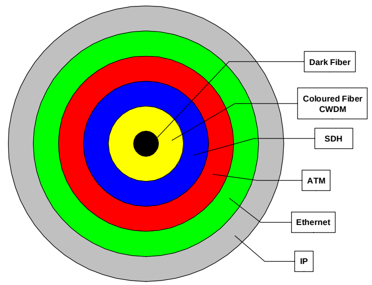
S-38.3192 Verkkopalvelujen tuotanto
S-38.3192 Network Service Provisioning

Lecture 2: L2 Network Technologies

Metropolitan Area Network

- MANs are build to connect urban locations with a high bandwidths
- Requires high bandwidth technologies with intermediate range passive components
  - Usually based to
    - Optical fibers
      - Single mode
    - Radio
      - LMDS, MMDS, WiMAX, FlashOFDM

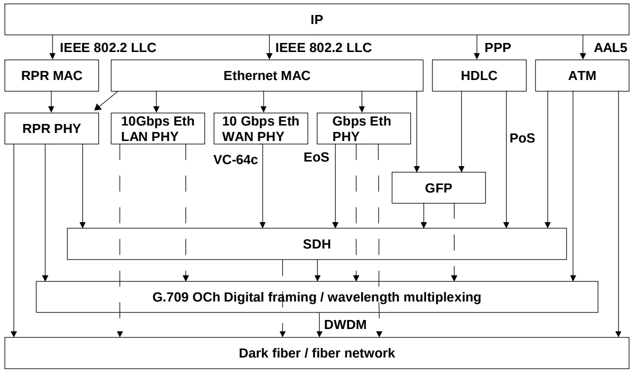


Core Network

- Connects MAN networks together
- Requires high bandwidth technologies with long range passive operation
  - Transmission speed and distance without repeaters tend to be inversely proportional
    - 1Gbps Ethernet -> 80-150km in SM-fiber with ZX-transmitter
    - 10Gbps Ethernet -> 10-40km in SM-fiber with ZX-transmitter
- Typical medias are
  - Fiber (Single Mode)
  - Radio (Microwave, Satellite)

Technologies

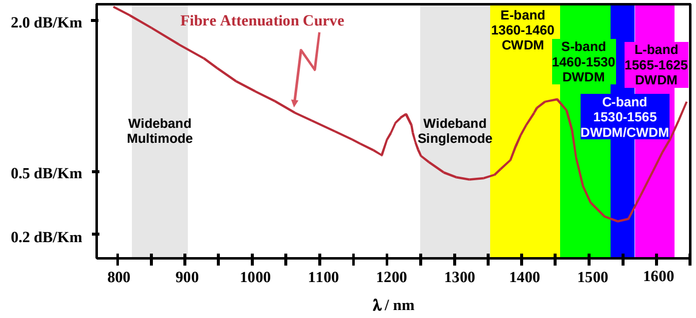
- High bandwidth requirements
- Transmission speeds are jumping up with constant rate
  - 1995: 155Mbps (SDH/ATM)
  - 2000: 2.4Gps (SDH)
  - 2004: 10 Gbps (SDH/Ethernet)
  - 2000-2004 wavelength technologies brought a new means to increase capacity
    - DWDM
    - CWDM
- Frame based multiplexing
  - Irrespective of low layer functionality
    - Fiber/Radio
  - Options today are
    - GMPLS
    - SDH
    - ATM
    - Ethernet
    - GFP
    - RPR



EoS Ethernet over SDH (Proprietary)
PoS Packet over SDH
RPR Resilient Packet Rings (IEEE 802.17)
GFP Generic Framing Procedure

Fiber communication

- Fiber optics offers wide spectrum of which only narrow part is used by conventional wideband fiber transmitters

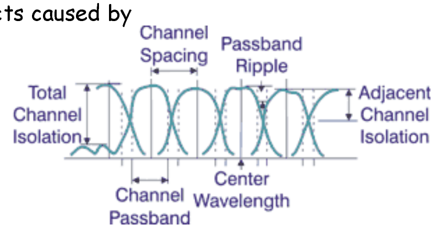


Modern fiber communication

- **The goal is to push the limits of wideband fiber communications**
  - Wideband transmitters are expensive and electrical part with high speed is error prone
    - Multiple narrowband transmitters achieve same performance on lower cost and lower error margin
  - To achieve longer transmission distances
    - Lower attenuation of lower frequencies serves this goal
      - Narrow transmission window in C-band
        - » Narrow spacing of transmission channels

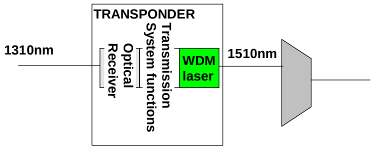
Modern fiber communication

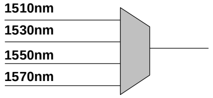
- Packing of several channels into a single media causes multiple problems related to interferences
  - Not just within fiber but also between channels
    - How to inject multiple closely spaced signals into a fiber
    - How to detect them in receiver
    - How to control their defects caused by
      - Dispersion
      - Attenuation



WDM

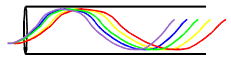
- Effectively N fold increase of transmission capacity from the same fiber infrastructure
  - Individual lambdas can be used independently
    - Usage depends on transponder unit





WDM

- Two operative versions
  - CWDM - Coarse Wavelength Division Multiplexing
    - Current spec 8 channels between (1470 - 1610nm / 20nm steps)
    - New spec 18 channels between (1271 - 1611nm / 20nm steps)
  - DWDM - Dense Wavelength Division Multiplexing
    - ITU Grid (100 Ghz resolution)
      - 50 channels between 1569.80nm to 1611.79nn
      - 50 channels between 1529.75nm to 1569.59nm
      - 50 channels between 1491.69nm to 1529.55nm

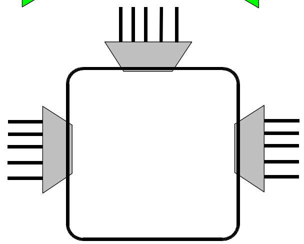


WDM

- DWDM
  - Narrow channel spacing
    - Components need to be compensated for temperature effects
      - Expensive
  - More channels to choose from
    - nonlinearities of fibers can be avoided by selecting proper wavelengths
- CWDM
  - Wide channel spacing
    - Component requirements are looser
      - Cheaper lasers and receivers
  - Less channels
    - Not suitable for long-haul networks
    - Suitable for MANs

WDM

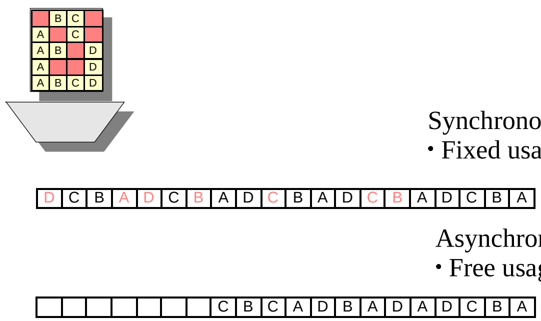
- Can be used as link or network technology
  - Link technology
    - Multiplexers at the ends of the links
  - Network technology
    - Optical switching components
      - Optical delay lines
      - Wavelength conversion
      - Photonic switching
    - Collision free routing
    - Crosstalk issues



WDM

- **Pros:**
  - Protocol independent
  - Virtual fiber
  - Multiplexing different traffic through different wavelengths
- **Cons:**
  - Depending on system pay as you go may not be possible
    - The number of required channels need to be estimated for the lifetime of systems
      - Filters are designed for certain amount of wavelengths and spacing

Frame Multiplexing

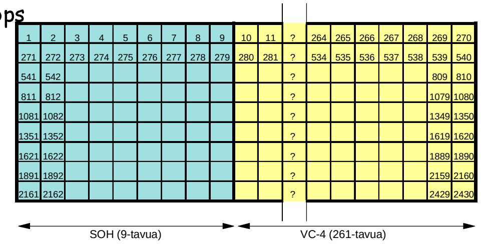


Frame Multiplexing

- Synchronous
  - Fixed usage of resources
  - Information does not need L2 addresses
  - Wastes resources if communication is not CBR
  - Easy to integrate
  - SDH
- Asynchronous
  - Free usage of resources
  - Information requires L2 addresses
  - Does not waste resources
  - Requires additional logics to control resource usage
  - ATM, Ethernet

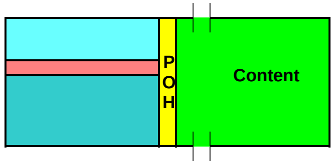
SDH

- **Synchronous frame based multiplexing** of transmitted signals
  - Link framing is done with 2430 byte frames
  - Generation interval is 125us -> reflects the original coding of speech with 8kHz sampling rate
    - Datarate = 155,52Mbps



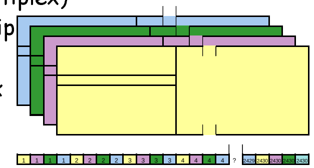
SDH

- Link frames contain virtual containers which carry the actual information
  - Header information (POH)
    - Flow and error control information between edge devices
  - Content
  - Virtual containers form point-to-point permanent connections through SDH network



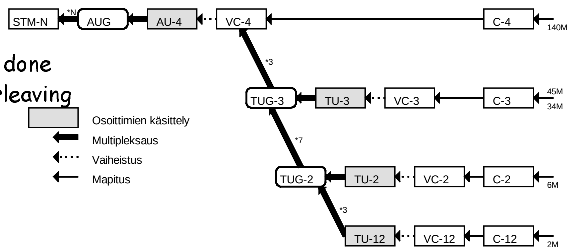
SDH

- **SDH hierarchy makes possible to use multiples and fractions of basic rate**
  - **Multiples** are generated by injecting multiple (factor of four) link frames within time-slot
    - STM-1: 155.52 Mbit/s (basic rate)
    - STM-4: 622.08 Mbit/s (first multiplex)
    - STM-16: 2488.32 Mbit/s (second multiplex)
    - STM-64: 9953.28 Mbit/s (third multip
  - Operation is byte synchronous
    - Timing of individual bytes in multiplex is same than in basic rate frame



SDH

- **Fractions** are generated by multiplexing different streams of content into individual frame
  - Several virtual containers destined to same or different points in network
  - Multiplexing is done with byte interleaving

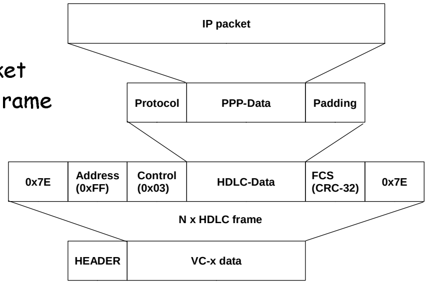


SDH

- SDH supports also **concatenation** of resources
  - **Old version – strict mode**
    - Clear channel operation (small 'c' after the virtual container type)
    - All VC:s in different frames form a single bit stream
    - Not feasible in SDH **networks**
    - Feasible if SDH is used as a point to point link technology
  - **New version – flexible mode**
    - Concatenation is used only in edge devices
      - Supports SDH networks
      - Concatenated VC:s need not be with same speeds
        - » Even over different fibers

SDH

- **IP can not be used directly with SDH**
  - Packet over Sonet (PoS) is method for delivering IP packets in SDH
    - Additional framing
      - IP packet into PPP-packet
      - PPP packet into HDLC frame
      - HDLC frame into SDH virtual container

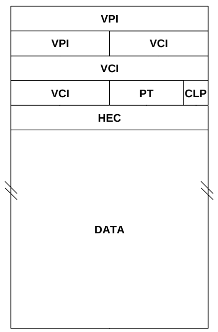


SDH

- **Pros:**
  - Optimized for TDM services (large income from leased line services)
  - Fully compatible with metro ring networks (SDH ADM rings)
  - Reliable and fast failure recovery (roughly 50ms with APS)
  - Price of SDH continuously coming down
- **Cons:**
  - Not cost effective for burst data traffic
    - Capacity in SDH network can only be allocated on multiples of 2Mbps
  - No multiple QoSs for different service charges
  - Expensive interfaces at routers

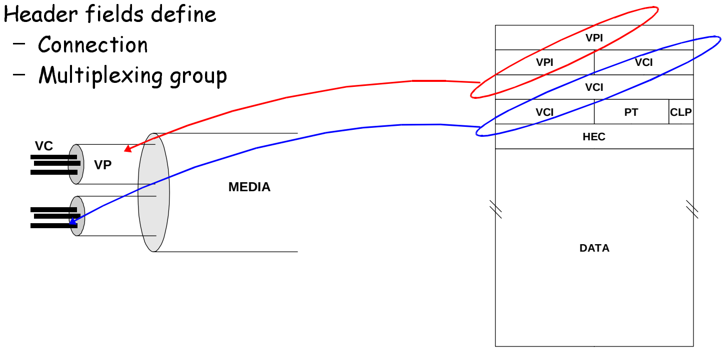
ATM

- **Asynchronous frame based multiplexing**
- Capabilities for dynamic switching
  - Not only PVP's or PVC's
- Connection oriented
- Fixed packet structure
  - 5 bytes of headers
    - Addresses (VPI, VCI)
    - Packet content type (PT)
    - Priority (CLP)
    - Checksum (HEC)
  - 48 bytes of data



ATM

- Header fields define
  - Connection
  - Multiplexing group

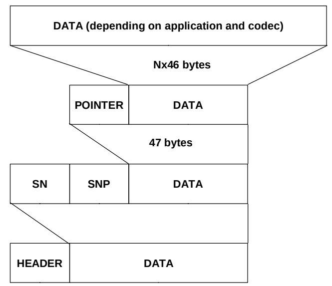


ATM

- Can be used
  - As is over the transmission media
    - Assumes low bit error ratio from the media
  - Over any other L2 protocol
    - Benefits from the error control of L2 media
- Why sensitivity to BER
  - Packet has not markers
    - Delineation is accomplished through state-machine which goes through packet bit by bit and looks header checksum matches
      - Sensitive to errors if high BER

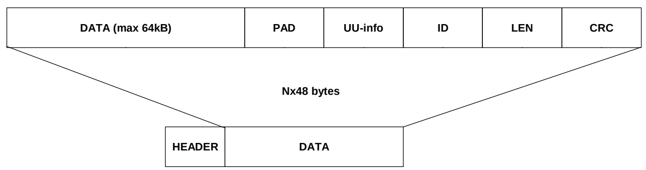
ATM

- 48 byte content field is too big for voice communications
  - Separate protocol layers to handle
    - Sub cell delineation
    - Timing
    - Sequencing
  - Clear channel communication for video applications



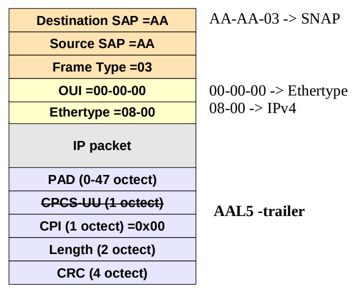
ATM

- 48 byte content field is too little for data networks
  - Fragmentation of data packets into multiple ATM cells
  - Separate protocol layer to handle the fragmentation and reassembly of protocol packets



ATM

- Framing options for IP traffic in ATM links:
  - RFC2684: Multiprotocol Encapsulation over ATM Adaptation Layer 5 (Classical IP)
    - Uses LLC/SNAP encapsulation of traffic within ATM adaption layer 5



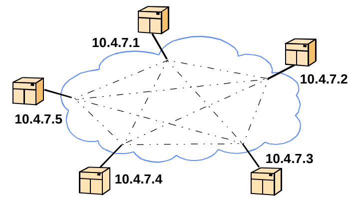
ATM

- Framing options for IP traffic in ATM links:
  - RFC2364: Point to Point Protocol over ATM
    - Uses in AAL5 frames either
      - raw PPP packets
      - PPP on LLC/NLPID packets

Destination SAP	LLC-otsikko
Source SAP	
Frame Type (UI)	
NLPID (PPP)	Network Layer Protocol ID
Protocol ID	
PPP Information	PPP
Padding	
PAD (0-47 octect)	AAL5 -trailer
CPCS-UU (1 octect)	
CPI (1 octect)	
Length (2 octect)	
CRC (4 octect)	

ATM

- **ATM network is from IP perspective**
  - **NBMA network**
    - Separate virtual connection between each and every router
      - Large number of connections and adjacencies in routing
    - Usually subinterface per connection



ATM

- Pros:
  - Easy capacity management
  - Virtual short-cuts without routing
  - MPLS ready
  - Fault tolerant if ATM-level dynamic routing is used
- Cons:
  - Additional layer of technology
    - Not good for framing itself
  - Expensive interfaces at routers
    - Subinterface structure in networked ATM

Ethernet

- **Technology has scaled to level where conventional core network technologies are**
  - STM-64 and 10GbE are the same
    - Even in optical interface level they are the same but ethernet is only 20% of the price
  - STM-256 will be the base for 40GbE ?
  - 1GbE is based on fiber channel but can be multiplexed in STM-16 networks by having two independent connections

Ethernet

- **10GbE**
  - IEEE 802.3ae
  - Full duplex
  - Adjustable MAC speed
    - 10Gb in LAN
    - 9.29Gb in WAN
  - Optical media
  - SDH WAN Phy
  - 10Gb LAN Phy
- **1GbE**
  - 802.3z
  - CSMA/CD + Full Duplex
  - Optical and copper media
  - Fiber channel Phy

Preamble (7 bytes)	Start (1 byte)	DST Add (6 bytes)	SRC Add (6 bytes)	Length (2 bytes)	Data	PAD	FCS (4 bytes)

Ethernet

- **Possibility to build transparent LAN services**
  - Majority of LAN networks are build with ethernet
  - Some applications benefit from the fact that ethernet headers are preserved
    - Possibility to have same IP subnet on both ends
    - WAN network is transparent for ethernet network
      - No PPP protocol in between SDH and Ethernet
  - Core network technologies are evolving
    - Metro VLAN separation
    - Core provider framing

Virtualisation of Ethernet

- **Virtual LAN is <u>a</u> network within <u>the</u> network**
  - It logically creates virtual networks on top of physical network
    - Virtuality is realized with additional fields in the frame
      - Tag Frame ID: 0x8100 for 802.1q
      - Priority: 802.1p priority
      - CFI: Canonical format indicator (MAC address can be or not)
      - VLAN ID: 4096 VLAN IDs
        - » Only few (read expensive) devices support simultaneously this number

Preamble (7 bytes)	Start (1 byte)	DST Add (6 bytes)	SRC Add (6 bytes)	(0x8100) (2 bytes)	Priority (3 bits)	CFI (1 bit)	VLAN ID (12 bits)	Length (2 bytes)	Data	PAD	FCS (4 bytes)

VLAN

- Separation of network resources to logical units is based on forwarding information databased (FID)
  - In **independent mode**, each VLAN has its own FID
    - Clients residing in different forwarding table are not able to communicate without external help
  - In **shared mode**, part of VLANs share a common FID
    - Clients residing (symmetrically) in same FID are able to communicate together
- Communication between VLANs is established with
  - 'Misconfigured' bridge that connects VLANs together
  - Router forwarding packets between VLANs

VLAN association

- Association of devices to VLANs is based on
  - Device tagging (end system is VLAN aware) - rarely
  - **Port based VLAN membership (switchports are assigned to particular VLAN) - commonly**
  - Protocol inspection (MAC address, ethertype, IP address, TCP/UDP port) - only in service switches

VLANs

- **PVID ~ Port VLAN identifier**
  - Each and every switchport is assigned to belong to particular VLAN
    - Incoming untagged traffic is forwarded by using this VLANs FID
      - Address learning is bound to that FID
    - Incoming tagged traffic is associated to VLAN based on VID or FID depending on ingress filtering rules
- **VID ~ VLAN identifier**
  - If frame is
    - Coming in from a trunk interface it contains 802.1q tag which carries VID
    - Going out to trunk link packet is coded to 802.1q tag mode
      - VID usually is PVID from the ingress port

Port based VLANs

- **Ingress filtering rules:**
  - **Received frame is untagged**
    - Forward using PVID
    - Discard
  - **Received frame is tagged**
    - Forward using VID
      - VID = 0:
        - » Use only P-bits, forward using PVID
      - VID = 1
        - » Default tree, all interfaces
    - Forward using PVID
    - Discard

Port based VLANs

- **Egress filtering rules**
  - **Interface is in untagged mode**
    - Forward untagged frame
      - Use configured priorities
  - **Interface is in tagged mode**
    - Set tag based on classification rules
      - Ingress VID
      - PVID
  - P-bits

Protocol VLANs

- Each port belongs to VLAN based on PVID
  - Non matching traffic is forwarded with PVID FID
  - Matching traffic e.g. Traffic having frametype + ethertype defined in protocol group database is forwarded with FID defined by protocol group VID
    - Depends on the
      - encapsulation: ethernet, LLC, SNAP etc
      - ethertype: ARP, IP, IPX, etc

Ethernet priority

- **802.1p is amendment in 802.1q**
  - Allow traffic prioritization within Ethernet networks
  - 3 bits -> 8 priorities
    - Number of queues dependent of HW
    - At minimum strict priority queuing between queues
  - Mapping traffic to queues is dependent on
    - Number of queues
    - Configured policy (egress filtering)
      - MAC address
      - Ethertype
      - DSCP
      - Address

Provider Backbone

- 802.1ad
  - Q-in-Q
  - Provider tagging
  - Enterprise addresses are carried by provider switches in customer dependent service VLANs

Enterprise Header	Provider Tag#1	Provider Tag#N	Enterprise Tag	Data

- 802.1ah
  - M-in-M
  - Provider encapsulation
  - Enterprise addresses are invisible for provider switches

Provider Header	Provider Q-Tag	Provider S-Tag	Enterprise Header	Enterprise Tag	Data

Q-in-Q Frame Format

- **Provider tagging cascades several Q-tags into the frame**
  - Ethertype 0x88a8
  - Priority is provider dependent not copied from customer settings
  - Provider Tag = S-Tag = Service Tag
    - P-VLAN ID = EVC ID

C-DST (6 bytes)	C-DST (6 bytes)	(0x88a8) (2 bytes)	P-Priority (3 bits)	P-CFI (1 bit)	P-VLAN ID (12 bits)	(0x8100) (2 bytes)	C-Priority (3 bits)	C-CFI (1 bit)	C-VLAN ID (12 bits)	Data	FCS (4 bytes)

M-in-M Frame Format

- **Provider encapsulation allows second layer of operation**
  - With or without Q-in-Q-tag
  - With or without Q-tag
- First tag in M-in-M header is traffic engineering tunnel tag
  - Ethernet traffic engineering
- Second tag in service tag = EVC ID

P-DST (6 bytes)	P-DST (6 bytes)	(0x8100) (2 bytes)	PQ-Priority (3 bits)	PQ-CFI (1 bit)	PQ-VLAN ID (12 bits)	(M-in-M) (2 bytes)	Reserved (7 bits)	PT (1 bit)	Service ID (24 bits)	P-FCS (4 bytes)

C-DST (6 bytes)	C-DST (6 bytes)	(0x8100) (2 bytes)	C-Priority (3 bits)	C-CFI (1 bit)	C-VLAN ID (12 bits)	Data	C-FCS (4 bytes)

M-in-M Frame Format

- M-in-M + Q-in-Q allows scalable provisioning of core + metro services
  - Q-in-Q: Small Metro encapsulation
  - M-in-M: Large Metro and Core aggregation
    - Traffic Engineering
    - State space reduction

P-DST (6 bytes)	P-DST (6 bytes)	(0x8100) (2 bytes)	PQ-Priority (3 bits)	PQ-CFI (1 bit)	PQ-VLAN ID (12 bits)	(M-in-M) (2 bytes)	Reserved (7 bits)	PT (1 bit)	Service ID (24 bits)	P-FCS (4 bytes)

C-DST (6 bytes)	C-DST (6 bytes)	(0x88a8) (2 bytes)	P-Priority (3 bits)	P-CFI (1 bit)	P-VLAN ID (12 bits)	(0x8100) (2 bytes)	C-Priority (3 bits)	C-CFI (1 bit)	C-VLAN ID (12 bits)	Data	C-FCS (4 bytes)

Ethernet

- **Pros:**
  - Optimized for burst data services
  - No protocol conversion for interfacing with routers and LAN switches
  - Plug-and-play ideology in operation
- **Cons:**
  - Support for TDM services has not matured
  - Poor in trouble isolation and network recovery
    - Spanning tree operation takes seconds to recover large networks

Carrier Grade Ethernet

- Also known as Metro Ethernet (interest group called: Metro Ethernet Forum)
  - **Technology perspective**
    - Customer side: semi-transparent Ethernet
    - Provider side: SDH, VPLS, Q-in-Q etc
  - **Service definitions**
    - Ethernet line (E-LINE)
      - point-to-point
    - Ethernet LAN (E-LAN)
      - multipoint- to-multipoint

Perspective

- **Ethernet is true service delivery layer**
  - IP is artificially brought to the middle while in many cases it is not even needed

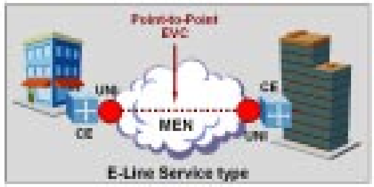
Enabled Service over Ethernet	Storage	Internet Access	IP VPN	CESoE	IP Telephony	Video on Demand
Ethernet Connectivity Service	E-Line and E-LAN (Virtual and Private, MAN and WAN)					
Service Delivery Technology	Ethernet over Fiber	Ethernet over SONET/SDH	Ethernet over RPR	Ethernet over MPLS	Ethernet over WDM	

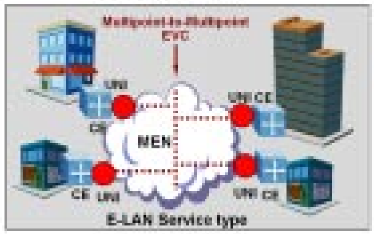
Major changes

- **Service concept**
  - E-LAN and E-Line
- **Connection orientation**
  - Ethernet Virtual Connection (EVC)
    - Filter word for distinguishing packets from different connections
- **QoS**
  - SLA is required for large scale deployment within corporate interconnections
  - Bandwidth control
    - Committed information rate control (inherited form FrameRelay)

E-Line and E-LAN Services

- E-Line Service used to create
  - Private Line Services
  - Ethernet Internet Access
  - Point-to-Point VPNs
- E-LAN Service used to create
  - Multipoint VPNs
  - Transparent LAN Service





Source: Metro Ethernet Forum

Service Attribute	Service Attribute Parameters
EVC Type	Point-to-Point or Multipoint-to-Multipoint
UNI List	A list of UNIs (identified via the UNI Identifier service attribute) used with the EVC
CE-VLAN ID Preservation	Yes or No. Specifies whether customer VLAN ID is preserved or not.
CE-VLAN CoS Preservation	Yes or No. Specifies whether customer VLAN CoS (802.1p) is preserved or not.
Unicast Service Frame Delivery	Specifies whether unicast frames are Discarded, Delivered Unconditionally or Delivered Conditionally
Multicast Service Frame Delivery	Specifies whether multicast frames are Discarded, Delivered Unconditionally or Delivered Conditionally
Broadcast Service Frame Delivery	Specifies whether broadcast frames are Discarded, Delivered Unconditionally or Delivered Conditionally
Layer 2 Control Protocol Processing	Discard or Tunnel per protocol
Service Performance	Specifies the Frame Delay, Frame Jitter and Frame Loss per EVC or frames within an EVC identified via their CE-VLAN CoS (802.1p) value

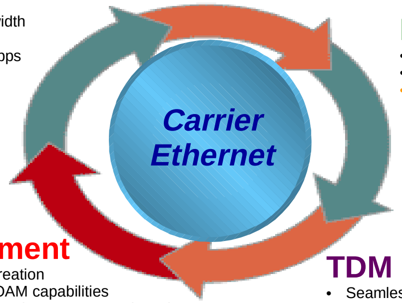
Source: Metro Ethernet Forum

Service Attribute	Service Attribute Parameters
UNI Identifier	A string used to identity of a UNI, e.g., NYCBldg12Rm102Slot22Port3
Physical Medium	Standard Ethernet PHY
Speed	10 Mbps, 100 Mbps, 1 Gbps or 10 Gbps
Mode	Full Duplex or Auto negotiation
MAC Layer	IEEE 802.3-2002
Service Multiplexing	Yes or No. Defines whether multiple services can be on the UNI
UNI EVC ID	A string used identify an EVC, e.g., NYCBldg1Rm102Slot22Port3EVC3
CE-VLAN ID / EVC Map	Mapping table of customer VLAN IDs to EVC
Max. Number of EVCs	The maximum number of EVCs allowed per UNI
Bundling	No or Yes. Specifies that one or more customer VLAN IDs are mapped to an EVC at the UNI
All to One Bundling	No or Yes (all customer VLAN IDs are mapped to an EVC at the UNI).
Ingress Bandwidth Profile Per Ingress UNI	None or <CIR, CBS, EIR, EBS>. This Bandwidth profile applies to all frames across the UNI.
Ingress Bandwidth Profile Per EVC	None or <CIR, CBS, EIR, EBS>. This Bandwidth profile applies to all frames over particular EVC.
Ingress Bandwidth Profile Per CoS ID	None or <CIR, CBS, EIR, EBS>. This Bandwidth profile applies to all frames marked with a particular CoS ID over an EVC.
Layer 2 Control Protocol Processing	Discard, Peer or Pass to EVC per protocol

Source: Metro Ethernet Forum

Protection
- 50ms Protection
- End to End Path Protection
- Aggregated Line & Node Protection

Scalability
- Services and Bandwidth
- 100,000's of EVC's
- From Mbps to x10Gbps



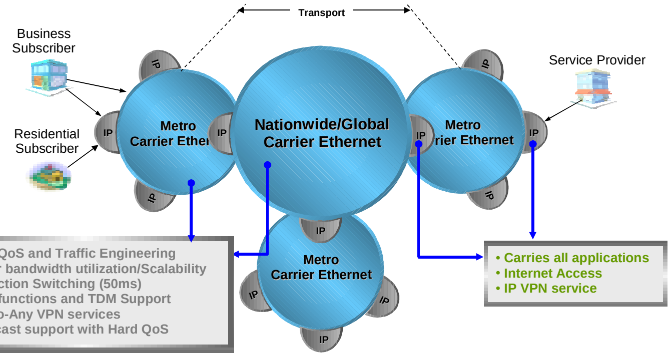
Hard QoS
- Guaranteed end to end SLA
- End to End CIR and EIR
- Business, Mobile, Residential

Service Management
- Fast service creation
- Carrier class OAM capabilities
- Customer Network Management (CNM)

TDM Support
- Seamless integration of TDM
- Circuit Emulation Services
- Support existing voice applications

Source: Metro Ethernet Forum

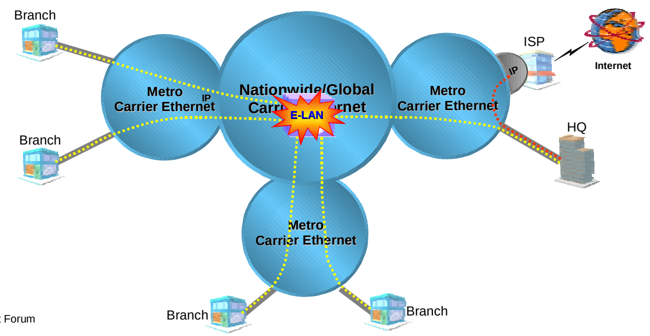
MEF Vision for NGN



Source: Metro Ethernet Forum

VPNs in CGE

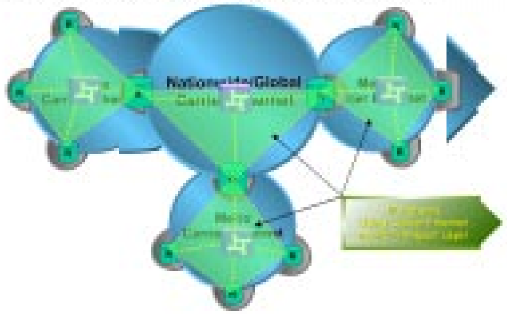
- Any-to-any Layer 2 VPN to interconnect multiple locations transparently and effectively (with high and guaranteed bandwidth)



Source: Metro Ethernet Forum

CGE and IP

- E-LAN connections among routers
  - Any logical IP topology, multiple EVC's on a single physical interface



Source: Metro Ethernet Forum

CGE and IP

- **IP networks are collapsed in to two layers**
  - Residential customer concentration routers
    - Integrated in DSLAM
  - Border routers
    - For address propagation and policy control
- **IP aggregation is vanishing**
  - Ethernet aggregation is taking the role of the true transport
    - Even in 3G networks
      - Look for latest ITU drafts for Ethernet aggregation in 3G networks