International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.212
(11/2006)

SERIES J: CABLE NETWORKS AND TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

Interactive systems for digital television distribution

Downstream external Physical layer interface for modular cable modem termination systems

ITU-T Recommendation J.212

ITU-T

ITU-T Recommendation J.212

Downstream external Physical layer interface for modular cable modem termination systems

Summary

This Recommendation defines an interface known as the Downstream External PHY Interface (DEPI) and associated protocol requirements for the transport of downstream user data between the "M-CMTS Core" and the EQAM. It describes the characteristics of the DEPI interface, provides requirements that must be met by the M-CMTS Core and the EQAM, and also describes various aspects of technical issues that are involved in the implementation and deployment of a DOCSIS system using the M-CMTS architecture.

Source

ITU-T Recommendation J.212 was approved on 29 November 2006 by ITU-T Study Group 9 (2005-2008) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

ITU-T Recommendation J.212

Downstream external Physical layer interface for modular cable modem termination systems

1 Scope

This Recommendation is part of the DOCSIS® family of Recommendations that define a Modular Cable Modem Termination System (M-CMTS™) architecture for head-end components that comply with DOCSIS.

The DOCSIS Recommendations [J.122] define the requirements for the two fundamental components that comprise a high-speed data-over-cable system: the cable modem (CM) and the cable modem termination system (CMTS). The M-CMTS architecture was designed as an extension to the DOCSIS Recommendations to allow for flexibility and independent scaling of certain CMTS functions, and to allow operators to more efficiently use available network resources.

One of the key elements of the M-CMTS architecture is the separation of the downstream physical layer QAM modulation and up-conversion functions from the CMTS, and the placement of that functionality into an "Edge-QAM" (EQAM) device. This separation allows for the development of EQAM products that support both video and DOCSIS, which in turn allows operators to use the same network resources to support multiple types of services such as data, voice and video.

This Recommendation defines an interface known as the Downstream External PHY Interface (DEPI) and associated protocol requirements for the transport of downstream user data between the "M-CMTS Core" and the EQAM. It describes the characteristics of the DEPI interface, provides requirements that must be met by the M-CMTS Core and the EQAM, and also describes various aspects of technical issues that are involved in the implementation and deployment of a DOCSIS system using the M-CMTS architecture.

Two different master clock frequency options are available within the M-CMTS architecture. The first option, a 10.24 MHz master clock is used in the Americas, Europe and other regions. The second option, a 9.216 MHz master clock is used in Japan. Compliance with this Recommendation requires compliance with the one or the other of these implementations, not with both. It is not required that equipment built to one option shall interoperate with equipment built to the other.

A list of the documents in the Modular CMTS Interface Recommendations family is provided below.

Designation	Title
J.212	Downstream External PHY Interface
J.211	DOCSIS Timing Interface
CM-SP-ERMI	Edge Resource Manager Interface, CM-SP-ERMI-I02-051209, 9 December 2005, Cable Television Laboratories, Inc.
CM-SP-M-OSSI	M-CMTS Operations Support System Interface, CM-SP-M-OSSI-I02-051209, 9 December 2005, Cable Television Laboratories, Inc.

2 References

2.1 Normative references

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[H.222.0] ITU-T Recommendation H.222.0 (2006) | ISO/IEC 13818-1:2007, *Information technology – Generic coding of moving pictures and associated audio information: Systems*.

[J.83] ITU-T Recommendation J.83 (1997), *Digital multi-programme systems for television, sound and data services for cable distribution*.

[J.122] ITU-T Recommendation J.122 (2002), *Second-generation transmission systems for interactive cable television services – IP cable modems.*

[J.210] ITU-T Recommendation J.210 (2006), *Downstream RF interface for Cable Modem Termination Systems*.

[J.211] ITU-T Recommendation J.211 (2006), *Timing interface for Cable Modem Termination Systems*.

[IANA-PORTS] IANA (2004), *Port Numbers*.

[IEEE-802.1Q] IEEE Std 802.1Q-2005, *IEEE Standard for Local and Metropolitan Area Networks – Virtual Bridged Local Area Networks*.

[IEEE-802.3] IEEE Std 802.3-2005, *IEEE Standard for information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 3: Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method And Physical Layer Specification*.

[RFC-IP] IETF RFC 791 (1981), *Internet Protocol – DARPA Internet Program – Protocol specification*.

[RFC-L2TP-DSCP] IETF RFC 3308 (2002), *Layer Two Tunnelling Protocol (L2TP) Differentiated Services Extension*.

[RFC-L2TPv3] IETF RFC 3931 (2005), *Layer Two Tunnelling Protocol – Version 3 (L2TPv3).*

[RFC-MTU] IETF RFC 1191 (1990), *Path MTU Discovery*.

[RFC-PHBID-AF] IETF RFC 2597 (1999), *Assured Forwarding PHB Group*.

[RFC-PHBID-EF] IETF RFC 3246 (2002), *An Expedited Forwarding PHB (Per-Hop Behaviour).*

[RFC-UDP] IETF RFC 768 (1980), *User Datagram Protocol*.

2.2 Informative references

[DVB-RF] ETSI EN 300 429 V1.2.1 (1998), *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for cable systems*.

[ERMI]	*Edge Resource Manager Interface*, CM-SP-ERMI-I02-051209, 9 December 2005, Cable Television Laboratories, Inc.
[ISO 8802-2]	ISO/IEC 8802-2:1998 (IEEE Std 802.2:1994) – *Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 2: Logical link control*.
[MOSSI]	*DOCSIS M-CMTS Operations Support Interface*, CM-SP-M-OSSI-I02-051209, 9 December 2005, Cable Television Laboratories, Inc.
[RFC-DSCP-1]	IETF RFC 2983 (2000), *Differentiated Services and Tunnels*.
[RFC-DSCP-2]	IETF RFC 3260 (2002), *New Terminology and Clarifications for Diffserv*.
[VCCV]	IETF draft-ietf-pwe3-vccv-12.txt, *Pseudo Wire Virtual Circuit Connectivity Verification (VCCV)*, January 2007.

2.3 Reference acquisition

– Cable Television Laboratories, Inc., Internet: http://www.cablelabs.com/; http://www.cablemodem.com/specifications/

– The Institute of Electrical and Electronics Engineers, Inc, Internet: http://standards.ieee.org

– Internet Engineering Task Force (IETF), Internet: http://www.ietf.org

– Internet Assigned Numbers Authority, IANA, Internet: http://www.iana.org

– European Telecommunications Standards Institute, ETSI, http://www.etsi.org

3 Terms and definitions

This Recommendation defines the following terms:

3.1 bonded channels: A logical channel comprising multiple individual channels.

3.2 cable modem (CM): A modulator-demodulator at subscriber locations intended for use in conveying data communications on a cable television system.

3.3 converged interconnect network: The network (generally gigabit Ethernet) that connects an M-CMTS Core to an EQAM.

3.4 customer premises equipment (CPE): Equipment at the end user's premises; may be provided by the service provider.

3.5 data rate: Throughput, data transmitted in units of time usually in bits per second (bit/s).

3.6 decibels (dB): Ratio of two power levels expressed mathematically as $dB = 10\log_{10}(P_{OUT}/P_{IN})$.

3.7 decibel-millivolt (dBmV): Unit of RF power expressed in decibels relative to 1 millivolt, where $dBmV = 20\log_{10}(\text{value in mV}/1\text{ mV})$.

3.8 downstream (DS):

1) Transmissions from CMTS to CM. This includes transmission from the M-CMTS Core to the EQAM as well as the RF transmissions from the EQAM to the CM.

2) RF spectrum used to transmit signals from a cable operator's headend or hub site to subscriber locations.

3.9 edge QAM modulator (EQAM): A headend or hub device that receives packets of digital video or data. It re-packetizes the video or data into an MPEG transport stream and digitally modulates the digital transport stream onto a downstream RF carrier using quadrature amplitude modulation (QAM).

3.10 flow: A stream of packets in DEPI used to transport data of a certain priority from the M-CMTS Core to a particular QAM channel of the EQAM. In PSP operation, there can exist several flows per QAM channel.

3.11 GigE (GE): Gigabit Ethernet (1 Gbit/s).

3.12 hybrid fibre/coax (HFC) system: A broadband bidirectional shared-media transmission system using optical fibre trunks between the headend and the fibre nodes, and coaxial cable distribution from the fibre nodes to the customer locations.

3.13 Institute of Electrical and Electronics Engineers (IEEE): A voluntary organization which, among other things, sponsors standards committees and is accredited by the American National Standards Institute (ANSI).

3.14 Internet Engineering Task Force (IETF): A body responsible for, among other things, developing standards used in the Internet.

3.15 Internet Protocol (IP): An Internet network-layer protocol.

3.16 L2TP access concentrator (LAC): If an L2TP Control Connection Endpoint (LCCE) is being used to cross-connect an L2TP session directly to a data link, we refer to it as an L2TP Access Concentrator (LAC). An LCCE may act as both an L2TP Network Server (LNS) for some sessions and LAC for others, so these terms must only be used within the context of a given set of sessions unless the LCCE is in fact single purpose for a given topology.

3.17 L2TP attribute value pair (AVP): The L2TP variable-length concatenation of a unique Attribute (represented by an integer), a length field, and a Value containing the actual value identified by the attribute.

3.18 L2TP control connection: An L2TP control connection is a reliable control channel that is used to establish, maintain, and release individual L2TP sessions as well as the control connection itself.

3.19 L2TP control connection endpoint (LCCE): An L2TP node that exists at either end of an L2TP control connection. May also be referred to as LAC or LNS, depending on whether tunnelled frames are processed at the data link (LAC) or network layer (LNS).

3.20 L2TP control connection ID: The Control Connection ID field contains the identifier for the control connection, a 32-bit value. The Assigned Control Connection ID AVP, Attribute Type 61, contains the ID being assigned to this control connection by the sender. The Control Connection ID specified in the AVP must be included in the Control Connection ID field of all control packets sent to the peer for the lifetime of the control connection. Because a Control Connection ID value of 0 is used in this special manner, the zero value must not be sent as an Assigned Control Connection ID value.

3.21 L2TP control message: An L2TP message used by the control connection.

3.22 L2TP data message: An L2TP message used by the data channel.

3.23 L2TP endpoint: A node that acts as one side of an L2TP tunnel.

3.24 L2TP network server (LNS): If a given L2TP session is terminated at the L2TP node and the encapsulated network layer (L3) packet processed on a virtual interface, we refer to this L2TP node as an L2TP Network Server (LNS). A given LCCE may act as both an LNS for some sessions and LAC for others, so these terms must only be used within the context of a given set of sessions unless the LCCE is in fact single purpose for a given topology.

3.25 L2TP pseudowire (PW): An emulated circuit as it traverses a packet-switched network. There is one Pseudowire per L2TP Session.

3.26 L2TP pseudowire type: The payload type being carried within an L2TP session. Examples include PPP, Ethernet, and Frame Relay.

3.27 L2TP session: An L2TP session is the entity that is created between two LCCEs in order to exchange parameters for and maintain an emulated L2 connection. Multiple sessions may be associated with a single Control Connection.

3.28 L2TP session ID: A 32-bit field containing a non-zero identifier for a session. L2TP sessions are named by identifiers that have local significance only. That is, the same logical session will be given different Session IDs by each end of the control connection for the life of the session. When the L2TP control connection is used for session establishment, session IDs are selected and exchanged as Local Session ID AVPs during the creation of a session. The Session ID alone provides the necessary context for all further packet processing, including the presence, size, and value of the Cookie, the type of L2-Specific Sublayer, and the type of payload being tunnelled.

3.29 MAC domain: A grouping of layer 2 devices that can communicate with each other without using bridging or routing. In DOCSIS, it is the group of CMs that are using upstream and downstream channels linked together through a MAC forwarding entity.

3.30 maximum transmission unit (MTU): The layer 3 payload of a layer 2 frame.

3.31 media access control (MAC): Used to refer to the layer 2 element of the system which would include DOCSIS framing and signalling.

3.32 modulation error ratio (MER): The ratio of the average symbol power to average error power.

3.33 physical media dependent (PMD) sublayer: A sublayer of the Physical layer which is concerned with transmitting bits or groups of bits over particular types of transmission link between open systems and which entails electrical, mechanical and handshaking procedures.

3.34 QAM channel (QAM ch): Analog RF channel that uses quadrature amplitude modulation (QAM) to convey information.

3.35 quadrature amplitude modulation (QAM): A modulation technique in which an analog signal's amplitude and phase vary to convey information, such as digital data.

3.36 radio frequency (RF): In cable television systems, this refers to electromagnetic signals in the range 5 to 1000 MHz.

3.37 radio frequency interface (RFI): Term encompassing the downstream and the upstream radio frequency interfaces.

3.38 request for comments (RFC): A technical policy document of the IETF; these documents can be accessed on the World Wide Web at http://www.rfc-editor.org/.

3.39 session: An L2TP data plane connection from the M-CMTS Core to the QAM channel. There must be one session per QAM Channel. There is one DEPI pseudowire type per session. There may be one MPT flow or one or more PSP flows per session. Multiple sessions may be bound to a single control connection.

3.40 StopCCN: L2TPv3 Stop-Control-Connection-Notification message.

3.41 upconverter: A device used to change the frequency range of an analog signal, usually converting from a local oscillator frequency to an RF transmission frequency.

3.42 upstream (US):

1) Transmissions from CM to CMTS. This includes transmission from the EQAM to M-CMTS Core as well as the RF transmissions from the CM to the EQAM.
2) RF spectrum used to transmit signals from a subscriber location to a cable operator's headend or hub site.

3.43 upstream channel descriptor (UCD): The MAC Management Message used to communicate the characteristics of the upstream physical layer to the cable modems.

3.44 video-on-demand (VoD) system: System that enables individuals to select and watch video content over a network through an interactive television system.

4 Abbreviations, acronyms and conventions

4.1 Abbreviations and acronyms

This Recommendation uses the following abbreviations:

ACK	L2TPv3 Explicit Acknowledgement message
AVP	L2TPv3 Attribute Value Pair
CDN	L2TPv3 Call-Disconnect-Notify message
CIN	Converged Interconnect Network
CLI	Command Line Interface
CM	Cable Modem
CMCI	Cable Modem CPE Interface
CMTS	Cable Modem Termination System
CPE	Customer Premises Equipment
CRC	Cyclic Redundancy Check
CRC-16	CRC of length 16
CSMA	Carrier Sense Multiple Access
DEPI	Downstream External PHY Interface
DOCSIS	Data-Over-Cable Service Interface Specifications
DOCSIS-MPT (D-MPT)	DOCSIS MPT Mode
DRFI	Downstream Radio Frequency Interface
DS	Downstream
DSCP	Differentiated Services Code Point
DTI	DOCSIS Timing Interface
DTS	DOCSIS Time Stamp, 32 bits
EQAM	Edge QAM
ERM	Edge Resource Manager
ERMI	Edge Resource Manager Interface
ETSI	European Telecommunications Standards Institute
FQDN	Fully Qualified Domain Name
GE	Gigabit Ethernet (GigE)
HELLO	L2TPv3 Hello message
HFC	Hybrid Fibre/Coax
ICCN	L2TPv3 Incoming-Call-Connected message

ICRP	L2TPv3 Incoming-Call-Reply message
ICRQ	L2TPv3 Incoming-Call-Request message
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv4	Internet Protocol version 4
ISO	International Organization for Standardization
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
L2TP	Layer 2 Transport Protocol
L2TPv3	Layer 2 Transport Protocol – Version 3
L3	Layer 3
LAC	L2TP Access Concentrator
LCCE	L2TP Control Connection Endpoint
LNS	L2TP Network Server
LSB	Least Significant Bit
M/N	Relationship of integer numbers M,N that represents the ratio of the downstream symbol clock rate to the DOCSIS master clock rate
MAC	Media Access Control
M-CMTS	Modular Cable Modem Termination System
MER	Modulation Error Ratio
MIB	Management Information Base
MPEG	Moving Picture Experts Group
MPEG-TS	Moving Picture Experts Group Transport Stream
MPT	MPEG-TS mode of DEPI
MPTS	Multi-Program Transport Stream
MSB	Most Significant Bit
MTU	Maximum Transmission Unit
OSSI	Operations Support System Interface
PCR	Program Clock Reference. A time stamp in the Video Transport Stream from which decoder timing is derived.
PHY	Physical Layer
PID	Packet Identifier; PID (system): A unique integer value used to identify elementary streams of a program in a single or multi-program Transport Stream as described in 2.4.3 of [H.222.0 | ISO 13818-1]
PMD	Physical Media Dependent Sublayer
PPP	Point-to-Point Protocol

PSI	Program Specific Information
PSP	Packet Streaming Protocol
PW	Pseudowire
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency
RFC	Request For Comments
RFI	Radio Frequency Interface
SCCCN	L2TPv3 Start-Control-Connection-Connected message
SCCRP	L2TPv3 Start-Control-Connection-Reply message
SCCRQ	L2TPv3 Start-Control-Connection-Request message
S-CDMA	Synchronous Code Division Multiple Access
SLI	L2TPv3 Set Link Info message
SPTS	Single Program Transport Stream
StopCCN	L2TPv3 Stop-Control-Connection-Notification message
TSID	MPEG2 Transport Stream ID
UCD	Upstream Channel Descriptor
UDP	User Datagram Protocol
US	Upstream
VoD	Video-on-Demand

4.2 Conventions

In this Recommendation, the following convention applies any time a bit field is displayed in a figure. The bit field should be interpreted by reading the figure from left to right, then from top to bottom, with the MSB being the first bit so read and the LSB being the last bit so read.

Throughout this Recommendation, the words that are used to define the significance of particular requirements are capitalized. These words are:

"MUST"	This word or the adjective "REQUIRED" means that the item is an absolute requirement of this Recommendation.
"MUST NOT"	This phrase means that the item is an absolute prohibition of this Recommendation.
"SHOULD"	This word or the adjective "RECOMMENDED" means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications should be understood and the case carefully weighed before choosing a different course.
"SHOULD NOT"	This phrase means that there may exist valid reasons in particular circumstances when the listed behaviour is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behaviour described with this label.
"MAY"	This word or the adjective "OPTIONAL" means that this item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because it enhances the product; for example, another vendor may omit the same item.

5 Technical overview

This clause is informative.

5.1 System architecture

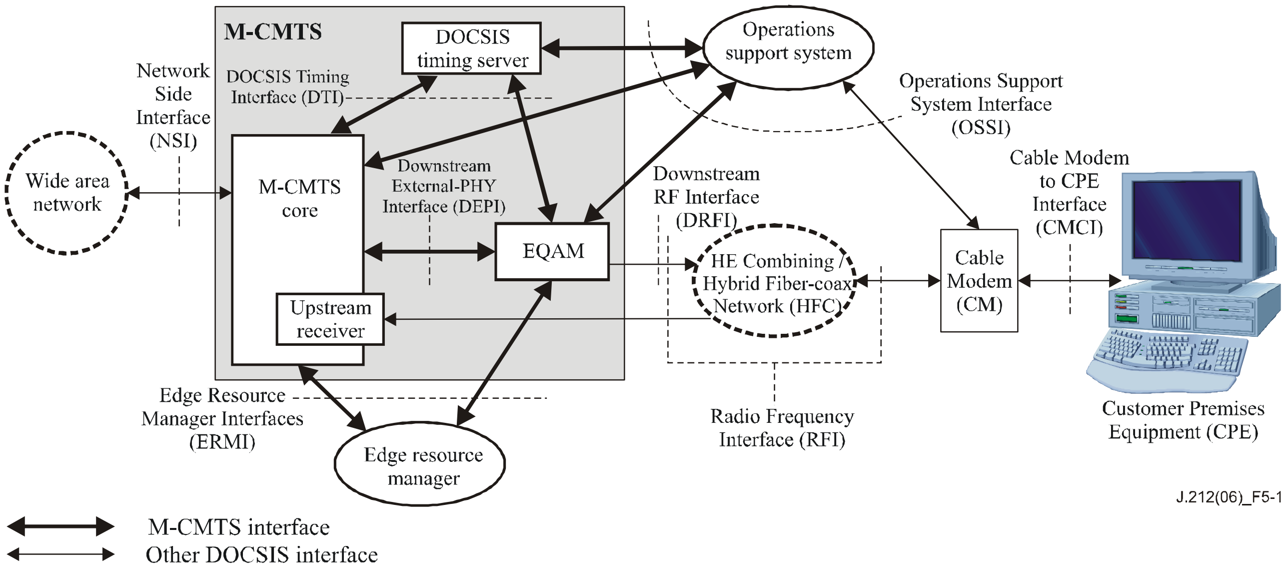

Figure 5-1/J.212 – Modular CMTS reference architecture

In the M-CMTS architecture, a device referred to as the M-CMTS Core contains the DOCSIS MAC. This includes all signalling functions, downstream bandwidth scheduling, and DOCSIS framing. The EQAM box contains mainly PHY related circuitry, such as QAM modulators, and tunnelling logic, to connect to the M-CMTS Core.

5.1.1 Reference architecture

The reference architecture for a Modular CMTS system is shown in Figure 5-1. This architecture contains several pieces of equipment along with interfaces between those pieces of equipment. This clause will briefly introduce each device and interface.

The Edge QAM device, or EQAM for short, has its origins in the VoD environment. This device typically has one or more gigabit Ethernet interfaces and multiple QAM modulators and RF upconverters. This EQAM is being adapted for use in a Modular CMTS environment. Each individual output of these devices is often referred to as QAM Channel rather than the combination "QAM Modulator and RF Upconverter".

The M-CMTS Core contains everything a traditional CMTS does, except for functions performed in the EQAM. The M-CMTS Core contains the downstream MAC and all the initialization and operational DOCSIS related software.

This diagram shows the Upstream Receivers of DOCSIS upstream channels located internally to the M-CMTS Core. However, there is nothing preventing an implementation of a Modular CMTS from using a proprietary external upstream receiver interface. In the future, an interface may be defined between the M-CMTS Core and external upstream receivers.

The DOCSIS Timing Interface (DTI) Server provides a common master clock frequency and a DOCSIS timestamp to other M-CMTS elements.

DEPI, the Downstream External PHY Interface, is the interface between the M-CMTS Core and the EQAM. More specifically, it is an IP Tunnel between the MAC and PHY in a Modular CMTS system which contains both a data path for DOCSIS frames, and a control path for setting up, maintaining, and tearing down sessions.

DRFI, or Downstream Radio Frequency Interface, is intended to capture all the current and future RF requirements for the downstream direction for both integrated DOCSIS CMTS systems, Modular DOCSIS CMTS systems, and VoD EQAM systems.

DTI, or DOCSIS Timing Interface, is a point-to-point interface from the DTI Server to other M-CMTS elements. The DTI Recommendation [J.211] defines DTI Server and DTI Client behaviours and protocols. The DTI Server is the Timing Signal Generator while each M-CMTS Core and EQAM has a DTI Client. The DTI Server distributes a DOCSIS master clock and a DOCSIS timestamp over unshielded twisted pair (UTP). The DTI protocol automatically compensates for cable length and ensures that all M-CMTS elements have the same sense of time and frequency.

ERMI, or Edge Resource Manager Interface [ERMI], involves three interfaces: a registration interface between an EQAM and ERM (Edge Resource Manager), a control interface between an EQAM and an ERM, and a control interface between an M-CMTS Core and an ERM. The first interface is used to register and unregister EQAM resources (i.e., QAM channels) with an ERM. The second interface is used by an ERM to request QAM channel resources from an EQAM, and by an EQAM to deliver resources to an ERM. The third interface is used by the M-CMTS Core to request specific QAM channel resources from the ERM, and by the ERM to respond to such requests with the location of QAM channel resources.

MOSSI, or Modular CMTS Operations Support System Interface [MOSSI], provides the management interface to each system component. This interface is an extension of the OSSI defined in the DOCSIS Recommendations for monitoring a management of CMTS functions. This interface could be used, in place of an ERM and the ERMI, to statically configure and associate QAM channel resources with M-CMTS Cores. This interface allows for the modification of a QAM channel's physical layer parameter by either the M-CMTS Core or the EQAM, and provides a mechanism by which the operator can "lock" certain parameters at the EQAM so that they can only be modified there. This Recommendation defines the mechanism to communicate these parameter settings to the other side.

NSI, or the Network Side Interface, is unchanged, and is the physical interface the CMTS uses to connect to the backbone network. Today, this is typically 100 Mbit/s or 1 Gbit/s Ethernet.

CMCI, or Cable Modem to Customer Premises Equipment Interface, is also unchanged, and is typically 10/100 Mbit/s Ethernet or USB.

5.1.2 DEPI operation

DEPI is an IP Tunnel that exists between the DOCSIS MAC in the M-CMTS Core and the DOCSIS PHY that exists in the EQAM. DEPI's job is to take either formatted DOCSIS frames or MPEG packets and transport them through a layer 2 or layer 3 network and deliver them to the EQAM for transmission.

The base protocol that is used for the DEPI is the Layer 2 Tunnelling Protocol Version 3, or L2TPv3 for short [RFC-L2TPv3]. L2TPv3 is an IETF protocol that is a generic protocol for creating a pseudowire. A pseudowire is a mechanism to transparently transport a layer 2 protocol over a layer 3 network. Examples of protocols supported by L2TPv3 include ATM, HDLC, Ethernet, Frame Relay, PPP, etc.

Clause 8.1 "L2TPv3 Transport Packet Format" shows the format of an L2TPv3 data packet. Each data packet contains a 32-bit session ID which is associated with a single QAM Channel on the EQAM. The UDP header is optional in the L2TPv3 protocol. In the DEPI protocol, both the EQAM and the M-CMTS core are required to support the ability to include the UDP header in both the control plane and the forwarding plane (see 7.3). Support for DEPI operation without the

UDP header is optional for both the EQAM and M-CMTS core, and can be used if both ends support it, as described in clauses 7.3 and 8.

L2TPv3 then permits a sub-header to exist whose definition is specific to the payload being carried. The control channel allows for signalling messages to be sent between the M-CMTS Core and EQAM. Typical control messages will set up a "control connection" between the M-CMTS Core and EQAM, and then set up multiple data sessions (one for each downstream QAM channel). Each session can utilize different DiffServ Code Points (DSCPs), and support different encapsulation protocols.

There are two basic tunnelling techniques defined by DEPI. The first technique, known as D-MPT mode, transports multiple 188-byte MPEG-TS packets by placing them into the L2TPv3 payload with a unique sub-header which contains a sequence number so packet drops can be detected. The encapsulation of DOCSIS Frames into MPEG-TS packets is performed in the M-CMTS Core. The second technique, known as the Packet Streaming Protocol (PSP), transports DOCSIS Frames in the L2TPv3 payload. The DOCSIS Frames are then encapsulated in MPEG-TS packets within the EQAM. PSP mode allows DOCSIS frames to be both concatenated, to increase network performance, and fragmented, in case the tunnelled packets exceed the network MTU size.

One of the technical considerations of the Modular CMTS architecture is its impact on the round trip request-grant delay time. The request-grant delay time is the time from when a CM requests bandwidth using an uncontended bandwidth request (REQ) to when it receives a MAP message with the granted transmit opportunity in it.

To prevent the MAP from being slowed down by other traffic in the CIN, the DOCSIS traffic (or a subset containing the MAP messages) may be sent in an independent L2TPv3 flow that has a unique DSCP. This DSCP will have a "per hop behaviour (PHB)" that will give MAPs the highest priority and lowest latency service.

5.1.3 EQAM Operation

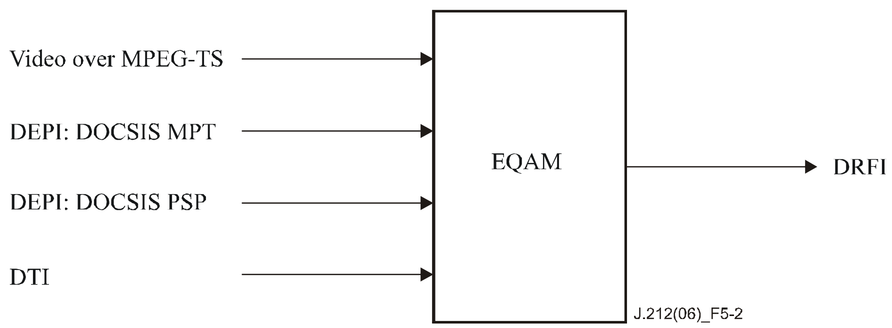

Figure 5-2/J.212 – EQAM block diagram

Figure 5-2 shows a high level block diagram of an EQAM that is capable of handling either Video MPEG traffic or DOCSIS traffic. The expression "D-MPT" is an acronym for DOCSIS MPEG Transport.

The first interface that is shown is the VoD transport. VoD SPTS or MPTS streams are received with a format of MPEG packets over UDP/IP. The video processing functions generally include de-jittering, re-multiplexing PID remapping, MPEG-2 PSI insertion, and PCR timestamp correction. These functions are not defined in this Recommendation.

The next set of interfaces are the DEPI interfaces.

The first interface defined is D-MPT. This is a mode where the EQAM must search the incoming D-MPT frames for DOCSIS SYNC messages and correct the timestamp value in these messages based upon the EQAM's internal DOCSIS timestamp which has been derived from DTI. The resulting D-MPT frames are then copied to the QAM channel without further interpretation or modification. This mode is intended for DOCSIS frames where the MAP is embedded into the stream and network latency or jitter is not a concern.

Because D-MPT mode encapsulates all DOCSIS traffic into a single DEPI flow, it does not allow for QoS differentiation among various types of traffic either across the CIN or within the EQAM. For example, acceleration of MAPs relative to other DOCSIS data is not possible when D-MPT mode is used. This may still give acceptable performance when the delay and jitter introduced onto DEPI packets by the CIN and EQAM are determined by the operator to be sufficiently low.

Certain network conditions and/or architectures may reduce network delay and/or jitter and make it more likely that DOCSIS MPT mode will give acceptable system performance. Some examples are:

– networks with a very small number of hops (e.g., 1 or 2 hops), especially if these hops are all switches (which generally introduce less delay and jitter than routers);
– networks which are largely shared with video traffic from VoD servers to EQAMs, on which most of the traffic is video and it is acceptable to prioritize all DEPI traffic on the network above all VoD traffic;
– networks which are lightly loaded so that delays due to congestion are highly unlikely.

Such conditions may exist on today's CINs. They may be less likely to occur as deployments become denser and total IP traffic increases. Evaluation of network conditions and decisions about acceptable performance levels with DOCSIS MPT mode are in the realm of operator discretion.

The next interface is DOCSIS PSP. This interface transports DOCSIS Data and MAPs in separate flows which are streamed together in a uniform byte stream by the M-CMTS Core. The PSP re-assembly engine removes this overhead and recovers the DOCSIS Frames. The PSP scheduler then allows MAPs to be placed in order, ahead of data and SYNC messages. In PSP mode, the EQAM generates all SYNC messages as detailed in 6.1.1. The output is then delivered to a transmission convergence layer which converts the results to a DOCSIS MPEG stream.

The last interface is the DTI interface which provides a common frequency and DOCSIS timestamp. The reference is used to synchronize the downstream symbol rate and DOCSIS timestamp for use with DOCSIS cable modems. The timestamp is used for the DOCSIS SYNC correction.

The output from the EQAM consists of a stream of MPEG frames which carry video and/or DOCSIS data, and are modulated onto an RF carrier in accordance with the DRFI Recommendation [J.210].

5.2 Bonding services model

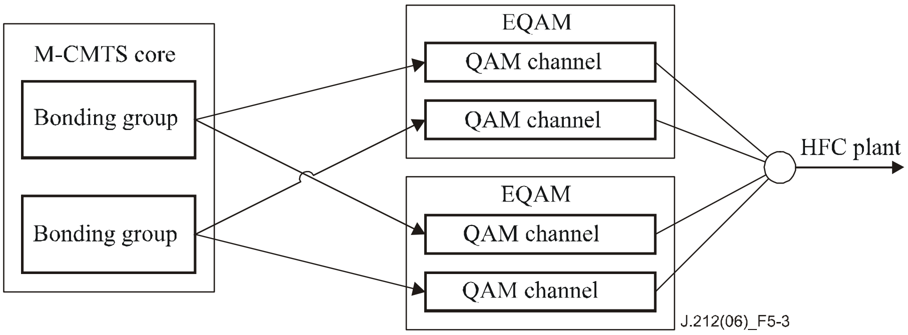

Figure 5-3/J.212 – Bonding services model

Downstream Channel Bonding refers to the forwarding of a stream of DOCSIS frames across multiple QAM carriers.

In the Modular CMTS architecture, the Downstream Channel Bonding is implemented in the M-CMTS Core. At the M-CMTS Core, packets from the IP backbone are placed in a DOCSIS frame. That DOCSIS frame is then sent to one of several QAM channels in a Bonding Group. The frame may be transported using D-MPT or PSP.

In this system, the EQAM is unaware that it is doing bonding. It is also unaware of any of the details of the Bonding protocol.

5.3 Multiple services model

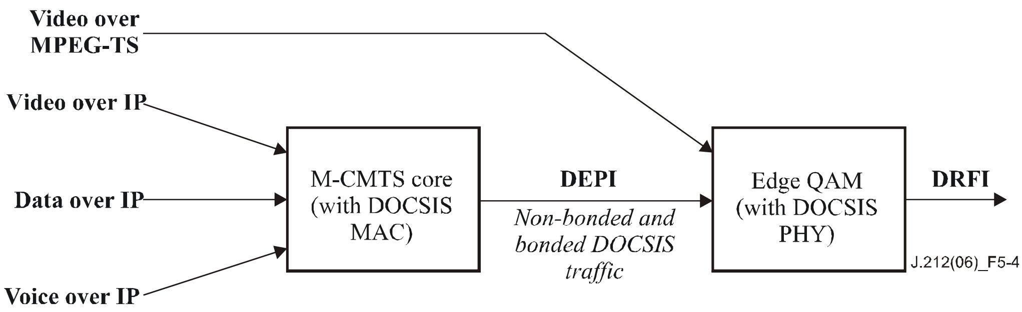

Figure 5-4/J.212 – Multi-service mode

The Modular CMTS (M-CMTS) architecture reflects an effort to merge the video-on-demand (VoD) and high speed data (HSD) applications. This effort is made in anticipation of achieving higher efficiencies than are possible with two separate transmission networks feeding into the cable plant. In particular, the M-CMTS effort has repartitioned the traditional CMTS architecture so that the transmission technology which is common to both the VoD environment and the HSD environment will be able to share the same EQAM devices.

In a video delivery system, the EQAM is used to deliver video streams in MPEG-TS format to set top boxes at the subscribers' premises. This functionality will continue to be used in the future and operates independently of other processing which is described here. The M-CMTS architecture adds new interfaces which are unique to HSD service. These interfaces support both traditional DOCSIS and multichannel (bonded) DOCSIS payloads as received from the M-CMTS Core device.

The M-CMTS Core provides the gateway functionality between the IP-based core network and the CIN. As such, the M-CMTS Core provides support for a multitude of services – including, but not limited to video over IP, voice-over-IP (VoIP), email, gaming, video telephony, etc.

6 DEPI architecture

This clause is normative.

6.1 DEPI Data Path

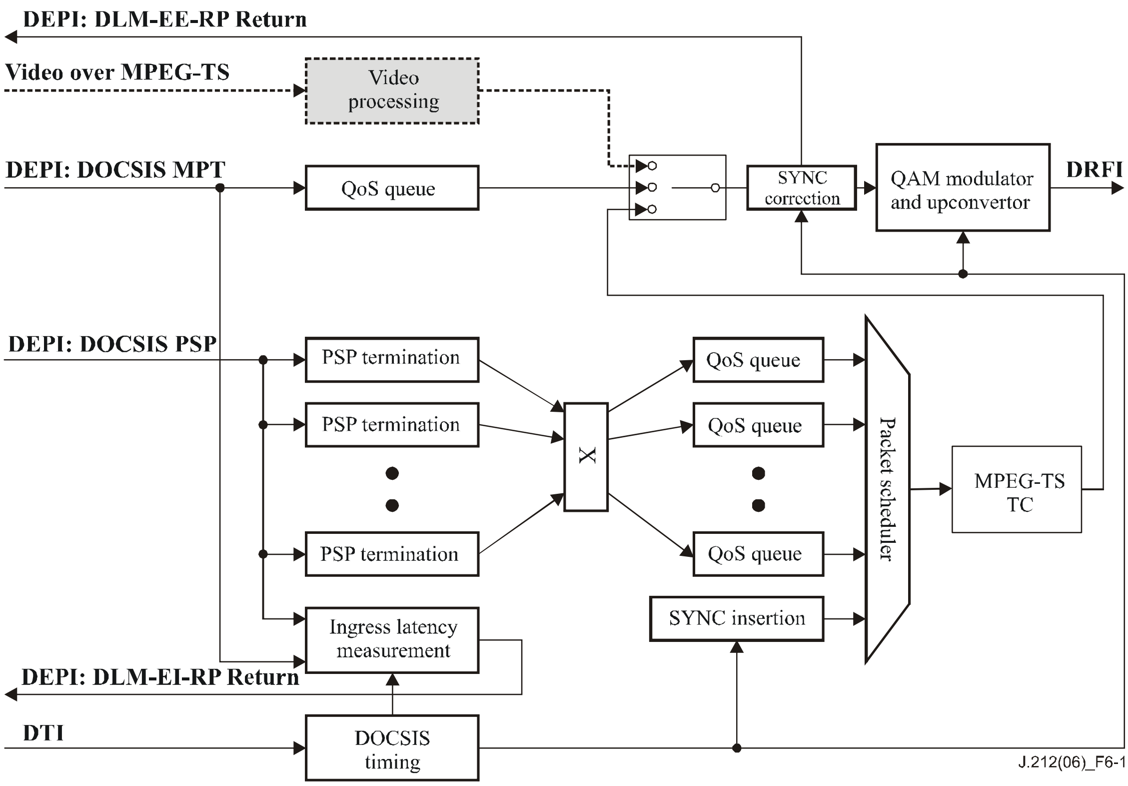

Figure 6-1/J.212 – Downstream EQAM block diagram

A simplified logical block diagram of the internal data path of the EQAM is shown in Figure 6-1.

It is recognized, but not specified by this Recommendation, that the EQAM may receive non-DOCSIS MPEG elementary streams which have been encapsulated in MPEG packets and placed in a UDP datagram. This Recommendation does not define requirements for transport of this type of traditional MPEG video. However, it is recognized that M-CMTS EQAM implementations may (and likely will) be capable of transporting traditional MPEG video (either interleaved with DOCSIS traffic on a single QAM channel, or on separate QAM channels within the same EQAM chassis).

The M-CMTS Core MUST support PSP mode, D-MPT mode, or both. The EQAM MUST support PSP mode, D-MPT mode, or both.

Within a session, the M-CMTS Core MUST support either one priority level of D-MPT or at least two priority levels of PSP, where each priority level would have a different DSCP. The M-CMTS Core MUST provide a mechanism to map DOCSIS traffic to the multiple priority levels of PSP. The M-CMTS Core MUST NOT attempt to establish a session which includes both PSP and D-MPT flows.

Within a session, the EQAM MUST support either one priority level of D-MPT or at least two priority levels of PSP, where each priority level would have a different DSCP. The EQAM is not intended to simultaneously support D-MPT and PSP within a single session, and MUST reject any attempt to establish such a session. Each priority level for each DEPI type, maps to one or more DEPI flows. The EQAM MUST support the establishment of a single DEPI flow per priority level. The EQAM MAY support the establishment of more than one DEPI flow per priority level. The mapping of individual flows to priority levels (QoS Queues) is vendor-specific (see Figure 7-3). The mapping of flows will be done through the local EQAM Command Line Interface (CLI) configuration.

For both D-MPT and PSP modes, the EQAM MUST insert null MPEG packets when it has no data to send. The EQAM SHOULD NOT insert a null MPEG packet if it has data to send. Note that MPEG null insertion should take place prior to the correction of the DOCSIS SYNC message (refer to 6.1.1).

6.1.1 DOCSIS D-MPT data path

The DEPI DOCSIS D-MPT flows contain DOCSIS frames using the format as described in 8.2. All DOCSIS frames, including packet based frames and MAC management based frames, are included within the one D-MPT flow. The EQAM searches the D-MPT payload for any DOCSIS SYNC messages and performs SYNC corrections as described in 6.1.3.2. It then forwards the D-MPT packet to the RF interface.

The intent of D-MPT mode is that MPEG packets can be received by the EQAM and forwarded directly to the RF interface without having to terminate and regenerate the MPEG framing. The only manipulation of the D-MPT payload is the SYNC correction.

6.1.2 PSP data path

The Packet Stream Protocol (PSP) is a layer-3 convergence layer protocol, which allows packets to be consecutively streamed together and fragmented at arbitrary boundaries. The intent of the PSP mode is to facilitate Quality of Service. This mode is to be used for transporting traditional DOCSIS data and signalling messages which use one or more DSCP values. For example, in order to reduce Request-Grant latency, MAP MAC management messages may be sent using a different DSCP on a different PSP flow than the rest of the DOCSIS channel. Refer to 6.2.1 for more information. The EQAM MUST support a minimum of two PSP receivers per QAM modulator. The intent of two receivers is to permit the implementation of a higher latency PSP flow and a lower latency PSP flow.

Each PSP flow is terminated, and the DOCSIS Frames within the flow are extracted. The DOCSIS frames are placed into corresponding output QoS queues. The output of the QoS queues go to a Packet Scheduler, which decides which queue is to be serviced based upon the PHB (negotiated between the M-CMTS Core and EQAM) of the PSP flow, which carried the DOCSIS frames. The Packet Scheduler is also responsible for inserting DOCSIS SYNC messages within the time interval specified by the DOCSIS SYNC Control AVP (see Figure 7-31). The Packet Scheduler SHOULD support a strict priority scheduler. The Packet Scheduler MAY support other queue scheduling disciplines.

The phrase "packet scheduler" is a general term that describes a method of applying priorities to different queues as packets are moved from different input queues to the output queue. An example of a typical Packet Scheduling algorithm would be Weighted Fair Queuing (WFQ) where some streams are given priority over other streams, but only up to some limit. This should not be confused with the more complex DOCSIS upstream scheduler, which deals with requests and grants.

The output of the Packet Scheduler goes to a Transmission Convergence engine which places the DOCSIS frames into MPEG packets according to the requirements in [J.210]. This includes the insertion of stuffing bytes and the DOCSIS SYNC message as described by 6.1.3. The output of Transmission Convergence (TC) engine is sent to the RF Interface.

PSP mode primarily provides acceleration of MAPs through the network, in an attempt to reduce Request Grant latency. The PSP mode is most useful when all or most traffic has migrated to DOCSIS, and therefore, marking MAPs with higher QoS, when compared to other DOCSIS traffic that is needed in order to provide lower latency for MAPs, when traversing a fully subscribed network. Therefore, PSP has value in the long term and is included to address the case where most or all traffic is carried to the home via DOCSIS.

6.1.3 DOCSIS SYNC message

6.1.3.1 SYNC message format

The DOCSIS Time Synchronization (SYNC) MAC message is transmitted by a Modular CMTS system at a periodic interval to establish MAC sublayer timing in cable modems. This message MUST use an FC field with FC_TYPE = MAC Specific Header and FC_PARM = Timing MAC Header. This MUST be followed by a Packet PDU in the format shown in Figure 6-2.

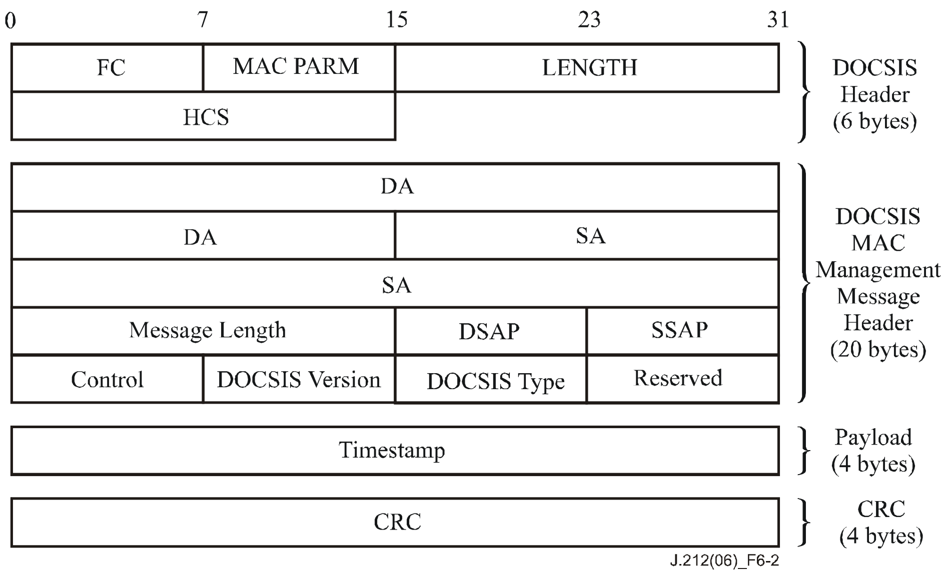

Figure 6-2/J.212 – Format of a DOCSIS SYNC MAC message

The fields MUST be as defined below:

FC, MAC PARM, LEN, HCS: Common MAC frame header with FC_PARM field to indicate a timing header – refer to J.122 for details.

Destination Address (DA): Set to the DOCSIS MAC multicast address of 01-E0-2F-00-00-01.

Source Address (SA): The MAC address of the M-CMTS Core. In PSP Mode, the EQAM learns the appropriate M-CMTS Core MAC address via explicit signalling during L2TPv3 session setup.

Msg Length: Length of the MAC message from DSAP to the end of the payload.

DSAP: The LLC null destination SAP (00) as defined by [ISO 8802-2].

SSAP: The LLC null source SAP (00) as defined by [ISO 8802-2].

Control: Unnumbered information frame (03) as defined by [ISO 8802-2].

DOCSIS Version: Set to 1.

DOCSIS Type: Set to 1 to indicate a SYNC message.

CMTS Timestamp: The count state of an incrementing 32-bit binary counter clocked with the master clock derived from [J.211].

The CMTS timestamp represents the count state at the instant that the first byte (or a fixed time offset from the first byte) of the Time Synchronization MAC Management Message is transferred from the Downstream Transmission Convergence Sublayer, to the Downstream Physical Media Dependent Sublayer, as described in [J.210].

6.1.3.2 Correction and insertion

The EQAM MUST derive a local DOCSIS timestamp from the DTI Client, specified in [J.211]. The EQAM MUST support either a 10.24 MHz master clock frequency or a 9.216 MHz master clock frequency depending on the region in which it will operate. In the D-MPT mode, the EQAM must be capable of correcting all embedded SYNC messages in the DOCSIS stream. For the DOCSIS PSP mode, the EQAM MUST be capable of inserting DOCSIS SYNC messages based upon its internal timestamp into the downstream MPEG-TS stream within the specifications in 6.1.3.3. In PSP Mode, the EQAM MUST insert the SYNC message starting at the 6th byte of the MPEG-TS frame (the fifth byte would be the MPEG pointer_field). When referencing a timebase to use for SYNC timestamp insertion or correction, the EQAM MUST utilize a timebase that is delayed by no less than 0 and no more than 100 clock cycles (approximately 10 µs) of the master clock compared to the time communicated via the DTI Client Test Port output.

This time difference between the time communicated via the DTI Client Test Port output and the time reference applied by an EQAM is expected to be essentially constant. All timing and jitter requirements that exist in this Recommendation and [J.210] still apply. The requirements here do not preclude an EQAM from taking longer than 100 clock cycles to process the DTI conveyed time, but in that case, the EQAM would need to internally adjust the applied timebase such that it is delayed between zero and 100 clock cycles from the DTI conveyed timebase.

When using the D-MPT mode, the M-CMTS Core MUST generate SYNC messages and include them in the D-MPT payload. The M-CMTS Core MUST insert the SYNC message starting at the 6th byte of the MPEG-TS frame (the fifth byte would be the MPEG pointer_field). This is intended to simplify the EQAM implementation by allowing the EQAM to check only the payload_unit_start_indicator bit and the fifth and sixth byte (which together will contain 0x00C0) of the MPEG-TS packet to locate a DOCSIS SYNC message. Note that the CMTS Timestamp in the SYNC messages generated by the M-CMTS core is not required to accurately reflect the current timestamp received via a DTI Client in the M-CMTS Core. For example, it is permissible for the M-CMTS Core to use a value of zero for the CMTS Timestamp in all SYNC messages. When using the DOCSIS PSP mode, the M-CMTS Core MUST NOT generate SYNC messages as part of the PSP payload.

6.1.3.3 Timestamp jitter

The DOCSIS timestamp jitter MUST be less than 500 ns peak-to-peak at the output of the Downstream Transmission Convergence Sublayer. This jitter is relative to an ideal Downstream Transmission Convergence Sublayer that transfers the MPEG packet data to the Downstream Physical Media Dependent Sublayer with a perfectly continuous and smooth clock at the MPEG packet data rate. Downstream Physical Media Dependent Sublayer processing MUST NOT be considered in timestamp generation and transfer to the Downstream Physical Media Dependent Sublayer.

Thus, any two timestamps N1 and N2 (N2 > N1) which were transferred to the Downstream Physical Media Dependent Sublayer at times T1 and T2 respectively must satisfy the following relationship:

$$| (N2 - N1)/f_{CMTS} - (T2 - T1) | < 500 \times 10^{-9}$$

In the equation, the value of (N2 – N1) is assumed to account for the effect of rollover of the timebase counter, and T1 and T2 represent time in seconds. f_{CMTS} is the actual frequency of the CMTS master timebase and may include a fixed frequency offset from the nominal master clock frequency. This frequency offset is bounded by a requirement further below in this clause.

The jitter includes inaccuracy in timestamp value and the jitter in all clocks. The 500 ns allocated for jitter at the Downstream Transmission Convergence Sublayer output MUST be reduced by any jitter that is introduced by the Downstream Physical Media Dependent Sublayer.

NOTE – Jitter is the error (i.e., measured) relative to the CMTS master clock.

6.1.4 Latency and skew requirements

6.1.4.1 Latency

For PSP DEPI sessions, latency is defined as the absolute difference in time from when the last bit of a DEPI packet containing the last bit of a single DOCSIS MAC frame enters the EQAM DEPI port to the time that the first bit of the DOCSIS MAC frame exits the EQAM RFI port. For D-MPT DEPI sessions, latency is defined as the absolute difference in time from when the last bit of a DEPI packet enters the EQAM DEPI port to the time that the first bit of the first MPEG packet contained within said DEPI packet exits the EQAM RFI port. At the EQAM input, the last bit of the arriving DEPI packet is used because the EQAM's layer 2 interface (e.g., GigE Ethernet port) must receive the entire packet before the EQAM can begin processing. At the EQAM output, the first bit of the first MPEG packet (in D-MPT mode) or DOCSIS frame (in PSP mode) inside the DEPI packet is used to guarantee that the data complies with the definition of "isolated packets" (see next paragraph). If this were not done, the measurement could be corrupted by delays incurred due to queuing of data behind other packets destined for the same RF interface. The EQAM SHOULD allow enough buffer memory per QAM channel to buffer at least 20 ms worth of traffic across all L2TPv3 sessions destined for that QAM channel.

In PSP DEPI sessions, the multiple flows supported in the EQAM provide for prioritized access to the modulator. In the absence of higher priority traffic, and regardless of load of lower priority traffic, the EQAM MUST forward isolated packets in each DEPI flow with a latency of less than 500 µs plus the delay of the interleaver. Isolated packets are spaced such that, at the nominal downstream data rate, the EQAM would complete transmission of the current packet before the arrival of the next packet.

6.1.4.2 Skew

Skew is defined as the difference between the maximum latency and the minimum latency through the EQAM, as measured from two reference bits on the network interface to the same bits on two separate RF outputs. Skew is to be measured with the PHY parameters set equal on the QAM Channels under measurement.

The skew between any two bonded QAM Channels within an EQAM MUST be less than 500 µs. The skew requirements for the EQAM are implicitly met when the EQAM meets the latency requirements set out above in 6.1.4.1. This requirement is intended for the transmission of skew sensitive traffic such as bonded traffic.

6.2 Networking considerations

6.2.1 Per hop behaviour usage

The Per Hop Behaviour Identifier (PHBID) is used by network devices to signal the correct Per Hop Behaviour (PHB) between one another. PHB permits Expedited Forwarding (EF) as described in [RFC-PHBID-EF], Assured Forwarding (AF) as described in [RFC-PHBID-AF], or best effort forwarding as described in [RFC-PHBID-AF] to be used. The M-CMTS Core MUST support the Expedited Forwarding PHBID for PSP mode. The M-CMTS Core MUST support the best effort

PHBID for both DEPI modes. The EQAM MUST support the EF PHBID for PSP mode. The EQAM MUST support the best effort PHBID for both DEPI modes.

The DEPI interface supports multiple traffic types including DOCSIS MAC and DOCSIS data traffic. Within both traffic types, there may be different levels of priority. For PSP operation, the M-CMTS Core SHOULD provide a mechanism to map traffic of different priorities to DEPI flows with different PHB values. The M-CMTS Core SHOULD NOT use the same PHB across multiple DEPI flows within a session.

The CIN should provide the appropriate Per Hop Behaviour for the differentiated traffic types. The level of granularity provided for differentiated traffic is determined by the network operator, but at a minimum, it would be expected that DOCSIS MAP messages and VoIP data traffic are prioritized over best effort data traffic.

The EQAM uses the PHB signalled in the establishment of the DEPI flow when scheduling multiple DEPI PSP flows onto one QAM Channel as described in 6.1.2.

6.2.2 DiffServ code point usage

The DSCP is a value located in the 6-bit DiffServ Field of the IP Header. The DSCP of the L2TPv3 packet SHOULD be assigned at the egress of the M-CMTS Core to provide Quality of Service for DEPI traffic within the CIN. The DSCP MAY be used at the ingress of the EQAM.

The M-CMTS Core MUST tag all packets within a DEPI flow of an L2TPv3 session with the same DSCP.

DOCSIS frames encapsulated in L2TPv3 packets may contain IP packets which also have a DSCP assigned. The EQAM is not required to schedule packets based upon the original DSCP contained within the DOCSIS frame.

6.2.3 Packet sequencing

For a stream of packets transmitted on a DEPI flow, the packet sequence number will be incremented by one for each packet sent, as described in 8.2 and 8.3. If the EQAM detects a discontinuity in the packet sequence numbers indicating that one or more packets were dropped or delayed, an error is logged and the EQAM SHOULD transfer the current packet to the QAM Channel without waiting for the missing packets. If the EQAM detects a discontinuity in the packet sequence numbers indicating that one or more packets have arrived late, then those packets SHOULD be discarded. The EQAM MUST NOT forward packets which were skipped due to a discontinuity in the sequence numbers. Storing and re-ordering of packets so that they can be delivered to the QAM Channel in the correct sequence is not prohibited by these requirements and the EQAM MAY perform such re-ordering as long as the latency requirements of clause 6.1.4 are met.

6.2.4 Network MTU

The network between the M-CMTS Core and the EQAM will have a certain Maximum Transmission Unit. If a maximum size DOCSIS frame were to be tunnelled from the M-CMTS Core to the EQAM without fragmentation, the size of the resulting packet may be greater than the CIN can handle. Both the D-MPT and PSP modes avoid this issue by offering streaming and fragmentation. As such, IP fragmentation is not required. IP fragmentation is also undesirable because the EQAM may forward packets based upon the destination UDP port, and the UDP port is only available in the first IP fragment.

Determining the MTU to use for the L2TPv3 tunnel between the M-CMTS Core and the EQAM is a two-step process. The first step is done as part of L2TPv3 session establishment (see clause 7) using the DEPI MTU AVPs. When the M-CMTS Core starts the session establishment with an ICRQ message, it MUST supply the DEPI Local MTU AVP with a payload size that is the lesser of its receive capabilities and the receive capabilities defined by its lower layer. The receive capabilities of the M-CMTS Core are defined by its internal constraints, and any configured maximums. The receive capabilities defined by its lower layer is a calculation based on referencing the payload size constraints of the interface below which this tunnel is being created, as defined in A.1. The M-CMTS Core MUST support an MTU size of at least 1500 bytes. The EQAM MUST send L2TPv3 frames with a payload size less than or equal to this maximum. If the EQAM cannot meet this criterion, then it MUST fail session creation by generating a CDN message. The EQAM needs to consider the same criteria in calculating its MTU. The EQAM MUST support an MTU size of at least 1500 bytes, as calculated in A.1. The EQAM MUST insert the DEPI Remote MTU AVP in the ICRP message with its MTU size. The M-CMTS Core MUST send L2TPv3 frames with a payload size less than or equal to this maximum. If the M-CMTS Core cannot meet this criterion, then it MUST fail session creation by generating a CDN message.

The second step is to determine the MTU of the path between the M-CMTS Core and the EQAM. The M-CMTS Core MUST support a mechanism to prevent sending packets larger than the network MTU. This SHOULD be done using Path MTU Discovery, as described in [RFC-MTU]. Clause A.3 gives a brief overview of the Path MTU discovery protocol. Alternatively, this MAY be done via a static configuration option. Both the M-CMTS Core and the EQAM MUST have a way to statically configure an MTU for each L2TPv3 session. To avoid IP fragmentation, the M-CMTS Core and the EQAM MUST set the Don't Fragment bit (DF) in the IPv4 header for all transmissions into the L2TPv3 pseudowire.

6.3 System timing considerations

To ensure proper system operation, the M-CMTS Core SHOULD utilize a timebase that is delayed by no less than 0 and no more than 100 clock cycles (approximately 10 µs) of the master clock compared to the time communicated via the DTI Client Test Port output. The M-CMTS Core MUST support either a 10.24 MHz master clock frequency or a 9.216 MHz master clock frequency depending on the region in which it will operate.

This time difference between the time communicated via the DTI Client Test Port output and the time reference applied by an M-CMTS Core is expected to be essentially constant. The requirements here do not preclude an M-CMTS Core from taking longer than 100 clock cycles to process the DTI conveyed time, but in that case, the M-CMTS Core would need to internally adjust the applied timebase such that it is delayed between zero and 100 clock cycles from the DTI conveyed timebase.

7 DEPI control plane

NOTE – This clause is normative.

The DEPI control plane is based upon L2TPv3 signalling. The intent of this Recommendation is to follow the provisions of [RFC-L2TPv3]. This clause includes some examples of how L2TPv3 signalling is used, and includes the extensions and interpretations of the [RFC-L2TPv3] specification as it applies to DOCSIS.

All requirements from [RFC-L2TPv3] MUST be met by both the M-CMTS Core and the EQAM unless this Recommendation explicitly states that a particular requirement from [RFC-L2TPv3] is not required.

7.1 Topology

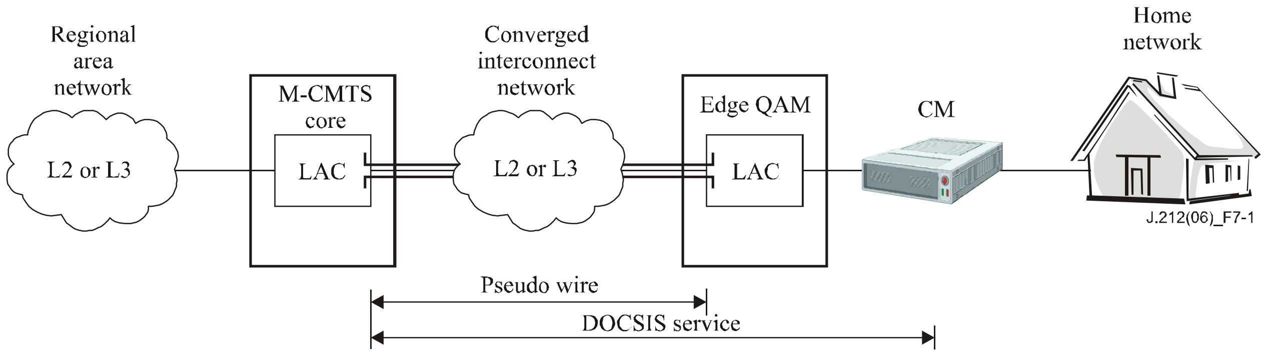

Figure 7-1/J.212 – L2TP topology for modular CMTS

Figure 7-1 shows how the Modular CMTS architecture maps to L2TP topology. In L2TPv3 nomenclature, the M-CMTS Core and the EQAM are known as a L2TP Access Concentrator (LAC). The M-CMTS Core and the EQAM are considered as peers and may also be referred to as L2TP nodes or as L2TP Control Connection Endpoints (LCCE). For the purposes defined in this Recommendation, each LCCE is identified on the network by a single IP address. The connections between two LCCEs are known as Pseudo wires (PW).

The L2TP supports both a data path and an inband control path. In L2TPv3 nomenclature, data messages are sent on the data channel, and control messages are sent on the control connection.

First a control connection is established between the two LCCEs, then a session is established. An L2TP session is established before L2TP begins to forward session frames for data. Multiple sessions may be bound to a single control connection.

7.2 Addressing

The M-CMTS Core SHOULD use the IP Address of the EQAM and the TSID of the QAM Channel to uniquely identify a QAM Channel within an EQAM during initial configuration.

The M-CMTS Core MUST establish at most one control connection for each LCCE pair. This control connection will manage all sessions between the M-CMTS Core and the EQAM. If the UDP Header is used, the control connection SHOULD use UDP port 1701.

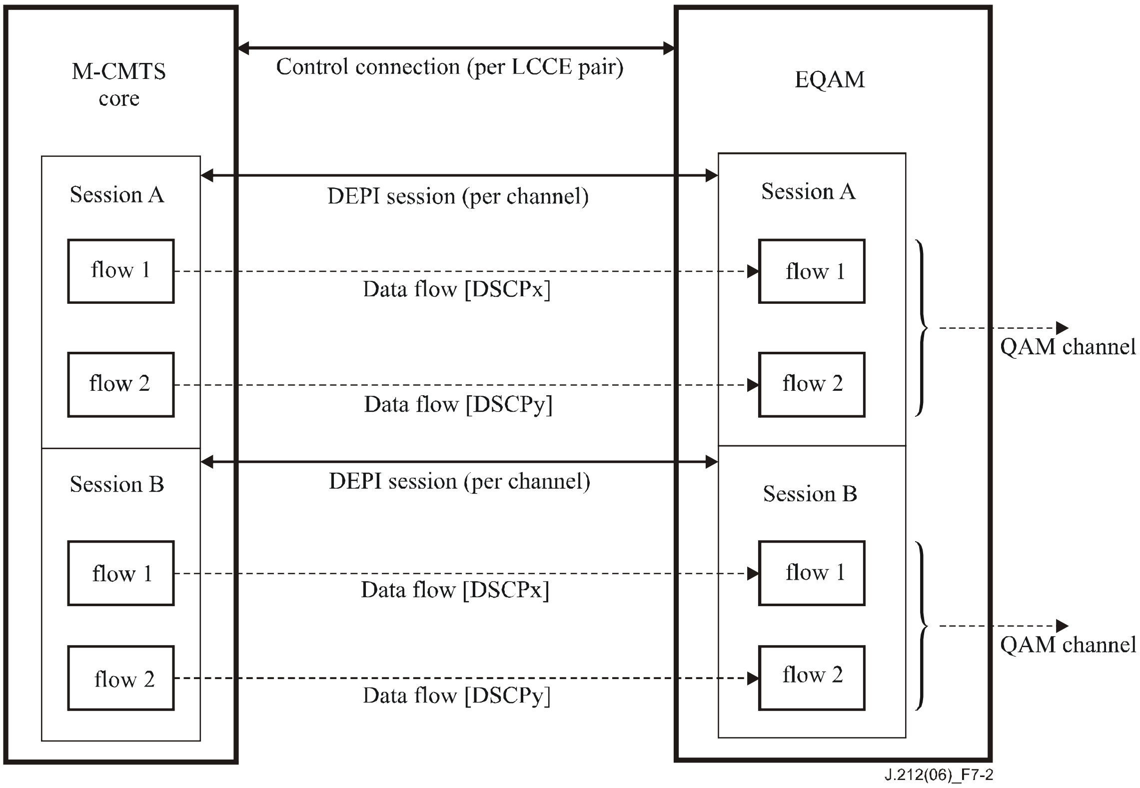

Figure 7-2/J.212 – DEPI addressing hierarchy

The M-CMTS Core MUST support the creation of a single session per QAM channel. The EQAM MUST support a single session per QAM channel. Each DEPI session uses one of the Pseudowire types described in 7.5.1.1. Per [RFC-L2TPv3], both the M-CMTS Core and the EQAM assign a unique L2TPv3 session ID for each session. The M-CMTS Core MUST NOT attempt to create a session to a QAM channel for which the M-CMTS Core already has an active session. Unless specifically configured otherwise, the EQAM MUST reject an attempt to establish a session to a QAM channel for which a session has already been established.

The M-CMTS Core MAY create multiple PSP flows per session. Different EQAM implementations may support different numbers of PSP flows on any given DEPI session, which is described further in 6.1. During L2TPv3 session setup, the EQAM MUST assign each flow a unique flow ID. Reassembly, if applicable, is done per flow ID at the EQAM. The EQAM MAY assign each flow a unique UDP Destination Port. The M-CMTS Core MUST address DEPI packets using the UDP Destination Port (if the UDP header is used), L2TPv3 Session ID, and flow ID assigned by the EQAM. This is shown in Figure 7-2 and in more detail in Figure 7-3.

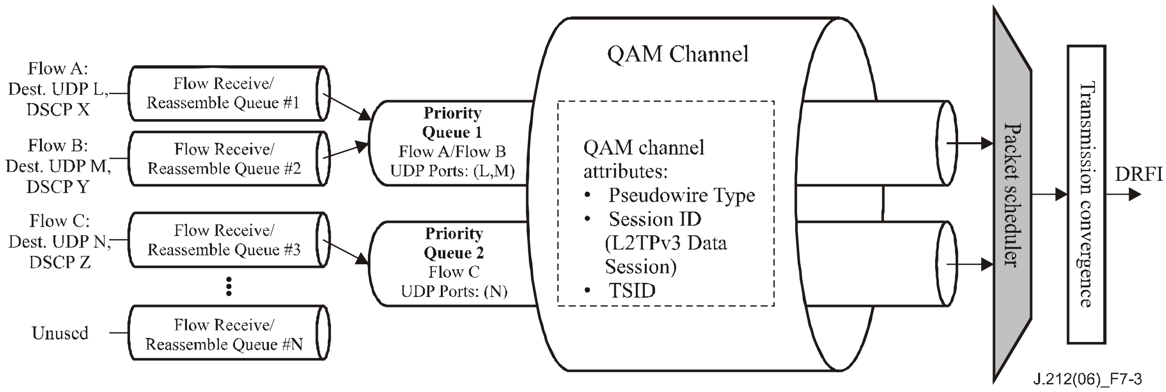

Figure 7-3/J.212 – DEPI addressing hierarchy

7.3 Control message format

The format of a DEPI Control Packet, as shown in Figures 7-4 and 7-5, is based on [RFC-L2TPv3], with extensions for DOCSIS. The fields which are common with the DEPI Data Packet are described in 8.1. Fields which are used differently or are new are described below. Unless otherwise indicated, all values are shown in decimal notation.

The choice of using or not using a UDP header is done through system configuration and is not a negotiated DEPI parameter. The UDP version of DEPI is intended for systems which use the UDP port to map flows to a QAM Channel within an EQAM. The non-UDP version of DEPI is intended for systems which use the L2TPv3 Session ID to map flows to a QAM Channel within an EQAM.

The M-CMTS Core MUST support DEPI with a UDP header. The M-CMTS MAY support DEPI without a UDP header. The EQAM MAY support DEPI without a UDP header. The EQAM MUST support DEPI with a UDP header.

7.3.1 Control message with a UDP header

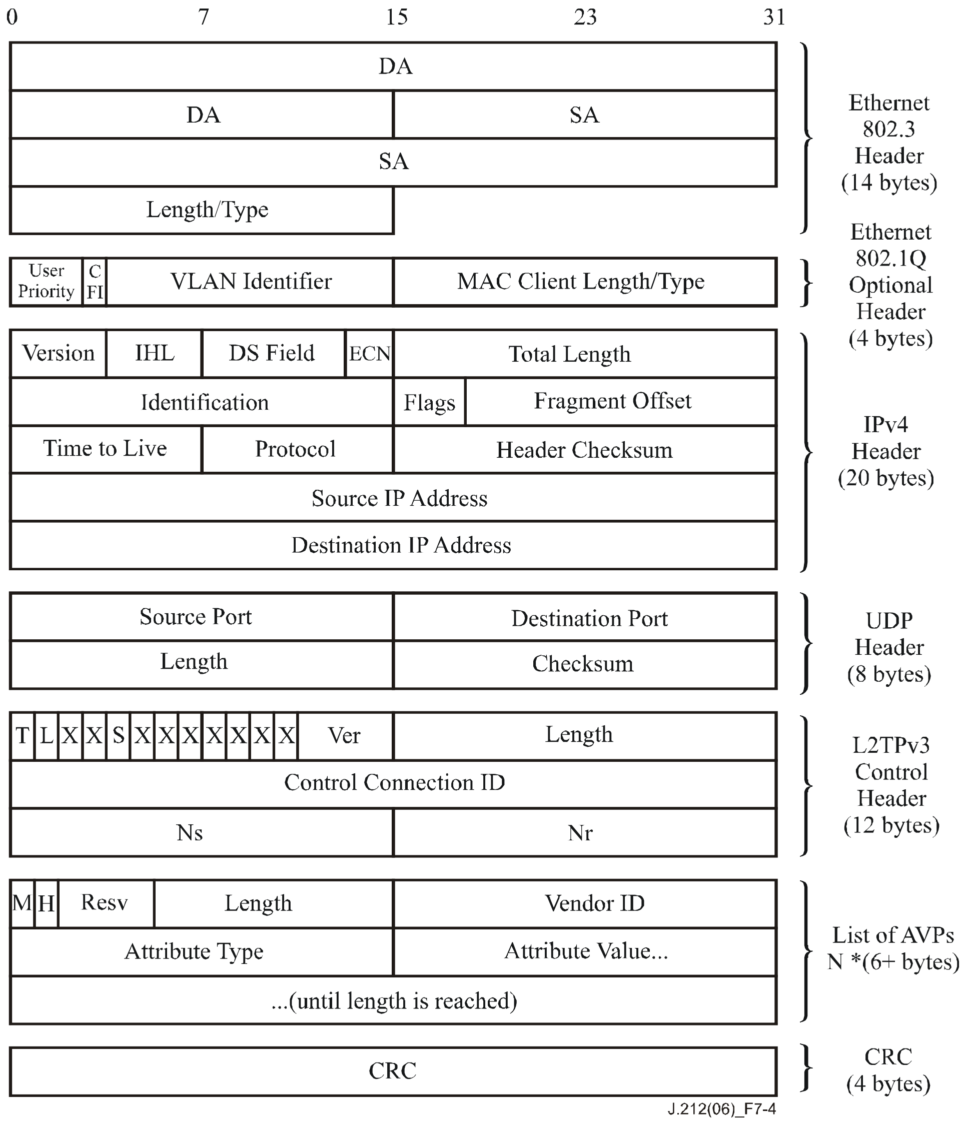

Figure 7-4/J.212 – DEPI control packet with UDP

7.3.2 Control message without a UDP header

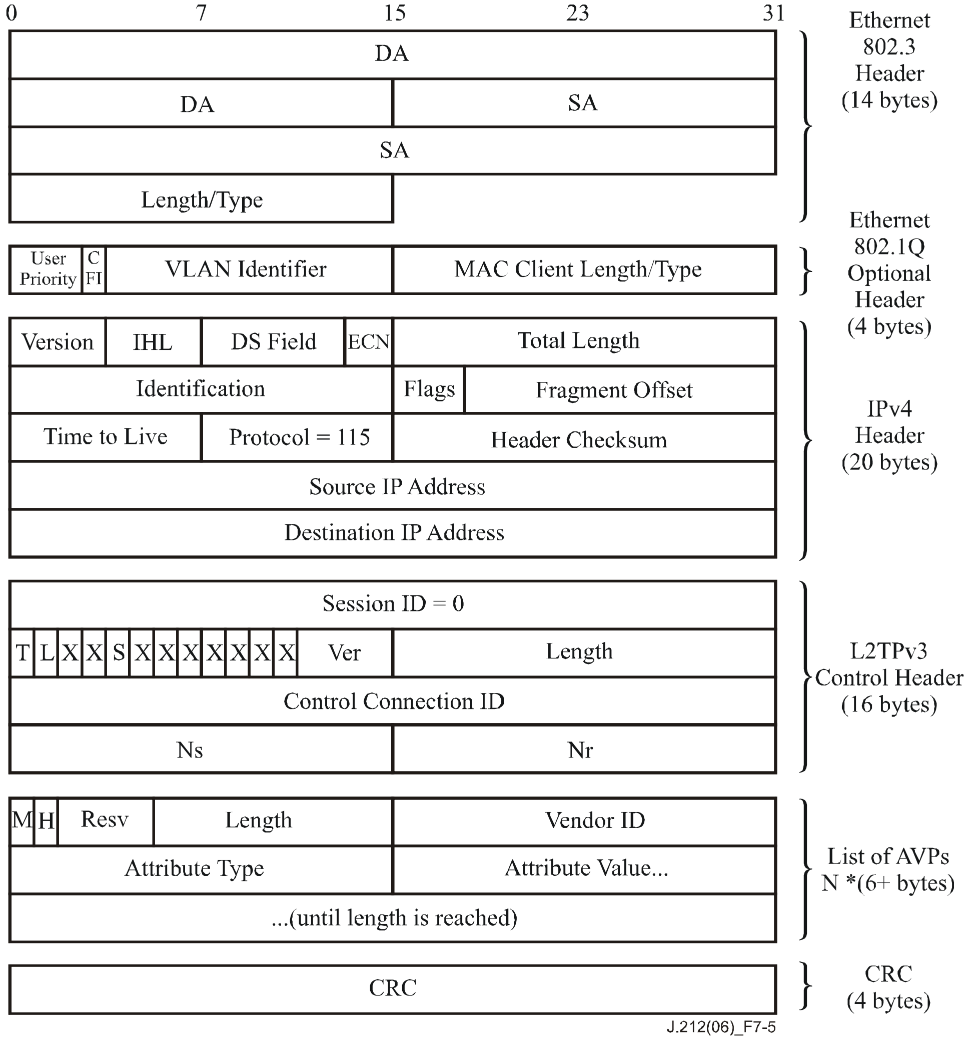

Figure 7-5/J.212 – DEPI control packet without UDP

7.3.3 Common headers for control and data messages

7.3.3.1 Ethernet 802.3 header

The Ethernet header is defined by [IEEE-802.3]. The Ethernet Destination Address is an individual address. Operation with DEPI over Ethernet group addresses is not specified at this time. The Ethernet Destination Address may be locally or globally administered.

Upon transmission of this frame by the M-CMTS Core, the Ethernet Destination Address will be the Ethernet address of the EQAM or of the next hop router. Upon reception of this frame by the EQAM, the Ethernet Source Address will be the Ethernet address of the output port of the M-CMTS Core or of the previous hop router.

If the networking interface is Ethernet, the M-CMTS Core MUST support the Ethernet header. If the networking interface is Ethernet, the EQAM MUST support the Ethernet header. If another physical layer interface is used instead of Ethernet, then the Ethernet headers would be replaced with the header format pertaining to that physical layer.

7.3.3.2 Ethernet 802.1Q header

The Ethernet 802.1Q header is defined by [IEEE-802.1Q]. The use of this header is optional and provides frame prioritization and VLAN support at layer 2. The M-CMTS Core MUST be able to support the Ethernet 802.1Q header. The M-CMTS Core SHOULD support mapping of 802.1Q user priority of tunnelled packets based on the DSCP value of a tunnel's IP packet. The EQAM SHOULD support the Ethernet 802.1Q header.

7.3.3.3 IPv4 header

The IP header is defined by [RFC-IP]. The IP Source Address is an IP address belonging to the M-CMTS Core. Currently, the IP Destination Address is unicast and is an IP Address belonging to the EQAM. Operation with DEPI over IP Multicast is not specified at this time.

For implementation considerations and for coexistence with network policies that are not amenable to IP fragmentation, EQAMs are not required to perform IP reassembly. The M-CMTS Core MUST NOT use IP fragmentation. The M-CMTS Core MUST set the IP DF (don't fragment) bit.

The M-CMTS Core MUST support a configurable 6-bit Differentiated Services Code Point (DSCP) to be used in the DS Field. The DS Field and DSCP are described in further detail by [RFC-DSCP-1] (RFC 2983) and [RFC-DSCP-2] (RFC 3260).

The M-CMTS Core MUST support the IPv4 header. The EQAM MUST support the IPv4 header.

7.3.3.4 UDP header

The UDP header is defined by [RFC-UDP]. The value of the UDP Source Port and the UDP Destination Port is determined through the L2TPv3 control plane between the M-CMTS Core and the EQAM. This value SHOULD conform to [IANA-PORTS].

When transmitting packets, both the EQAM and M-CMTS Core MUST support the generation of UDP checksums as defined in [RFC-UDP]. The sender MAY choose to set the UDP checksum to 0 for L2TPv3 data messages. This value is reserved by [RFC-UDP] to mean no checksum has been calculated. The sender MUST NOT set the UDP checksum to 0 for L2TPv3 control messages. The receiver MUST support validation of the UDP checksum field, as per [RFC-UDP].

The M-CMTS Core MUST support the UDP header. The EQAM MUST support the UDP header.

7.3.3.5 CRC

The CRC is CRC-32 and is defined by [IEEE-802.3].

The M-CMTS Core MUST support the CRC field. The EQAM MUST support the CRC field.

7.3.4 Specific headers for control messages

7.3.4.1 L2TPv3 control header

These fields are defined in [RFC-L2TPv3] and are repeated here for reference. The fields have the following significance:

T Type bit. The T bit MUST be set to 1, indicating that this is a control message.

L Length bit. The L bit MUST be set to 1, indicating that the Length field is present.

S Sequence bit. The S bit MUST be set to 1 indicating that sequence numbers (Ns and Nr) are present.

X Reserved bits. All reserved bits MUST be set to 0 on outgoing messages and ignored on incoming messages.

Ver Version. 4 bits. Set to 3.

Length 2 bytes. The Length field indicates the total length of the message in octets, always calculated from the start of the control message header itself, beginning with the T bit. It does not include the Session ID (shown in Figure 7-5) when present.

CCID Control Connection Identifier. 4 bytes. Negotiated per control connection.

Ns Sending Sequence Number. 2 bytes. Indicates sending sequence of this control message.

Nr Received Sequence Number. 2 bytes. Indicates next expected received sequence number.

7.3.4.2 Attribute Value Pairs (AVP)

There can be one or more attribute value pairs (AVPs) per DEPI control message. The fields have the following significance:

M Mandatory bit. If this bit is set to a 1 and this AVP is rejected, the control connection or session setup in which the AVP is carried will be torn down.

H Hidden bit. This bit is set to a 1 when the contents of the AVP message are encrypted and a 0 when the contents of the AVP message are not encrypted. Encryption of AVP messages is not required for DEPI.

Resv Reserved. 4 bits. Set to all zeros on transmit. Ignored on receive.

Length 10 bits. Equal to the length of the attribute value field plus 6 bytes.

Vendor ID 2 bytes. For AVPs defined by [RFC-L2TPv3], this field is set to 0. For AVPs defined by this Recommendation, this field is set to the IANA assigned vendor ID of 4491 (0x118B). For AVPs defined outside the scope of this Recommendation, this field may be set to a vendor specific ID.

Attribute Type 2 bytes.

Attribute Value N bytes.

Reserved 8 bits. If an AVP has a reserved field, the bits in this field should be set to 0 on transmit and ignored on the receive.

If an LCCE receives an AVP with a Vendor ID that it does not recognize, it MUST silently discard the AVP or tear down the session depending on the value of the Mandatory bit.

7.4 Signalling

The supported L2TPv3 messages for the DEPI control plane are shown in Table 7-1:

Table 7-1/J.212 – DEPI control messages

#	Mnemonic	Name
Control connection management		
1	SCCRQ	Start-Control-Connection-Request
2	SCCRP	Start-Control-Connection-Reply
3	SCCCN	Start-Control-Connection-Connected
4	StopCCN	Stop-Control-Connection-Notification
6	HELLO	Hello
20	ACK	Explicit Acknowledgement
Session management		
10	ICRQ	Incoming-Call-Request
11	ICRP	Incoming-Call-Reply
12	ICCN	Incoming-Call-Connected
14	CDN	Call-Disconnect-Notify
16	SLI	Set Link Info

L2TPv3 outgoing call messages (OCRQ, OCRP, OCCN) and the WAN-Error-Notify (WEN) messages are not required to be supported.

There is a reliable control message delivery mechanism that is accomplished either by sending an Explicit Acknowledgement (ACK) message after any of the control messages or by piggybacking an acknowledgment with the Nr and Ns fields in a later control message. If Control Messages are not acknowledged within the time Control Message Timeout (refer to Annex B), then the Control Message MUST be retransmitted up to the value of Control Message Retry Count (refer to Annex B). For example, the Control Message must be retransmitted a total of 10 times using an exponential back-off value starting at 1 second and growing to a maximum value of 8 seconds.

NOTE – There will be 7 intervals of 8 seconds in this scheme.

Control message authentication MAY be supported. If Control Message Authentication is supported, the methods described in [RFC-L2TPv3], section 5.4.1, should be followed.

The following flow diagrams show typical DEPI message exchanges along with the required AVPs from L2TPv3 and DEPI. Optional AVPs are not shown but may also be present.

7.4.1 Control connection signalling

7.4.1.1 Control connection setup

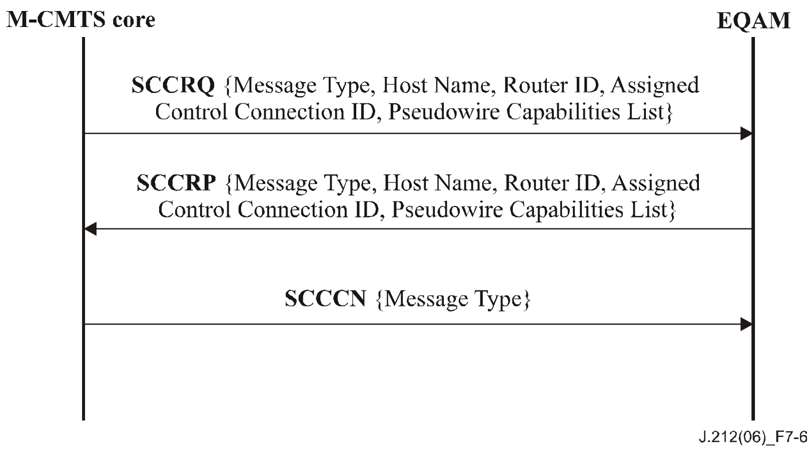

Figure 7-6/J.212 – DEPI control connection setup

In order to tunnel DOCSIS frames over IP using L2TPv3, an L2TPv3 Control Connection is first established as described in [RFC-L2TPv3]. Establishment of the control connection involves an exchange of AVPs that identifies the peer and its capabilities. Each control connection has a Control Connection ID which is assigned by the recipient, and negotiated with Connection ID AVPs during the creation of a control connection.

The M-CMTS Core MUST support the ability to initiate control connection signalling (L2TPv3 caller). The EQAM MUST support the ability to receive incoming control connection requests from the M-CMTS Core (L2TPv3 callee). Establishment of control connections by the EQAM is not required for DEPI and is outside the scope of this Recommendation.

7.4.1.2 Control connection teardown

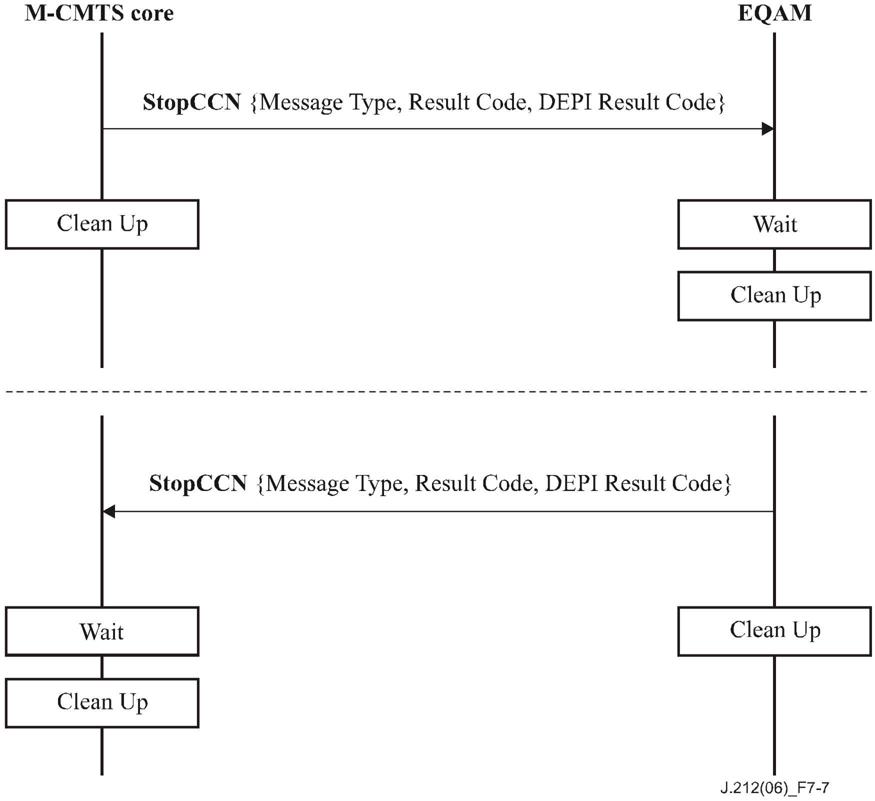

Figure 7-7/J.212 – DEPI control connection teardown

Control connection teardown may be initiated by either LCCE and is accomplished by sending a single StopCCN control message. An implementation may shut down an entire control connection and all sessions associated with the control connection by sending the StopCCN. Thus, it is not necessary to clear each session individually when tearing down the whole control connection. The peer receiving the StopCCN message MUST maintain session and control state for a period of time equal to StopCCN Timeout (Annex B) after acknowledging the StopCCN. This provision is for managing lost acknowledgments.

7.4.1.3 Control connection keep-alive

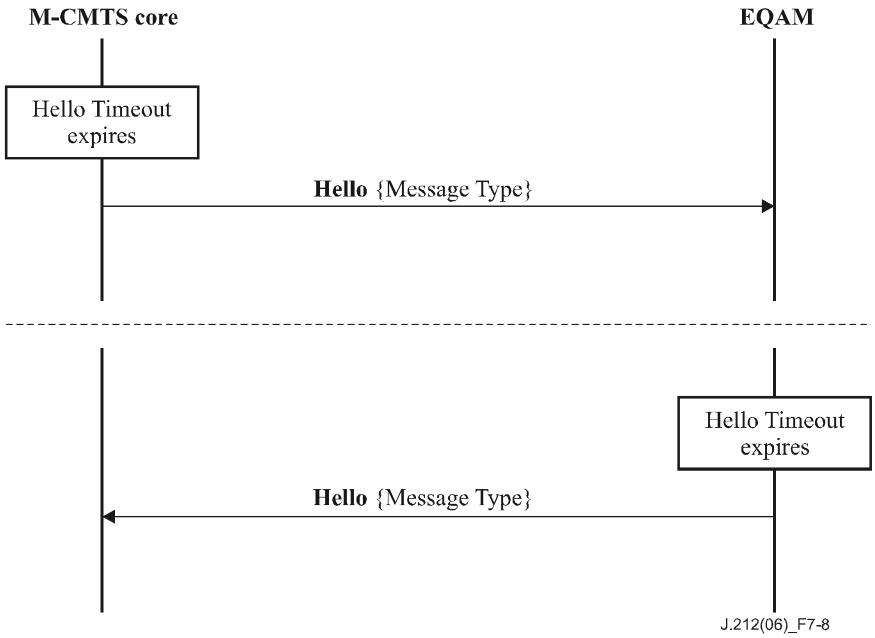

Figure 7-8/J.212 – DEPI keep-alive

A periodic keep-alive for the control connection is implemented by sending a Hello message if a period of time known as the Hello Timeout (refer to Annex A) has passed without receiving any message (data or control) from the peer.

7.4.2 Session signalling

7.4.2.1 Session setup

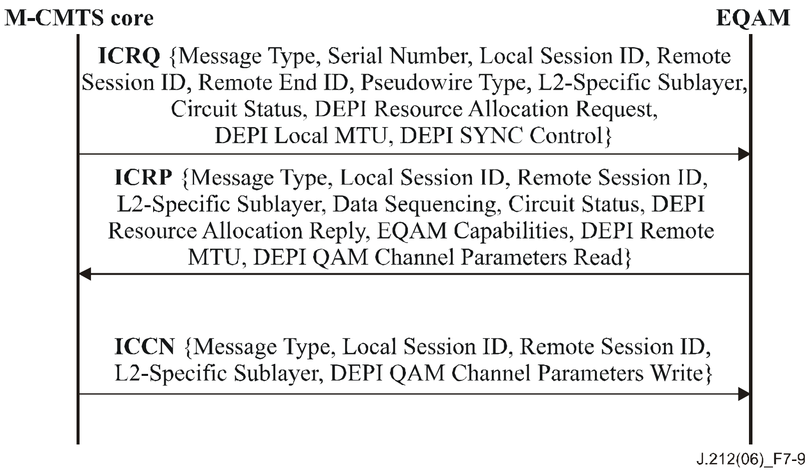

Figure 7-9/J.212 – DEPI session setup

After successful control connection establishment, individual sessions may be created. Each session corresponds to a single data stream between the two LCCEs. In addition to the mandatory and optional AVPs in [RFC-L2TPv3], the following DEPI specific AVPs are used as part of the session setup.

ICRQ contains the Remote End ID AVP which contains the TSID of the QAM Channel for which the session is intended.

ICRP contains the Remote Session AVP which indicates the session ID that the EQAM wants to use. ICRP also contains a suite of QAM Channel AVPs (refer to 7.5.2) which indicate the EQAM's current configuration, which parameters can be changed, EQAM capabilities, and assigned values such as the UDP port values. If these values are not acceptable to the M-CMTS Core, the M-CMTS Core will return a CDN message with the appropriate error code.

ICCN contains the parameters that the M-CMTS Core wants to change. If these parameters are acceptable to the EQAM, the EQAM will return an ACK (either explicit or implicit). If these parameters are not acceptable to the EQAM, the EQAM will return a CDN message with the appropriate error code.

The receipt and processing of the ICCN triggers the initiation of data forwarding in the EQAM for the session. The EQAM MUST NOT transmit data on the QAM channel until the session is configured according to the parameters in the ICCN message. The EQAM SHOULD NOT buffer data while the session is being configured.

The M-CMTS Core MUST support the ability to generate session establishment signalling. The EQAM MUST support the ability to receive incoming session establishment requests from the M-CMTS Core. Establishment of L2TP sessions by the EQAM is outside the scope of this Recommendation.

7.4.2.2 Session teardown

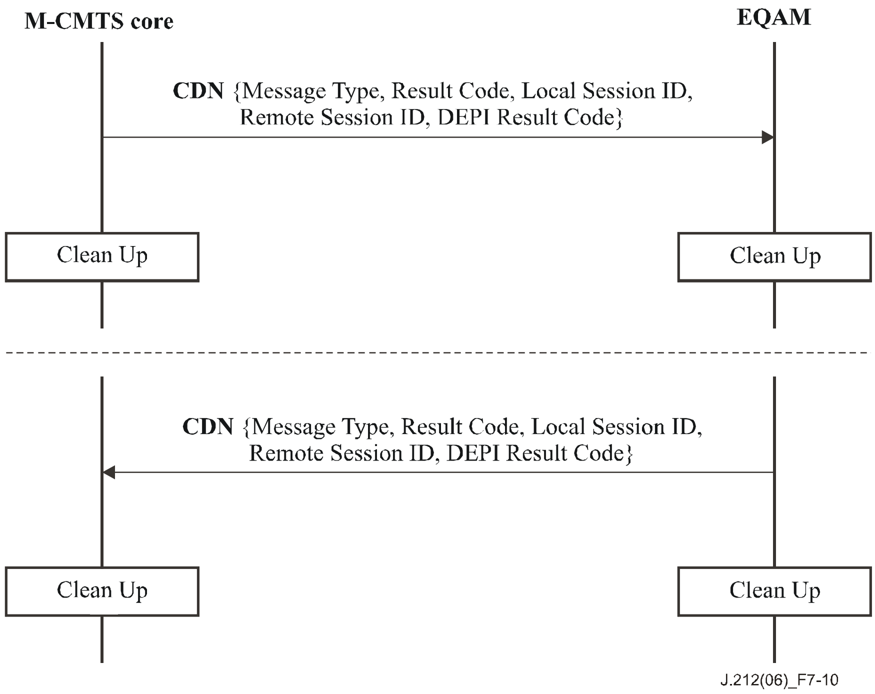

Figure 7-10/J.212 – DEPI session teardown

Session teardown may be initiated by either LCCE and is accomplished by sending a single CDN control message. An implementation may shut down an entire control connection and all sessions associated with the control connection by sending the StopCCN. Thus, it is not necessary to clear each session individually when tearing down the whole control connection.

7.4.2.3 Session updates

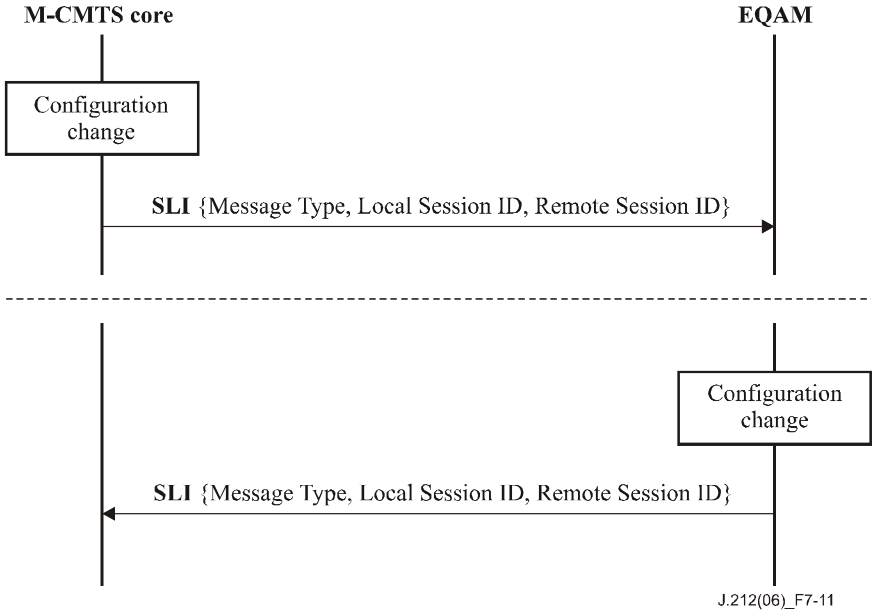

Figure 7-11/J.212 – DEPI session updates

If there is a configuration change to one of the EQAM parameters which is described by a DEPI AVP, the M-CMTS Core MUST send the updated AVP to the EQAM with the Set-Link-Info (SLI) message. If there is a configuration change to one of the EQAM parameters which is described by a DEPI AVP, the EQAM MUST send the updated AVP to the M-CMTS Core using an SLI message.

7.4.3 Required and optional AVPs

In addition to the mandatory and optional AVPs listed in [RFC-L2TPv3] and modified by Table 7-3, the following DEPI AVPs listed in Table 7-2 MUST be present in the DEPI control message if mandatory, and MAY be present in DEPI control messages if optional.

Table 7-2/J.212 – DEPI mandatory and optional AVPs

DEPI control message	DEPI mandatory AVPs	DEPI optional AVPs
ICRQ	DEPI Resource Allocation Request DEPI Local MTU DS QAM Channel DOCSIS SYNC Control	Local UDP Port
ICRP	DEPI Resource Allocation Reply DEPI Remote MTU EQAM Capabilities DS QAM Channel Frequency DS QAM Channel Power DS QAM Channel Modulation DS QAM Channel J.83 Annex DS QAM Channel Symbol Rate DS QAM Channel Interleaver Depth DS QAM Channel RF Mute	Downstream QAM Channel TSID Group
ICCN		DS QAM Channel Frequency DS QAM Channel Power DS QAM Channel Modulation DS QAM Channel J.83 Annex DS QAM Channel Symbol Rate DS QAM Channel Interleaver Depth DS QAM Channel RF Mute
CDN		DEPI Result Code
SLI		DS QAM Channel DOCSIS SYNC Control DS QAM Channel Frequency DS QAM Channel Power DS QAM Channel Modulation DS QAM Channel J.83 Annex DS QAM Channel Symbol Rate DS QAM Channel Interleaver Depth DS QAM Channel RF Mute

7.5 AVP definitions

7.5.1 Conventional L2TPv3 AVPs

The AVP types from [RFC-L2TPv3] and [RFC-L2TP-DSCP] that are supported as part of this Recommendation are shown in Table 7-3.

Table 7-3/J.212 – DEPI supported L2TPv3 AVPs

Attribute type	Control, Session	Description	Required	Not required
0	C, S	Message Type	●	
1	S	Result Code	●	
5	C, S	Control/Session Tie Breaker		●
7	C	Host Name	●	
8	C	Vendor Name		●
10	C	Receive Window Size		●
15	S	Serial Number	●	
25	S	Physical Channel ID		●
34	S	Circuit Errors		●
36	C, S	Random Vector		●
47	C	Control Connection DSCP		●
48	S	Session DSCP		●
58	C, S	Extended Vendor ID AVP		●
59	C, S	Message Digest		●
60	C	Router ID	●	
61	C	Assigned Control Connection ID	●	
62	C	Pseudowire Capabilities List	●	
63	S	Local Session ID	●	
64	S	Remote Session ID	●	
65	S	Assigned Cookie		●
66	S	Remote End ID	●	
68	S	Pseudowire Type	●	
69	S	L2-Specific Sublayer	●	
70	S	Data Sequencing	●	
71	S	Circuit Status	●	
72	C	Preferred Language		●
73	C	Control Message Authentication Nonce		●
74	S	Tx Connect Speed		●
75	S	Rx Connect Speed		●

Conventional AVPs whose usage is specific to DEPI are described below. A more complete description as well as requirements for the conventional AVPs is in [RFC-L2TPv3].

7.5.1.1 Message type (All messages)

0		7	15	23 31
M	H	Resv	Length = 8	Vendor ID = 0
Attribute Type = 0				Message Type

Figure 7-12/J.212 – Message type AVP

This identifies the particular L2TPv3 Control Message. It is always the first AVP.

7.5.1.2 Result code (StopCCN, CDN)

0			7	15	23	31
M	H	Resv	Length = 8 + N		Vendor ID = 0	
Attribute Type = 1					Result Code	
Error Code (Optional)					Error Message ...	
... Error Message (Optional)						

Figure 7-13/J.212 – Result code AVP

This message contains results codes, optional error codes, and optional error messages when tearing down a Control Connection or a Session.

7.5.1.3 Host name (SCCRQ, SCCRP)

0			7	15	23	31
M	H	Resv	Length = 6 + N		Vendor ID = 0	
Attribute Type = 7					Host ...	
... Name ...						

Figure 7-14/J.212 – Host name AVP

The host name is typically the fully qualified domain name (FQDN) of each device.

7.5.1.4 Vendor name (SCCRQ, SCCRP)

0			7	15	23	31
M	H	Resv	Length = 6 + N		Vendor ID = 0	
Attribute Type = 8					Vendor ...	
... Name ...						

Figure 7-15/J.212 – Vendor name AVP

The M-CMTS Core SHOULD identify itself with an ASCII Vendor ID string during the SCCRQ message. The EQAM SHOULD identify itself with an ASCII Vendor ID string during the SCCRP message. Note that this is an optional AVP in [RFC-L2TPv3].

7.5.1.5 Serial number (ICRQ, OCRQ)

0			7	15	23	31
M	H	Resv	Length = 10		Vendor ID = 0	
Attribute Type = 15					Serial ...	
... Number						

Figure 7-16/J.212 – Serial number AVP

This number is assigned by the originator of the message and is similar in concept to a transaction ID. Its main use is to aid in debugging message flows.

7.5.1.6 Router ID (SCCRQ, SCCRP)

<table>
<tr><td colspan="3">0</td><td>7 15</td><td>23 31</td></tr>
<tr><td>M</td><td>H</td><td>Resv</td><td>Length = 10</td><td>Vendor ID = 0</td></tr>
<tr><td colspan="4">Attribute Type = 60</td><td>Router ...</td></tr>
<tr><td colspan="4">... ID</td><td></td></tr>
</table>

Figure 7-17/J.212 – Router ID AVP

The Router ID is typically the IP address of each endpoint.

7.5.1.7 Assigned control connection ID (SCCRQ, SCCRP, StopCCN)

<table>
<tr><td colspan="3">0</td><td>7 15</td><td>23 31</td></tr>
<tr><td>M</td><td>H</td><td>Resv</td><td>Length = 10</td><td>Vendor ID = 0</td></tr>
<tr><td colspan="4">Attribute Type = 61</td><td>Control Connection ...</td></tr>
<tr><td colspan="4">... ID</td><td></td></tr>
</table>

Figure 7-18/J.212 – Control connection ID AVP

This is the control connection ID for DEPI. During SCCRQ, the M-CMTS Core uses this AVP to inform the EQAM what value of Control Connection ID to use in the L2TPv3 Control Header for control messages originated by the EQAM. During SCCRP, the EQAM uses this AVP to inform the M-CMTS Core what value of Control Connection ID to use in the L2TPv3 Control header for control messages originated by the M-CMTS Core. Since the M-CMTS Core has not been informed by the EQAM prior to the first SCCRQ message, the M-CMTS core uses a value of 0 for the control connection ID in the L2TPv3 Control Header of the SCCRQ message.

7.5.1.8 Pseudowire capabilities list (SCCRQ, SCCRP)

<table>
<tr><td colspan="3">0</td><td>7 15</td><td>23 31</td></tr>
<tr><td>M</td><td>H</td><td>Resv</td><td>Length = 6 + 2N</td><td>Vendor ID = 0</td></tr>
<tr><td colspan="4">Attribute Type = 62</td><td>Pseudowire Type 0</td></tr>
<tr><td colspan="4">Pseudowire Type 1</td><td>Pseudowire Type N</td></tr>
</table>

Figure 7-19/J.212 – Pseudowire capabilities list AVP

The Pseudowire Capabilities List indicates the capabilities of the M-CMTS Core and EQAM. There are two PW Types defined for DEPI. See Table 7-4.

Table 7-4/J.212 – Pseudowire types

Pseudowire type	Mnemonic	Value
MPT Pseudowire	MPTPW	0x000C*
PSP Pseudowire	PSPPW	0x000D*
* These are to be considered temporary values, pending official assignments from IANA.		

The M-CMTS Core MUST indicate its support for PSP and D-MPT mode by including one or both of the DEPI PW Types in the Pseudowire Capabilities List. The EQAM MUST indicate its support for PSP and D-MPT mode by including one or both of the DEPI PW Types in the Pseudowire Capabilities List.

7.5.1.9 Local session ID (ICRQ, ICRP, ICCN, CDN, SLI)

0			7	15	23	31
M	H	Resv	Length = 10		Vendor ID = 0	
Attribute Type = 63					Local Session ...	
... ID						

Figure 7-20/J.212 – Local session ID AVP

When a session is established, the M-CMTS Core and EQAM each choose their own Session ID and advertise it to each other with this AVP. This means that a single session setup sets up two unidirectional sessions, one in each direction.

7.5.1.10 Remote session ID (ICRQ, ICRP, ICCN, CDN, SLI)

0			7	15	23	31
M	H	Resv	Length = 10		Vendor ID = 0	
Attribute Type = 64					Remote Session ...	
... ID						

Figure 7-21/J.212 – Remote session ID AVP

When the M-CMTS Core or EQAM sends a session message to each other, the Remote Session ID is set to the session ID learned previously from the Local Session ID. If the Remote Session ID is not yet known, it is set to 0.

7.5.1.11 Remote end ID (ICRQ)

0			7	15	23	31
M	H	Resv	Length = 8		Vendor ID = 0	
Attribute Type = 66					Remote ID = TSID	

Figure 7-22/J.212 – Remote end ID AVP

DEPI uses the TSID from the QAM Channel as the Remote End ID. The TSID is an unsigned two-octet integer and is used to bind a session to a QAM Channel.

7.5.1.12 Pseudowire type (ICRQ)

0			7	15	23	31
M	H	Resv	Length = 8		Vendor ID = 0	
Attribute Type = 68					Pseudowire Type	

Figure 7-23/J.212 – Pseudowire type AVP

DEPI uses the Pseudowire Type values defined in 7.5.1.8 to indicate the type of DEPI session requested.

7.5.1.13 L2-specific sublayer (ICRQ, ICRP, ICCN)

0			7	15	23	31
M	H	Resv	Length = 8		Vendor ID = 0	
Attribute Type = 69					L2-Specific Sublayer Type	

Figure 7-24/J.212 – L2-specific sublayer AVP

The M-CMTS Core MUST include the L2-Specific Sublayer AVP in the ICRQ and ICCN messages, indicating the L2-Specific Sublayer Header Type consistent with the Pseudowire type for the DEPI flow. The EQAM MUST include the L2-Specific Sublayer AVP in the ICRP message, indicating the L2-Specific Sublayer Header Type consistent with the Pseudowire type for the DEPI flow.

Table 7-5/J.212 – L2-Specific sublayer types

L2-specific sublayer type	Value
MPT Specific Sublayer	3*
PSP Specific Sublayer	4*
* These are to be considered temporary values, pending official assignments from IANA.	

7.5.1.14 Data sequencing (ICRP)

0			7	15	23	31
M	H	Resv	Length = 8		Vendor ID = 0	
Attribute Type = 70					Data Sequencing Level = 2	

Figure 7-25/J.212 – Data sequencing AVP

The EQAM MUST include the Data Sequencing AVP in the ICRP message, indicating Data Sequencing Level 2 is required (all incoming data packets require sequencing).

7.5.1.15 Circuit status (ICRQ, ICRP, ICCN, SLI)

0			7	15	23		31
M	H	Resv	Length = 8		Vendor ID = 0		
Attribute Type = 71					Reserved	N	A

Figure 7-26/J.212 – Circuit status AVP

N 1 bit – The New bit indicates whether the status indication is for a new DEPI session (1) or an existing DEPI session (0). The New bit SHOULD be set the first time the DEPI session is established after provisioning.

A 1 bit – The Active bit indicates whether the DEPI session is up (1) or down (0). Once the M-CMTS Core is aware that the DEPI session is down, the M-CMTS Core MUST NOT attempt to pass data traffic on the DEPI session.

DEPI uses the Circuit Status AVP to indicate whether the DEPI session is up and able to pass data traffic, or down and unable to pass data traffic. The Circuit Status ID does not control the RF output of the QAM Channel. Note that the Circuit Status AVP is sent by both the M-CMTS Core and the EQAM.

7.5.2 DEPI specific AVPs

The AVP types defined specifically for DEPI are shown in Table 7-6. The range of Attribute Types from 0-99 is reserved for DEPI Session Specific AVPs. These AVPs are only used in L2TP session messages.

Table 7-6/J.212 – DEPI defined general session AVPs

Attribute type	Description
0	Reserved
1	DEPI Result Code
2	DEPI Resource Allocation Request
3	DEPI Resource Allocation Reply
4	DEPI Local MTU
5	DOCSIS SYNC Control
6	EQAM Capability Bits
7	DEPI Remote MTU
8	DEPI Local UDP Port

7.5.2.1 DEPI result and error code (CDN, SLI)

<table>
<tr><td>0</td><td></td><td>7</td><td>15</td><td>23</td><td>31</td></tr>
<tr><td>M</td><td>H</td><td>Resv</td><td>Length = 8 + N</td><td colspan="2">Vendor ID = 4491</td></tr>
<tr><td colspan="4">Attribute Type = 1</td><td colspan="2">Result Code</td></tr>
<tr><td colspan="4">Error Code (Optional)</td><td colspan="2">Error Message ...</td></tr>
<tr><td colspan="6">... Error Message (Optional)</td></tr>
</table>

Figure 7-27/J.212 – DEPI result and error code AVP

The format of the field of this AVP is identical to the field in the standard L2TPv3 result and error code AVP except that the Vendor ID field has the value of 4491 rather than 0. The Result and Error Codes for this AVP are unique to DEPI, and are in addition to the standard Result and Error codes in [RFC-L2TPv3].

The following new Result Codes, Error Codes, and Error Messages are required specifically for use with DEPI:

Result code	Result description
0	Session not established – Bad DEPI AVP Reference
1	Session not established – Bad PHY AVP Request
2	Session not established or disconnected for the reason indicated in Error Code.

Error code	Error description
0	Device is not yet ready or configured properly.
1	Attempt to modify a locked PHY Parameter.
2	Attempt to modify PHY Parameter failed – out of range value.
3	Requested PSP Flow PHBIDs not supported.
4	Incorrect pseudowire type used in session.
NOTE – One cause of Error Code 0 would be if the EQAM did not have the values of M and N configured for the QAM Channel Symbol Rate.	

7.5.2.2 DEPI resource allocation request (ICRQ)

0 7 15 23 31

M	H	Resv	Length = 6 + N	Vendor ID = 4491					
Attribute Type = 2				X	X	PHBID 1	X	X	PHBID N

Figure 7-28/J.212 – DEPI resource allocation request AVP

M 1 bit. The Mandatory bit MUST be set to a 1.

Length 10 bits. 6 bytes plus one additional byte for each flow that is requested.

Attribute Type 2 bytes. Set to 2.

Each flow request entry consists of:

PHBID 6 bits. Per Hop Behaviour Identifier being requested by the M-CMTS Core. Per hop behaviours identifiers are defined in 6.2.1.

In the ICRQ message, the M-CMTS Core requests for a number of flows for a session. Each byte in the attribute payload represents a request for one unique flow. Each request contains the PHBID that will be used for that flow. For D-MPT mode operation, the M-CMTS Core MUST request a single flow.

7.5.2.3 DEPI resource allocation reply (ICRP)

0 7 15 23 31

M	H	Resv	Length = 8 + 4 * N		Vendor ID = 4491
Attribute Type = 3					Reserved
X	X	PHBID 1	Reserved	Flow ID 1	UDP Destination Port 1
X	X	PHBID N	Reserved	Flow ID N	UDP Destination Port N

Figure 7-29/J.212 – DEPI resource allocation reply AVP

M 1 bit. The Mandatory bit MUST be set to a 1.

Length 10 bits. 8 bytes plus 4 additional bytes for each flow.

Attribute Type 2 bytes. Set to 3.

Each flow response entry consists of:

PHBID 6 bits. Per Hop Behaviour Identifier that was requested by the M-CMTS Core. Per hop behaviours identifiers are defined in 6.2.1.

Flow ID 3 bits. This is the flow ID assigned by the EQAM. The flow ID is unique within a session.

UDP Dest Port 2 bytes. This is the UDP Destination Port as specified by the EQAM that the M-CMTS Core MUST use for the session header if the M-CMTS and EQAM have been configured to use a UDP header with L2TPv3. This value MUST be unique for each session. This value MAY be unique for each flow. If L2TPv3 has been configured to not use UDP headers, then this field MUST be set to all 0's by the EQAM and MUST be ignored by the M-CMTS Core.

In the ICRP message, the EQAM responds by creating flows that match the flows requested. The EQAM specifies the Flow ID and the UDP Destination Port for each flow. The PHBID is unchanged from the flow request field. If the EQAM does not support a PHBID referenced in the request from the M-CMTS Core, the EQAM can signal that by not including the PHBID in this response. If the EQAM cannot support any of the PHBIDs requested by the M-CMTS Core, then the EQAM MUST tear down the session by issuing a CDN message.

7.5.2.4 DEPI local MTU (ICRQ)

0		7	15	23	31
M	H	Resv	Length = 8	Vendor ID = 4491	
Attribute Type = 4				DEPI Local MTU	

Figure 7-30/J.212 – DEPI local MTU AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 4.

DEPI Local MTU 2 bytes. In the ICRQ message, this is the MTU (Maximum Transmission Unit) that the M-CMTS can receive from the EQAM on the CIN interface.

The MTU is Layer 3 payload of a Layer 2 frame. For DEPI, the MTU would include the L2TPv3 header and payload, the UDP header if present, and the IP header, but would not include the Ethernet Header or CRC. For example, a 1518-byte Ethernet frame (1522 bytes if VLAN tags are present) would support an MTU of 1500 bytes.

7.5.2.5 Downstream QAM channel DOCSIS SYNC control (ICRQ, SLI)

0		7	15	23	31
M	H	Resv	Length = 14	Vendor ID = 4491	
Attribute Type = 5				E	DOCSIS SYNC Interval
MAC SA					
MAC SA					

Figure 7-31/J.212 – DOCSIS SYNC AVP

Length 10 bits. Set to 14.

Attribute Type 2 bytes. Set to 5.

E 1 bit. SYNC Enable. Operation is described below.

Interval 15 bits. Nominal interval between SYNC messages in 200 µs steps.

MAC SA 48 bits. IEEE 802 MAC address to be used in the source address field.

This AVP is used differently for D-MPT Mode and PSP Mode.

In D-MPT Mode, if E = 0, then the EQAM MUST NOT modify the timestamp values in DOCSIS SYNC messages. If E = 1, then the EQAM MUST find and correct timestamp values in SYNC messages. The DOCSIS SYNC Interval field is set to all zeros by the M-CMTS Core and ignored by the EQAM. The M-CMTS Core SHOULD set the MAC SA field of this AVP to the MAC address of the DOCSIS interface it has associated with the session. The MAC SA field is ignored by the EQAM.

In PSP Mode, if E = 0, then the EQAM MUST NOT transmit a DOCSIS SYNC message. If E = 1, then the EQAM MUST insert and send a DOCSIS SYNC message at a nominal interval specified by the Interval field and with a source MAC address specified by the MAC SA field. The EQAM MUST support DOCSIS SYNC Interval values between 0x000A (2 ms) and 0x03E8 (200 ms). Although the measured time between any two SYNC messages might vary depending on traffic, the measured time MUST be within ±2.5 ms of the nominal value, and MUST NOT exceed the maximum value defined in Annex B. The M-CMTS Core SHOULD set the MAC SA field of this

AVP to the MAC address of the DOCSIS interface it has associated with the session. The EQAM MUST use the address in the MAC SA field as the source address in the MAC management header for all subsequent SYNC messages.

The M-CMTS Core MUST use this AVP to convey the treatment of DOCSIS SYNC messages by the EQAM.

7.5.2.6 EQAM capabilities AVP (ICRQ)

0		7	15	23	31
M	H	Resv	Length = 8	Vendor ID = 4491	
Attribute Type = 6				EQAM Capabilities Field	

Figure 7-32/J.212 – EQAM capabilities AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 6.

Capabilities 2 bytes. EQAM Capabilities Field. Default value of all bits is 0.

– Bit 0: A "1" indicates that the EQAM supports DLM-EE-RQ and DLM-EE-RP packets. A "0" indicates the EQAM does not support these two DLM packets.

– Bits 1 through 15: Reserved. Sender must set to 0. Receiver must ignore.

7.5.2.7 DEPI remote MTU (ICRP)

0		7	15	23	31
M	H	Resv	Length = 8	Vendor ID = 4491	
Attribute Type = 7				DEPI Remote MTU	

Figure 7-33/J.212 – DEPI remote MTU Max payload AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 7.

DEPI MTU 2 bytes. In the ICRP message, this is the MTU that the EQAM can receive from the M-CMTS Core on the CIN interface. The MTU is Layer 3 payload of a Layer 2 frame.

7.5.2.8 Local UDP port (ICRQ)

0		7	15	23	31
M	H	Resv	Length = 8	Vendor ID = 4491	
Attribute Type = 8				Local UDP Port	

Figure 7-34/J.212 – Local UDP port AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 8.

Connect Speed 16 bits. UDP port to be used for session packets that are being sent to the local LCP.

The M-CMTS Core would issue this AVP during the session setup if UDP is enabled and if the M-CMTS Core wants a data session to use a different UDP port for sending session packets from the EQAM to the M-CMTS Core other than the UDP port that was negotiated during the Control Connection setup.

The M-CMTS Core MAY support the Local UDP Port AVP. The EQAM MAY support the Local UDP Port AVP.

7.5.3 QAM channel PHY AVPs

The QAM Channel Physical Layer specific AVP types defined for DEPI are shown in Table 7-7. The range of Attribute Types from 100-199 is reserved for QAM Channel PHY AVPs. These AVPs are only used in L2TP session messages.

Table 7-7/J.212 – DEPI defined QAM channel PHY AVPs

Attribute type	Description
100	Downstream QAM Channel TSID Group
101	Downstream QAM Channel Frequency
102	Downstream QAM Channel Power
103	Downstream QAM Channel Modulation
104	Downstream QAM Channel J.83 Annex
105	Downstream QAM Channel Symbol Rate
106	Downstream QAM Channel Interleaver Depth
107	Downstream QAM Channel RF Mute

These AVPs define the generic PHY parameters for a QAM Channel. These AVPs are sent from the EQAM to the M-CMTS Core to inform the M-CMTS Core of the EQAM current configuration and which values are allowed to be changed. This AVP is then sent from the M-CMTS Core to the EQAM to configure selected PHY level parameters.

The following fields have a common meaning across this group of AVPs, and are described once here:

M 1 bit. The Mandatory bit MUST be set to a 1 for both ICRP and ICCN (if applicable).

L 1 bit. Lock bit. This bit allows the EQAM to indicate which elements of the configuration have been locked down in configuration. For ICRP (from EQAM to M-CMTS Core), a value of 0 indicates the parameter described in the Attribute Value field is Read-Only. A value of 1 indicates the parameter is Read/Write. For ICCN (M-CMTS Core to EQAM), this value is set to 0 by the M-CMTS Core and ignored by the EQAM.

TSID Group ID 7 bits. If the referenced attribute is common to other QAM Channels, this field is set to a TSID Group as defined by a TSID Group AVP. Otherwise, this field is set to all zeros.

The AVP programming options listed below in this Recommendation may not be available in all EQAM products. To be considered compliant to the DEPI Recommendation, a M-CMTS Core or an EQAM MUST support a particular QAM Channel AVP attribute only if that attribute represents a feature available on that particular platform. For example, if an EQAM does not support Annex C/J.83 as a feature, then it does not have to support the QAM Channel AVP Annex C attribute value.

7.5.3.1 Downstream QAM channel TSID group (ICRP)

0 7 15 23 31

M	H	Resv	Length = 8 + 2N	Vendor ID = 4491		
Attribute Type = 100				L	TSID Group ID	Reserved
TSID #1				TSID #2		
TSID #3				TSID #N		

Figure 7-35/J.212 – TSID group AVP

Length 10 bits. Variable. Set to (8 + 2 * number of TSID entries).

Attribute Type 2 bytes. Set to 100.

L 1 bit. Lock bit. Not used. Set to 0.

TSID Group ID 7 bits. This is the TSID Group ID that the TSIDs in this list belong to.

TSID #1-N 16 bits. List of TSIDs.

Some PHY level attribute types may be common to more than one QAM Channel. As such, changing that attribute on one QAM Channel may change that attribute on other QAM Channels. The EQAM indicates this dependency by defining TSID groups. This AVP MAY be repeated to define more than one TSID Group. Each TSID group is associated with one or more PHY level parameters by including the TSID Group ID within the PHY parameter AVP.

7.5.3.2 Downstream QAM channel frequency (ICRP, ICCN, SLI)

0 7 15 23 31

M	H	Resv	Length = 12	Vendor ID = 4491		
Attribute Type = 101				L	TSID Group ID	Reserved
Frequency						

Figure 7-36/J.212 – Frequency AVP

Length 10 bits. Set to 12.

Attribute Type 2 bytes. Set to 101.

Frequency 4 bytes. This specifies the downstream frequency of a QAM Channel. This is the centre frequency of the downstream channel in Hz stored as a 32-bit binary number.

7.5.3.3 Downstream QAM channel power (ICRP, ICCN, SLI)

0 7 15 23 31

M	H	Resv	Length = 10	Vendor ID = 4491		
Attribute Type = 102				L	TSID Group ID	Reserved
Power						

Figure 7-37/J.212 – Power AVP

Length 10 bits. Set to 10.

Attribute Type 2 bytes. Set to 102.

Power 2 bytes. TX Power in dBmV (unsigned 16-bit, 0.1 dB unit)

7.5.3.4 Downstream QAM channel modulation (ICRP, ICCN, SLI)

<table>
<tr><td>0</td><td></td><td>7</td><td>15</td><td></td><td>23</td><td></td><td>31</td></tr>
<tr><td>M</td><td>H</td><td>Resv</td><td>Length = 8</td><td colspan="4">Vendor ID = 4491</td></tr>
<tr><td colspan="4">Attribute Type = 103</td><td>L</td><td>TSID Group ID</td><td>Resv</td><td>Modulation</td></tr>
</table>

Figure 7-38/J.212 – Modulation AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 103.

Modulation 4 bits. This indicates the modulation type of the downstream QAM Channel. The values of this field are:

0 = 64 QAM

1 = 256 QAM

2-15 = reserved

7.5.3.5 Downstream QAM channel J.83 Annex (ICRP, ICCN, SLI)

<table>
<tr><td>0</td><td></td><td>7</td><td>15</td><td></td><td>23</td><td></td><td>31</td></tr>
<tr><td>M</td><td>H</td><td>Resv</td><td>Length = 8</td><td colspan="4">Vendor ID = 4491</td></tr>
<tr><td colspan="4">Attribute Type = 104</td><td>L</td><td>TSID Group ID</td><td>Resv</td><td>J.83 Annex</td></tr>
</table>

Figure 7-39/J.212 – J.83 Annex AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 104.

J.83 4 bits. This indicates the Annex of [J.83] to be used on the downstream QAM Channel. The J.83 Annex defines the values of:

- Alpha (which in turn is dependent upon the choice of modulation)
- Forward Error Correction Frame Sync on/off
- Forward Error Correction Parity Bytes
- Trellis Coding enabled/disabled

The values of this field are:

0 = Annex A / DVB EN-300429

1 = Annex B

2 = Annex C

3-15 = reserved

Note that a particular EQAM might only support a subset of the attribute values listed above. For more information, refer to [J.210].

7.5.3.6 Downstream QAM channel symbol rate (ICRP, ICCN, SLI)

0		7	15		23	31
M	H	Resv	Length = 8 + 4 * N	Vendor ID = 4491		
Attribute Type = 105				L	TSID Group ID	Reserved
M				N		
M				N		

Figure 7-40/J.212 – Symbol Rate AVP

Length 10 bits. Set to 8 plus 4 times the number of M/N pairs.

Attribute Type 2 bytes. Set to 105.

M 2 bytes. Numerator of frequency to symbol coefficient.

N 2 bytes. Denominator of frequency to symbol coefficient.

The downstream symbol rate is set by choosing the appropriate values for M and N such that:

Symbol Rate (Msymb/s) = f * M/N

where f denotes the system master clock frequency.

In the ICRP message, the EQAM MUST list all the M and N value pairs that it has been configured to support. The EQAM MAY include a M/N pair, each with a value of 0xFFFF, to indicate that the EQAM has a variable symbol rate capability. In that case, the M-CMTS Core can request a value of M and N that the M-CMTS Core has been configured to use. Note that the lock bit is asserted when the M/N values are pre-configured, and de-asserted when a variable symbol rate capability is indicated.

If the EQAM has not been configured with M and N values and does not support a variable symbol rate option, then the EQAM MUST reject the session setup, and return the appropriate error code.

In the ICCN message, the M-CMTS Core MUST select one of those M and N pairs to indicate to the EQAM what symbol rate to use. The M-CMTS Core will subsequently use the values of M and N in the UCD MAC management message.

It should be noted that in operational scenarios where multiple downstreams may be used as sources of synchronization for CMs that drive a common upstream, those multiple downstreams will have to provide the same M/N values since the single M-CMTS Core upstream receiver can only operate with one M/N value.

7.5.3.7 Downstream QAM channel interleaver depth (ICRP, ICCN, SLI)

0		7	15		23	31
M	H	Resv	Length = 10	Vendor ID = 4491		
Attribute Type = 106				L	TSID Group ID	Reserved
I			J			

Figure 7-41/J.212 – Interleaver depth AVP

Length 10 bits. Set to 10.

Attribute Type 2 bytes. Set to 106.

I 1 byte. This indicates the I value of the interleaver depth of the downstream QAM Channel.

J 1 byte. This indicates the J value of the interleaver depth of the downstream QAM Channel.

7.5.3.8 Downstream QAM Channel RF mute (ICRP, ICCN)

0		7	15		23	31
M	H	Resv	Length = 8	Vendor ID = 4491		
Attribute Type = 107				L	TSID Group ID	QAM Ch Status

Figure 7-42/J.212 – RF mute AVP

Length 10 bits. Set to 8.

Attribute Type 2 bytes. Set to 107.

QAM Ch Status 1 byte. Bit 0 = 0 to unmute the RF output of the QAM Channel. Bit 0 = 1 to mute the RF output of the QAM Channel. Bits 7-1 are reserved. They should be set to 0 when transmitted and ignored when received.

8 DEPI forwarding plane

NOTE – This clause is normative.

The DEPI protocol uses the L2TPv3 protocol over IP either with or without a UDP header. The choice of whether to use a UDP header is based upon system configuration and is the same choice for both the control messages and data messages.

Within the L2TPv3 payload, there are two types of payloads used for DEPI. The first is a MPEG Transport Stream (D-MPT) based format and the second is a Packet Streaming Protocol (PSP) based format. The choice of which format to use is based upon the type of traffic being carried and the capabilities negotiated between the M-CMTS Core and the EQAM.

8.1 L2TPv3 transport packet format

This clause describes the various fields of the L2TPv3 packet as it applies to DEPI. The outer encapsulation of the L2TPv3 datagram is shown with a UDP header in Figure 8-1 and without a UDP header in Figure 8-2.

8.1.1 Data message with a UDP header

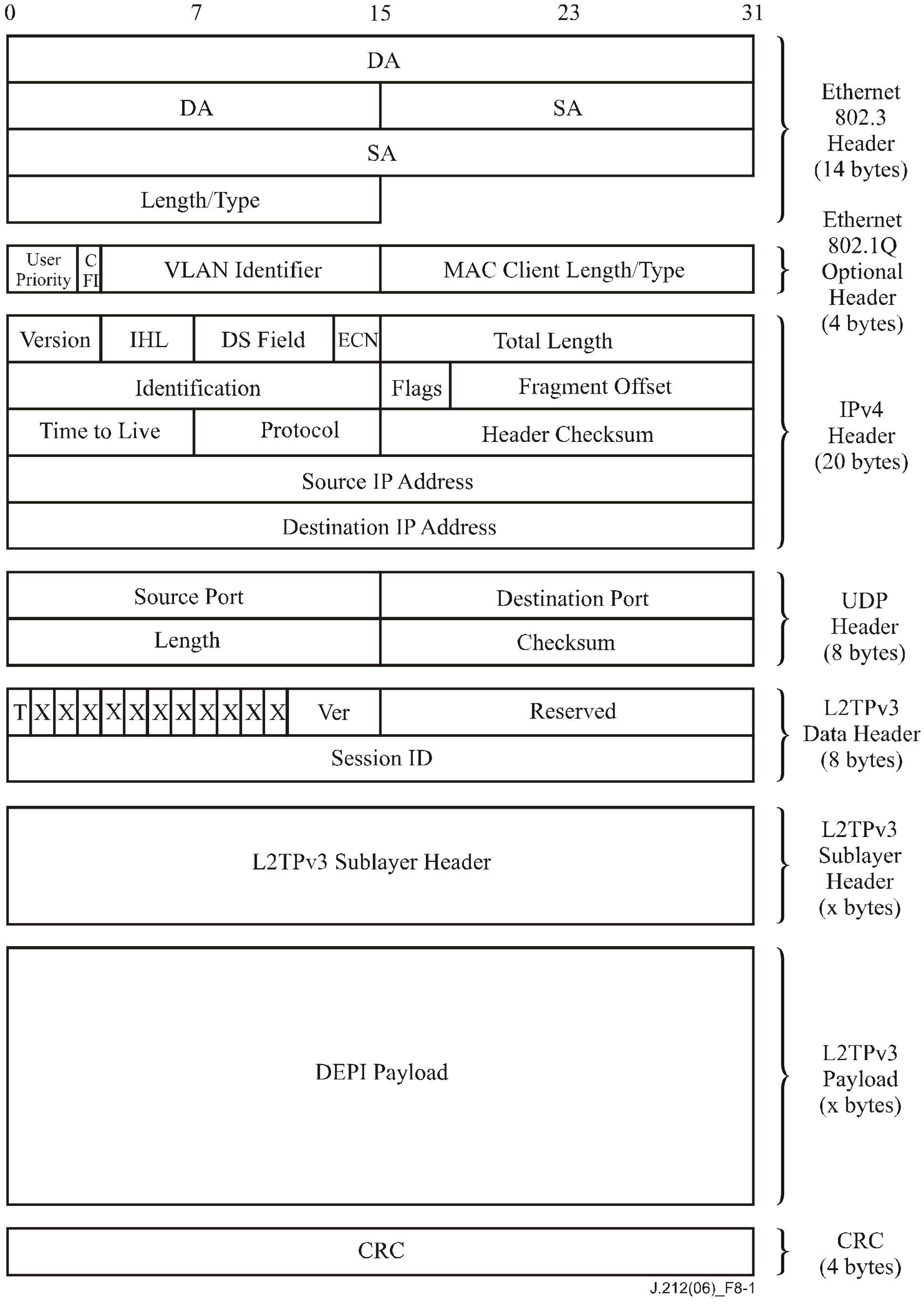

Figure 8-1/J.212 – L2TPv3 data packet outer encapsulation with UDP

8.1.2 Data message without a UDP header

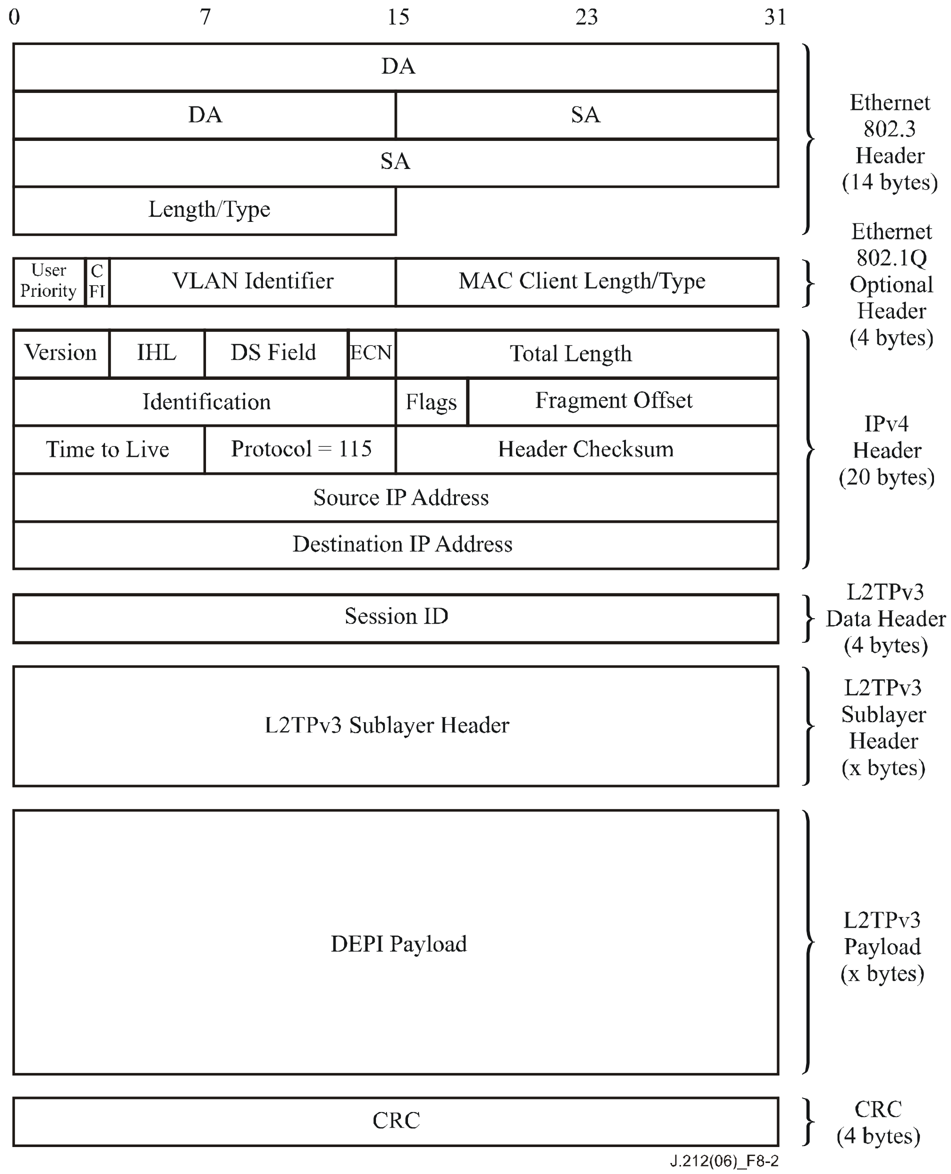

Figure 8-2/J.212 – L2TPv3 data packet outer encapsulation without UDP

8.1.3 Specific headers for data messages

8.1.3.1 L2TPv3 data header

The L2TPv3 fields as defined by [RFC-L2TPv3] are used as follows:

T Transport bit. 1 bit. Set to 0 to indicate that this is a data message.

X Reserved bits. 11 bits. Set to 0 by M-CMTS Core; ignored by EQAM.

Version Version Field. 4 bits. Set to 3.

Reservation Reserved field. 16 bits. Not used. Set to 0 by M-CMTS Core; ignored by EQAM.

Session ID Session Identifier. 32 bits. This value is negotiated by the L2TPv3 control plane.

The L2TPv3 cookie field is not required to be supported in DEPI.

The M-CMTS Core MUST support the L2TPv3 data header. The EQAM MUST support the L2TPv3 data header.

8.1.3.2 L2TPv3 DEPI sublayer header

DEPI supports two pseudowire types. The first type, known as D-MPT, is used for transporting MPEG packets. The second type, known as PSP, is used for transporting DOCSIS frames. Each pseudowire type has a unique L2TPv3 DEPI sublayer header format. The fields of these sublayer headers are defined in 8.2 and 8.3. Additionally, both pseudowire types support a latency measurement sublayer header, the fields of which are defined in 8.4.

The M-CMTS Core MUST use the appropriate DEPI Sublayer header for the pseudowire type of the DEPI session.

The EQAM MUST accept packets from the M-CMTS Core that contain the appropriate DEPI sublayer header for the negotiated pseudowire type. The EQAM MUST send a CDN message to tear down a session in which packets are received with the wrong pseudowire type. The EQAM MUST ignore packets received that do not comply with these L2TPv3 DEPI sublayer header definitions.

8.1.3.3 DEPI payload

The payload contains one or more segments. In D-MPT mode, each segment is a 188-byte MPEG packet. In PSP mode, a segment contains either a full DOCSIS frame or a partial DOCSIS frame.

8.2 DOCSIS MPT mode

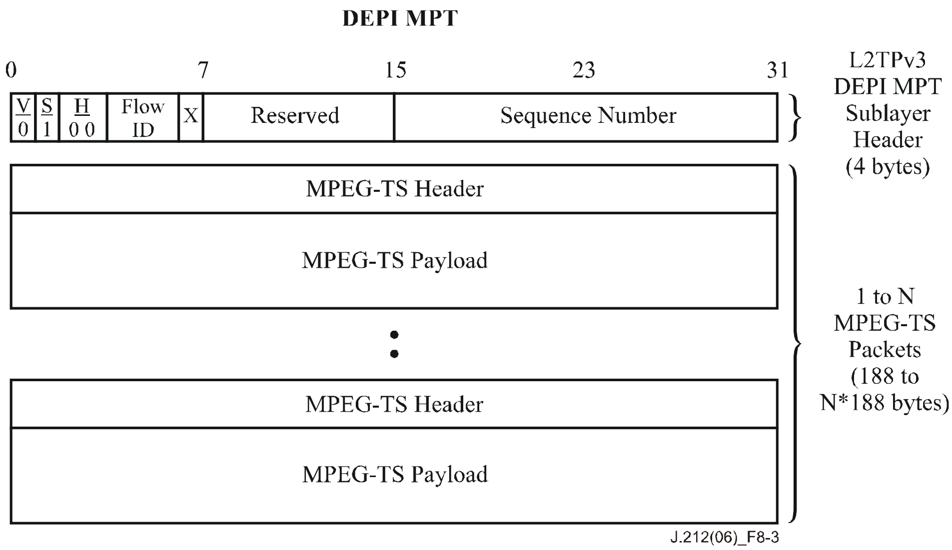

Figure 8-3/J.212 – DOCSIS MPT sublayer header and payload

V	1 bit. VCCV bit. Set to 0. Reserved for compatibility with [VCCV].
S	1 bit. Sequence bit. Set to 1 to indicate that the sequence number field is valid. Set to 0 to indicate that the sequence number field is not valid.
H	2 bits. Extended Header bits. Set to '00' to indicate no extended header.
X	1 bit. Reserved bit. Set to all zeros at the transmitter, ignored at the receiver.
Flow ID	3 bits. Flow Identifier.
Reserved	1 byte. Reserved field. Set to all zeros at the transmitter, ignored at the receiver.

Seq Num 2 bytes. Sequence Number. The sequence number increments by one for each data packet sent, and may be used by the receiver to detect packet loss. The initial value of the sequence number SHOULD be random (unpredictable).

The M-CMTS Core MUST NOT put stuffing bytes between the UDP header and the first MPEG-TS header or between consecutive MPEG-TS packets. The M-CMTS Core MUST support all bits in the MPT sublayer header.

The EQAM MUST accept one to seven MPEG-TS packets in a L2TPv3 payload when the path MTU is 1500 bytes in length. The length of an Ethernet frame containing seven MPEG-TS packets with L2TPv3 with a D-MPT L2TPv3 sublayer, UDP, IPv4, 802.1Q headers is 1378 bytes. If the EQAM, the M-CMTS Core, and the network between them all allow larger MTU sizes, the M-CMTS Core MAY increase the total number of MPEG-TS packets transmitted per L2TP packet.

The M-CMTS Core MAY insert null MPEG packets into the D-MPT stream. A null MPEG packet is 188 bytes in length with a reserved PID value of 0x1FFF as defined in [H.222.0]. The EQAM MAY discard these null MPEG packets. The EQAM is only required to support one flow for MPT mode.

8.3 PSP mode

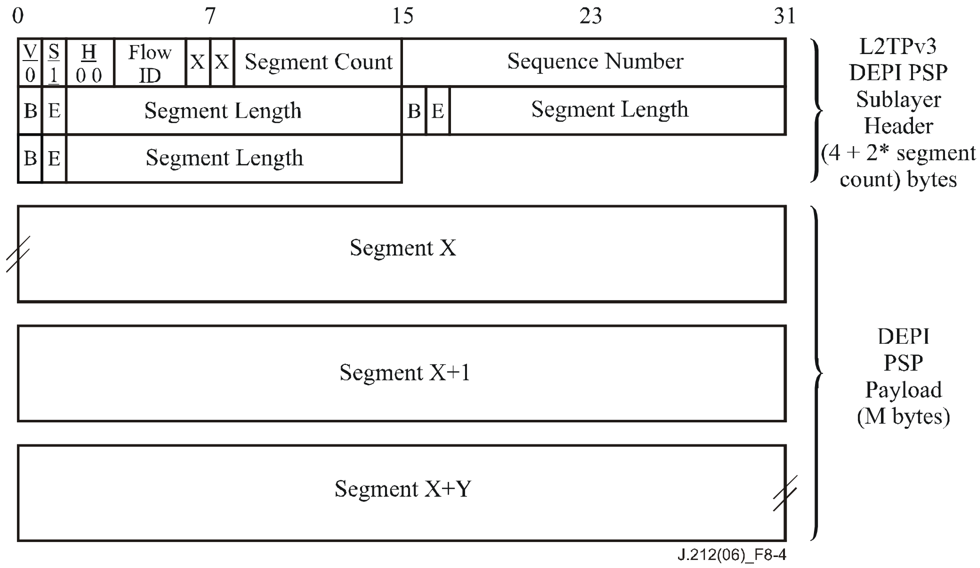

Figure 8-4/J.212 – DEPI PSP sublayer header and payload

V 1 bit. VCCV bit. Set to 0. Reserved for compatibility with [VCCV].

S 1 bit. Sequence bit. Set to 1 to indicate that the sequence number field is valid. Set to 0 to indicate that the sequence number field is not valid.

H 2 bits. Extended Header bits. Set to '00' to indicate no extended header.

X 1 bit. Reserved bit. Set to all zeros at the transmitter, ignored at the receiver.

Flow ID 3 bits. Flow Identifier.

Segment Count 7 bits. This is the number of segments in the DEPI PSP Payload, and this is also the number of 2 byte entries in the PSP Segment Table.

Seq Num 2 bytes. Sequence Number. The sequence number increments by one for each data packet sent, and may be used by the receiver to detect packet loss. The initial value of the sequence number SHOULD be random (unpredictable).

B Begin bit. 1 bit. Set to a 1 to indicate that the PSP Frame contains the beginning of a DOCSIS frame. Otherwise, set to 0.

E End bit. 1 bit. Set to a 1 to indicate that the PSP Frame contains the end of a DOCSIS frame. Otherwise, set to 0.

Segment Length 14 bits. Length of DEPI segment in bytes.

The Packet Streaming Protocol can take a series of DOCSIS frames, assemble them into a stream of back to back DOCSIS frames, and then split that stream up into PSP PDUs. In doing so, the first and last DOCSIS frame of a PSP PDU may be split into segments across PSP PDUs. DOCSIS frames which are not the first or last frame in a PSP PDU will not be split. A DOCSIS frame may be split into more than two segments and therefore may be spread across more than two PSP PDUs.

The Segment Table provides information on the contents in each of the subsequent PSP frames. This includes signifying if the frame is the beginning, middle, end, or an entire DOCSIS frame.

8.4 DEPI latency measurement (DLM) sublayer header

The DEPI Latency Measurement (DLM) Sublayer Header is used by the M-CMTS Core and the EQAM to measure the delay and latency of the CIN. This measurement is important as the transition network has the potential to affect the latency budgets already established in [J.122] for legacy DOCSIS 1.x and 2.0 devices. To perform this measurement, a packet is sent using the DEPI Latency Measurement Sublayer Header (shown in Figure 8-5) to which the receiver responds. This sublayer header is designed to be used within any active L2TPv3 session between the M-CMTS and the EQAM. It may be used with both MPT and PSP Pseudowire types. It is anticipated that this particular message exchange may occur between hardware mechanisms at either end of the DEPI interface.

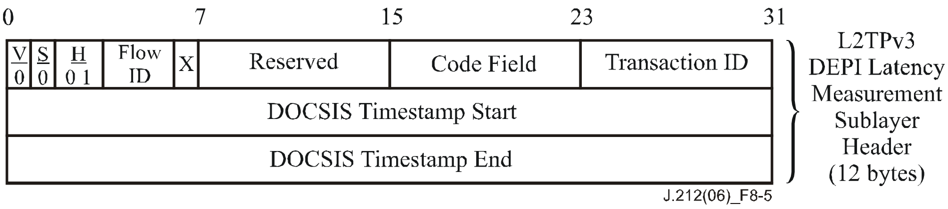

Figure 8-5/J.212 – DLM sublayer header

V 1 bit. VCCV bit. Set to 0. Reserved for compatibility with [VCCV].

S 1 bit. Sequence bit. Set to 0.

H 2 bits. Extended Header bits. Set to '01' to indicate the presence of the latency measurement extended header.

X 1 bit. Reserved bit. Set to all zeros at the transmitter, ignored at the receiver.

Flow ID 3 bits. Flow Identifier.

Reserved 1 byte. Reserved field. Set to all zeros at the transmitter, ignored at the receiver.

Code Field 1 byte. The permitted values are described below:

- A value of 0 indicates a DLM-EI-RQ (DLM EQAM Ingress Request) packet originated by the M-CMTS Core requesting a measurement to be made at reference point adjacent to the DEPI ingress port of the EQAM.
- A value of 1 indicates a DLM-EI-RP (DLM EQAM Ingress Reply) packet originated by the EQAM with a Timestamp End value calculated at reference point adjacent to the DEPI ingress port of the EQAM.

- A value of 2 indicates a DLM-EE-RQ (DLM EQAM Egress Request) packet originated by the M-CMTS core requesting a measurement to be made at reference point adjacent to the DEPI egress port of the EQAM.
- A value of 3 indicates a DLM-EE-RP (DLM EQAM Egress Reply) packet originated by the EQAM with a Timestamp End value calculated at reference point adjacent to the DEPI egress port of the EQAM.
- The values of 4 through 255 are reserved.

Transaction ID 1 byte. This is a unique ID assigned by the sender and returned by receiver. The transaction ID is unique per session.

Timestamp Start 4 bytes. Timestamp sent by sender.

Timestamp End 4 bytes. Timestamp existing at the receiver.

The M-CMTS Core MUST support the sending of a DLM-EI-RQ packet and the receiving of a DLM-EI-RP packet. The M-CMTS Core MAY support the sending of a DLM-EE-RQ packet and the receiving of a DLM-EE-RP packet. The CMTS Core (Sender) MUST send the DLM packet to the EQAM on an existing DEPI Session flow (either PSP or MPT). The M-CMTS Core MUST set the proper DSCP values for the DLM packet based on the DEPI Session being measured. The M-CMTS Core MUST provide a configuration mechanism to set the sampling interval through the M-CMTS Core MIB.

The M-CMTS Core MUST report delta measurements that exceed a configured threshold through the M-CMTS Core MIB. The M-CMTS Core MUST NOT send a DLM-EI-RQ or DLM-EE-RQ packet to a particular DEPI session, if there is already a DLM-EI-RQ or DLM-EE-RQ outstanding for that session, or if a DOCSIS frame on that flow has been segmented and the complete DOCSIS frame has not been sent. The M-CMTS Core MUST place the current DOCSIS 32-bit timestamp value in the Timestamp Start field of the message. The M-CMTS Core SHOULD use a timestamp value that is accurate to within a 100 µs of the current timestamp as derived from DTI.

The EQAM MUST support the receiving of a DLM-EI-RQ packet and the sending of a DLM-EI-RP packet. The EQAM MAY support the receiving of a DLM-EE-RQ packet and the sending of a DLM-EE-RP packet. The EQAM MUST support the use of the DLM packet within any active DEPI session. The EQAM is not required to support more than one concurrent latency measurement per session. The EQAM MUST ensure that timestamp value inserted in the Timestamp End field is accurate to within 100 µs of the current timestamp used for SYNC insertion/correction.

For DLM-EI-RQ packets, the EQAM MUST perform the Timestamp insertion prior to queuing the DEPI frame on the MPT or PSP QoS queues. For DLM-EE-RQ packets, if supported, the EQAM MUST perform the Timestamp insertion at the point where the SYNC message is originated. This is outlined in Figure 6-1. The EQAM MUST send the completed Latency Measurement Packet back to the M-CMTS Core that originated the measurement request, and do so with the EQAM DLM Timer value specified in Annex B.

An EQAM that does not support a DLM-EE-RQ packet MUST silently discard the DLM packet without generating a response packet.

8.5 M-CMTS core output rate

It is possible that if the M-CMTS Core was to send at data rate which exactly equalled the payload rate of the downstream QAM, and those packets were subject to jitter, that the EQAM would insert an MPEG-TS Null. Once the packets from the M-CMTS Core arrived, there would be an output queue delay that would equal the worst case jitter present at the DEPI input of the EQAM. That queuing delay would not be removed until the input to the EQAM was interrupted and the internal queue was drained. Further, the rate at which this delay is removed is related to the amount of jitter the input stream has, the peak input rate, and the maximum burst size of the input stream.

The M-CMTS Core MUST be able to rate limit the aggregate of all DEPI sessions, including any null MPEG packets that the M-CMTS Core may have inserted, that are destined to the same QAM Channel within an EQAM. The peak rate of this aggregate MUST be configurable to be a percentage of the QAM Channel payload. The burst size of this aggregate MUST be configurable. The default burst size of the aggregate MUST be the equivalent of three frames per DEPI session. (For a frame size of 1522 bytes, this would be 4566 bytes.)

Annex A

DEPI MTU

A.1 L2TPv3 lower layer payload size

Typically an interface calculates its default maximum payload size by asking the interface below it in the interface column what its maximum payload size and considering its own encapsulation. For example, by default, Ethernet has a frame size of 1518 (without VLANs). The Ethernet encapsulation is 18 bytes, leaving 1500 bytes of payload (MTU) to its upper layer. IP then subtracts the IP header size (typically 20 bytes) to arrive at 1480 bytes available to its upper layer. UDP must then subtract the UDP header size of 8 bytes, arriving at a payload of 1472 bytes. L2TPv3 must subtract the L2TPv3 Data Header size in bytes (8) and calculate its payload as the remainder (typically 1464). This remainder is the defined maximum payload size for an L2TPv3 session that would exist over Ethernet/IP/UDP.

A.2 Maximum frame size for DEPI

This annex documents the maximum frame size of DEPI when a PSP Pseudowire is used without fragmenting or concatenation.

Table A.1/J.212 – MTU of DEPI

<table>
<tr><th colspan="4">Field</th><th>Size</th></tr>
<tr><td rowspan="12">DEPI frame</td><td colspan="3">Ethernet Header</td><td>14 bytes</td></tr>
<tr><td colspan="3">802.1Q Header</td><td>4 bytes</td></tr>
<tr><td rowspan="9">DEPI MTU</td><td colspan="2">IPv4 Header*</td><td>20 bytes</td></tr>
<tr><td colspan="2">UDP Header</td><td>8 bytes</td></tr>
<tr><td colspan="2">L2TPv3 Header</td><td>8 bytes</td></tr>
<tr><td colspan="2">DEPI-PSP Header</td><td>6 bytes</td></tr>
<tr><td rowspan="5">DOCSIS frame</td><td>DOCSIS Header</td><td>6-246 bytes</td></tr>
<tr><td>Ethernet Header</td><td>14 bytes</td></tr>
<tr><td>802.1Q Header</td><td>4 bytes</td></tr>
<tr><td>Ethernet PDU</td><td>1500 bytes</td></tr>
<tr><td>Ethernet CRC</td><td>4 bytes</td></tr>
<tr><td colspan="3">Ethernet CRC</td><td>4 bytes</td></tr>
<tr><td colspan="4">Total</td><td>1592-1832 bytes</td></tr>
<tr><td colspan="5">* Currently, this Recommendation only requires support for IPv4. If IPv6 were to be used for transport, then value should be increased by an additional 20 bytes plus the length of any IPv6 Extension Headers.</td></tr>
</table>

A.3 Path MTU discovery

Path MTU Discovery relies on the fact that the network elements between the M-CMTS Core and the EQAM all support this protocol [RFC-MTU]. If these network elements do not support Path MTU Discovery, then this mechanism cannot be used and the static configuration option should be used instead.

Path MTU Discovery works when the path between the M-CMTS Core and the EQAM is less than the total frame size generated when using the payload size negotiated during L2TPv3 session establishment. In this circumstance, when the M-CMTS Core sends packets larger than the network can support, then network elements between the M-CMTS Core and the EQAM will generate an ICMP Destination Unreachable message with the code "fragmentation needed and DF set", as per [RFC-IP], toward the source of the tunnelled packet. This error message includes the IP header and UDP header. The M-CMTS Core and the EQAM should have a way to map a UDP port to an L2TP Session ID. The M-CMTS Core and EQAM should reduce the Max Payload of the session referenced in the ICMP Destination Unreachable message to the size requested in the Next-Hop MTU field of the message. The M-CMTS Core and EQAM may periodically attempt to increase the Max Payload of the session to its negotiated maximum and restart this process in case the path through the network has changed and large MTUs are allowed. This technique is described in [RFC-MTU]. The Max Payload size learned through this process will never be greater than the negotiated maximum learned during session establishment.

Annex B

Parameters and constants

Table B.1/J.212 – Parameters and constants

System	Name	Time reference	Minimum value	Default value	Maximum value
M-CMTS Core, EQAM	SYNC Interval	Nominal time between transmission of SYNC messages			200 ms
M-CMTS Core, EQAM	HELLO Timer	L2TPv3 section 4.4		60 s	
M-CMTS Core, EQAM	Control Message Timeout	L2TPv3 section 4.2	First attempt: 1 s, then second attempt 2 s, then third attempt: 4 s, then each additional attempt: 8 s.		
M-CMTS Core, EQAM	Control Message Retry Count	L2TPv3 section 4.2			10
M-CMTS Core, EQAM	StopCCN Timeout	L2TPv3 section 3.3.2	31 s		
EQAM	EQAM DLM Timer	Time between receipt and retransmission of a DLM packet.			100 ms

Appendix I

DEPI and DOCSIS system performance

I.1 Introduction

The M-CMTS architecture provides interoperability among different types of equipment performing various sub-functions of a complete CMTS. In previous architectures, all these functions were contained in a single physical card or chassis, so that the delays in communication between sub-functions were effectively zero. M-CMTS introduces non-zero delays between sub-functions. In some instances these delays may impact system performance. Specifically, interposing a CIN between the M-CMTS Core and the EQAM increases the *round-trip time* of the DOCSIS system.

I.2 Round-trip time and performance

Broadly speaking, round-trip time is the time from a CM's request to the time the CM transmits the data corresponding to that request. The quicker all this happens, the sooner the CM can transmit another request (e.g., a piggyback request), thereby transmitting more data, etc.

Round-trip time limits the performance of a single modem by limiting the number of grants the modem can receive in a given time. For instance, if the system round-trip time is 10 ms, it is not possible for a modem to receive more than 100 grants per second. If every grant were the size of the modem's allowed maximum burst size (as configured by, e.g., Maximum Concatenated Burst, [J.122]), an upper bound on the modem's performance could be found by the following simple calculation:

$$\text{max throughput (bits/s)} = \text{max burst (bits)} \cdot 1/[\text{round trip time (s)}]$$

In practice, due to the need to share bandwidth among many users and services, the maximum burst size must be limited to reasonable values, and the CMTS generally cannot grant the maximum burst size in every grant, even if a modem requests it.

A longer round-trip time also increases the access latency seen by a single modem – i.e., the time it takes for a modem to gain access to the upstream to begin transmission of new data after an idle period. Conversely, if reducing round-trip time enables higher throughput to be achieved or speeds the opening of the TCP window, the modem's transactions (e.g., download of an FTP file or HTTP web page) may be completed more quickly (assuming plant bandwidth is available to do so). These factors may in turn affect the overall bandwidth efficiency of the system.

I.3 Elements of round-trip time

It is convenient to begin measurement of round-trip time at the instant when a modem begins transmission of a request. Round-trip time can then be measured as the time from this initial request to the instant when the modem begins transmission of its next request. These events can easily be captured on a network sniffer.

The elements of round-trip time may be categorized as follows:

– *Upstream propagation delay*: Time occupied by plant delays in the upstream direction.

– *Upstream reception and request parsing time*: Time from the start of burst arrival at the CMTS until reception and parsing of the request is complete. Some CMTSs require that the entire burst be received before it can be parsed. Others can recognize piggyback requests near the beginning of a burst, even if the end of that burst has not yet arrived. If FEC is enabled for the burst, at least one full FEC block must be received and decoded before any MAC-layer parsing can be performed.

– *Scheduler queueing and processing delay*: Time from arrival of the request at the scheduler until completion of the MAP message containing a grant for the request. If the request arrives just after the scheduler has finished creating a MAP, the request is delayed by the time interval until the next MAP. On the other hand, if the request arrives just before the scheduler finishes creating a MAP, the request may see nearly zero delay. In general, the actual queueing delay is a random variable between zero and the maximum MAP interval. Under some lab conditions involving only one or a few CMs, this delay may appear to be constant, but this cannot generally be assumed in a real system. Some scheduler implementations may vary the MAP interval to optimize this delay.

  The time required for the scheduler to make scheduling decisions and actually create the MAP message is also included here. This factor is highly implementation-dependent.

– *MAP delivery time (to the EQAM DOCSIS PHY layer)*: The time from the completion of the MAP message creation, to delivery of the MAP to the PHY layer. This includes any time consumed by the M-CMTS Core's MAC function; encapsulation of the MAP into a DEPI packet; queueing and transmission of the DEPI packet at the egress of the M-CMTS Core; delay and jitter of the CIN; queueing and processing delays inside the EQAM; and any delay in inserting the MAP into the MPEG-encapsulated DOCSIS stream (e.g., due to the need to wait for a previous packet to complete transmission).

– *Downstream physical-layer delays*: This includes the latency of the downstream modulator, downstream interleaver delay, and physical propagation delay between the EQAM and the CM.

– *CM MAP processing time*: The time from arrival of the first bit of the MAP at the CM, until the MAP becomes effective. The minimum value is specified in the Relative Processing Delay clause of [J.122]. It accounts for all internal CM processing delays.

– *Time until grant*: If the first grant in the MAP is not to this CM, the CM's actual transmission will be "delayed" until the actual time of the grant.

– *Margin*: In practice, the CMTS cannot precisely control all delays to guarantee that MAPs arrive at the modem at exactly the right instant. Thus, the CMTS must add margin to account for worst-case propagation delays to the farthest modems, variations in MAP creation time and CIN delays, etc.

Table I.1 lists example values for the round-trip time components described above. These values are given ONLY by way of example and should not be interpreted as typical values applying to any particular system and/or implementation.

Table I.1/J.212 – DOCSIS Request-Grant round trip worksheet

(All time values in usec)

Delay source:	**subtotal**	**total**	**Remarks**
Upstream propagation time		**800**	
Physical HFC plant delay	800		apprx 100 miles
Upstream reception/parsing time		**469**	
upstream PHY processing time	369		time for 1 FEC block (236,200) at 5.12 Mbit/s
upstream MAC processing time	100		
Scheduler queueing and processing		**1000**	
MAP delivery to PHY layer		**1338**	
DEPI packetization at Core	383	**plus CIN delay**	for MPT mode, transmit time of 7 MPEG packets
CIN delay	varies		
EQAM latency	500		per J.212
Queueing behind max-length packet during MPEG encaps.	455		transmit time of 1518 bytes (64QAM Annex B)
Downstream transmission time		**1841**	
MAP duration on wire	61		200B MAP at 64QAM
downstream FEC/interleaver delay	980		(I, J) = (32,4) at 64QAM
downstream prop. delay	800		apprx 100 miles
MAP Advance Margin		**500**	
CM processing time		**200**	TDMA, no byte interleaving
Total round trip time		**6148 plus CIN delay**	

I.4 CIN characteristics

In an M-CMTS system, the CIN adds delay which may be a noticeable contributor to the total system round-trip time. The CIN is configured and managed by the operator. Its extent may vary widely from one system to the next. It may be as simple as a short Ethernet cable between one M-CMTS Core and one EdgeQAM. Alternatively, it may be an IP network consisting of multiple switch and/or router hops which is shared with other (non-DEPI) traffic to and from other nodes (e.g., IP-encapsulated video from VoD servers).

All but the most trivial of packet-based networks will experience variations in delay. These variations arise due to variability in length and arrival time of packets, changes in instantaneous loading, queuing of packets within network elements, differences in QoS parameters applied to different packets, and other factors. Thus, network delay is often modelled as a random variable.

One common way of describing network delay is as a sum of two components: a "typical" delay which is treated as a constant, plus a "jitter" term which may differ from packet to packet. Another approach is to describe the network as having a "minimum" delay and a "maximum" or "worst-case" delay, with delays on individual packets distributed between the two extremes in some way. These approaches are often convenient ways of approximating network behaviour.

One approach to rigorously modelling the network involves determining a Cumulative Distribution Function (CDF) of packet delays. This CDF is a function or graph in which the x-axis shows delay D_0, while the y-axis shows the probability that the actual delay D on any particular packet is less than D_0. To be useful, this CDF may be limited in scope: for example, it may be from some specified source node to a particular destination node and for a particular class of traffic. It can then be stated that a certain fraction of applicable packets will experience delay less than a particular amount (e.g., "99.995% of packets will experience delay less than 2 ms", or some similar statement). In some networks, it may be impossible to guarantee an upper bound on delay for 100% of all packets.

In addition to the CIN itself, the "network" may include switching and/or queuing that occurs inside an M-CMTS Core or an EQAM. For example, an M-CMTS Core supporting multiple downstream QAM channels may have multiple Gigabit Ethernet output ports, and thus might include an internal "switch" to connect various DEPI flows to different output ports. If present, this switching function could also add implementation-dependent queuing delays which could be modelled as part of the CIN.

To aid in characterizing the CIN, DEPI specifies the Latency Measurement Subheader (clause 8.4). This subheader takes advantage of the fact that both the M-CMTS Core and the EQAM contain a DTI interface which provides a common, high precision clock. The M-CMTS Core transmits a Latency Measurement Packet, on any given DEPI session, containing its current timestamp value. Upon receipt of this message, the EQAM adds its own timestamp value and returns both values to the M-CMTS Core. The result is a single measurement of CIN delay (plus any M-CMTS Core egress and EQAM input queueing delays in the path between timestamp insertion points) between the CMTS and the EQAM. The CMTS may make use of this information in several different ways including CDF calculations or adjusting specific DOCSIS network parameters.

Oftentimes the design of the CIN is the one variable in the M-CMTS architecture that the operator has complete control over. Decisions made here regarding the size and number of network hops will have a direct effect on overall performance of the DOCSIS network.

I.5 Queueing delays in network elements

Many layer 2 and layer 3 switch products today support line-rate switching in a non-blocking fashion. "Non-blocking" means that packets being switched across particular source and destination ports will not interfere with packets being switched across different source and destination ports. If all traffic within the switch is non-blocking (as is typically the case in laboratory tests per IETF RFCs), switching will be very fast. (Exact values for this "intrinsic switch delay" are implementation-specific.) However, if packets simultaneously arriving from several different source ports are directed to a single destination port, these packets will suffer queueing delay within the switch.

For very simple switched topologies where only DEPI traffic is present, worst-case queuing delay at a given destination port may be calculated as a function of packet size and number of source ports switching to that destination port. This calculation only applies to a single node within the network.

If every source port delivered a packet to the destination port at the same instant, the last of these packets to actually exit the destination port would be delayed by the time required to transmit all preceding packets. The calculation is as follows:

Max delay (s) = (# of source ports – 1) · packet size (bits) / line rate (bits/s)

The result of this calculation plus the intrinsic switch delay (which may, by way of example, be on the order of 10-100 μs) gives a reasonable upper bound on the total worst-case delay through the switch when the traffic at the source ports is rate-limited by the M-CMTS Core. As a result, the aggregate traffic out the destination port of the switch does not exceed the port's capacity. This would typically be the case for a switch internal to a CMTS or EQAM, or one that carries only DEPI traffic in a one- or two-hop network. For more complex systems, additional queue build-up could occur and the latency through the switch could easily exceed 1 ms. As such, it is important to apply QoS policies to any traffic that is latency sensitive so that it might bypass queue build-ups. More information on the use of QoS in the CIN is provided in the clauses below on "Traffic Prioritization" and "PSP Mode".

In recent years products have evolved so that the distinction between a "switch" and a "router" is not always clear. A "switch" may incorporate extensive QoS functionality, while a "router" may perform large amounts of processing in hardware and be as fast as some switches. Thus, in more complex CIN topologies, delays may vary widely and do not readily lend themselves to calculation. Even routers or switches capable of line-rate performance are subject to congestion and internal queueing and processing delays which may vary depending on the load, traffic type, QoS functions supported, and many other conditions.

I.6 Traffic prioritization and network delays

In general, it is possible to reduce network delay for certain packets by instructing routers and/or switches to give these packets higher priority than other packets. The mechanism for accomplishing this is usually the IEEE 802.1Q VLAN tag for Ethernet, or the DSCP field and associated per-hop behaviour for IP.

In the simplest case, packets tagged as high priority are sent by the router or switch ahead of all other packets of lower priority which may already be queued within the device. More complex behaviours also exist, but fundamentally, they all reduce delay on the high-priority packets at the expense of lower-priority packets. In order for this to be effective, high-priority packets must represent a relatively small fraction of the total traffic on the network.

In a CIN, traffic prioritization may be useful for reducing round-trip time when PSP mode is in use by prioritizing DEPI flows containing MAPs ahead of flows containing other types of data. This concept may be extended to add more priority levels (e.g., for voice vs best-effort data) if the M-CMTS Core and EQAM both support the desired number of PSP flows. In MPT mode, this approach cannot be used since all data for a given QAM channel, regardless of type or priority, are carried in a single flow.

If the CIN also carries non-DEPI traffic, it may be possible to reduce DEPI delays by giving DEPI higher priority. However, this will result in increased delay on the non-DEPI packets. The consequences of this should be carefully considered by the operator.

I.7 Queue persistence in a DEPI flow

DEPI requires the M-CMTS Core to rate-shape its output to match the rate of the EQAM's RF output. If the path between the two devices had fixed delay (i.e., zero jitter), DEPI data arriving at the EQAM would immediately be transmitted on the HFC network. In practice, a real DEPI path will always contain variable delays (i.e., jitter). The presence of jitter in this path requires data to be queued inside the EQAM.

To see this, consider a simple example involving a single MPT-mode DEPI flow. Suppose that packets are being sent on a 1 Gbit/s Ethernet network. The flow itself occupies only about 40 Mbit/s and is rate-shaped, so there will be "space" on the network between packets. The interval on the CIN between the start of one packet and the start of the next packet will correspond to the same interval on the EQAM's RF output. As long as every packet is being delivered from M-CMTS Core to EQAM with constant delay D, each packet will arrive at the EQAM just as the EQAM completes transmission of the previous packet. However, suppose a single packet P_0 experiences higher-than-normal delay $D + J$. The time on the network between this packet and the previous one (packet P_{-1}) is extended, while the time between this packet and the following one (packet P_1) is compressed (since P_1 experiences only delay D). Thus, the EQAM completes transmission of P_{-1} before P_0 arrives. Since the EQAM now has no DOCSIS data to send, it must transmit MPEG Null packets until P_0 arrives. At that time, it must complete the MPEG Null packet it is currently transmitting and then begin transmission of P_0. However, since the time between P_0 and P_1 has been shortened, P_1 will arrive at the EQAM before transmission of P_0 has completed. P_1 must now be queued by the EQAM and will experience queueing delay J', where:

$$J' = ceiling(J/T_{MPEG}), \text{ where } T_{MPEG} = \text{duration of one MPEG packet}$$

i.e., the amount of delay that was added to P_0 due to jitter on the network rounded up to the next MPEG packet boundary. In addition, as long as the M-CMTS Core continues to output data at the exact rate of the EQAM's RF output, all subsequent packets in the flow will also experience queueing delay J'. This can be a serious problem for an M-CMTS system, since even if jitter J occurs very infrequently, once it does occur, the additional queueing delay could persist indefinitely.

To prevent this infinite queue persistence, DEPI requires that the M-CMTS Core be capable of rate-shaping its output to some configurable fraction of the EQAM's actual output data rate. This does not prevent queueing delays from occurring due to network jitter events, but it does mean that, in the aggregate, the EQAM can drain the queue (by transmitting data) faster than it fills.

It is possible to calculate the time T required for the queue depth to return to zero after a jitter event increases the delay on a single packet from D to $D + J$. Suppose that the CMTS is rate-shaping to a rate r which is some "rate-shaping ratio" ρ of the EQAM's output rate R (i.e., $\rho = r/R$). During time J', it will send MPEG Null packets when, had the jitter event not occurred, it would have ordinarily received from the M-CMTS Core and immediately transmitted on the wire $J' \cdot r$ bits of data. This number of bits will instead be queued inside the EQAM and, once available, transmitted and hence removed from the queue at rate R. Meanwhile, data continues to arrive and be added to the queue at rate r. Thus, over a time interval T, the total queue size decreases by $T \cdot (R - r)$. The queue size will return to zero when:

$$J' \cdot r = T \cdot (R - r)$$

Solving for T and expressing in terms of ρ (the ratio configured in the M-CMTS Core) gives:

$$T = J' \cdot \rho / (1 - \rho)$$

For example, if network jitter on a single packet results in 2 ms of queueing delay (J'), and the M-CMTS Core continues to send traffic at a rate equal to 98% of the EQAM's actual output rate, the time to drain the resulting queue at the EQAM will be (2 ms) [0.98/(1 – 0.98)], or 98 ms. This time will be reduced if the CMTS does not have enough downstream traffic available during this period to "fill the pipe" to 98% of nominal. Conversely, if another network jitter event occurs before this time has elapsed, the queue size will again suddenly increase. The impacts of continuous jitter is not additive. Rather, the worst-case queue build-up will equal the worst-case jitter. If network jitter events are frequent enough and traffic load remains relatively high, the queue may never completely drain.

For MPT operation, a single flow contains all DOCSIS data for a given QAM channel, including MAP messages. The above calculation of *T* can be useful in estimating how many MAP messages will experience increased delay due to network jitter events. If a MAP interval is typically 2 ms, then during time $T = 98$ ms, 49 MAP messages on each upstream channel will see increased delay as a result of each network jitter event adding $J = 2$ ms delay to a single DEPI packet (whether or not that packet contains a MAP message). Note that this calculation may not be exact, since the CMTS may vary the MAP interval.

Depending on how much MAP Advance margin the CMTS has provided (see round-trip time calculations in Clause I.3), some of these 49 MAPs may still be usable by CMs. To account for this, the amount of margin provided may be subtracted from *J* and the result used to calculate *T*, i.e., if margin is represented by μ, then when a jitter event occurs which adds jitter *J*, the number of MAPs on each upstream channel which are delayed enough to be lost is given by:

$$\#\text{MAPs lost} = [(J - \mu) \cdot \rho / (1 - \rho)] / (\text{MAP interval})$$

In the current example, if the CMTS has added 500 μs of MAP Advance margin to each MAP, the queue would drain to the point where MAPs are usable by CMs after $(J - 500\ \mu s) \cdot [\rho/(1 - \rho)] = (1.5\ ms)\ (49) = 73.5$ ms. If the MAP interval is 2 ms, only 37 MAPs would be lost for each such network jitter event, as opposed to the 49 MAPs that would be lost if zero margin had been provided. These 37 lost MAPs translate into $(37 \cdot 2\ ms) = 74$ ms of time during which CMs cannot transmit as a result.

To minimize the number of MAPs lost due to network jitter when MPT mode is in use, the scheduler should add margin sufficient to give the desired reliability level, based on the CDF describing the network. For instance, suppose it has been determined that 99.9999% of the time (i.e., $1 - 10^{-6}$), the network delay will be less than D_1. The scheduler could then add a margin of D_1 when creating MAPs. It could then be said that an event resulting in lost MAPs will occur once for every 10^6 packets sent. On a relatively loaded downstream carrying, e.g., 50,000 packets per second, this would result in a "lost MAP event" about once every 20 seconds. The number of MAPs lost for each event, and the corresponding time for which CMs cannot transmit due to missing MAPs, would depend on the amount by which the actual delay exceeded D_1. In the previous example, if the MAP Advance margin had been set to 2 ms to account for the worst case CIN jitter, then the 49 MAP messages that were delayed would all still be usable.

The effect of this added margin is to increase system round-trip time by D_1 for all packets. However, it helps to minimize the negative system impacts of frequently lost MAPs.

I.8 PSP mode

DEPI provides PSP mode as a way of mitigating CIN delays affecting round-trip time. PSP mode does not reduce the delay of the network itself. Instead, it separates MAP messages into a separate DEPI flow and uses traffic prioritization (clause I.6) to minimize delay on this flow. PSP mode also allows the EQAM to transmit MAP messages ahead of lower-priority DOCSIS data, even if the lower-priority data is already queued within the EQAM. These two modifications in turn mean that, in PSP mode:

a) MAP messages are only delayed in the network by jitter events directly affecting packets containing MAPs; and

b) MAP messages can only be queued behind (and hence delayed by) other MAPs. Thus, the CDF for network delay for MAP messages in PSP mode may show significantly lower delays than a similar CDF for MPT mode on the same network.

Use of PSP mode should be considered by the operator whenever the impact of the CIN on system round-trip time is a concern. More specifically, PSP mode should be considered if the amount of margin which must be added to round-trip time to give a low enough rate of lost MAPs is unacceptably large. What exactly constitutes "unacceptably large" depends on many factors and may be subjective. For example, a system in which the plant is physically short has less propagation delay and therefore may tolerate larger CIN delays. As another example, on a plant where high modem performance is desired but the presence of UGS traffic or other factors compel the use of a small maximum burst size, even relatively small CIN delays may degrade performance.

In general, a very small CIN (one or possibly two switch hops) probably has little impact on system round-trip time, while a CIN which has multiple hops and/or on which non-DEPI traffic is present probably has a much more significant impact on system round-trip time. To accurately predict the actual improvement in round-trip time that can be gained by using PSP mode on a given CIN, it is necessary to characterize the delays of that CIN for different DSCPs.

Appendix II

Early adoption and evolving use of EQAM devices

Prior to the development of the M-CMTS Recommendations, EQAM devices in existing HFC digital CATV systems were responsible for video processing functions such as multiplexing and modulation (QAM). The M-CMTS Recommendations place additional requirements on EQAM devices to support DOCSIS functionality.

It is possible that certain EQAM devices may be able to support a subset of the DOCSIS functionality by adhering to parts of the M-CMTS Recommendations (including DEPI). These devices would not be considered M-CMTS compliant EQAMs, but nonetheless might be a valuable component of a transition strategy for operators.

The development of initial M-CMTS EQAM devices that are first on the market (early adopters) will be divided into two categories. The first category (category A) is EQAM vendors with existing video-only EQAM products (no support for the DTI). These vendors may want to offer their customers a software upgrade path for already deployed EQAM products to support bonded DOCSIS channels that do not contain timing information (no SYNC messages). The second category (category B) is EQAM vendors that intend to initially offer new EQAM products that have fully M-CMTS capable hardware including the DTI, but where full support for M-CMTS features is not offered initially. In both categories, support for additional M-CMTS features would typically be added via software upgrade using a phased approach.

II.1 EQAM development: Category A (no DTI)

Vendors falling into Category A will provide software upgrades to allow their existing deployed video EQAM products to be upgraded in the field to support bonded DOCSIS channels that do not contain DOCSIS timing information (no SYNC messages). This provides a level of investment protection for operators with already deployed EQAM products in the field. These upgraded video EQAM products would not support the DTI and would thus not be M-CMTS compliant, but they could be leveraged by operators when deploying DOCSIS channel bonding. This would help to minimize the number of new EQAM devices that the operators would be required to purchase when upgrading their systems to support DOCSIS channel bonding.

Upgrade Phases for Category A (no DTI):

– Phase 1: S/W upgrade existing "video" EQAM to support {L2TP control plane and DEPI MPT mode}
– Phase 2: Add (to phase 1) support for L2TP data plane.
– Phase 3: Add (to phase 2) support for ERMI.

II.2 EQAM development: Category B (with DTI)

Vendors falling into Category B will initially develop EQAM devices with hardware that has the capability of achieving full M-CMTS compliance. This mainly includes the physical DTI interface hardware. Support for M-CMTS features will then be phased in by the vendor over time by offering software upgrades. This allows operators to purchase M-CMTS EQAM devices up front knowing that only software upgrades are required to eventually achieve full M-CMTS compliance. An EQAM with M-CMTS compliant hardware would support the DTI allowing it to operate in either SCDMA mode or ATDMA mode.

Upgrade Phases for Category B (with DTI):

– Phase 1: Support DTI, L2TP control plane, and DEPI MPT mode.
– Phase 2: Add (to phase 1) support for L2TP data plane.
– Phase 3: Add (to phase 2) support for ERMI.
– Phase 4: Add (to phase 3) support for DEPI: PSP mode.

II.3 Possible M-CMTS feature phasing

In order to minimize time to market, EQAM vendors will likely add in support for M-CMTS system features using a phased approach by offering software upgrades. In both development categories described above (A and B), the software upgrade path towards full M-CMTS compliance may be very similar.

Development: Phase 1

The initial M-CMTS products will support the L2TP control plane, and they will likely support the DEPI MPT mode only (not DEPI PSP mode). EQAMs with DTI hardware may or may not support the DTI. If the DTI is not supported, the EQAM can only be used for video applications and or to support DOCSIS bonded channels with no DOCSIS timing information.

The L2TP control plane is required to allow the M-CMTS core to provision and configure the EQAM, so support for the L2TP control plane is mandatory. It is assumed that the L2TP control plane can be developed in software with no required hardware changes to either an existing "video" EQAM or an EQAM with M-CMTS capable hardware.

Support for the L2TP data plane may require hardware assistance or more extensive software development than the L2TP control plane, so it is likely that EQAM vendors would defer support of the L2TP data plane until later phases. Existing network processors typically do not support the L2TP, so any L2TP data plane processing in early adopter EQAM products would require custom software development. It should also be noted that the M-CMTS core will always include the L2TP data plane, but since the L2TP layers are fixed length, an early adopter EQAM product could be designed to simply skip over the L2TP headers when receiving DEPI data. In this case the UDP layer is required to support the routing of data packets to the appropriate QAM channel outputs.

It is likely that the early adopter EQAM devices will initially support the DEPI MPT mode, but not the DEPI PSP mode. The DEPI MPT mode delivers MPEG packets to the EQAM, and requires processing in the EQAM that is similar to the processing that is performed on video MPEG packets. The DEPI PSP mode requires the EQAM to perform additional functions such as transmission convergence (MPEG packetization) and DOCSIS SYNC insertion so it will likely be deferred to a later phase.

Development: Phase 2

It is likely that the next step (phase 2) after offering initial EQAM products as described above in phase 1 would be to offer support for the L2TP data plane. Support of the L2TP data plane could be realized in software, or network processors that support the L2TP could be leveraged if available.

The use of L2TP capable network processors could be achieved via software upgrade and or an FPGA firmware upgrade, or hardware changes could be required.

Development: Phase 3

Support for the M-CMTS ERMI will also be phased in via software upgrade. The M-CMTS system is designed to function without an ERM (without the use of the ERMI) by relying on the L2TP control plane and static system configuration. In this case, EQAM resources are manually (and statically) assigned to the M-CMTS core. As a result, support for the ERMI in the EQAM can be deferred until a later phase.

Development: Phase 4

The final phase of M-CMTS EQAM development is likely to be software upgrades to support the DEPI PSP mode. The DEPI PSP mode requires the EQAM to perform DOCSIS specific functions such as transmission convergence (MPEG packetization) and DOCSIS SYNC insertion. In addition, the data plane processing for PSP mode is more complex than MPT mode in that it requires support of variable size PDUs with multiple DOCSIS frames per PDU. As a result, it is likely that support for DEPI PSP mode will be offered by EQAM vendors in late or final phases of development. Support for DEPI PSP mode would only be applicable to EQAM products that support the DTI (category B EQAM products).

II.4 Optional UDP layer

The UDP layer is optional in the L2TP. It is desirable to use the L2TP protocol without the UDP layer since the L2TP provides built in mechanisms (session IDs) to bind payload data to software functions analogous to the use of a UDP port. If two devices that fully support the L2TP are communicating with each other (using L2TP), then there is no need to use a UDP layer below the L2TP layer since the session ID field provides the necessary data routing function. If one of the devices does not support the L2TP in the data plane, then the UDP layer can be used to provide the data routing function.

Since off-the-shelf network processors that support the L2TP will likely not be available when EQAM vendor begin M-CMTS development, early adopter EQAM products will likely not support the L2TP in the data plane. As a result, early adopter EQAM products will likely make use of the UDP layer to provide the data routing function since existing network processors are capable of parsing packets up to and including the transport layer in hardware.

Eventually, if/when L2TP capable network processors become available, or alternatively if EQAM vendors develop support for the L2TP layer in the data plane, the devices communicating via L2TP can stop using the UDP layer if desired.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	**Cable networks and transmission of television, sound programme and other multimedia signals**
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2007