144. Multipoint MAC Control for 100G–EPON

144.1 Overview

This clause deals with the mechanism and control protocols required in order to reconcile the MultiPoint Reconciliation Sublayer (MPRS, defined in Clause 143) into the Ethernet framework. The P2MP medium is a passive optical network (PON), an optical network with no active elements in the signal's path from source to destination. The only interior elements used in a PON are passive optical components, such as optical fiber, splices, and splitters. When combined with the Ethernet protocol, such a network is referred to as Ethernet passive optical network (EPON).

P2MP is an asymmetric medium based on a tree (or tree-and-branch) topology. The DTE connected to the trunk of the tree is called optical line terminal (OLT) and the DTEs connected at the branches of the tree are called optical network units (ONU). The OLT typically resides at the service provider's facility, while the ONUs are located at the subscriber premises.

In the downstream direction (from the OLT to an ONU), signals transmitted by the OLT pass through a 1:N passive splitter (or cascade of splitters) and reach each ONU. In the upstream direction (from the ONUs to the OLT), the signal transmitted by an ONU would only reach the OLT, but not other ONUs. To avoid data collisions and increase the efficiency of the subscriber access network, the ONU's transmissions are arbitrated. This arbitration is achieved by allocating a transmission window (grant) to each ONU. An ONU defers transmission until its grant arrives. When the grant arrives, the ONU transmits frames at wire speed during its assigned time slot.

A simplified P2MP topology example is depicted in Figure 144–1. Clause 67 provides additional examples of P2MP topologies.

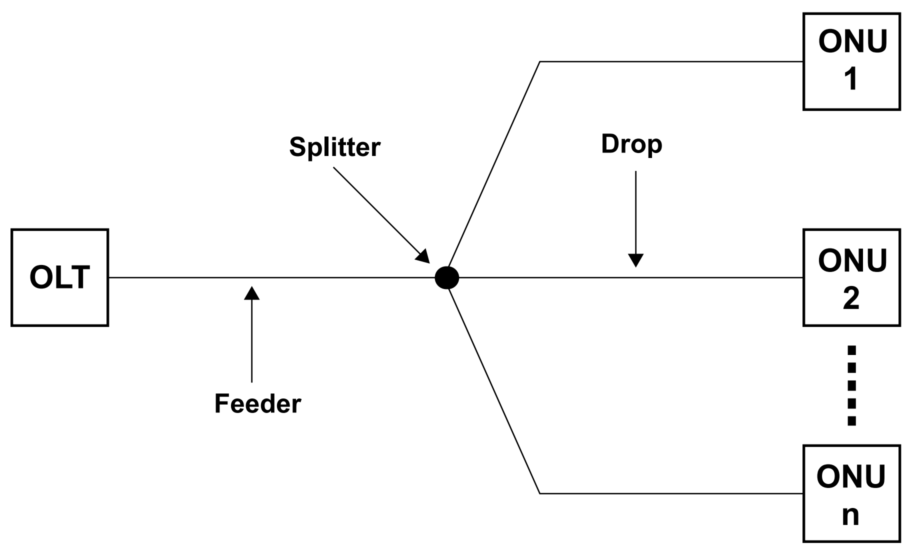


Figure 144–1—PON topology example

Topics dealt with in this clause include allocation of upstream transmission resources to different ONUs, discovery and registration of ONUs into the network, and reporting of congestion to higher layers to allow for dynamic bandwidth allocation schemes and statistical multiplexing across the PON.

This clause does not deal with topics including bandwidth allocation strategies, authentication of end-devices, quality-of-service definition, provisioning, or management.

This clause specifies the multipoint control protocol (MPCP) to operate an optical multipoint network by defining a Multipoint MAC Control sublayer as an extension of the MAC Control sublayer defined in Clause 31, and supporting current and future operations as defined in Clause 31 and annexes.

Each PON consists of a node located at the root of the tree assuming the role of OLT, and multiple nodes located at the tree leaves assuming roles of ONUs. The network operates by allowing only a single ONU to transmit in the upstream direction at a time. The MPCP located at the OLT is responsible for timing the different transmissions. Reporting of congestion by the different ONUs may assist in optimally allocating the bandwidth across the PON.

Automatic discovery of end stations is performed, culminating in registration through binding of an ONU to an OLT port by allocation of a Logical Link ID (see LLID in <TBD, Clause 143>), and dynamic binding to a MAC connected to the OLT.

The Multipoint MAC Control functionality shall be implemented for subscriber access devices containing point-to-multipoint Physical Layer devices defined in Clause 141.

144.1.1 State diagram conventions

The body of this standard comprises state diagrams, including the associated definitions of variables, constants, and functions. In case of any discrepancies between a state diagram and descriptive text, the state diagram prevails.

The notation used in the state diagrams follows the conventions of 21.5. State diagram timers follow the conventions of 14.2.3.2 augmented as follows:

a) [start x_timer, y] sets expiration of *y* to timer *x_timer*.
b) [stop x_timer] aborts the timer operation for *x_timer* asserting *x_timer_not_done* indefinitely.

The notation ++ after a variable indicates it is to be incremented by 1. The notation – – after a variable indicates it is to be decremented by 1. The notation –= after a variable indicates that the counter value is to be decremented by the following value. The notation += after a variable indicates that the counter value is to be incremented by the following value. Code examples given in this clause adhere to the style of the “C” programming language.

The state diagrams use an abbreviation MACR as a shorthand form for MA_CONTROL.request, MACI as a shorthand form for MA_CONTROL.indication, MADR as a shorthand for MA_DATA.request, and MADI as a shorthand for MA_DATA.indication primitives.

The vector notations used in the state diagrams for bit vector use 0 to mark the first received bit and so on (for example data<15:0>), following the conventions of 3.1 for bit ordering.

$a < b$: A function that is used to compare two values. Returned value is true when b is larger than a allowing for wrap around of a and b. The comparison is made by subtracting b from a and testing the MSB. When MSB($a–b$) = 1 the value true is returned, else false is returned. In addition, the following functions are defined in terms of $a < b$:

$a > b$ is equivalent to !($a < b$ or $a = b$)
$a \geq b$ is equivalent to !($a < b$)
$a \leq b$ is equivalent to !($a > b$)

144.1.2 Position of Multipoint MAC Control within the IEEE 802.3 hierarchy

Multipoint MAC Control defines the MAC control operation for optical point-to-multipoint networks. Figure 144–2 depict the architectural positioning of the Multipoint MAC Control sublayer with respect to the MAC and the MAC Control Clients. The Multipoint MAC Control sublayer takes the place of the MAC Control sublayer to extend it to support multiple clients and additional MAC control functionality. The Multipoint MAC Control sublayer does not interface with any MAC Clients.

Multipoint MAC Control is defined using the mechanisms and precedents of the MAC Control sublayer. The MAC Control sublayer has extensive functionality designed to manage the real-time control and manipulation of MAC sublayer operation. This clause specifies the extension of the MAC Control mechanism to manipulate multiple underlying MACs simultaneously. This clause also specifies a specific protocol implementation for MAC Control.

The Multipoint MAC Control sublayer is specified such that it can support new functions to be implemented and added to this standard in the future. MultiPoint Control Protocol (MPCP), the management protocol for P2MP is one of these protocols. Non-real-time, or quasi-static control (e.g., configuration of MAC operational parameters) is provided by Layer Management. Operation of the Multipoint MAC Control sublayer is transparent to the MAC.

As depicted in Figure 144–2, the layered system instantiates multiple MAC entities, using a single Physical Layer. The individual MAC instances offer a point-to-point emulation service between the OLT and the ONU. Additional Single Copy Broadcast (SCB) MAC instances are also created to communicate to all groups of Nx25G–EPON ONUs at once, for example, all 25/25G-EPON ONUs, all 50/50G-EPON ONUs. Such broadcast MAC instances takes maximum advantage of the broadcast nature of the downstream channel by sending a single copy of a frame that is received by all target Nx25G–EPON ONUs.

Editor's Note (to be removed prior to publication): Discussion and decision needed on how many SCB MACs we need: one per speed combination? Do we also need multicast MAC instances, associated with MLID?

The ONU only requires one MAC instance since frame filtering operations are done at the MPRS layer before reaching the MAC. Therefore, MAC and layers above are emulation-agnostic at the ONU.

Editor's Note (to be removed prior to publication): The above statement comes from .3av - Nx25G-EPON ONU might need more than one MAC when using 2x25G channels. Also, it might need extra MACs for MLID association.

Although Figure 144–2 and supporting text describe multiple MACs within the OLT, a single unicast MAC address may be used by the OLT. Within the EPON Network, MACs are uniquely identified by their LLIDs, which are dynamically assigned by the registration process.

144.1.3 Functional block diagram

Figure 144–3 and Figure 144–4 provide a functional block diagram of the Multipoint MAC Control architecture for the ONU and the OLT, respectively.

144.1.4 Service interfaces

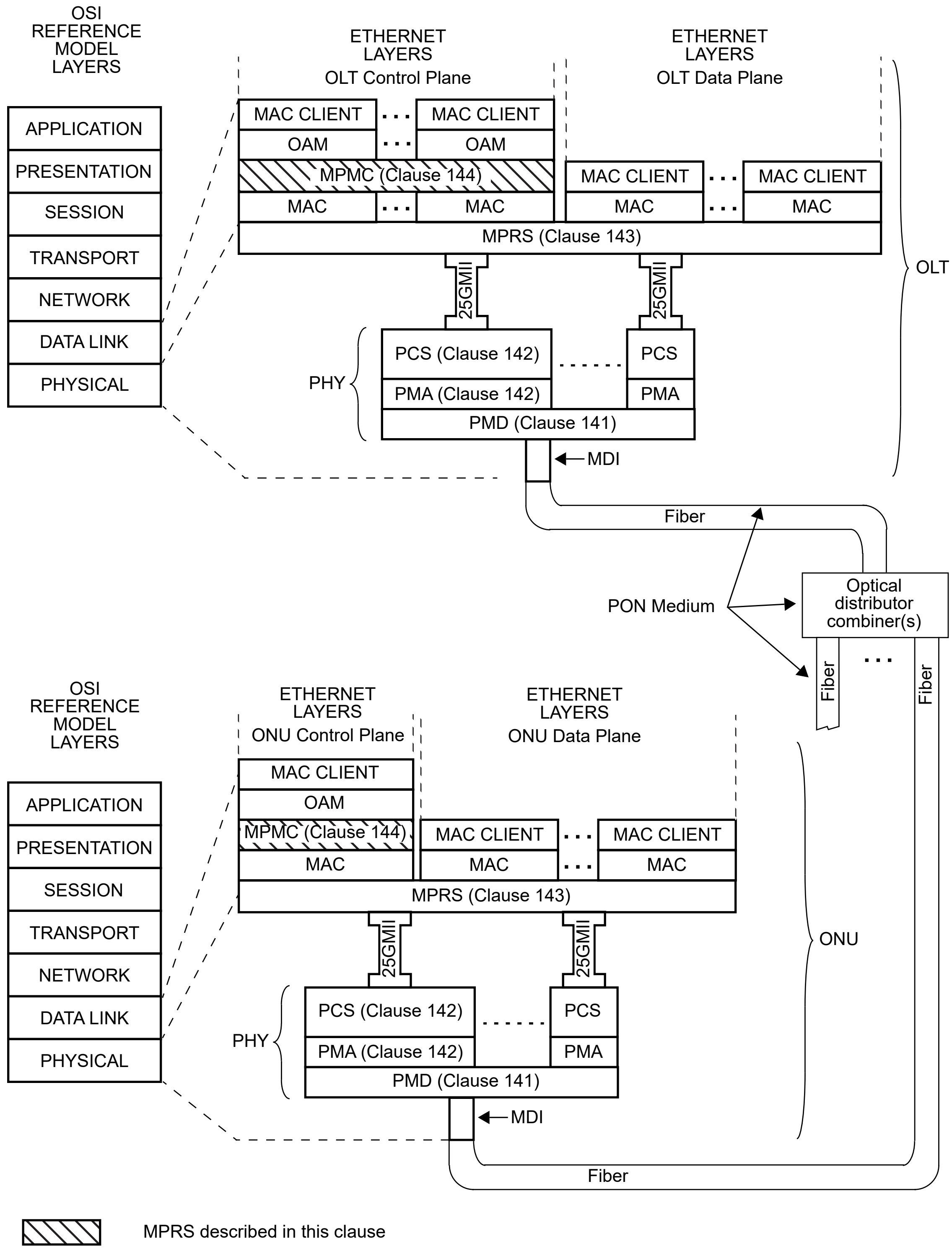
The MAC Clients communicate directly with dedicated MAC instances using the standard service interface specified in 2.3. The Multipoint MAC Control does not interface with any MAC Clients.

The MAC Control Clients communicate with Multipoint MAC Control instances using service interface defined in this clause. Each Multipoint MAC Control instance communicates with the underlying MAC sublayer using the standard service interface specified in Annex 4A.3.2. Similarly, Multipoint MAC Control communicates internally using primitives and interfaces consistent with definitions in Clause 31.

144.2 Multipoint MAC Control operation

As depicted in Figure 144–3 and Figure 144–4, the Multipoint MAC Control comprises the following functional blocks:

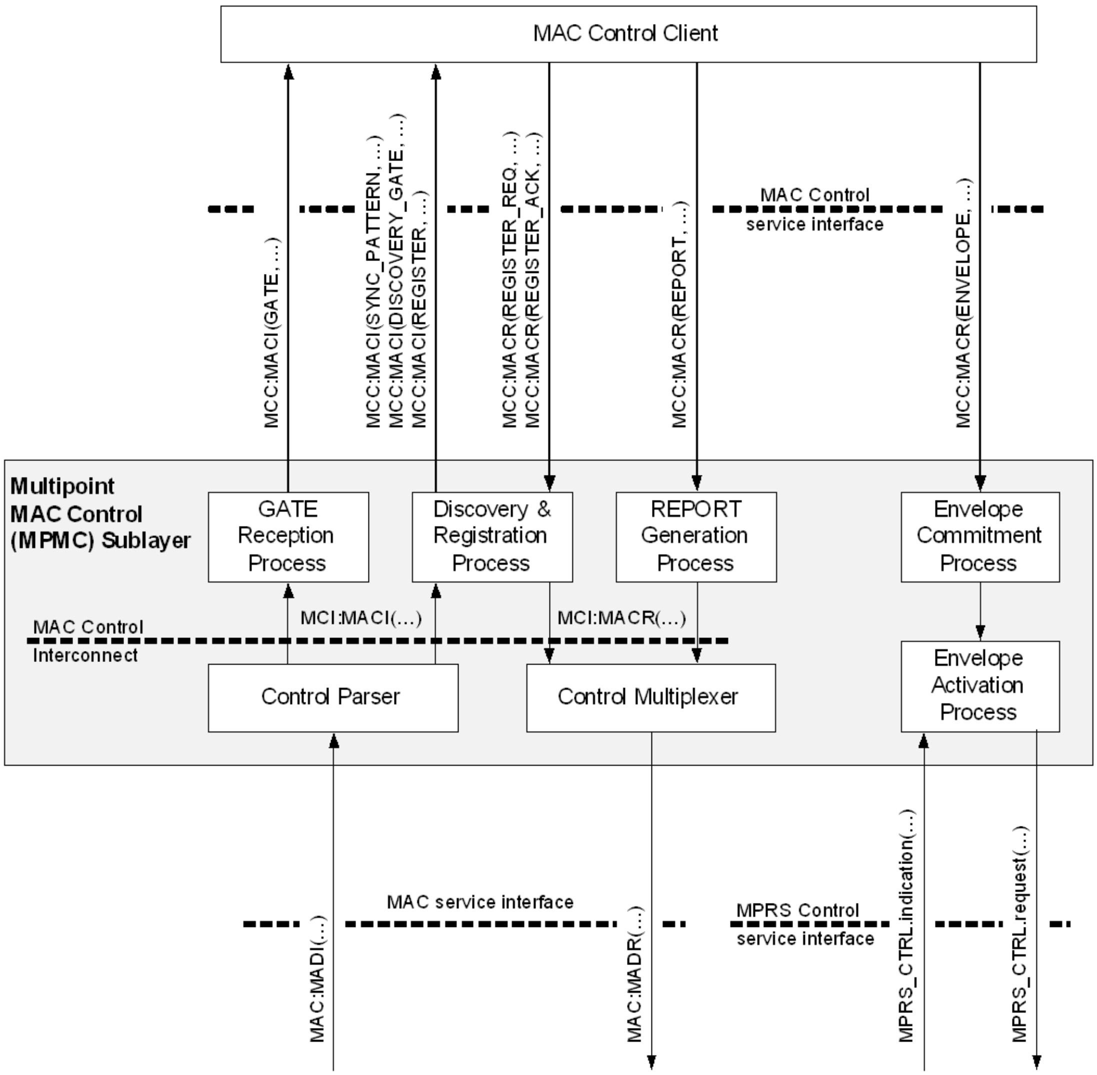
a) *Control Parser*. This block is responsible for parsing MAC Control frames, as well as interfacing with Clause 31 entities, and opcode specific blocks.
b) *Control Multiplexer*. This block is responsible for selecting the source of the forwarded frames.
c) *Discovery & Registration Processing*. This block is responsible for handling the MPCP in the context of the MAC.
d) *GATE Generation/Reception Process*. This block is responsible for generating and processing GATE messages.
e) *REPORT Generation/Reception Process*. This block is responsible for generating and processing REPORT messages.
f) *Envelope Commitment and Activation Process*. These blocks are responsible for committing and activating transmission envelopes.



MPRS described in this clause

25GMII=25 GIGABIT MEDIA INDEPENDENT INTERFACE
MDI = MEDIUM DEPENDENT INTERFACE
OAM = OPERATIONS, ADMINISTRATION & MAINTENANCE
OLT = OPTICAL LINE TERMINAL
MPRS= MULTI-POINT RECONCILIATION SUBLAYER
MPMC= MULTI-POINT MAC CONTROL
ONU = OPTICAL NETWORK UNIT
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 144–2—Relationship of EPON P2MP PMD to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model



Instances of Service Interface:
MAC = interface to MAC client
MCC = interface to MAC Control client
MCI = interface to MAC Control Interconnect client

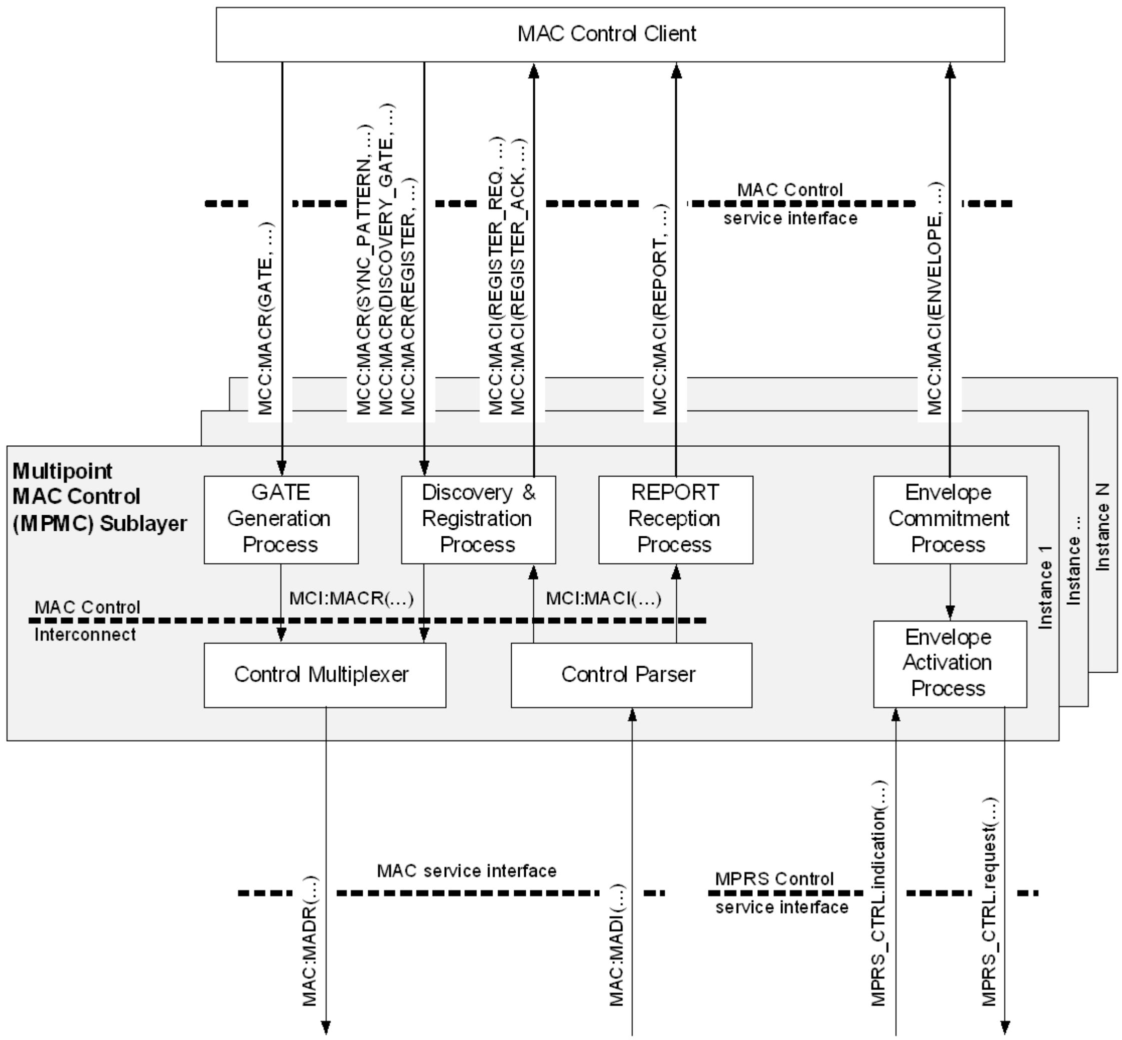
Figure 144–3—Multipoint MAC Control functional block diagram, ONU

144.2.1 Principles of Multipoint MAC Control

As depicted in Figure 144–4, the Multipoint MAC Control sublayer in the OLT may instantiate multiple Multipoint MAC Control instances in order to interface multiple MAC and MAC Control Clients above with multiple MACs below. A unique unicast MAC instance is used at the OLT to communicate with each ONU. The individual MAC instances utilize the point-to-point emulation service between the OLT and the ONU as defined in <TBD, Clause 142>. The Multipoint MAC Control sublayer in the ONU instantiates a single Multipoint MAC Control instance, as shown in Figure 144–3.

The scheduling algorithm is implementation dependent, and is not specified for the case where multiple transmit requests happen at the same time.

The information of the enabled interfaces is stored in the controller state variables, and accessed by the Control Multiplexer block.



Instances of Service Interface:
MAC = interface to MAC client
MCC = interface to MAC Control client
MCI = interface to MAC Control Interconnect client

Figure 144–4—Multipoint MAC Control functional block diagram, OLT

The Multipoint MAC Control sublayer uses the services of the underlying MAC sublayer to exchange both data and control frames.

Receive operation (MAC:MADI) at each Multipoint MAC Control instance:

a) A frame is received from the underlying MAC,
b) The received frame is parsed according to Length/Type field,
c) MAC Control frames are demultiplexed according to opcode and forwarded to the relevant processing functions
d) Data frames (see 31.5.1) are forwarded to the MAC Client outside of the Multipoint MAC Control sublayer.

Transmit operation (MAC:MADR) at each Multipoint MAC Control instance:

a) A protocol processing block attempts to issue a frame, as a result of a previous MACR or as a result of an MPCP event that generates a frame.

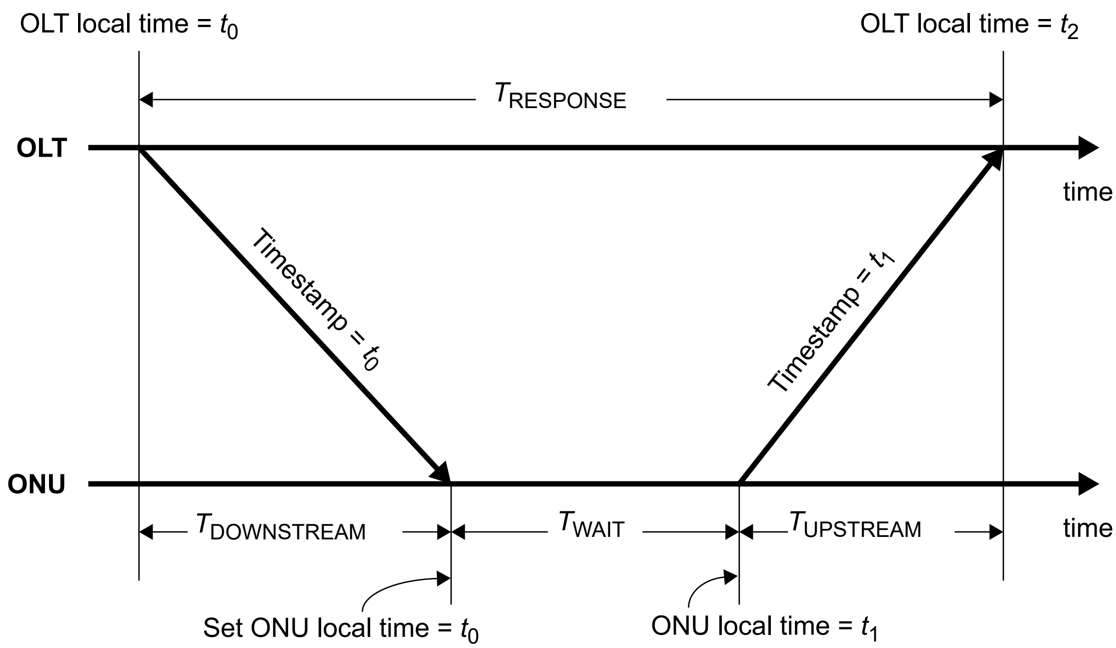
b) When allowed to transmit, the resulting frame is forwarded.

144.2.1.1 Ranging and timing process

Both the OLT and the ONU have 32-bit counters that increment every 1 EQ. These counters provide a local time stamp. When either device transmits an MPCPDU, the device maps its counter value into the timestamp field. The time of transmission of the first octet of the MPCPDU frame from the MAC Control to the MAC is taken as the reference time used for setting the timestamp value.

When the ONU receives an MPCPDU, the ONU sets its counter according to the value in the timestamp field in the received MPCPDU.

When the OLT receives an MPCPDU, it uses the received timestamp value to calculate or verify a round trip time between the OLT and the ONU. The Round Trip Time (RTT) is equal to the difference between the timer value and the value in the timestamp field. The calculated RTT is notified to the MAC Control Client via the MACI primitive. The MAC Control Client can use this RTT for the ranging process. The RTT calculation process is shown in Figure 144–5.



$T_{DOWNSTREAM}$ = downstream propagation delay

$T_{UPSTREAM}$ = upstream propagation delay

T_{WAIT} = wait time at ONU = $t_1 - t_0$

$T_{RESPONSE}$ = response time at OLT = $t_2 - t_0$

$$RTT = T_{DOWNSTREAM} + T_{UPSTREAM} = T_{RESPONSE} - T_{WAIT} = (t_2 - t_0) - (t_1 - t_0) = t_2 - t_1$$

Figure 144–5—Round trip time calculation

A condition of *timestamp drift error* occurs when the difference between OLT's and ONU's clocks exceeds some predefined threshold. This condition can be independently detected by the OLT and/or an ONU. The OLT detects this condition when an absolute difference between new and old RTT values measured for a given ONU exceeds the value of *guardThresholdOLT* (see 144.2.3.1), as shown in Figure 144–8. An ONU detects the timestamp drift error condition when absolute difference between a timestamp received in an MPCPDU and the localTime counter exceeds *guardThresholdONU* (see 144.2.3.1), as is shown in Figure 144–8.

144.2.1.2 PAUSE operation

Even though MPCP is compatible with flow control, optional use of flow control may not be efficient in the case of large propagation delay. If flow control is implemented, then the timing constraints in Annex 31B supplement the constraints found at 144.2.1.5.

NOTE—MAC at an ONU can receive frames from unicast channel and SCB channel. If the SCB channel is used to broadcast data frames to multiple ONUs, the ONU's MAC may continue receiving data frames from SCB channel even after the ONU has issued a PAUSE request to its unicast remote-end.

144.2.1.3 Optional Shared LAN emulation

By combining P2PE, suitable filtering rules at the ONU, and suitable filtering and forwarding rules at the OLT, it is possible to emulate an efficient shared LAN. Support for shared LAN emulation is optional, and requires an additional layer above the MAC, which is out of scope for this document. Thus, shared LAN emulation is introduced here for informational purposes only.

Specific behavior of the filtering layer at the MPRS is specified in <TBD, Clause 143?>.

144.2.1.4 Multicast and single copy broadcast support

In the downstream direction, the PON is a broadcast medium. In order to make use of this capability for forwarding broadcast frames from the OLT to multiple recipients without multiple duplication for each ONU, the SCB and multicast LLID support is introduced.

The OLT has at least one MAC associated with every ONU. In addition one more MAC at the OLT is marked as the SCB MAC. Moreover, the OLT has a multicast MAC associated with each defined multicast LLID. The SCB MAC handles all downstream broadcast traffic, but is never used in the upstream direction for client traffic, except for client registration. Similarly, the multicast MACs handle downstream multicast traffic, but are never used in the upstream direction for client traffic. Optional higher layers may be implemented to perform selective broadcast and multicast of frames. Such layers may require additional MACs (multicast MACs) to be instantiated in the OLT for some or all ONUs increasing the total number of MACs beyond the number of ONUs + 1.

When connecting the SCB MAC or a multicast MAC to an IEEE 802.1D bridge port it is possible that loops may be formed due to the broadcast or multicast nature of the associated LLIDs. Thus it is recommended that this MAC not be connected to an IEEE 802.1D bridge port.

Configuration of SCB channels as well as filtering and marking of frames for support of SCB at the MPRS is specified in <TBD, Clause 143?>.

144.2.1.5 Delay requirements

The MPCP protocol relies on strict timing based on distribution of timestamps. A compliant implementation needs to guarantee a constant delay through the MAC and PHY in order to maintain the correctness of the timestamping mechanism. The actual delay is implementation dependent; however, a complying implementation shall maintain a delay variation of no more than <TBD EQs> through the MAC.

The OLT shall not grant less than <TBD EQs> into the future, in order to allow the ONU processing time when it receives a gate message. The ONU shall process all messages in less than this period. The OLT shall not issue more than one message every <TBD EQs> to a single ONU. The unit of EQ is defined in <TBD>.

144.2.2 Control Parser and Control Multiplexer

The Control Parser (see Figure 144–6) is responsible for opcode-independent parsing of MAC frames in the receive path. By identifying MAC Control frames, demultiplexing into multiple entities for event handling is possible. Interfaces are provided to existing Clause 31 entities and multiple functional blocks associated with MPCP. There are no interfaces connecting the Control Parser to MAC Clients.

The Control Multiplexer (see Figure 144–7) is responsible for forwarding frames from multiple functional blocks associated with MPCP and prioritizing them accordingly. There are no interfaces connecting the Control Multiplexer to MAC Clients.

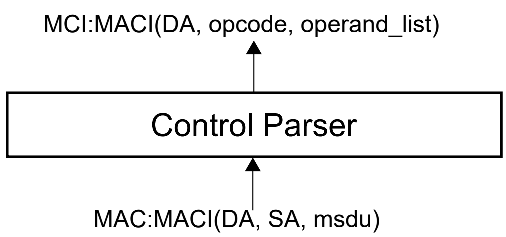


Figure 144–6—Control Parser service interfaces

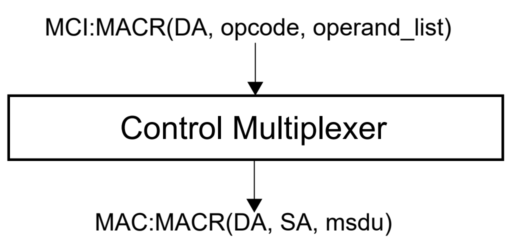


Figure 144–7—Control Multiplexer service interfaces

144.2.2.1 Constants

DRIFT_THOLD
TYPE: integer
This constant holds the maximum amount of drift allowed for a timestamp received at the given device. This value is measured in units of 1 EQ.
VALUE: <TBD> (for OLT) or <TBD> (for ONU)

144.2.2.2 Counters

localTime
TYPE: 32-bit unsigned
This variable holds the value of the local timer used to control MPCP operation. This variable is advanced by a timer at 390.625 MHz, and counts in the units of 1 EQ. At the OLT the counter shall track the transmit clock, while at the ONU the counter shall track the receive clock. For accuracy of

receive clock see <TBD, Clause 142>. This variable is reloaded with the received timestamp value (from the OLT) by the Control Parser (see Figure 144–9). Changing the value of this variable while running using Layer Management is undesirable and unspecified.

144.2.2.3 Variables

newRTT
TYPE: 16-bit unsigned integer
This variable holds the newly calculated Round Trip Time for the given ONU. The RTT value is represented in units of 1 EQ.

RTT
TYPE: 16-bit unsigned integer
This variable holds the measured Round Trip Time to the ONU. The RTT value is represented in units of 1 EQ.

timestamp
TYPE: 32-bit unsigned integer
This variable holds the value of times tamp of the last received MPCPDU frame (copied from the *TimeStamp* field).

timestampDrift
TYPE: Boolean
This variable is used to indicate whether an uncorrectable timestamp drift was detected (when set to True) or not (when set to False).

144.2.2.4 Functions

abs (n)
This function returns the absolute value of the parameter n.

ProcessTimestamp (timestamp)
This function takes the timestamp value from a received MPCPDU and checks whether the timestamp drift has exceeded the predefined device-specific threshold DRIFT_THOLD. In the OLT, this function also measures and updates the RTT of a given ONU.
In the OLT, the *ProcessTimestamp* function is defined as follows:

```
ProcessTimestamp( timestamp )
{
   newRTT = localTime – timestamp
   timestampDrift = abs(newRTT – RTT) > DRIFT_THOLD
   RTT = newRTT
}
```

In the ONU, the *ProcessTimestamp* function is defined as follows:

```
ProcessTimestamp( timestamp )
{
   timestampDrift = abs(localTime – timestamp) > DRIFT_THOLD
}
```

144.2.2.5 Timers

144.2.2.6 Messages

144.2.2.7 State diagrams

The OLT and ONU shall implement the Control Multiplexer state diagram shown in Figure 144–8. The OLT and ONU shall implement the Control Parser state diagram shown in Figure 144–9.

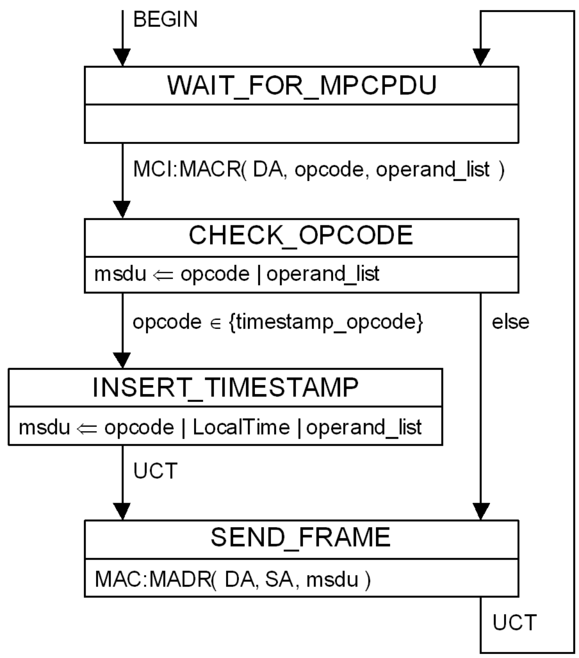


Figure 144–8—Control Multiplexer state diagram

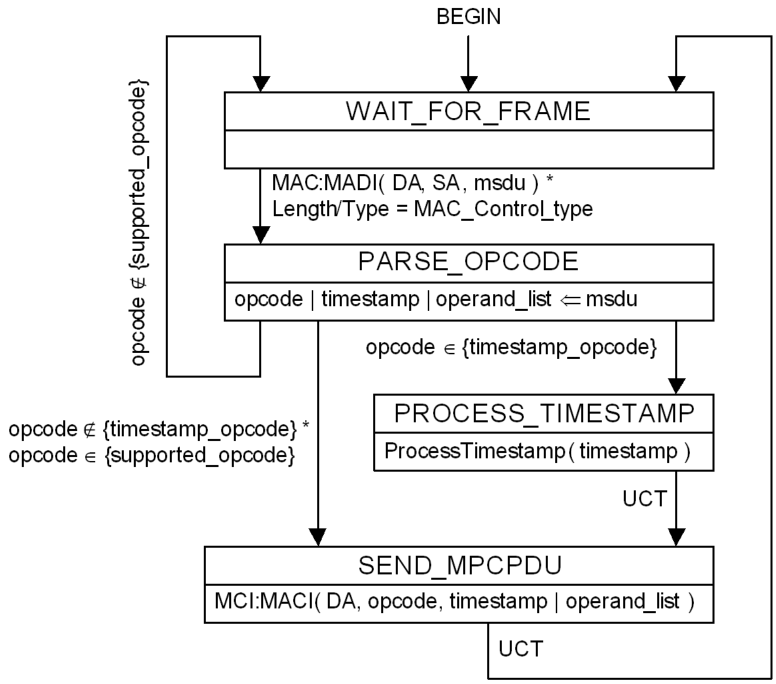


Figure 144–9—Control Parser state diagram