Carrier Ethernet

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides and audio/video recordings of this class lecture are at:

http://www.cse.wustl.edu/~jain/cse570-19/

Overview

1. Enterprise vs Carrier Ethernet
2. UNI vs Peer-to-Peer Signaling
3. Metro Ethernet
4. Ethernet Provider Bridge (PB)
5. Provider Backbone Network (PBB)
6. Connection Oriented Ethernet

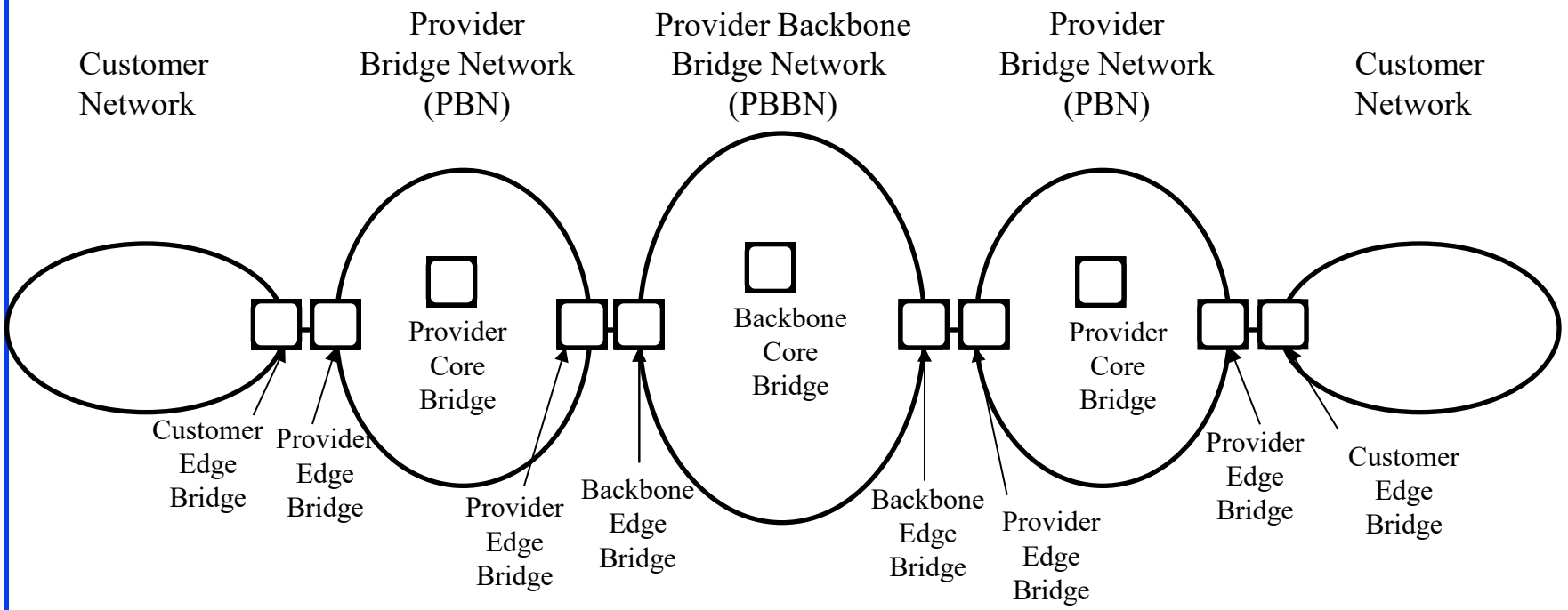
Note: Although these technologies were originally developed for carriers, they are now used inside multi-tenant data centers (clouds)

Enterprise vs. Carrier Ethernet

Enterprise	Carrier
❑ Distance: up to 2km	❑ Up to 100 km
❑ Scale: ➢ Few K MAC addresses	❑ Millions of MAC Addresses
➢ 4096 VLANs	❑ Millions of VLANs Q-in-Q
❑ Protection: Spanning tree	❑ Shortest Path Routing
❑ Path determined by spanning tree	❑ Traffic engineered path
❑ Simple service	❑ SLA
❑ Priority ⇒ Aggregate QoS	❑ Need per-flow QoS
❑ No performance/Error monitoring (OAM)	❑ Need performance/BER

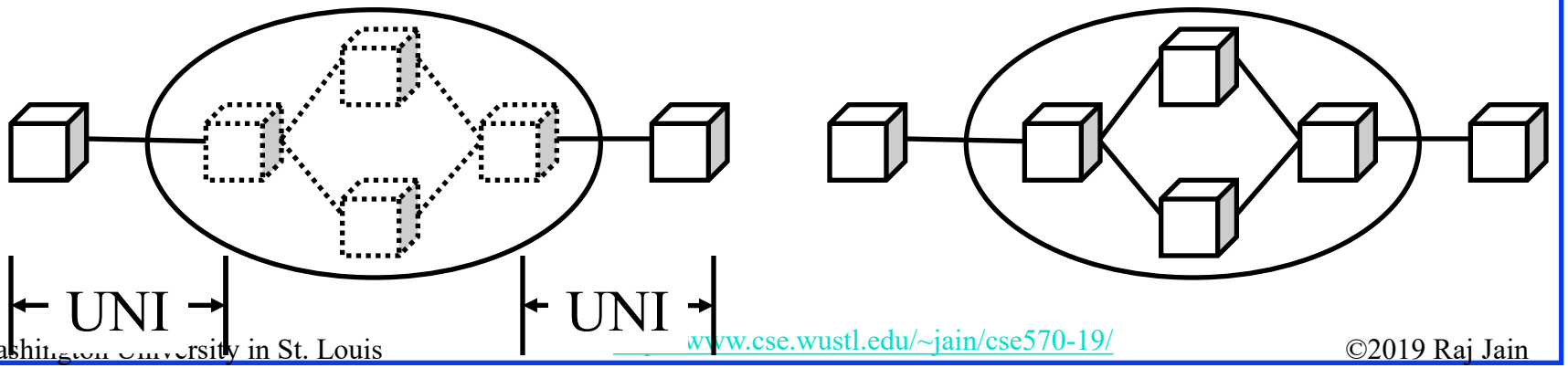
Carriers vs. Enterprise

Network Hierarchy



Issue: UNI vs Peer-to-Peer Signaling

- Two Business Models:
  - Carrier: Overlay or cloud
    - Network is a black-box
    - User-to-network interface (UNI)
  - Enterprise: Peer-to-Peer
    - Complete exchange of information



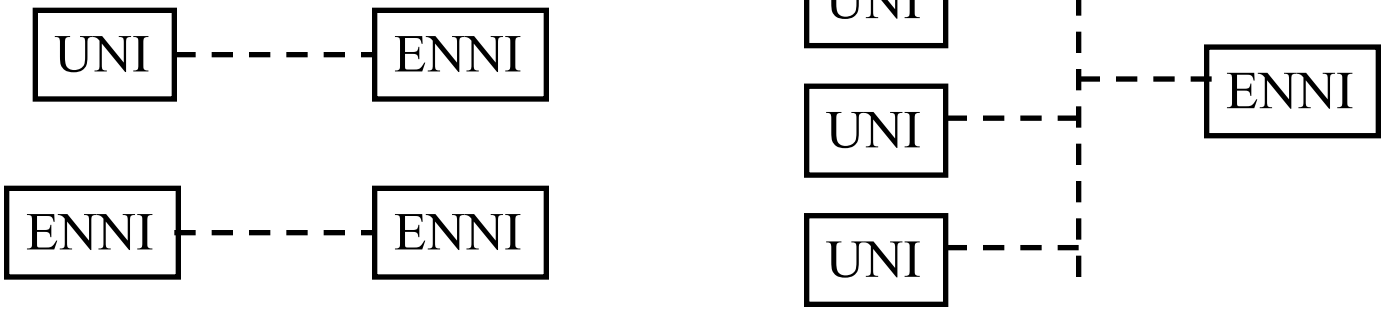
UNI vs. ENNI

- **User to Network Interface (UNI)**:
  - Separates responsibilities between the user and the provider. (Troubleshooting, failures etc).
  - Like the wired phone box outside your home.
  - Only one customer's traffic.
- **External Network to Network Interface (ENNI)**:
  - Separates responsibilities between two providers.
  - Many customer's traffic passes through an ENNI
  - Tier 2 *operators* sell services to Tier 3 service providers.

Ref: Fujitsu, "Carrier Ethernet Essentials," http://www.fujitsu.com/downloads/TEL/fnc/whitepapers/CarrierEthernetEssentials.pdf

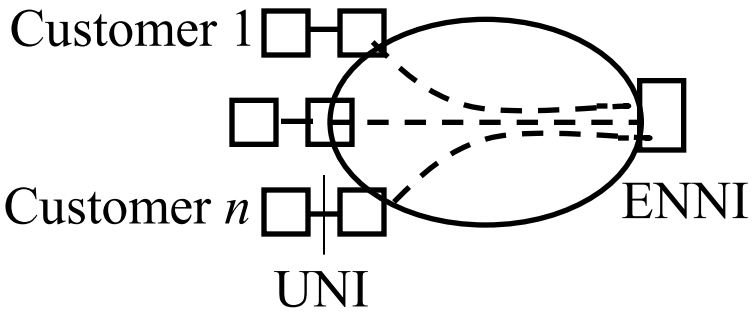
Operator Virtual Connection (OVC)

- Between UNI and ENNI or between two ENNIs.
- For wholesale service providers
- Two types: Point-to-Point and Multipoint-to-Multipoint
- Untagged or single tagged frames at NNI. Q-in-Q at ENNI
- UNIs may be 10 to 100 Mbps. ENNIs at 1 to 10 Gbps.



Metro Access Ethernet Private Line

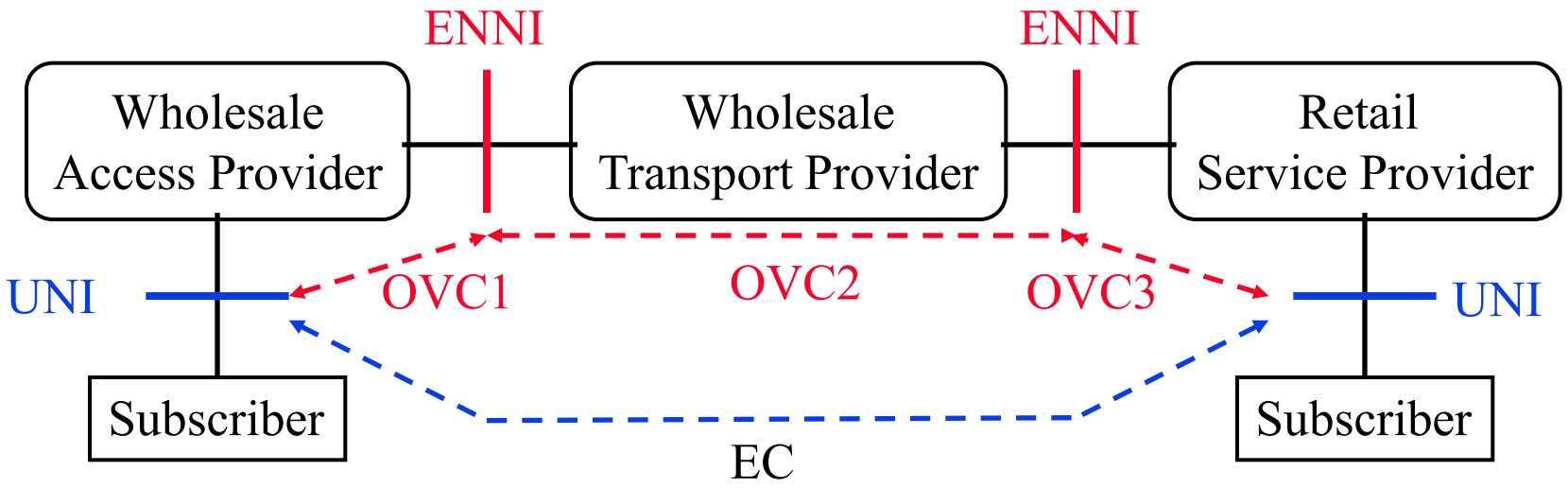
- **Access Ethernet Private Line (Access-EPL)**:
  - Port-based service for Internet access
    Like the service at your home.
  - Ends at your access provider, where many other Access-EPLs may end
  - Access provider has only one interface
    Shared by many Access-EPLs ⇒Different from p2p EPL.



E-Access

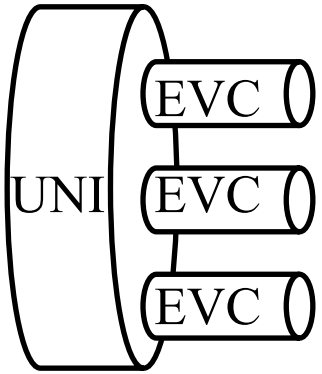
End-to-End Metro Ethernet Connection

- An EC may go through multiple service providers
  ⇒ Multiple OVCs can be concatenated to create an EC



Ethernet Virtual Connections (EVCs)

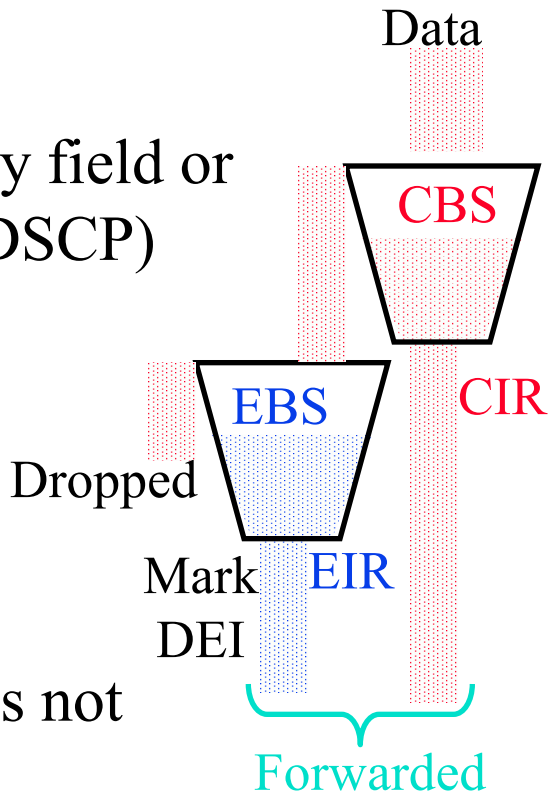
- **Port-based ECs**: Forwarding not based on VLANs.
  Frames delivered to remote UNI/ENNI for P2P or
  Based on destination address for P2MP
- **VLAN-based ECs**: Forwarding based on VLAN tag.
  ⇒ Multiple Virtual UNIs
  ⇒ Ethernet *Virtual* Connection (*EVC*)
  More cost-effective for Enterprise customers



- Types of EVCs:
  1. Ethernet Virtual Private Line (EVPL)
  2. Ethernet Virtual Private Tree (EVP-Tree)
  3. Ethernet Virtual Private LAN (EVPLAN)
  4. Access Ethernet Virtual Private Line (Access EVPL)
- Note: Service providers always share an ENNI for multiple connections ⇒ OVCs are always virtual

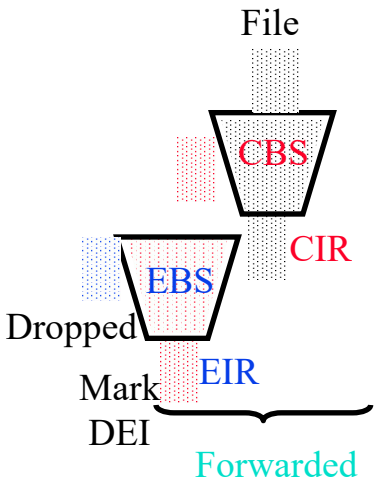
Metro Ethernet Service Attributes

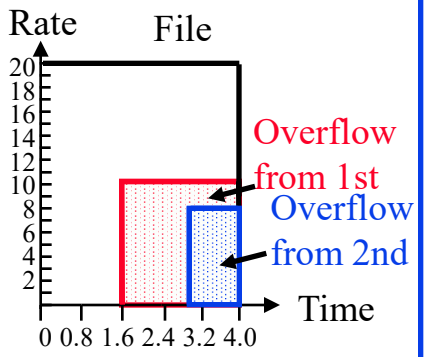
- Bandwidth Profiles: Limits on data dates
  - Ingress Profile: Incoming data rate
  - Egress Profile: Outgoing data rate
- Per UNI, Per EVC or OVC, or Per EVC/OVC per Class of Service (CoS)
- CoS is indicated by the 3-bits in the priority field or 4-bit Differentiated Services Code Point (DSCP)
- Rate specified by 5 parameters
  1. Committed Information Rate (CIR)
  2. Committed Burst Size (CBS)
  3. *Excess* Information rate (EIR)
  4. Excess Burst Size (EBS)
  5. Color Mode (CM): Customer does/does not mark drop eligibility indicator (DEI)



Example

- A user with CIR=10 Mbps, CBS=2 MB, EIR=2 Mbps, EBS=1MB, sends a 10 MB file at 20 Mbps.
  What percent of packets will be dropped?
  What percent will be marked?
- CBS = 2 MB = 16 Mb, EBS = 1 MB = 8 Mb,
  File size = 80 Mb
- File time = 80 Mb/20 Mbps = 4 s
- 1st Bucket fill rate = 20-10 = 10 Mbps,
  1st Bucket fill time = 16/10 = 1.6s
  Total output from 1st bucket = 4*10 Mbps + 16 Mb
  = 56 Mb (committed) = 70%
- 1st Bucket Overflow time = 4-1.6 = 2.4s
  1st bucket Overflow amount = 2.4 s*(20-10) Mbps
  = 24 Mb
- 2nd bucket fill Rate = 10-2 = 8 Mbps
  2nd bucket fill time = Size/Rate = 8 Mb/8 Mbps = 1 s
  Total output from 2nd bucket = 0*1.6 + 2*(4 - 1.6) + 8
  = 12.8 Mb (Marked) = 16%
- 2nd Bucket overflow time = 4 - 1.6 - 1 = 1.4s
  2nd bucket overflow amount = (10-2)*1.4 = 11.2 Mb (Dropped)
  = 14%
- Check: Total = 56+12.8+11.2 = 80 Mb



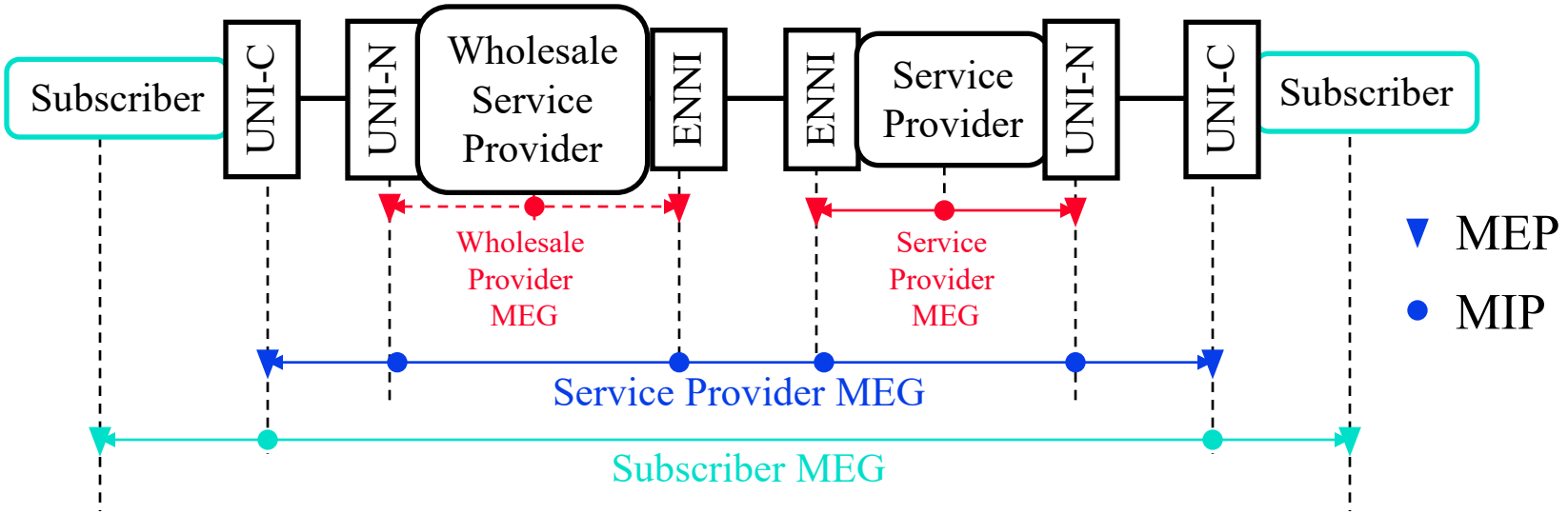


Homework 6

- A user with CIR=25 Mbps, CBS=2 MB, EIR=2 Mbps, EBS=1MB, sends a 15 MB file at 30 Mbps.
  What percent of packets will be dropped?
  What percent will be marked?

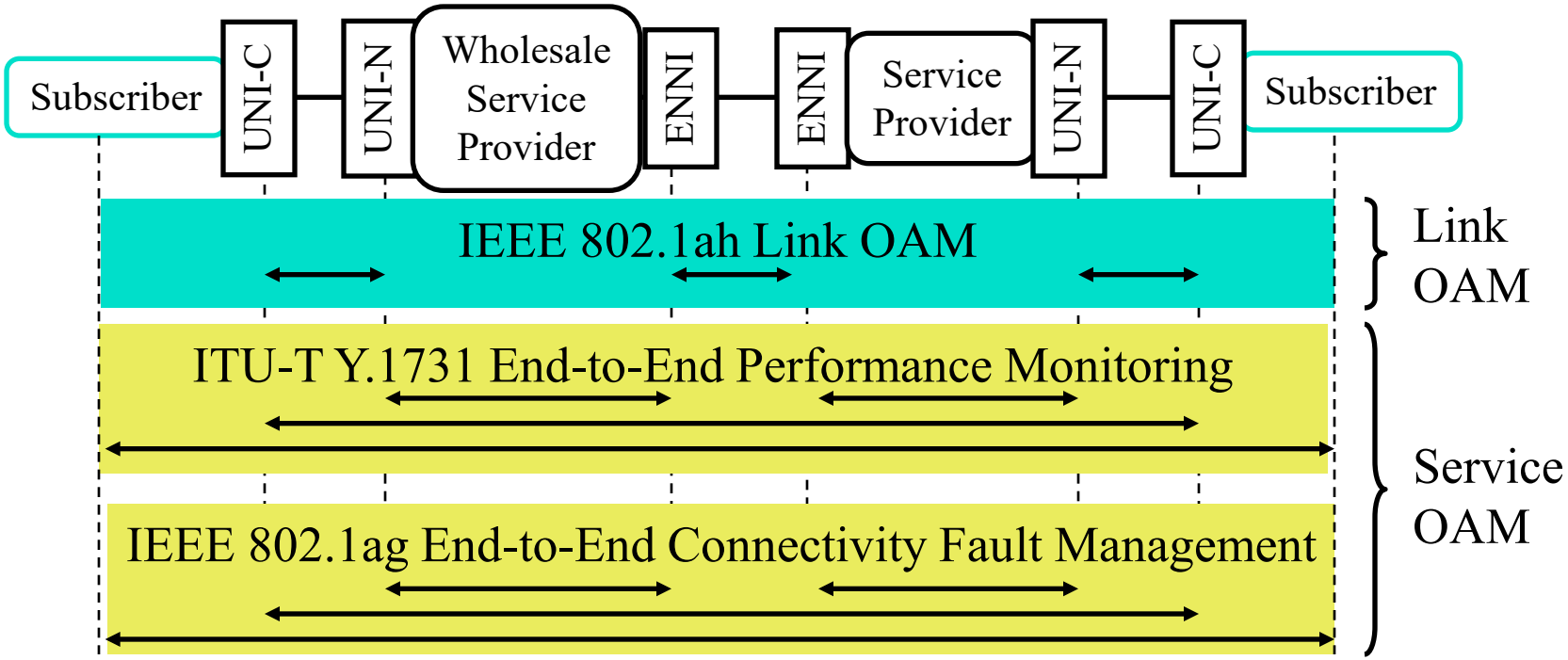
Metro Ethernet OAM

- Operation, Administration and Maintenance (OAM)
- Defined in IEEE 802.1ag, IEEE 802.1ah, and ITU Y.1731
- Maintenance End Points (MEPs)
- Maintenance Intermediate Points (MIPs)
- Maintenance Entity Group (MEG): Level of Administration



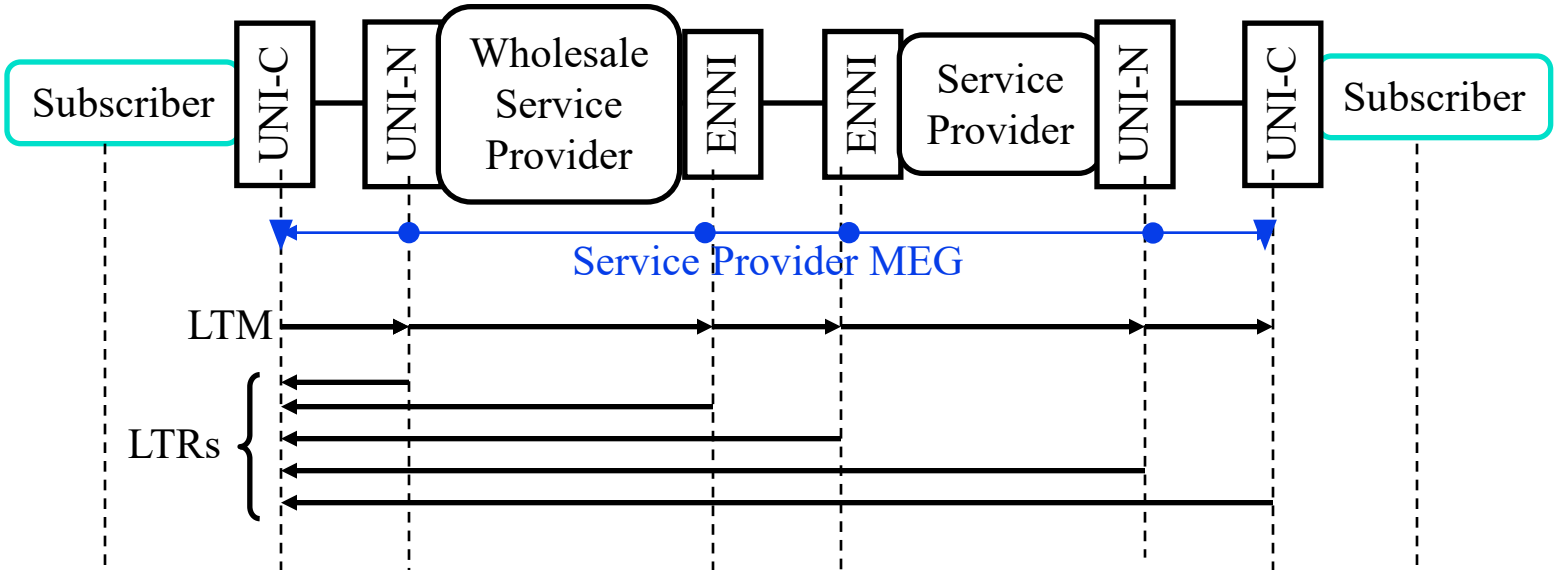
Metro Ethernet OAM (Cont)

- Performance Monitoring: Measure throughput and latency
- Connectivity Fault Management: Monitor downtime
  - Service Fault Management
  - Link Fault Management



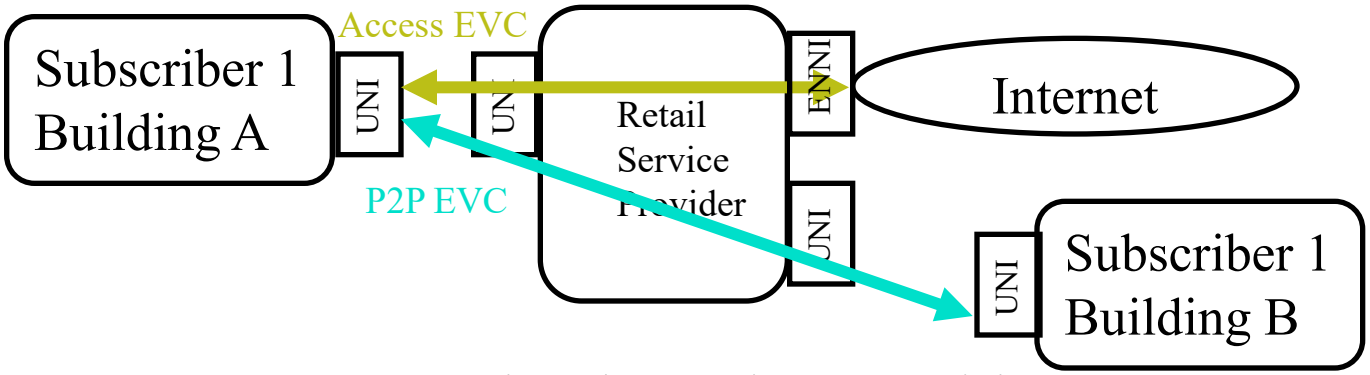
Metro Ethernet OAM Messages

- Continuity Check Message (CCM) in both directions (Similar to IP Ping)
- Link Trace Message (LTM): Locates fault. Link Trace Response (LTR) is returned by each end point and intermediate point (similar to IP trace route)
- Fault, Configuration, Accounting, Performance, Security (FCAPS)

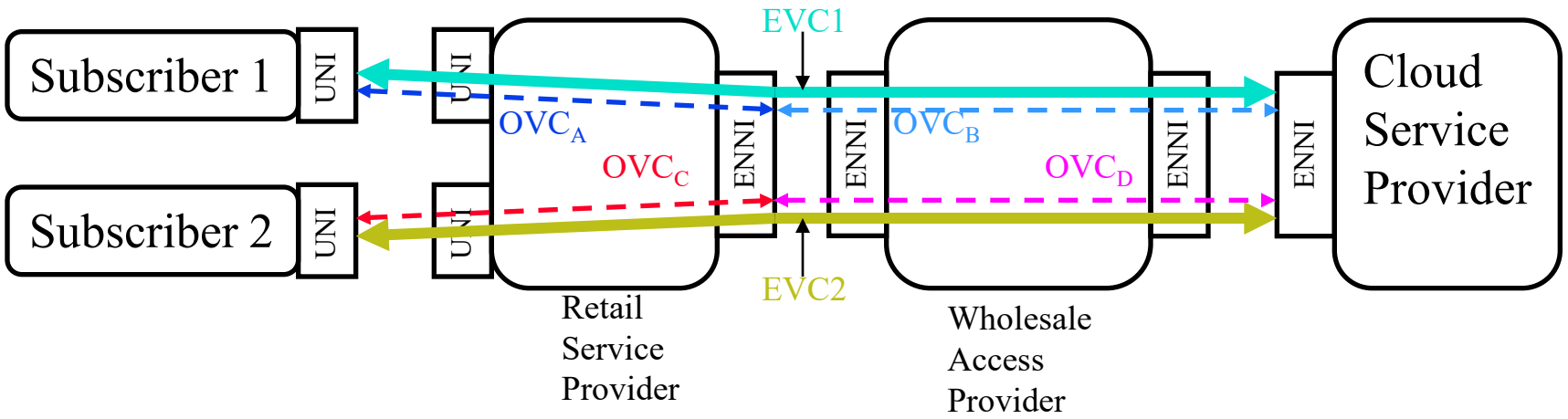


Metro Ethernet Use Cases

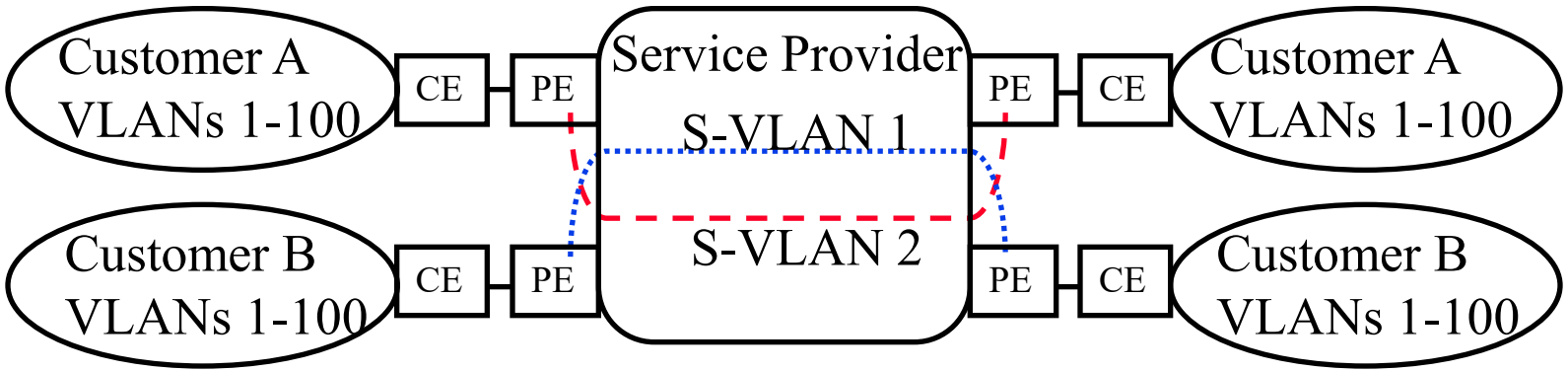
1. Head office to Satellite offices and/or Internet



2. Customers to Cloud Service Provider



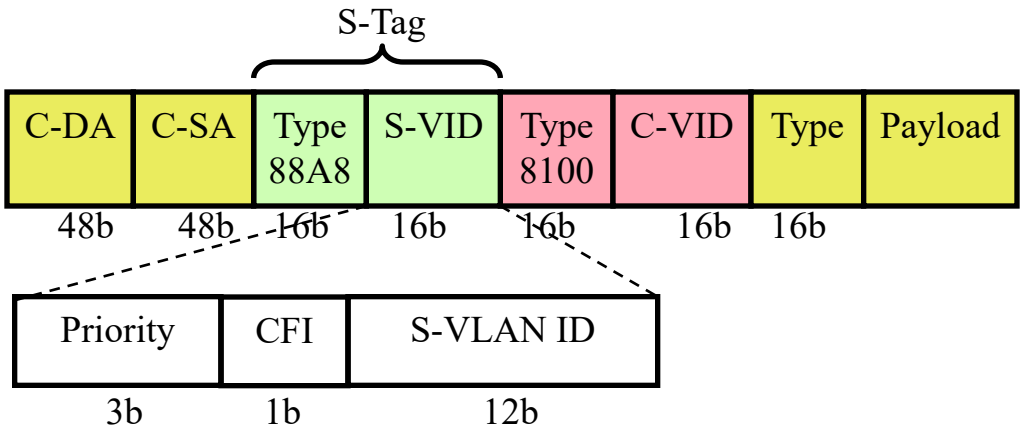
Ethernet Provider Bridge (PB)



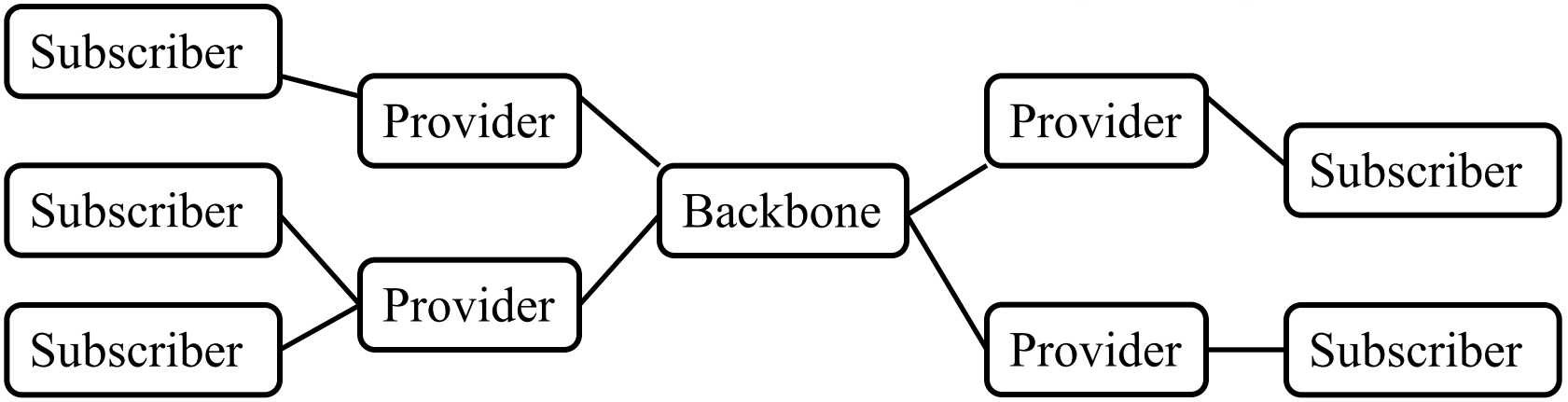
- IEEE 802.1ad-2005 incorporated in IEEE 802.1Q-2011
- Problem: Multiple customers may have the same VLAN ID. How to keep them separate?
- Solutions:
  1. VLAN translation: Change customer VLANs to provider VLANs and back
  2. VLAN Encapsulation: Encapsulate customer frames

Provider Bridge (Cont)

- Q-in-Q Encapsulation: Provider inserts a service VLAN tag VLAN translation Changes VLANs using a table
- Allows 4K customers to be serviced. Total 16M VLANs
- 8 Traffic Classes using Differentiated Services Code Points (DSCP) for Assured Forwarding



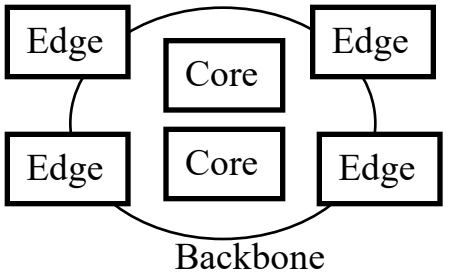
Provider Backbone Network (PBB)

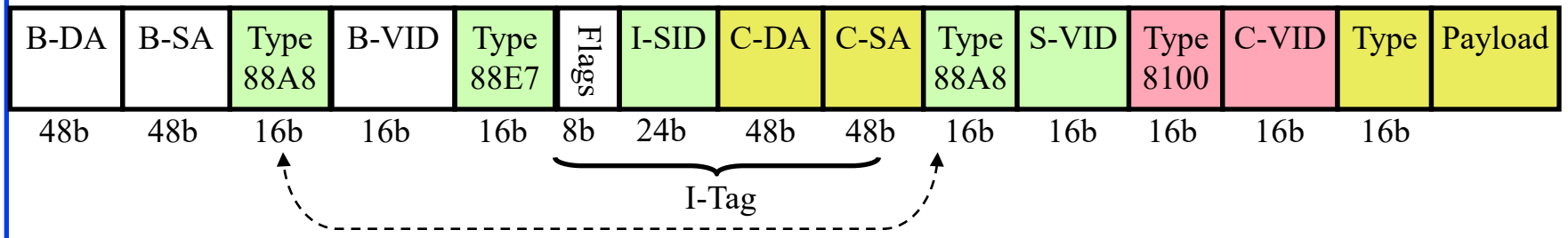


- Problem: Number of MAC addresses passing through backbone bridges is too large for all core bridge to remember
  Broadcast and flooded (unknown address) frames give unwanted traffic and security issues
- Solution: IEEE 802.1ah-2008 incorporated in 802.1Q-2011
- Add new source/destination MAC addresses pointing to ingress backbone bridge and egress backbone bridge
  ⇒Core bridges only know edge bridge addresses

MAC-in-MAC Frame Format

- Backbone edge bridges (BEB) forward to other BEB's and learn customer MAC addresses
  ⇒ Backbone *core* bridges (BCB) do not learn customer MACs
- B-DA = Destination backbone bridge address Determined by Customer Destination Address
- Backbone VLANs delimit the broadcast domains in the backbone



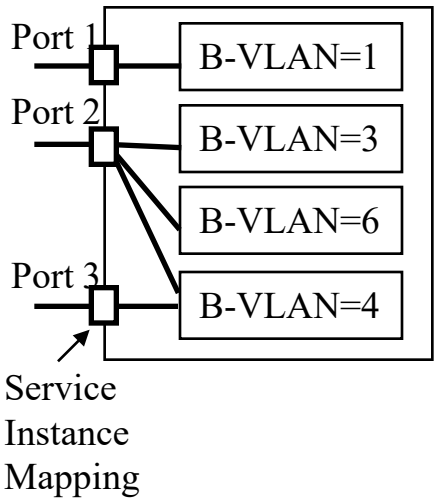


- Core switches forward based on Backbone Destination Bridge Address and Backbone-VLAN ID (60 bits)
  Similar to 802.1ad Q-in-Q. Therefore, same EtherType.

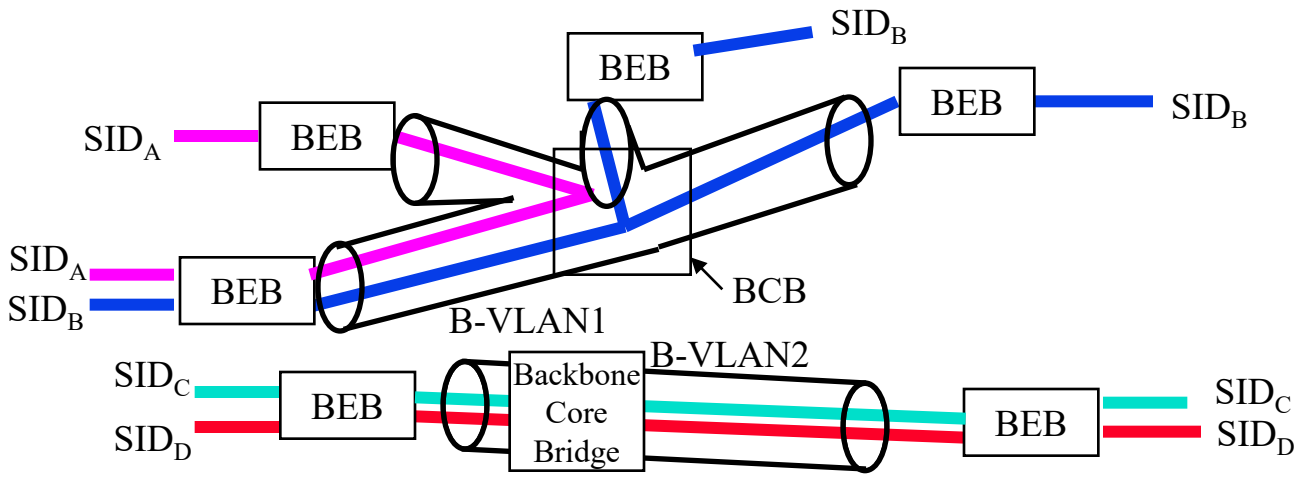
PBB Service Instance

- 24-bit Service instance ID (I-SID) indicates a specific flow
  - All frames on a specific port, or
  - All frames on a specific port with a specific *service* VLAN, or
  - All frames on a specific port with a specific service VLAN and a specific *customer* VLAN

SID	Definition	B-VLAN
1	Port 1	1
20	Port 2, S-VLAN=10	3
33	Port 2, S-VLAN=20	6
401	Port 2, S-VLAN=30, C-VLAN=100	4
502	Port 3, S-VLAN=40, C-VLAN=200	4



MAC-in-MAC (Cont)



- Each Backbone VLANs (B-VLAN) can carry multiple services
- 24-bit SID $\Rightarrow 2^{24}$ Service Instances in the backbone
- I-Tag format: I-Tag not looked at in the core.
  Includes C-DA+C-SA.
  UCA=1 $\Rightarrow$ Use customer addresses (used in CFM in the Edge)

Priority Code Point (I-PCP)	Drop Eligibility Indicator (I-DEI)	Use Customer Address (UCA)	Reserved 1	Reserved 2	Service Instance ID (I-SID)	Customer Destination Address (C-DA)	Customer Source Address (C-SA)
3b	1b	1b	1b	2b	24b	48b	48b

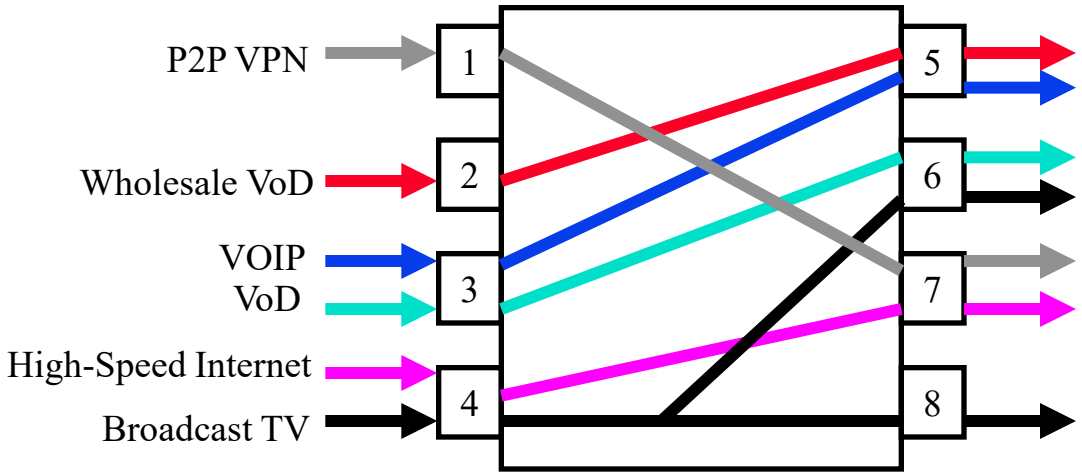
Connection Oriented Ethernet

- Connectionless: Path determined at forwarding
  ⇒ Varying QoS
- Connection Oriented: Path determined at provisioning
  - Path provisioned by management ⇒ Deterministic QoS
    - No spanning tree, No MAC address learning,
    - Frames forwarded based on VLAN Ids and Backbone bridges addresses
    - Path not determined by customer MAC addresses and other customer fields ⇒ More Secure
  - Reserved bandwidth per EVC
  - Pre-provisioned Protection path ⇒ Better availability

VLAN Cross-Connect

- Cross-connect ⇒ Circuit oriented
- Connection Oriented Ethernet with Q-in-Q
- Forward frames based on VLAN ID and Input port
  ⇒ No MAC Learning

Input Port	VLAN ID	Output Port
1	200	7
2	201	5
3	20	5
3	21	6
4	100	7
4	101	8

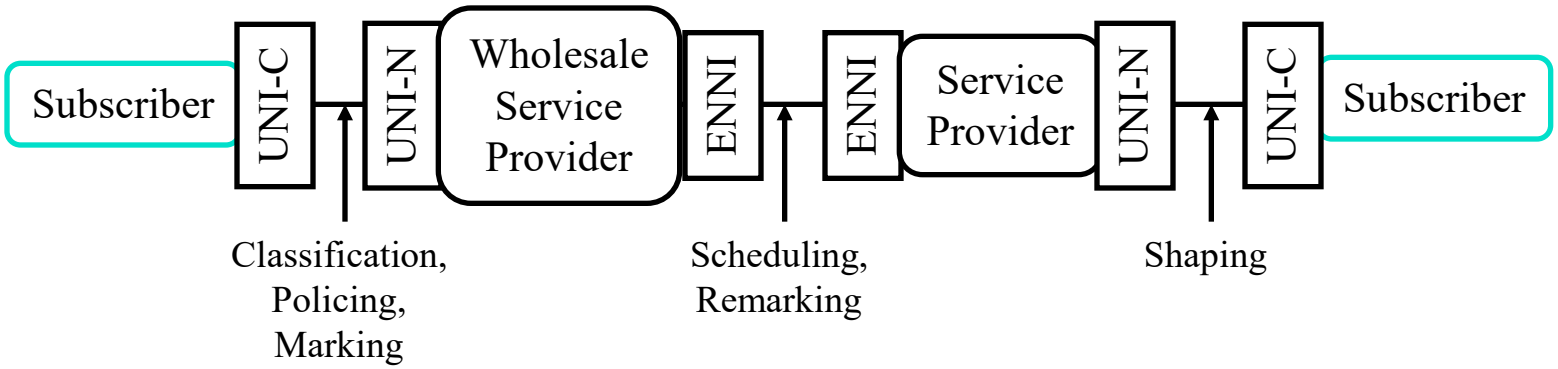


PBB-TE

- Provider Backbone Bridges with Traffic Engineering (PBB-TE)
- IEEE 802.1Qay-2009 incorporated in 802.1Q-2011
- Provides connection oriented P2P (E-*Line*) Ethernet service
- For PBB-TE traffic VLANs:
  - Turn off MAC learning
  - Discard frames with unknown address and broadcasts.
    $\Rightarrow$ No flooding
  - Disable Spanning Tree Protocol.
  - Add protection path switching for each direction of the trunk
- Switch forwarding tables are administratively populated using management
- Same frame format as with MAC-in-MAC. No change.

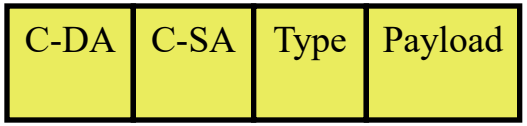
PBB-TE QoS

- Guarantees QoS ⇒ No need for MPLS or SONET/SDH
- UNI traffic is classified by Port, Service VLAN ID, Customer VLAN ID, priority, Unicast/Multicast
- UNI ports are *policed* ⇒ Excess traffic is dropped
  No policing at NNI ports. Only remarking, if necessary.
- Traffic may be marked and remarked at both UNI and NNI

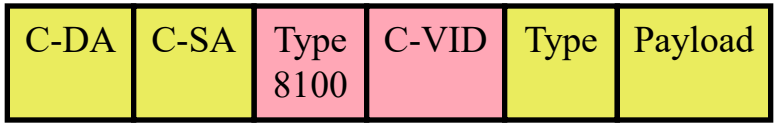


Ethernet Tagged Frame Format Evolution

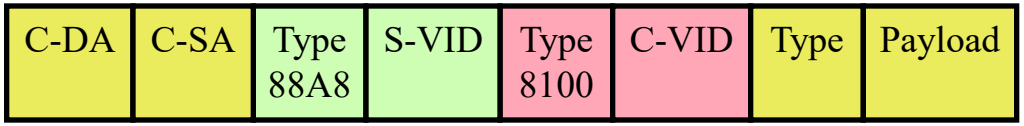
- Original Ethernet



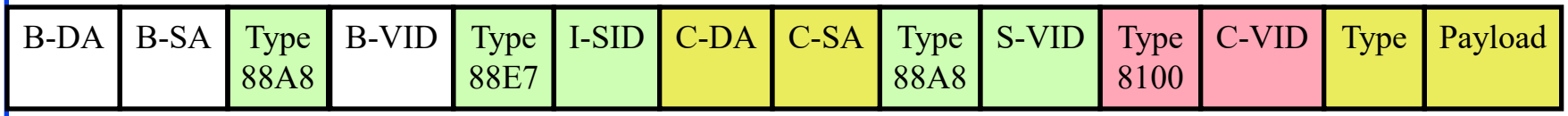
- IEEE 802.1Q VLAN



- IEEE 802.1ad PB



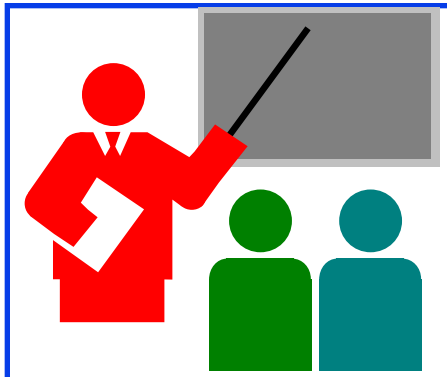
- IEEE 802.1ah PBB or 802.1Qay PBB-TE



Tag Type	Value
Customer VLAN	8100
Service VLAN or Backbone VLAN	88A8
Backbone Service Instance	88E7

Comparison of Technologies

	Basic Ethernet	MPLS	PB	PBB-TE
Resilience	No	Protection Fast Reroute	SPB/LAG	Protection Fast Reroute
Security	No	Circuit Based	VLAN	Circuit Based
Multicast	Yes	Inefficient	Yes	No. P2P only
QoS	Priority	Diffserve	Diffserve+ Guaranteed	Diffserve+ Guaranteed
Legacy Services	No	Yes (PWE3)	No	No
Traffic Engineering	No	Yes	No	Yes
Scalability	Limited	Complex	Q-in-Q	Q-in-Q+ Mac-in-MAC
Cost	Low	High	Medium	Medium
OAM	No	Some	Yes	Yes



Summary

1. Carriers use User-to-Network Interface (UNI) signaling rather than peer-to-peer signaling
2. Metro Ethernet allows E-Line, E-Access, E-Tree, and E-LAN services
3. Q-in-Q allows service providers to carry customer VLAN tags in their Ethernet Frames
4. MAC-in-MAC extension allows very large Ethernet networks spanning over several backbone carriers
5. PBB-TE extension allows connection oriented Ethernet with QoS guarantees and protection

Reading List

- Fujitsu, "Carrier Ethernet Essentials," http://www.fujitsu.com/downloads/TEL/fnc/whitepapers/CarrierEthernetEssentials.pdf
- R. Santitoro, "Metro Ethernet Services – A Technical Overview," The Metro Ethernet Forum, 2003, V2.7, https://www.mef.net/Assets/White_Papers/Metro-Ethernet-Services.pdf

References

- G. Santana, "Datacenter Virtualization Fundamentals," Cisco Press, 2014, ISBN: 1587143240 (Safari Book)
- H. Saboowala, M. Abid, S. Modali, "Designing Networks and Services for the Cloud: Delivering business-grade cloud applications and services," Cisco Press 2013, ISBN:1587142945 (Safari Book)

Wikipedia Links

- http://en.wikipedia.org/wiki/Carrier_Ethernet
- http://en.wikipedia.org/wiki/Connection-oriented_Ethernet
- http://en.wikipedia.org/wiki/Ethernet_Private_Line
- http://en.wikipedia.org/wiki/Ethernet_Virtual_Private_Line
- http://en.wikipedia.org/wiki/IEEE_802.1ad
- http://en.wikipedia.org/wiki/IEEE_802.1ag
- http://en.wikipedia.org/wiki/IEEE_802.1ah-2008
- http://en.wikipedia.org/wiki/Metro_Ethernet
- http://en.wikipedia.org/wiki/Metro_Ethernet_Forum
- http://en.wikipedia.org/wiki/Network-to-network_interface
- http://en.wikipedia.org/wiki/Operations,_administration_and_management
- http://en.wikipedia.org/wiki/Provider_Backbone_Bridge_Traffic_Engineering
- http://en.wikipedia.org/wiki/Traffic_policing
- http://en.wikipedia.org/wiki/Traffic_shaping
- http://en.wikipedia.org/wiki/User%E2%80%93network_interface
- http://en.wikipedia.org/wiki/Virtual_Private_LAN_Service

Acronyms

- B-VID Backbone VLAN Identifier
- BER Bit Error Rate
- C-VID Customer VLAN Identifier
- CBS Committed Burst Size
- CCM Continuity Check Message
- CE Customer Edge
- CFI Canonical Form Indicator
- CFM Connectivity Fault Management
- CIR Committed Information Rate
- CM Color Mode
- CoS Class of Service
- DA Destination Address
- DEI Drop Eligibility Indicator
- DSCP Differentiated Services Code Points
- EBS Excess Burst Size
- EC Ethernet Connection

Acronyms (Cont)

- EIR Excess Information rate
- ENNI External Network to Network Interface
- EPL Ethernet Private Line
- EVC Ethernet Virtual Connection
- EVP-Access Ethernet Virtual Private Access
- EVP-LAN Ethernet Virtual Private Local Area Network
- EVP-Line Ethernet Virtual Private Line
- EVP-Tree Ethernet Virtual Private Tree
- EVPL Ethernet Virtual Private Line
- I-SID Instance Service ID
- ID Identifier
- IEEE Institution of Electrical and Electronic Engineers
- IETF Internet Engineering Task Force
- IP Internet Protocols
- ITU International Telecommunications Union

Acronyms (Cont)

- LAN Local Area Network
- LTM Link Trace Message
- LTR Link Trace Response
- MAC Media Access Control
- MEG Maintenance Entity Group
- MEP Maintenance End Points
- MIP Maintenance Intermediate Points
- MP Multi-Point
- MPLS Multi-Protocol Label Switching
- NNI Network-to-Network Interface
- OAM Operation, Administration and Maintenance
- OC Optical Carrier
- OIF Optical Interoperability Forum
- OVC Operator Virtual Connection

Acronyms (Cont)

- PB Provider Bridge
- PBB-TE Provider Backbone Bridge with Traffic Engineering
- PBB Provider Backbone Bridge
- PBBE Provider BackBone Edge
- PBBN Provider Backbone Network
- PBEB Provider backbone edge bridges
- PBN Provider Bridging network
- PBX Private Branch Exchange
- PCP Priority Code Point
- PDH Plesiochronous Digital Hierarchy
- PE Provider Edge
- PW Pseudo-Wire
- PWE3 Pseudo-Wire Emulation Edge-to-Edge
- QoS Quality of Service
- S-VID Service (Provider) VLAN ID
- SA Source Address
- SDH Synchronous Digital Hierarchy

Acronyms (Cont)

- SID Service Identifier
- SLA Service Level Agreement
- SONET Synchronous optical network
- TE Traffic Engineering
- TV Television
- UCA Use Customer Address (flag)
- UNI User to Network Interface
- VID VLAN Identifier
- VLAN Virtual Local Area Network
- VoD Video on Demand
- VoIP Voice over IP
- VPN Virtual Private Network

Scan This to Download These Slides

Raj Jain
http://rajjain.com

Related Modules

CSE567M: Computer Systems Analysis (Spring 2013),
https://www.youtube.com/playlist?list=PLjGG94etKypJEKjNAa1n_1X0bWWNyZcof

CSE473S: Introduction to Computer Networks (Fall 2011),

https://www.youtube.com/playlist?list=PLjGG94etKypJWOSPMh8Azcgy5e_10TiDw

Wireless and Mobile Networking (Spring 2016),
https://www.youtube.com/playlist?list=PLjGG94etKypKeb0nzyN9tSs_HCd5c4wXF

CSE571S: Network Security (Fall 2011),

https://www.youtube.com/playlist?list=PLjGG94etKypKvzfVtutHcPFJXumyyg93u

Video Podcasts of Prof. Raj Jain's Lectures,
https://www.youtube.com/user/ProfRajJain/playlists