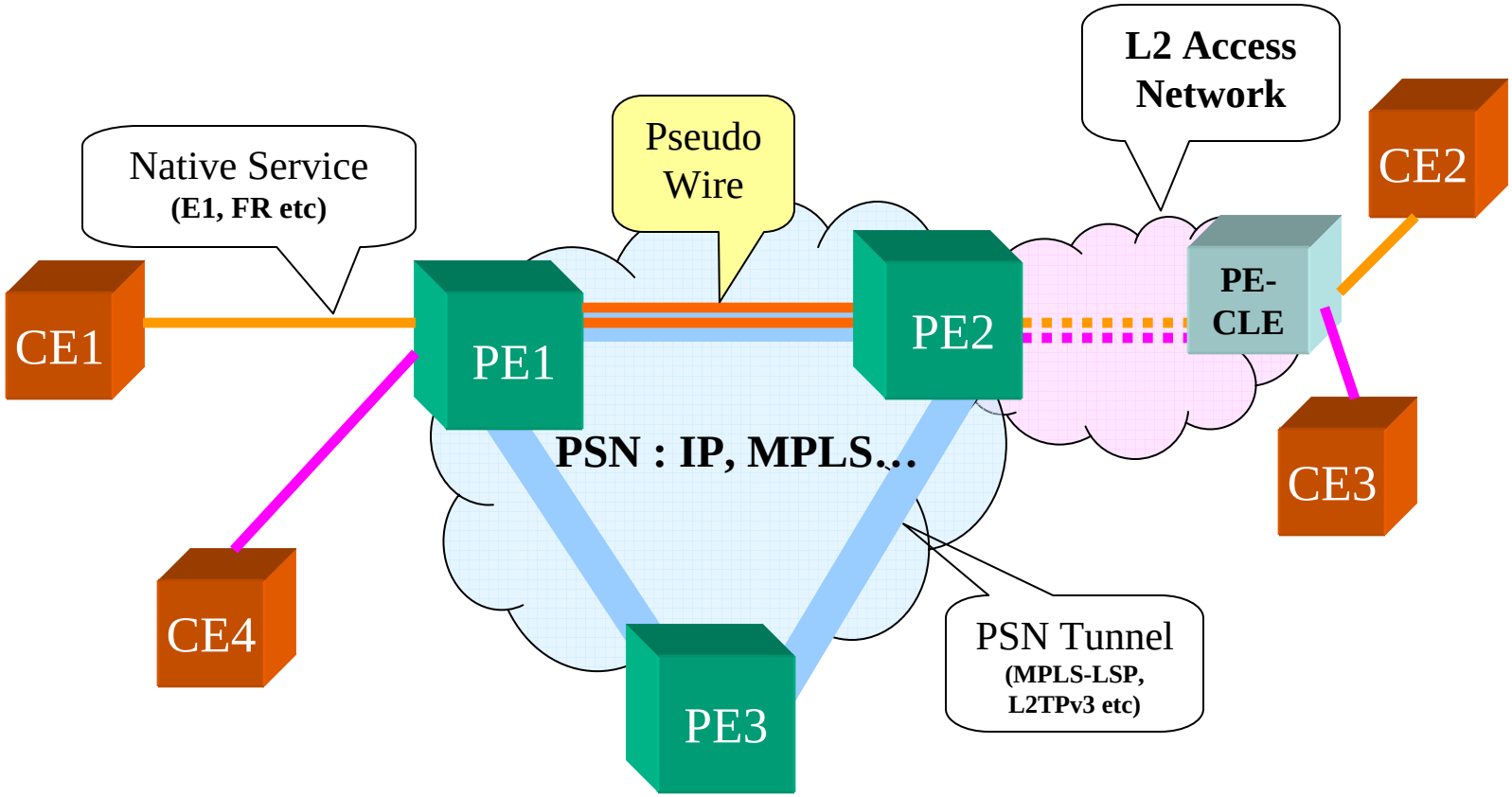
PWE3: Protocol Layers



VPLS (Virtual Private LAN Service)

- IETF's version of EPLn (Ethernet Private LAN Service)
- Ethernet only Extension to VPWS
- Adds Multipoint Capabilities through MAC address learning and forwarding

PWE3 Network Model with Ethernet Access Network



midas

Requirements for Control Plane and Forwarding in Core

- Mapping the Subscriber to the Virtual Connection over PSN
- Establish Virtual Circuit in the PSN
  - To Exchange Labels (VC-Id)
- To Transport Circuit Status
  - Link down, SVC link down etc
- In Sequence Delivery
  - Required for ATM & Frame Relay

midas

Control & Forwarding in Core

- Two Mechanisms are widely used
  - MPLS (Multiprotocol Label Switch)
  - L2TPv3 (Layer-2 Tunneling protocol) over the IP core

midas

Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

MPLS : Introduction

Destination Address	Next Hop Router	Port
206.103.10.x	60.10.1.1	1
203.199.x.x	60.10.1.1	1

Address Prefix	In Label	Out Label	Port
206.103.10.x	10	20	1
203.199.x.x	11	21	1

R1

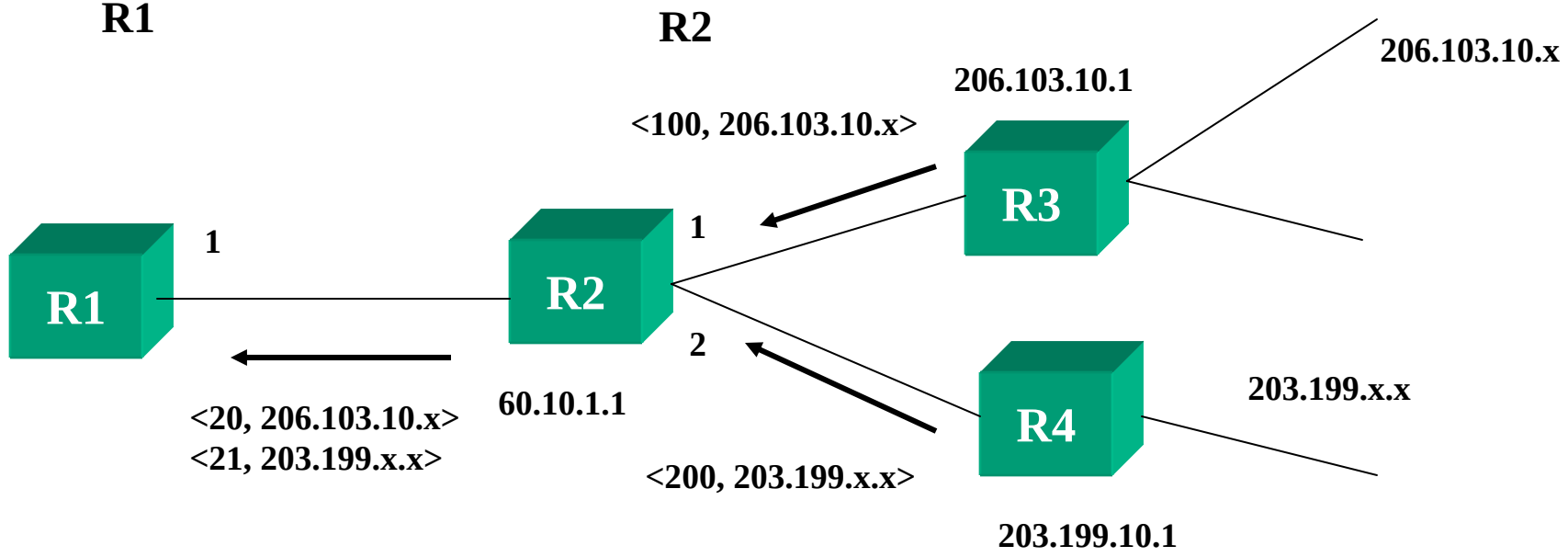
Destination Address	Next Hop Router	Port
206.103.10.x	206.103.10.1	1
203.199.x.x	203.199.10.1	2

Routing Table

Address Prefix	In Label	Out Label	Port
206.103.10.x	20	100	1
203.199.x.x	21	200	2

Label Table

R2



midas

MPLS : Introduction

Destination Address	Next Hop Router	Port
206.103.10.x	60.10.1.1	1
203.199.x.x	60.10.1.1	1

Address Prefix	In Label	Out Label	Port
206.103.10.x	10	20	1
203.199.x.x	11	21	1

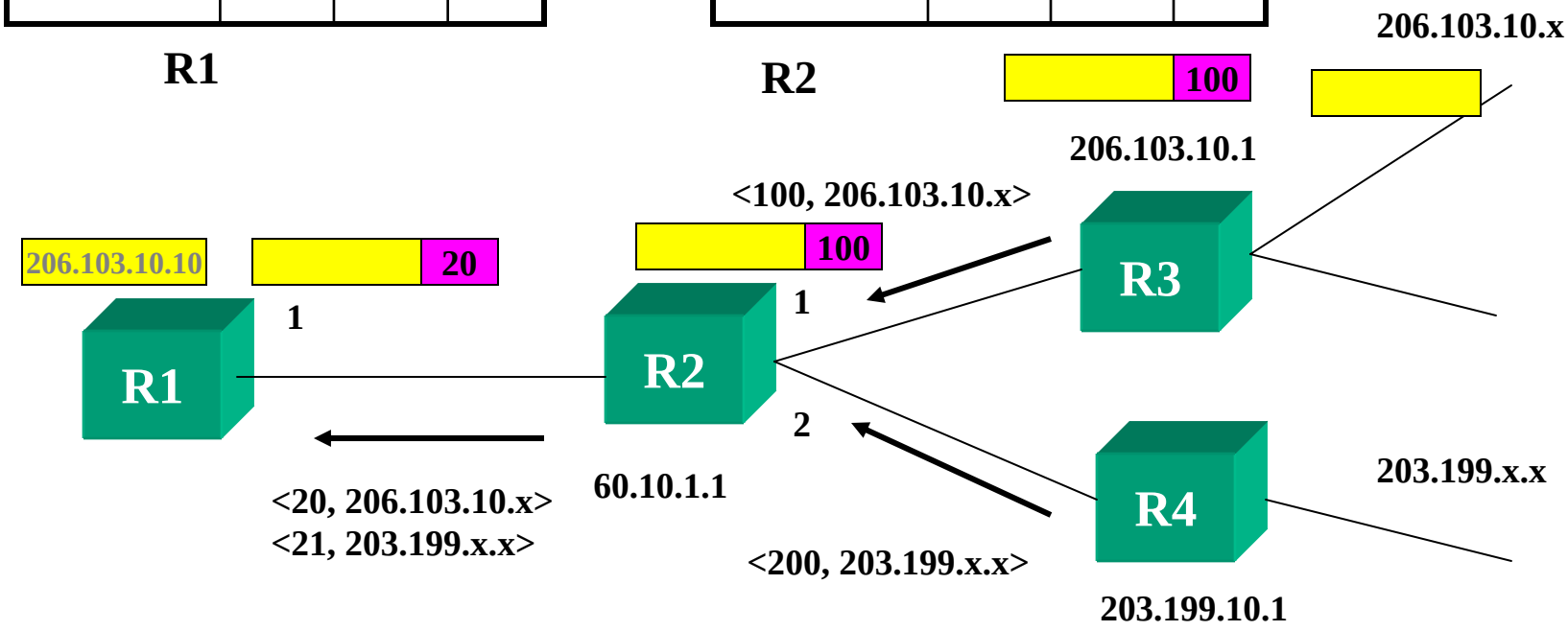
R1

Destination Address	Next Hop Router	Port
206.103.10.x	206.103.10.1	1
203.199.x.x	203.199.10.1	2

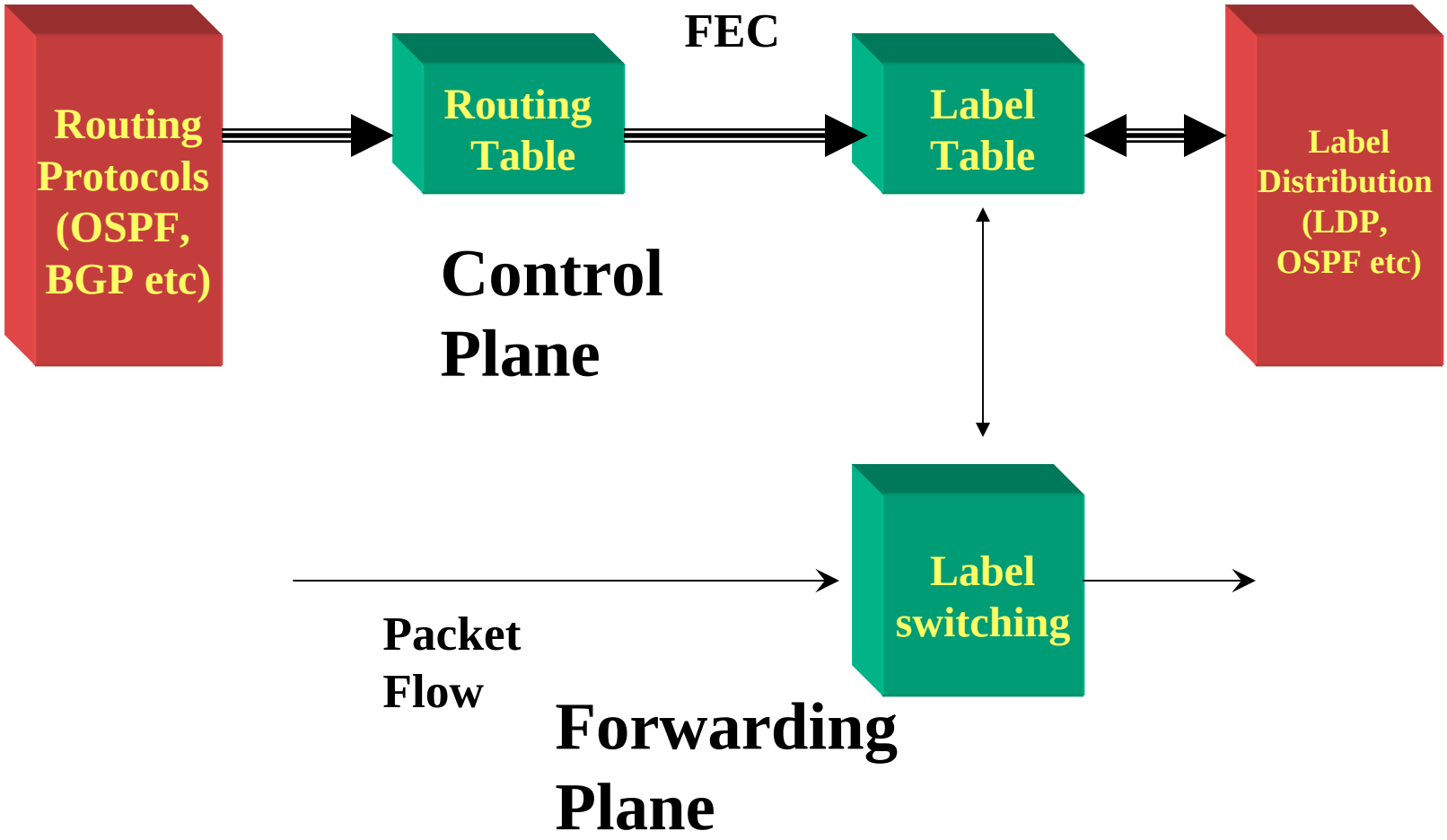
Routing Table

Address Prefix	In Label	Out Label	Port
206.103.10.x	20	100	1
203.199.x.x	21	200	2

Label Table



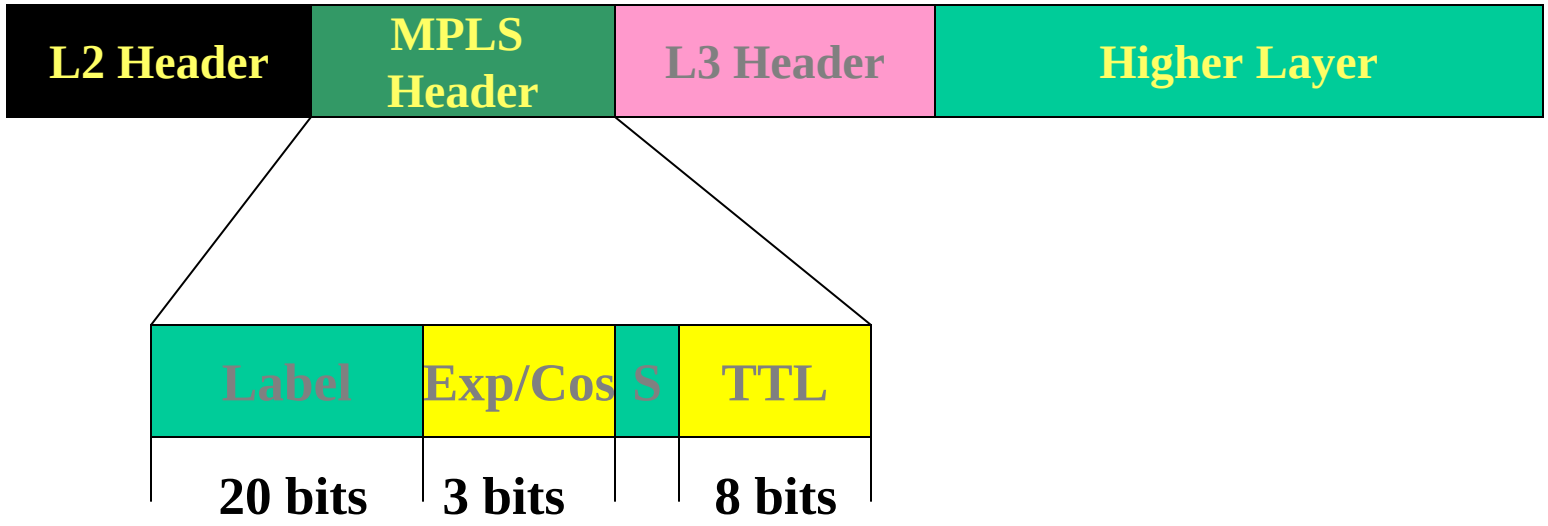
MPLS Control & Forwarding



midas

MPLS : Label Forwarding

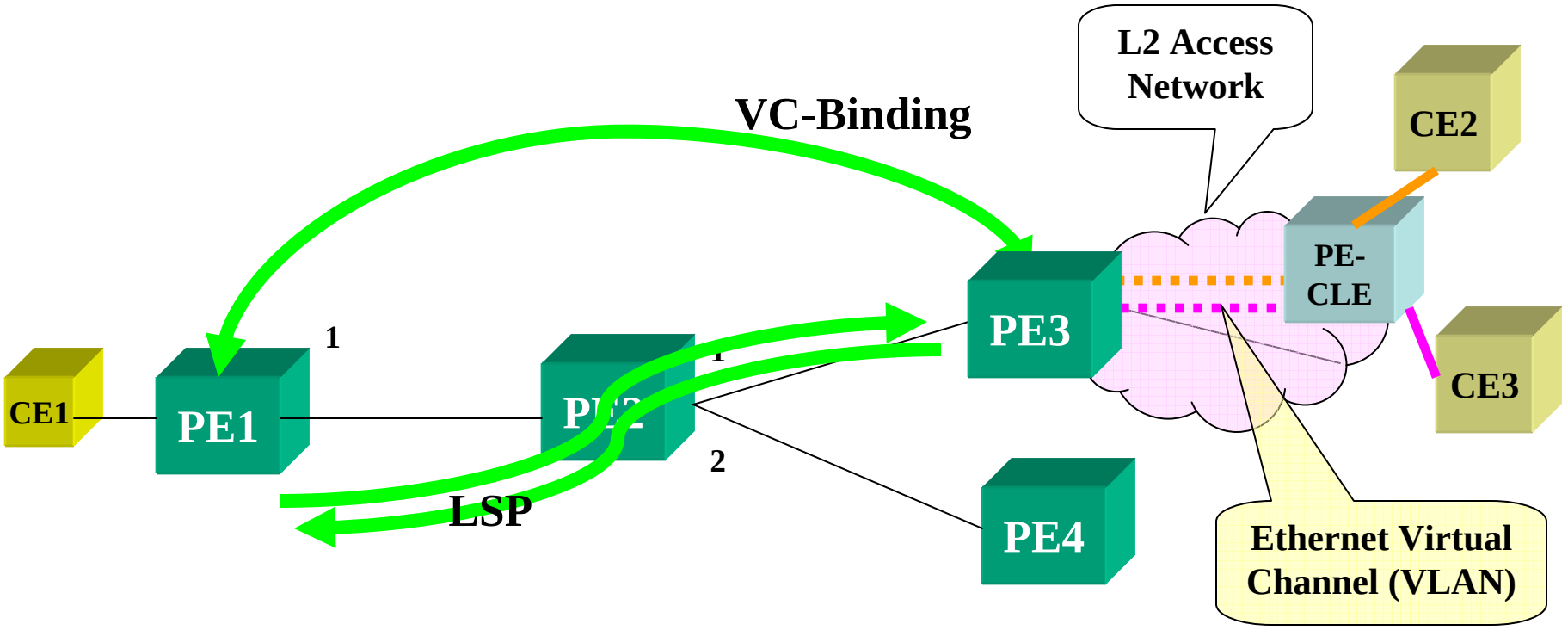
- Uses Data Link Label (ATM, Frame Relay etc)
- MPLS “Shim Header” between Data Link and IP Header



Tunneling in MPLS

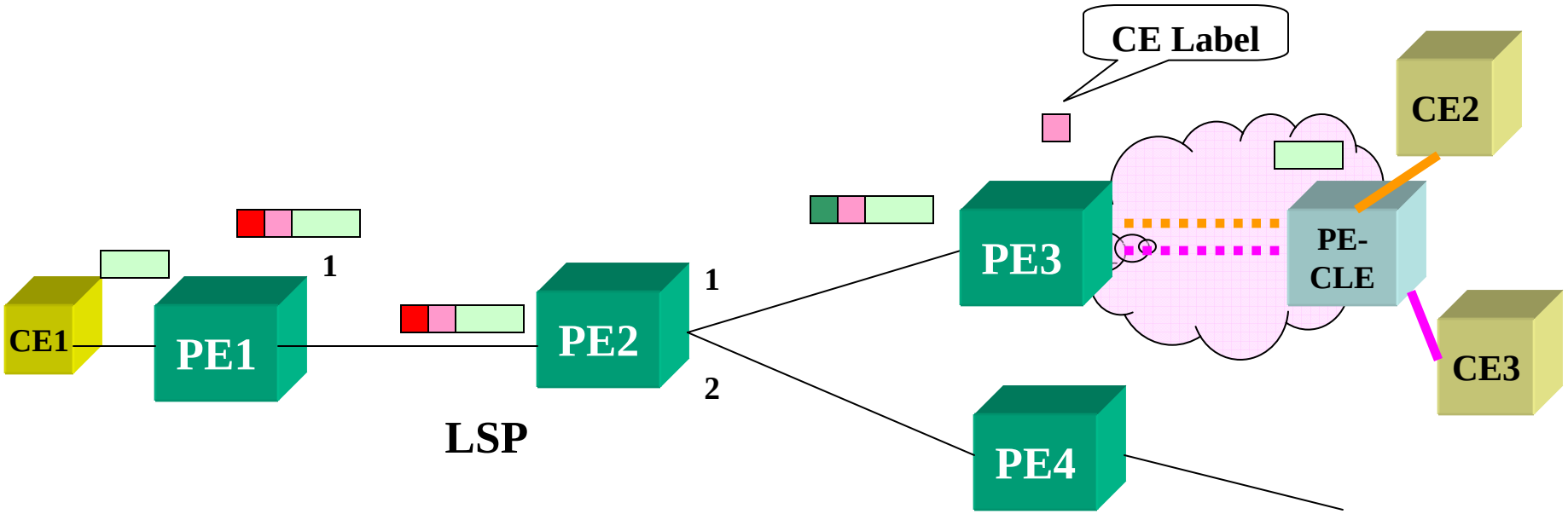
- L2 Transport over MPLS is implemented through label stacking (Tunnel Label & VC Label)
- The Egress PE allocates a VC label and binds the L2 Egress interface to the VC and signals the label to the ingress PE via targeted LDP session.
- PE router on detecting any failure, withdraws the corresponding VC label through LDP signaling.

MPLS : VC Label Binding



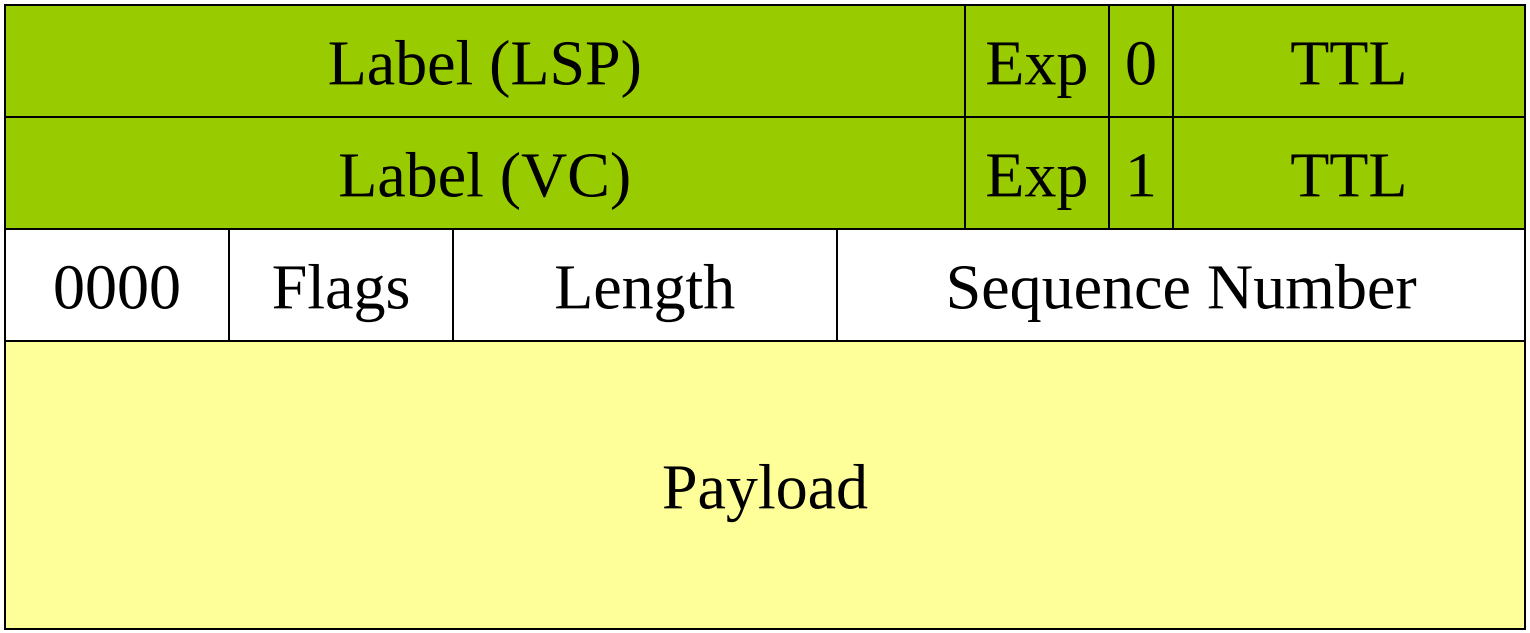
VC-Label for CE2 is signaled to PE1 through LDP

MPLS : Packet Flow



VC-Label for CE2 is signaled to PE1 through LDP

MPLS : Packet Format

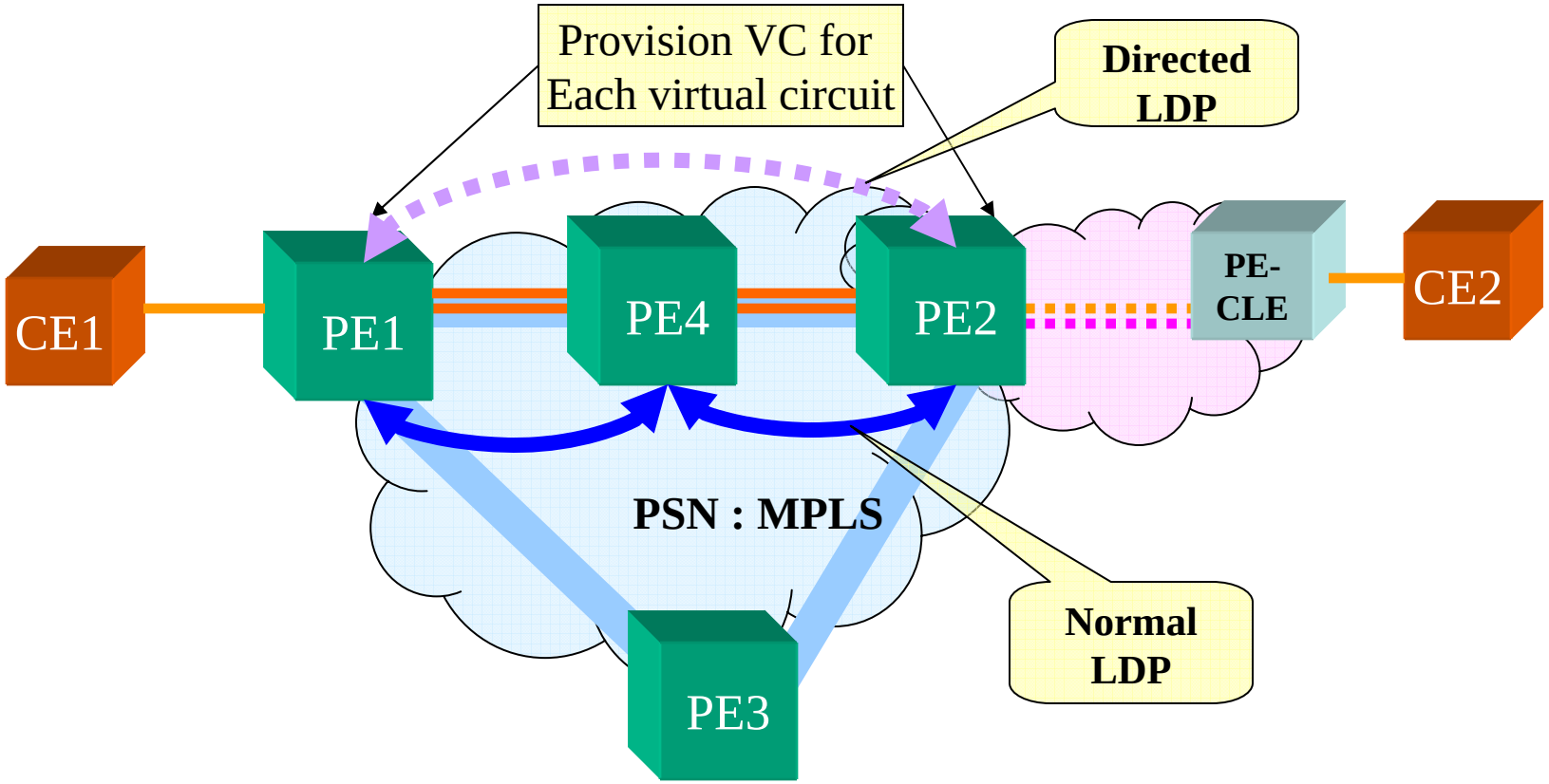


midas

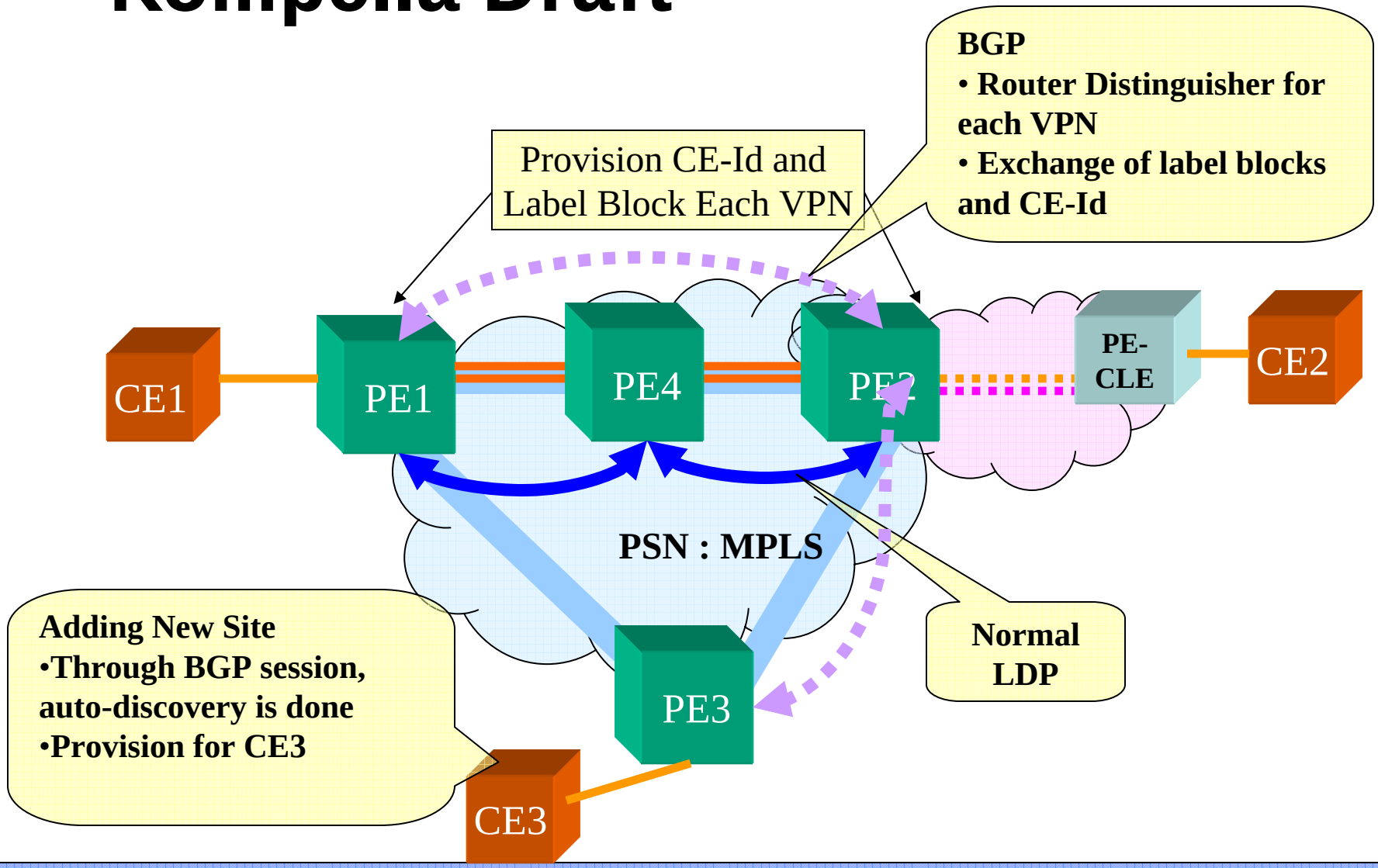
Control Plane Variants

- Martini & Kompella Drafts
  - Forwarding Plane & Packet Format Similar in Both Cases
  - Differ in provisioning the VPN and Signaling of the information to the other side
  - Martini supports only Point-to-Point, Kompella supports Multipoint services
  - Martini is more popular for its scalability

Martini Draft



Kompella Draft



Agenda

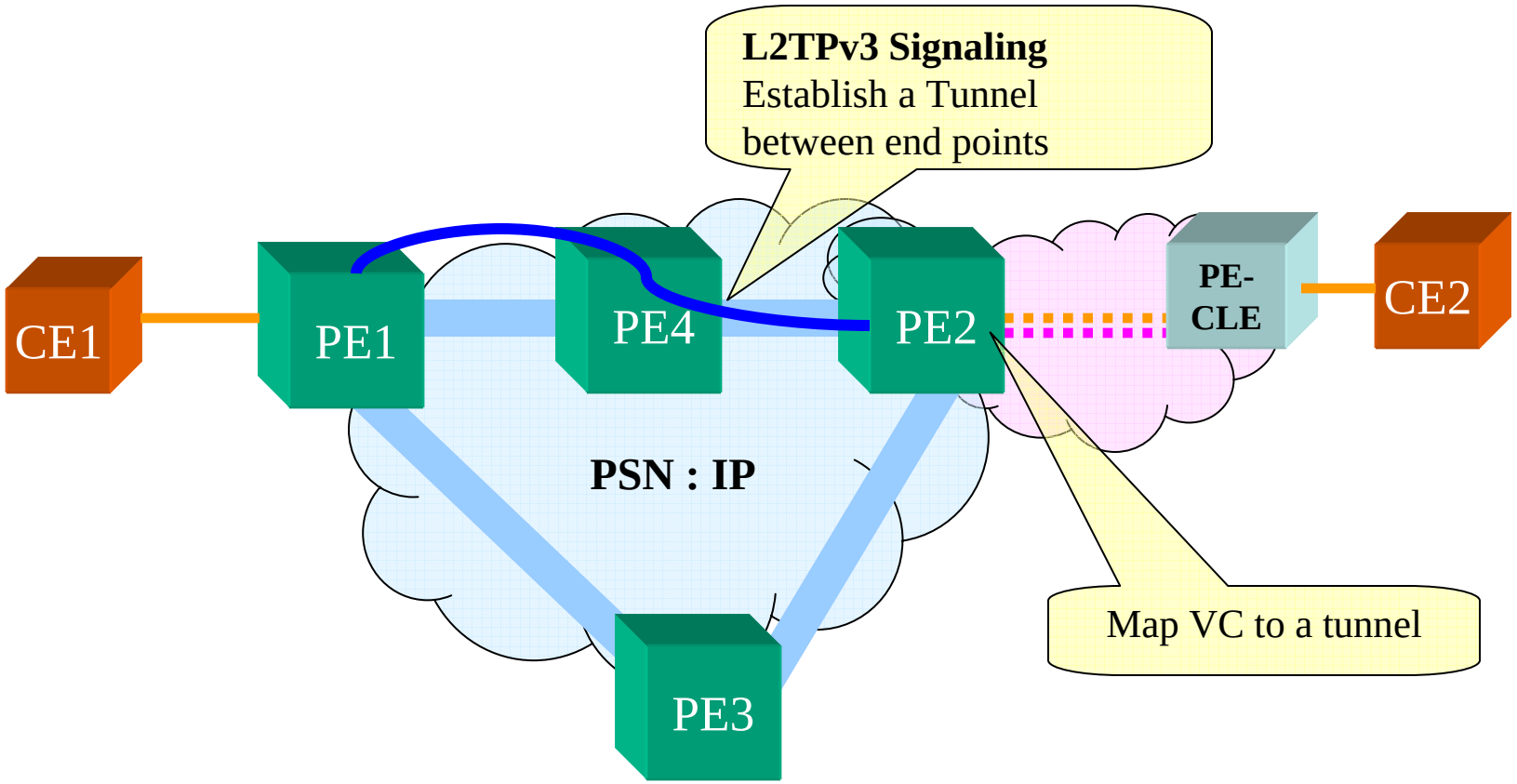
- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

midas

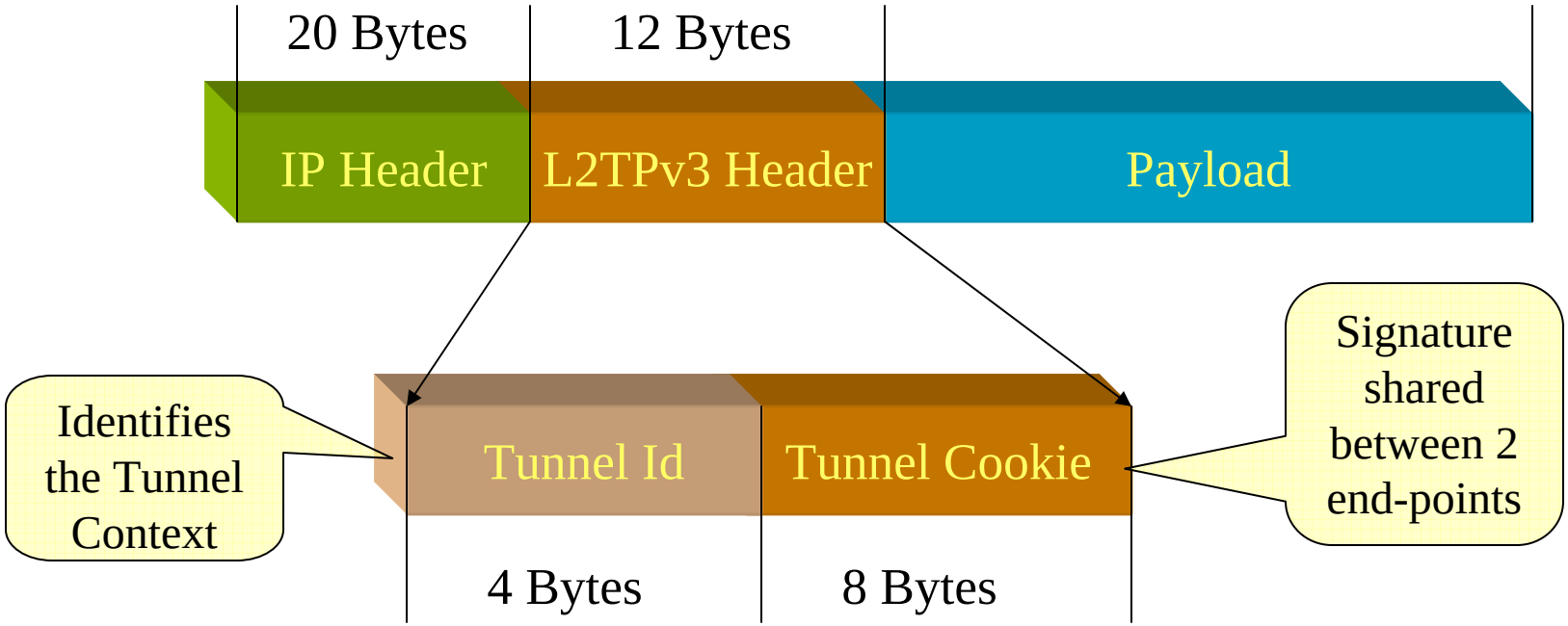
L2TPv3

- L2TPv3 is the Signaling and Encapsulation Protocol commonly used in the IP Core
- L2TPv3 is the IETF version of Cisco's UTI
- Extended to tunnel Layer-2 packets over IP core network
- Optimized for lookup
  - In GRE, the lookup is based on Source + Destination IP address pair

L2TPv3 Operation



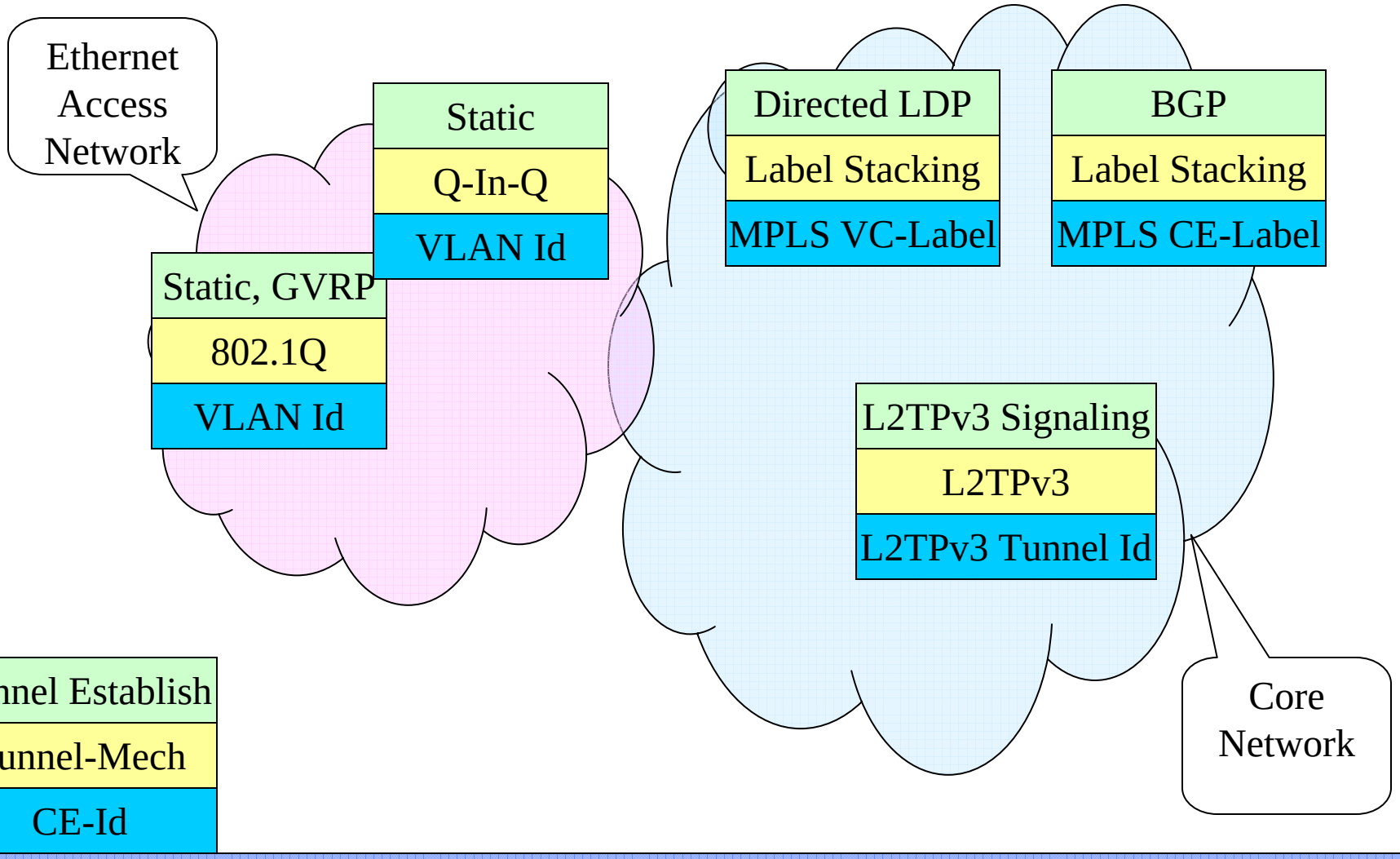
L2TPv3 Packet Format



Protocol Components

- CE Identifier
  - Identifies the subscriber in the Metro Network/Core Network
- Tunneling Mechanism
  - Mechanism to establish an end to end tunnel
- Tunnel Establishment Control Protocol
  - Mechanism to establish an end to end tunnel
  - Method for exchanging the CE-Id between end-points of the L2VPN network.

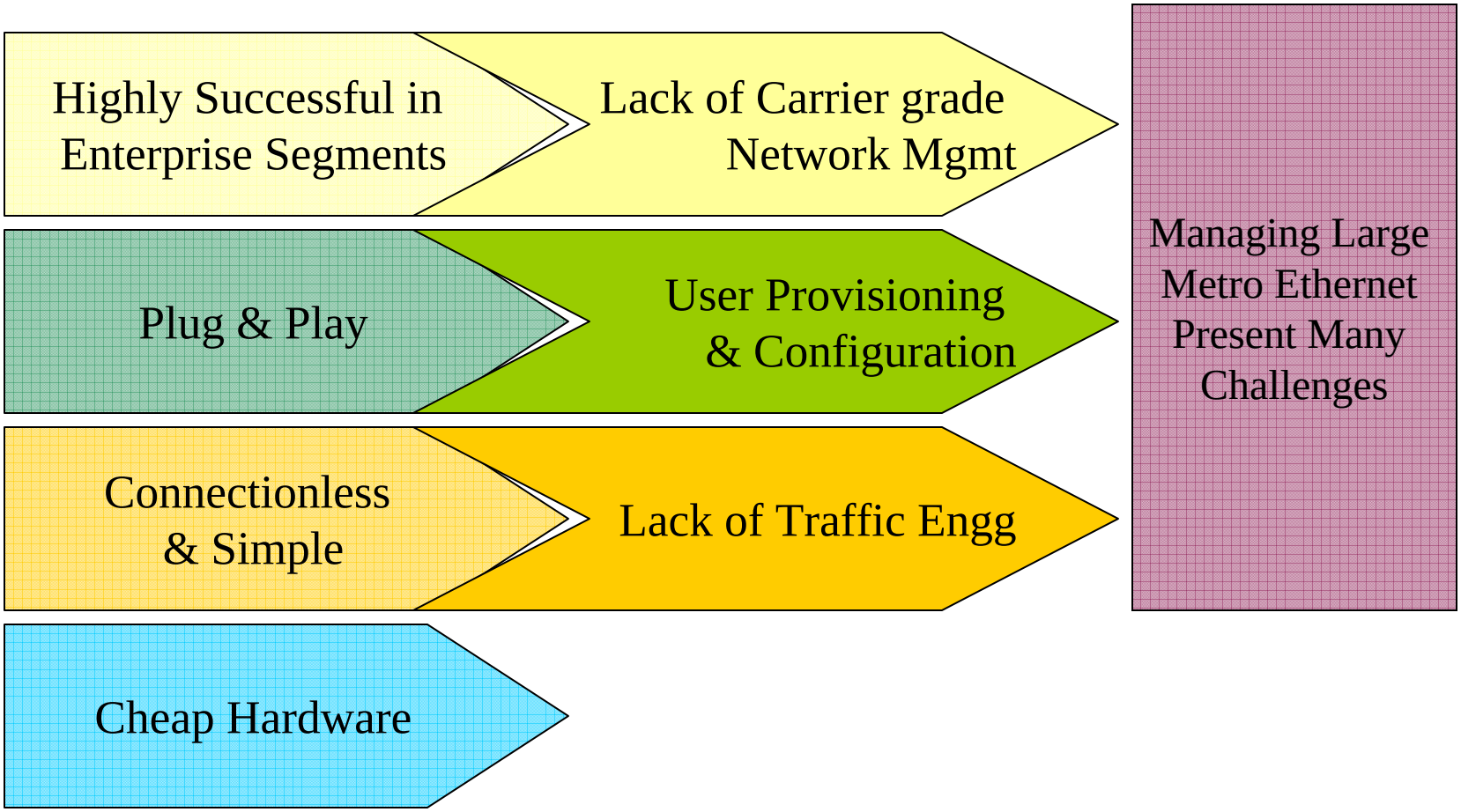
All About Metro Networks



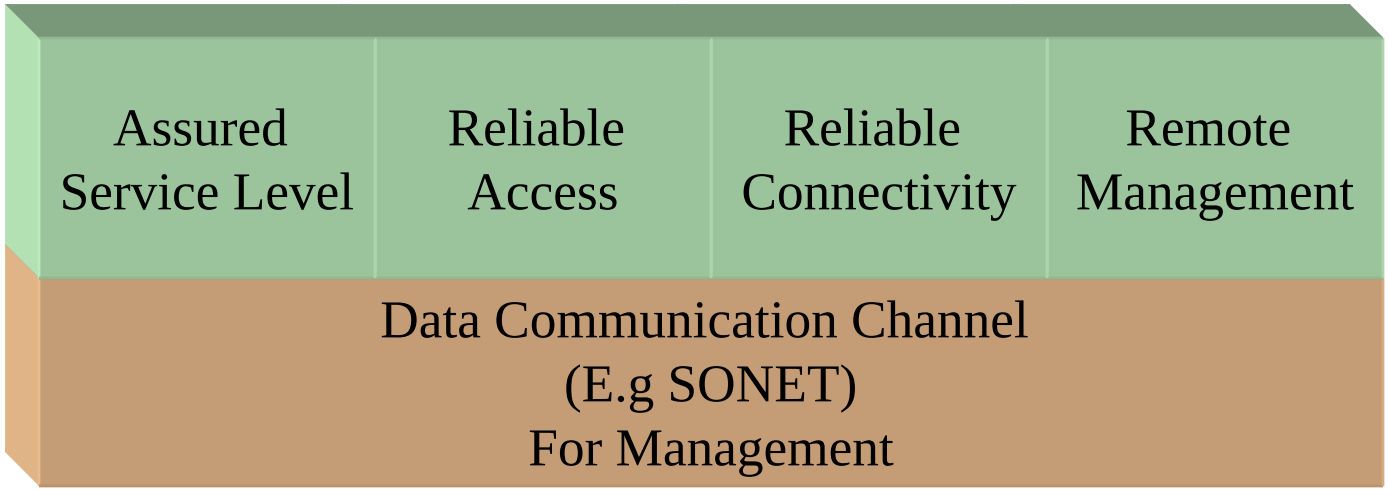
Agenda

- Service Model
- Ethernet in Access Networks
- Core Networks
  - IETF Initiatives
  - MPLS
  - L2TPv3
- Management

Management


Highly Successful in Enterprise Segments
Lack of Carrier grade Network Mgmt
Plug & Play
User Provisioning & Configuration
Connectionless & Simple
Lack of Traffic Engg
Cheap Hardware
Managing Large Metro Ethernet Present Many Challenges

Requirements



DCC type of Channel for Mgmt is required for keeping Ethernet in MAN

Performance Monitoring

- Monitoring Traffic Engineering Parameters
  - RMON Group
- Storm Monitoring
  - Alarm on exceeding the threshold bandwidth limit for broadcast & multicast traffic
- Monitoring other SLA Parameters
  - Packet loss & Jitter statistics on a interface
- L2 Control Packet
  - Monitoring the Bandwidth consumed by the L2 Control Packets in the network.

midas

Fault Management

- Link Errors
  - Remote Loopback Tests
- Environment Monitoring
  - Voltage level to line cards
  - Temperature
- Optical Interface
  - Quality of Optical Amplitude
    - Degraded due to cracks or connector contamination such as dust and moisture

midas

Subscriber Provisioning

- Associating CE-Id (e.g VLAN)
- Associating STP Instance to the VLAN
- VLAN Transparency
- Authentication Mechanism
- QoS Configuration
- Ingress Bandwidth
- Storm Control

midas

Summary

- Metro Ethernet can provide the “Real Broadband” access network
- Inherent Broadcast nature of Ethernet makes video streaming efficient
- Ethernet in Access and MPLS/IP in the core is highly scalable
- WAN like management capability is currently lacking
- Opens the door for many new applications

midas

References

- draft-ietf-pwe3-arch-01.txt
- draft-kompella-ppvpn-vpls-01.txt
- draft-ietf-ppvpn-bgpvpn-auto-03.txt
- draft-martini-l2circuit-encap-mpls-04.txt
- draft-martini-l2circuit-trans-mpls-10.txt
- draft-kompella-ppvpn-l2vpn-02.txt
- draft-ietf-ppvpn-l2-framework-01.txt
- draft-ietf-l2tpext-l2tpbase-02.txt

midas

Thank you..