



Technology Backgrounder

Ethernet Environment

1.	Introduction to Ethernet Transmission	1
2.	Ethernet LAN Topologies	1
3.	Ethernet Communication Protocol	2
	Media Access Method.....	2
	Basic Ethernet Frame Structure.....	3
4.	Bridging.....	5
	Communication between Nodes on Same LAN.....	5
	Communication between Nodes on Different LANs	5
	Using Virtual Bridged LANs	5
5.	Transporting IP Traffic over Ethernet	6
	Encapsulation in Ethernet Frames	6
	ARP Protocol	6
	AAL5 CPCS PDU Fields	7
6.	Transporting Ethernet over SDH/SONET Networks	8
	LAPS Encapsulation.....	8
	GFP Encapsulation	9
	GFP Protocol Stack	10
	GFP Multiplexing.....	10

1. Introduction to Ethernet Transmission

One of the most successful digital transmission technologies is referred to by the generic term **Ethernet**. The Ethernet technology is suitable for a wide range of physical media: coaxial cable, twisted pairs and optical fibers. The current standards for Ethernet transmission cover rates from 10 Mbps to 10 Gbps. In many office LANs, Ethernet runs at 10 Mbps and 100 Mbps; network interface cards (NICs) for PCs usually support both of these rates.

The basic standard covering Ethernet LANs is IEEE Standard 802.3, which is very similar to the original Ethernet V2.0 specification (ISO/IEC also have a similar standard). In addition to the aspects covered by IEEE 802.3 standards, there is a wide range of LAN standards (the IEEE 802 family) that cover other aspects of LAN transmission, for example, bridging, with particular emphasis on Ethernet LANs.

Ethernet standards (in their broadest interpretation) cover the physical and data link control layers (layers 1 and 2 in the OSI model; IP is a layer 3 protocol). The data link control layer is split into two sublayers: media access control (MAC) and logical link control (LLC).

2. Ethernet LAN Topologies

Ethernet LANs use a multidrop topology. The LANs can be implemented either in bus or star (hub-based) topology.

Figure 1 shows the general structure of a LAN using the star topology.

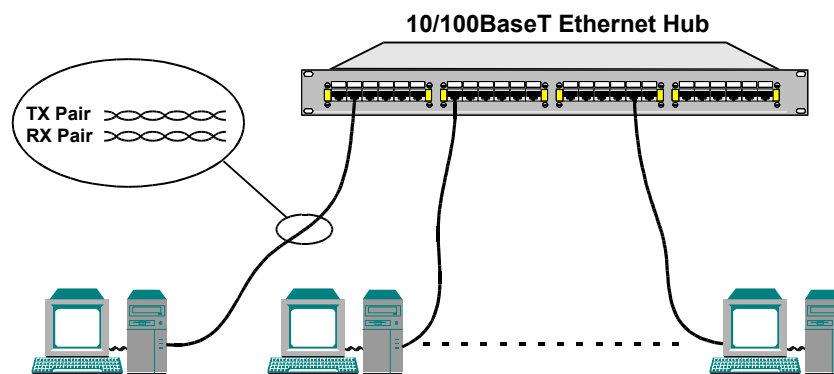


Figure 1. Star (Hub-Based) Ethernet LAN Topology

In the star topology, all the nodes on the LAN are connected to a common unit, which serves as the hub of the LAN. The hub can be implemented in two ways:

- Simple Ethernet hub, which detects the transmitting node and transparently distributes its signal to all the other nodes. A hub supports only half-duplex communication (the same as in a bus topology).

- Ethernet switch: the switch includes more sophisticated circuits that enable both half-duplex and full-duplex operation and prevent collisions.

The LAN cables are usually made of two twisted pairs (one transmit pair and one receive pair). The standard connector type is RJ-45, and its pin assignment has also been standardized. However, because of the need to use separate transmit and receive pairs, two types of port pin assignments have developed: station ports and hub ports (the difference is interchanging of the transmit and receive pins in the connector). This permits to interconnect connectors of different types by a cable wired pin-to-pin (straight cable). To interconnect ports of same type, a crossed cable (a cable wired to interconnect the transmit pair at one end to the receive pair at the other end) is necessary.

Interfaces operating on twisted pairs are designated in accordance with data rate: 10BaseT (10 Mbps) or 100BaseTX (100 Mbps, where X is the number of pairs). Interfaces that support both rates are identified as 10/100BaseT.

3. Ethernet Communication Protocol

Today, Ethernet is used as a generic term for a LAN transmission technology that uses Carrier Sense and Multiple Access with Collision Detection (CSMA/CD) to enable the transmission of short bursts of data (called **frames**) between two or more stations (**nodes**).

All the users have permanent access to the full bandwidth of the transmission medium but can only use it for short times, by transmitting short data bursts. Each data burst has a fixed structure, called a frame. The frame structure is explained below. The connection point of each user to the transmission media is called a node. For identification purposes, each LAN node has its own unique number, called MAC address.

Media Access Method

Media access is performed by means of the carrier sense, multiple access protocol (CSMA) with collision detection (CD), defined by IEEE Standard 802.3. The protocol defines three basic steps:

- A node that wants to transmit checks that the LAN is free. If another node is already transmitting, the node waits until the LAN is free.
- When the LAN is free, the node starts transmission and sends its frame. Each node has equal access rights, therefore the first node that starts transmitting is the one that seizes the LAN.
- When two nodes start transmitting at the same instant, a collision occurs. In this case, the transmitting nodes will continue to transmit for some time, in order to ensure that all transmitting nodes detected the collision (this is called “jamming”). After the jamming period, all transmitting nodes stop the transmission and wait for a random period of time before trying again.

The delay times are a function of collision numbers and random time delay, therefore there is a good chance that an additional collision between these nodes will be avoided, and the nodes will be able to transmit their messages.

The basic procedure described above has been developed for half-duplex communication, because it declares a collision whenever data is received during a local transmission. However, when using twisted pairs, separate pairs are used for the transmit and receive directions. Therefore, each node is capable of simultaneously transmitting and receiving (full-duplex operation), thereby doubling the effective data rate on the LAN.

Modern Ethernet interfaces, designated 10/100BaseT, are also capable of operation at the two basic rates, 10 Mbps and 100 Mbps. Therefore, the rate and operating mode (half-duplex or full-duplex) are user-configurable options.

When connecting equipment from different vendors to a common LAN, four operating modes are possible. These modes are listed below in ascending order of capabilities:

- Half-duplex operation at 10 Mbps.
- Full-duplex operation at 10 Mbps.
- Half-duplex operation at 100 Mbps.
- Full-duplex operation at 100 Mbps.

To ensure interoperability (which practically means to select the highest transport capability supported by all the equipment connected to the LAN), two approaches can be used: manual configuration of each equipment interface, or automatic negotiation (autonegotiation) in accordance with IEEE Standard 802.3.

The autonegotiation procedure enables automatic selection of the operating mode on a LAN, and also enables equipment connecting to an operating LAN to automatically adopt the LAN operating mode (if it is capable of supporting that mode).

When autonegotiation is enabled on all the nodes attached (or trying to attach) to a LAN, the process is always successful. However, even if the nodes on an operating LAN are manually configured for operation in a fixed mode, a late-comer node with autonegotiation capability can still resolve the LAN operating rate can be resolved, thereby enabling it to adopt the LAN rate. Under these conditions, an autonegotiating node cannot detect the operating mode (half or full duplex), and therefore they will default to half-duplex. Therefore, as a practical configuration rule, do not enable the full-duplex mode without enabling autonegotiation, except when all the nodes have been manually configured for the desired operating mode (which may of course be full duplex).

Basic Ethernet Frame Structure

The frame transmitted by each node contains routing, management and error correction information. For Ethernet LANs, the characteristics of frames are defined by IEEE Standard 802.3.

Basic frame lengths can vary from 72 to 1526 bytes and have the typical structure shown in *Figure 2*.

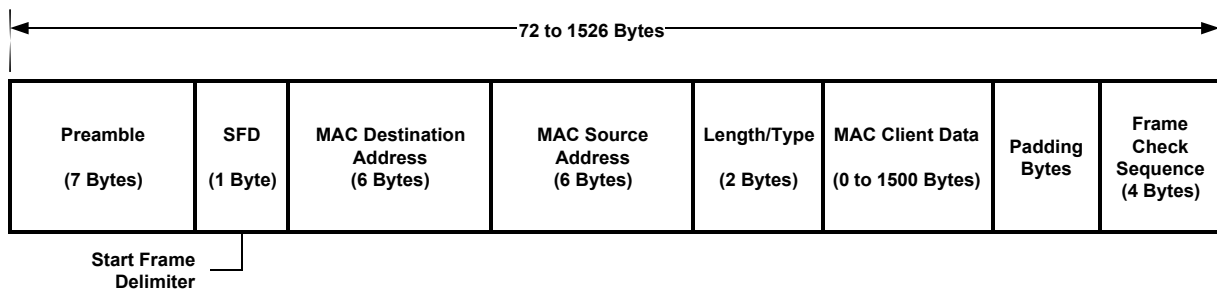


Figure 2. Basic Ethernet Frame Structure

- **Preamble.** Each frame starts with a preamble of seven bytes. The preamble is used as a synchronizing sequence for the interface circuits, and helps bit decoding.
- **Start-Frame Delimiter (SFD)** field – consists of one byte. The SFD field indicates where the useful information starts.
- **Medium-Access (MAC) Destination Address (DA)** field – consists of six bytes. The MAC DA field carries the address of the destination node.
- **Medium-Access (MAC) Source Address (SA)** field – consists of six bytes. The MAC SA field carries the address of the source node.

Note *In conventional notation MAC addresses are represented as 6 pairs of hexadecimal digits, separated by dashes, for example, 08-10-39-03-2F-C3.*

- **Length/Type** field – consists of two bytes that indicate the number of bytes contained in the logical link control (LLC) data field. In most Ethernet protocol versions, this field contains a constant indicating the protocol type (in this case, this field is designated **EtherType**).
- **MAC Client Data** field. The MAC client data field can contain 0 to 1500 bytes of user-supplied data.
- **Padding** field. The optional padding field contains dummy data that is used to increase the length of short frames to at least 64 bytes.
- **Frame Check Sequence (FCS)** field – contains four check bytes generated by a cyclic redundancy check (CRC) code. The FCS field is used to detect errors in the data carried in the frame.

4. Bridging

Communication between Nodes on Same LAN

A MAC address is unique and identifies a single physical port. Therefore, two Ethernet nodes attached to the same LAN exchange frame directly, by specifying the desired MAC destination address, together with the source MAC address.

The node that identifies its MAC address in the destination field can send a response by copying the source address of the frame to the destination address field.

Communication between Nodes on Different LANs

To enable nodes on different LANs to communicate, it is necessary to transfer frames between the two LANs. The device used for this purpose is called **MAC bridge**, or just **bridge**. Two types of bridges are used:

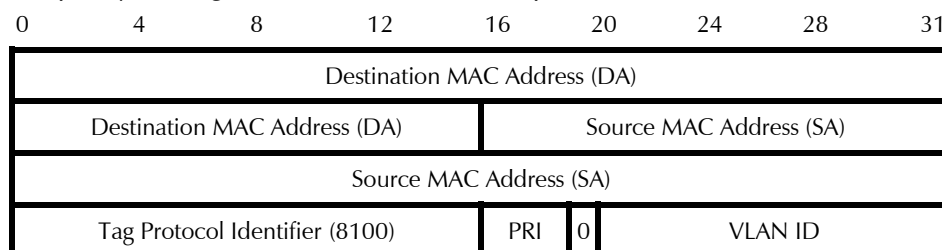
- Local bridges, which have Ethernet ports attached to the two LANs. The bridge control mechanism learns the nodes attached to each LAN by reading the source MAC addresses of the frames generated by the nodes. When the destination address of a frame is not on the LAN from which it was received, the bridge transfers it to the other LAN.
- Remote bridges, which are used in pairs. A basic remote bridge has one LAN port and one WAN port. The WAN port communicates through a link with the WAN port of the remote bridge connected to the desired remote LAN. In this case, the traffic addressed to destinations not located on the local LAN is transferred through the WAN link to the remote bridge.

Using Virtual Bridged LANs

VLAN can be used to provide separation between traffic from different sources sharing the same physical transmission facilities, and provide information on the relative priority the user assigns to each frame. The characteristics and use of virtual LANs (VLANs) and of the MAC bridges capable of handling tagged frames are defined in IEEE Standard 802.1Q.

VLANs are made possible by a slight modification to the Ethernet frame structure shown in [Figure 2](#).

The structure of an Ethernet frame with VLAN support is shown in [Figure 3](#) (for simplicity, the figure does not include the preamble and SFD fields).



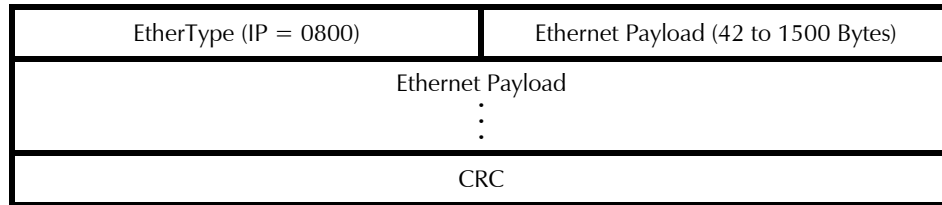


Figure 3. Structure of Ethernet Frame with VLAN Support

Ethernet frames with VLAN support include a tag header immediately after the source MAC address (therefore, such frames are also referred to as **tagged frames**).

The tag header comprises 4 bytes:

- Two bytes for the tag protocol identifier. For Ethernet-encoded tags in accordance with IEEE802.1Q (these are the tags used by ML-IP modules), these bytes carry the equivalent of 8100.
- Priority (PRI) specified by the user (3 bits: 7 is the highest priority and 0 is the lowest priority).
- One bit for the canonical format indicator (always 0 as shown in [Figure 3](#)).
- VLAN ID (12 bits), used to indicate the VLAN to which the frame belongs.

5. Transporting IP Traffic over Ethernet

Encapsulation in Ethernet Frames

IP traffic is carried in the LLC data field of the Ethernet frame (see [Figure 2](#)). This is called **encapsulation**. The EtherType value for the IP protocol is 0800.

Whenever possible, the whole IP packet (including the header) is inserted in one Ethernet frame. However, IP packets can be much longer than the LLC data field of Ethernet frames: in this case, it is necessary to fragment the IP packets in accordance with the desired size of data field, and transfer each fragment in a separate frame. The receiving IP host then reassembles the original packet from its fragments.

Note TDMoIP packets are never fragmented. This is not necessary anyway, because a TDMoIP packet is relatively short.

ARP Protocol

When sending IP packets over Ethernet, it is necessary to determine the MAC address of the destination, to insert it in the Ethernet destination MAC address of the packet. Actually, this is necessary for any physical transmission technology that is not limited to point-to-point topologies.

This is performed by means of the ARP (Address Resolution Protocol), part of the IP suite of protocols. ARP is used to generate a look-up table that translates IP

addresses to MAC addresses for any transmission technology. The translation is done only for outgoing IP packets, because this is when the IP header and the Ethernet header are created.

The ARP table contains one row for each IP host: each row has two columns, one listing the IP address and the other listing the corresponding MAC (Ethernet) address. When translating an IP address to an Ethernet address, the table is searched for the row corresponding to the destination IP address, and the corresponding Ethernet address is then found in the same row.

Whenever a packet must be sent to a new IP destination, that is, a destination whose MAC address is not known, the IP host sends an ARP request packet, listing its own IP address and MAC address, the destination IP address, but no destination MAC address. When the packet reaches the destination address (using the IP routing process), the destination returns an ARP response packet, in which its own MAC address field is filled. The packet eventually returns to the sender, thereby providing the missing information.

The structure of ARP packets is shown in *Figure 4*.

0	4	8	12	16	20	24	28	31
Hardware Address Type (Ethernet = 1)				Protocol Address Type (IP = 2048)				
Length of Hardware (MAC) Address Field (Ethernet = 6 Bytes)		Length of Protocol Address Field (IP = 4 Bytes)		Operation: ARP Request = 1 ARP Response = 2				
Source MAC Address (Bytes 0 through 3)								
Source MAC Address (Bytes 4, 5)				Source IP Address (Bytes 0, 1)				
Source IP Address (Bytes 2, 3)				Destination MAC Address (Bytes 0, 1)				
Destination MAC Address (Bytes 2 through 5)								
Destination IP Address (Bytes 0 through 3)								

Figure 4. Structure of ARP Packets

AAL5 CPCS PDU Fields

Ethernet traffic can be transferred over the ATM network using the AAL5 layer, in accordance with RFC 1483 (multiprotocol encapsulation over ATM). The Ethernet frames are transferred using LLC-SNAP encapsulation. All the Ethernet traffic is transferred over a single connection, irrespective of the MAC addresses of the frames.

Figure 5 shows the structure of the AAL5 CPCS PDU payload field used to carry Ethernet frames using the LLC-SNAP encapsulation method.

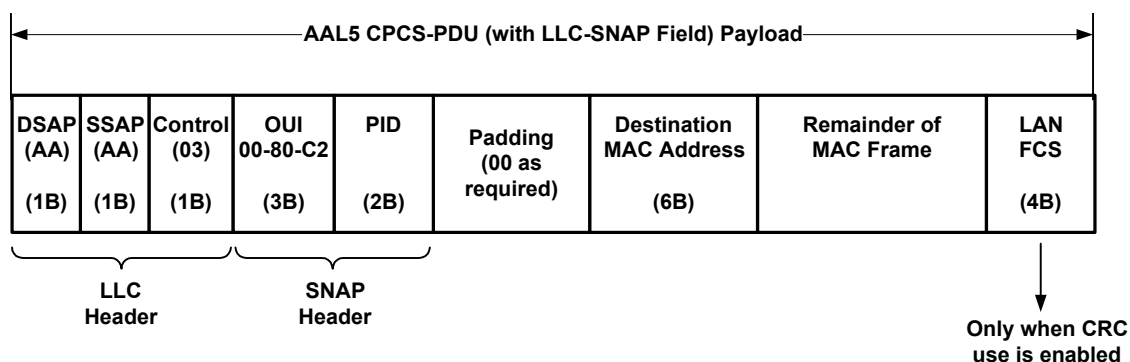


Figure 5. Structure of Payload Field of AAL5 CPCS PDU Frame

The payload field includes the original LLC and SNAP headers. The PID field in the SNAP header can assume two values:

- 0001 (hexa) to indicate the bridged IEEE 802.3 protocol, with end-to-end transmission of the FCS field of the Ethernet frame.
- 0007 (hexa) to indicate the bridged IEEE 802.3 protocol, without end-to-end transmission of the FCS field.

After the SNAP header, the AAL5 CPCS PDU includes optional padding bytes (0 through 47, as required to ensure that the length of the payload field is an integer multiple of 48 bytes (the ATM cell payload length), the destination MAC address and then the other parts of the MAC frame (see figure C-10). When the PID is 0001, the AAL5 CPCS PDU also includes the 4 bytes of the LAN frame FCS field.

6. Transporting Ethernet over SDH/SONET Networks

Ethernet payloads (10 Mbps or 100 Mbps) can be carried over SDH/SONET using several encapsulation protocols. The two encapsulation protocols supported by some RAD's devices are the following:

- Link Access Protocol – SDH (LAPS) in accordance with ITU-T Rec. X.86
- Generic Framing Procedure (GFP) in the framed mode, in accordance with ITU-T Rec. G.7041. With GFP, it is possible to increase the bandwidth utilization efficiency using the GFP multiplexing method.

LAPS Encapsulation

With LAPS, each Ethernet frame is encapsulated in the frame structure shown in [Figure 6](#). The LAPS frame is delineated by flags, followed by HDLC information (address and control), and by a LAPS service access point identifier (SAPI). The Ethernet frame is followed by a LAPS frame checksum (FCS), for error detection.

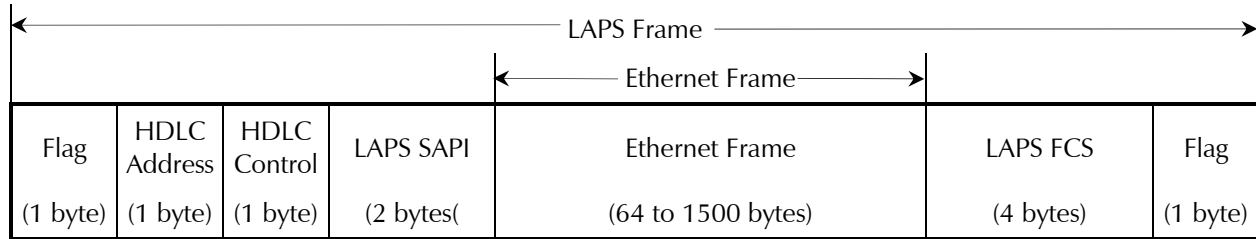


Figure 6. LAPS Encapsulation Format

GFP Encapsulation

The GFP encapsulation method uses the basic frame structure of [Figure 7](#).

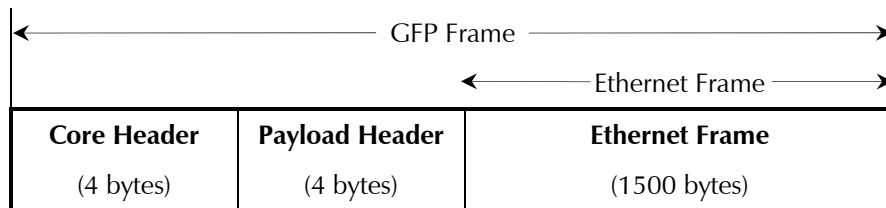


Figure 7. Basic GFP Encapsulation Format

[Figure 8](#) shows the detailed structure of the basic GFP frame. The frame includes the following fields:

- PLI** Payload length indicator
- cHEC** Core header CRC (calculated using ITU-T CRC-16 polynomial)
- Payload Area** Carries a framed PDU
- Payload Header** Header used for client PDU management
- pFCS** Optional payload FCS (calculated using ITU-T CRC-32 polynomial)

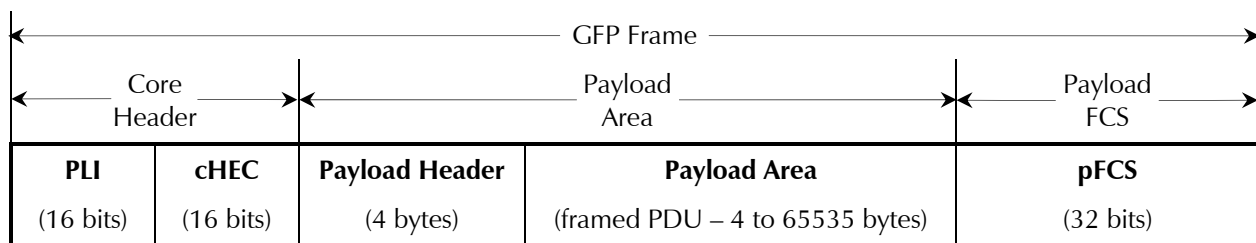


Figure 8. Detailed Structure of Basic GFP Frame

All the GFP OAM&P functions are handled by the GFP core header.

The payload header supports the payload-specific adaptation functions, which depend on the client application. The payload header also supports multiplexing (using extension headers), and any application-dependent link management functions (using dedicated client management frames)

Protection against errors (on a per frame basis) is provided by the optional payload frame checksum (FCS) field.

Idle frames are used for asynchronous rate adaptation.

GFP Protocol Stack

The encapsulation process can be described by means of a GFP protocol stack, illustrated in [Figure 9](#):

- The top layer is the client application, in this case Ethernet.
- The link layer is divided into two sections:
 - GFP client-specific aspects section: handles the frames received from the client application, and presents them in a standardized format to the GFP common aspects section.
 - GFP common aspects section: performs adaptation to the physical layer.
- The physical layer is provided by the SDH or SONET path.

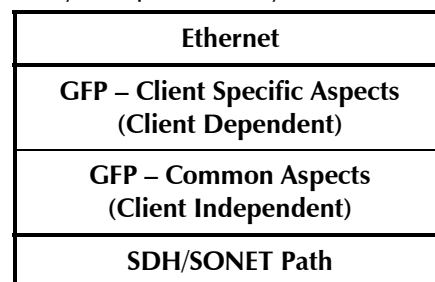


Figure 9. GFP Protocol Stack

GFP Multiplexing

Internal Ethernet switches provide up to four independent data streams. Each stream may be mapped to a different VC or SPE (or to a user-defined virtually concatenated group of VCs or SPEs), however in this case bandwidth may be wasted, because this approach always occupies the full SDH/SONET bandwidth, irrespective of the actual traffic load generated by the client application, which in general varies randomly with time.

To take advantage of the statistical distribution of traffic, multiplexing can be used. GFP provides support for multiplexing, using two approaches (see [Figure 10](#)):

- Frame multiplexing in accordance with the frame type, using the PTI (payload type identifier) field:
 - Client data frames have priority over client management frames
 - Client management frames have priority over idle frames.
- Client multiplexing, implemented by adding extension headers. For this purpose, the extension header includes a customer, or channel, identifier (CID). The CID enables discriminating among various data sources using the same transmission path.

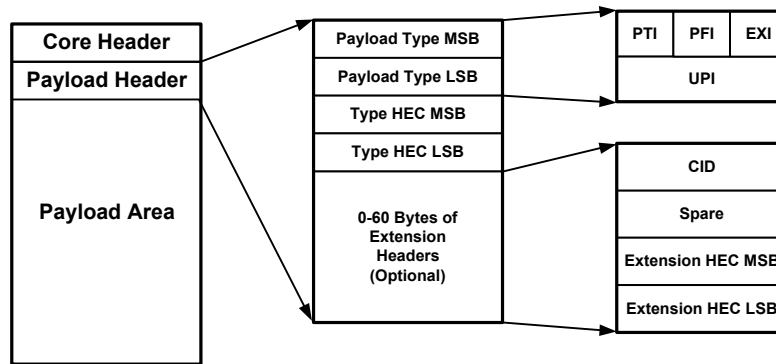


Figure 10. Support for GFP Multiplexing at the GFP Frame Level

To take advantage of the GFP client multiplexing capability, the user can configure:

- The virtual groups to be multiplexed
- The fraction of physical layer bandwidth to be guaranteed to each virtual group, in 12.5% steps
- The VCs/SPEs to carry the multiplexed groups, and the specific mapping.

With GFP multiplexing, the bandwidth available to each virtual group is never less than the specified minimum. However, if the instantaneous traffic carried by a virtual group does not fully utilize the reserved bandwidth, any unused bandwidth is made available to the other virtual groups. This ensures the best possible utilization of the available bandwidth, without degrading the quality of service.