Enhanced Transmission Selection for Half-Duplex Media

Peter Jones – Cisco Systems - 10-Sep-26

Revision History:

D1 – from 1-26-0022-00-ICn: Enhanced Transmission Selection for Half-Duplex Media

Overview

The problem statement presented in Consensus Building for 10BASE-T1S Differentiated Services: PLCA Multidrop Latency, Jitter, and scalability enhancements is

> 10BASE-T1-S multidrop are being widely adopted in automotive Ethernet applications across e.g., body, chassis, and powertrain control applications, whose functions span a broad range of latency and jitter requirements, from tightly bounded control loops to more timing-tolerant comfort and convenience functions.
> The Physical Layer Collision Avoidance (PLCA) layer developed as a part of 802.3cg and refined in 802.3da provides fair and equal access to stations on the mixing segment while providing bounded maximum access latency. The maximum access latency is not sufficiently bounded for many control loops. Likewise, the variable length of the PLCA cycle results in high jitter (lack of determinism) that such applications cannot tolerate. PLCA as defined today cannot treat traffic differently based on latency and jitter requirements. Additionally, adding traffic to multidrop networks engineered to meet latency and jitter requirements can change the latency and jitter of existing traffic in ways that are difficult to predict resulting in scalability issues and the need to frequently reengineer networks.

I restate this as:

> The current set of 802.1 and 802.3 tools do not provide sufficient control over the selection of the next frame to be transmitted on the shared media.

802.1Q provides extensive tools for a station to choose which frame to transmit next (8.6.8 Transmission selection), but the selection happens independent of the station gaining access to the media. Most of these tools have been designed with full duplex media in mind.

802.3 Media Access Control is priority unaware (802.1AC-2016 - Ethernet convergence function and 802-3-2022 Detailed service specification). Using 802.3 half duplex media with:

- CSMA/CD: provides no guarantee of when a given station will get access to the media
- CSMA/CD and PLCA: provides bounded access time for a station with worst case delay in the order of (station count * MTU transmission time).

Modifying PLCA can change when a specific station gets access to the media, but not which frame is sent by that station. Which frame is sent is controlled by 802.1Q Transmission selection which operates independently of media access. A frame is selected for transmission regardless of potential significant media access delay. This is an issue for all half duplex media. The frame transmitted may not be the one that would have been selected when media access was obtained. Full duplex media using 802.2 PAUSE function (802-3-2022 Annex 31B) has the same behavior.

Addressing the variable delay between transmission selection and media access is key to addressing the problem statement, and the solution space includes 802.1 and 802.3 half duplex CSMA/CD, PLCA and 802.3 PAUSE.

Table of Contents

Transmission Selection for Half Duplex Networks

Background

Half duplex networks create issues with 802.1Q Transmission Selection(8.6.8). Transmission Selection supports several scheduling algorithms as shown in Table 8-6.

Table 8-6—Transmission selection algorithm identifiers

Transmission selection algorithm	Identifier
Strict priority (8.6.8.1)	0
Credit-based shaper (8.6.8.2)	1
Enhanced Transmission Selection (ETS) (8.6.8.3)	2
ATS Transmission Selection (8.6.8.5)	3
Reserved for future standardization	4–254
Vendor-specific Transmission Selection algorithm value for use with DCBX (D.2.8.8)	255
Vendor-specific	A four-octet integer, where the most significant 3 octets hold an OUI or CID value, and the least significant octet holds an integer value in the range 0–255 assigned by the owner of the OUI or CID.

Figure 8-16 shows the structure used by Transmission Selection including transmission gates as specified in Enhancements for scheduled traffic(8.6.8.4):

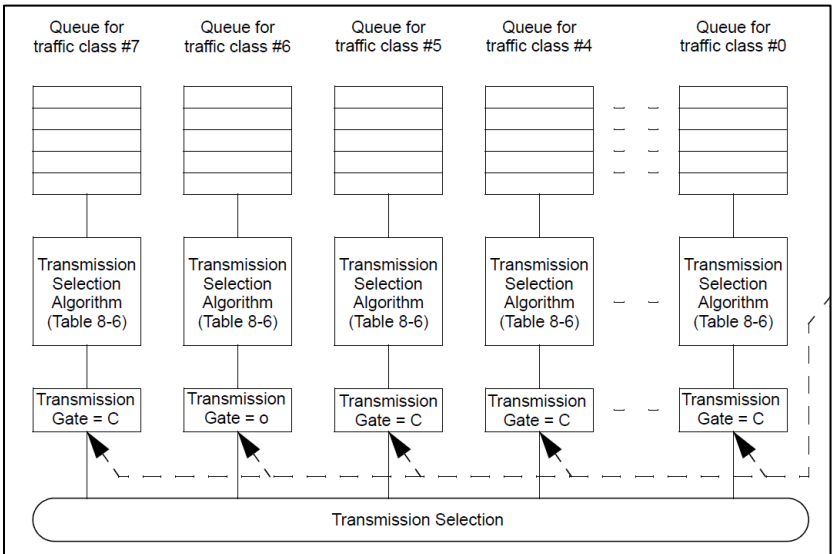


Figure 8-16—Transmission selection with gates

Frame transmission on half duplex media can produce unexpected results because the transmission selection decision is made before the station has access to the media. Consider a half-duplex point-to-point 100Mb/s Ethernet link. If the peer is transmitting when a frame is selected, the frame will not be sent to the media until the peer is done transmitting. This can delay transmission by up to an MTU worth of delay.

This situation is exacerbated in multidrop 802.3 segments (e.g., 10BASE-T1S/M) using CSMD/CD with multiple peers. Multiple nodes beginning transmission around the same time can cause collisions for an extended period. Each transmitting node has already performed transmission selection, and the system is controlled by the MAC back-off/retransmit function.

PLCA prevents collisions by delaying access to the media. Using PLCA with 10BASE-T1S/M bounds the worst case media access delay to ((nodes – 1) * (MTU transmission delay)).

Processing of PAUSE frames (802.3-2022 - Annex 31B MAC Control PAUSE operation) creates a similar situation to point-to-point half-duplex media. When the station is paused, MAC Control suspends processing `MCF:MA_DATA.request` primitives until the pause timer expires resulting in transmission selection happening significantly before transmission on the wire.

Transmission Selection and Media Availability

Transmission selection conducted when a node has access to the media produces results much closer to the desired behavior. This can be achieved by passing a "Media Available" signal to transmit selection as an additional constraint on the selection process.

This is analogous to the behavior of Priority-based Flow Control(PFC) as specified in 802.1Q Clause 36. PFC controls a vector of `Priority_Paused[n]` bits which act as an additional qualification in Transmission selection (8.6.8):

> In a port of a Bridge or station that supports PFC, a frame of priority n is not available for transmission if that priority is paused, i.e., if `Priority_Paused[n]` is TRUE (see 36.1.3.2) on that port.

PFC is passed between the MAC Client and the MAC using M_CONTROL PFC primitives. These primitives are specified in 802.1Q 36.1.2 "PFC primitives" and 802.3 Annex 31D "MAC Control PFC operation"[1].

To support "Media Available" signaling:

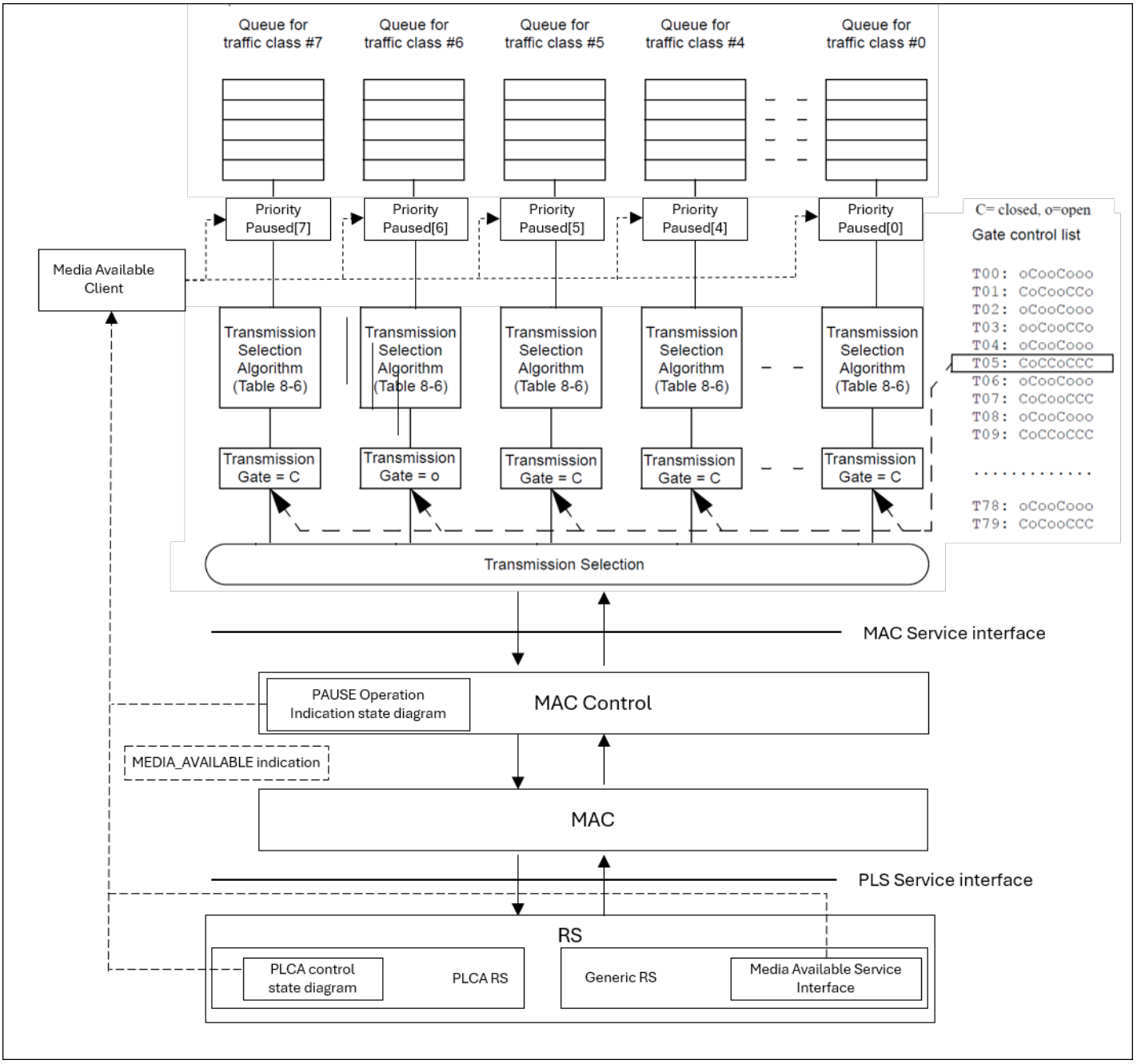
1. Define a new MA_CONTROL.indication `MEDIA_AVAILABLE` opcode which notifies the MAC Client of half duplex media availability using an 8 bit `priority_available_vector` . Each bit reports media availability for a specific priority value.
2. Define a new M_CONTROL MEDIA_AVAILABLE client in 802.1Q which sets the `Priority_Paused` vector directly from the `MEDIA_AVAILABLE priority_available_vector`

Use of MEDIA_AVAILABLE is incompatible with the use of PFC (defined for full duplex only). MEDIA_AVAILABLE passes an 8 bit vector as it's argument to align with 802.1Q Priority-based Flow Control (PFC) and to allow an instance of a MAC to selectively enable media access for a subset of priorities.

NOTE: Although 802.1AC-2016 mentions the use of M_CONTROL.request and an M_CONTROL.indication in the ISS to implement Priority-Based Flow Control, this text is only in a note. 802.1AC-2016 does not need to be updated when defining the new control functions.

[1] The PFC opcode includes both a per-priority enable vector, and a per-priority time vector containing the pause_quanta (512 bit times) to pause transmission. It does not include a code point for "pause indefinitely", so can't be directly reused for "Media Available" signaling.

Overview Figure

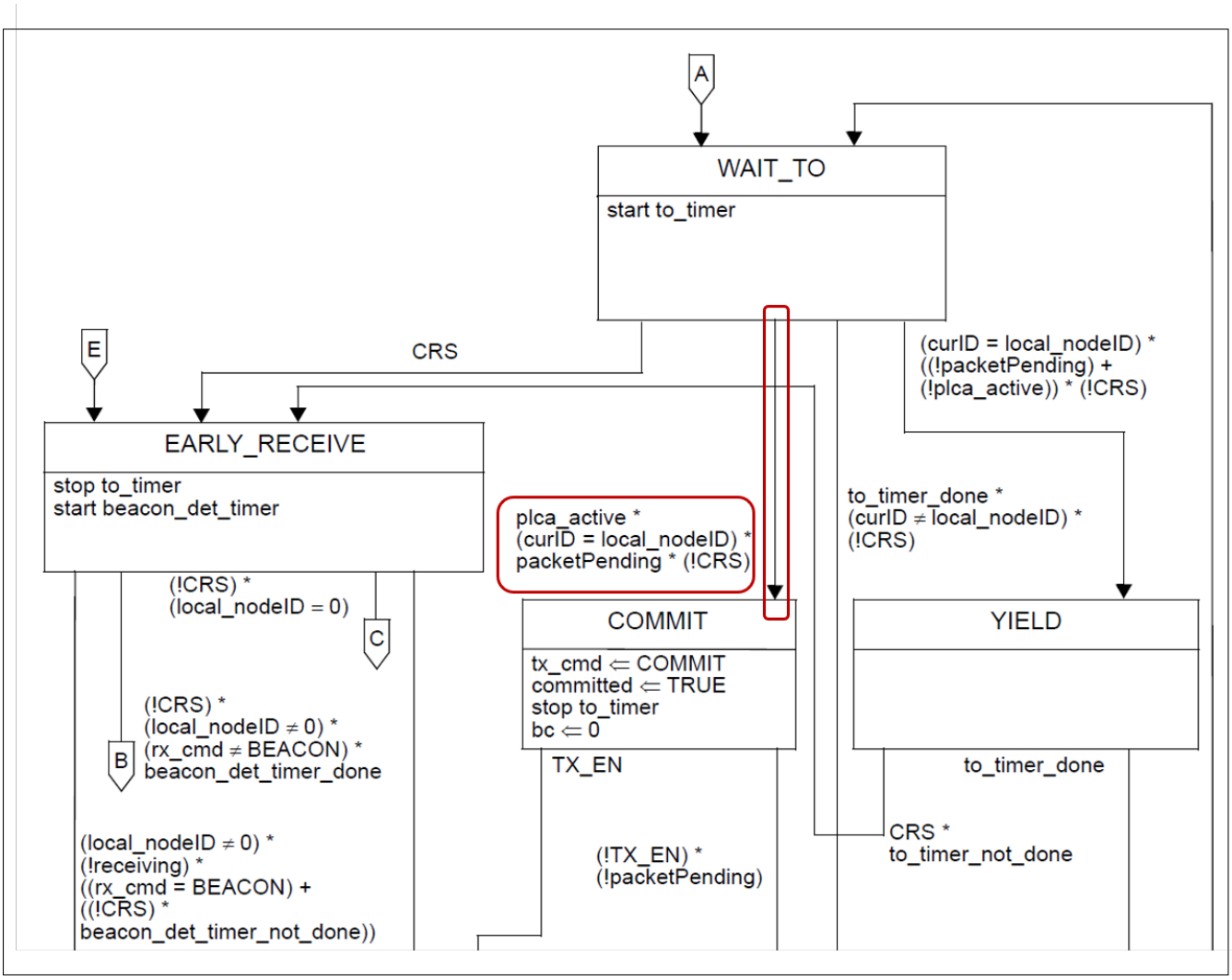


Enhanced Transmission Selection for PLCA

Change the PLCA Control state diagram to signal the media is:

1. not available when it enters WAIT_TO
2. available when it's this node's TO

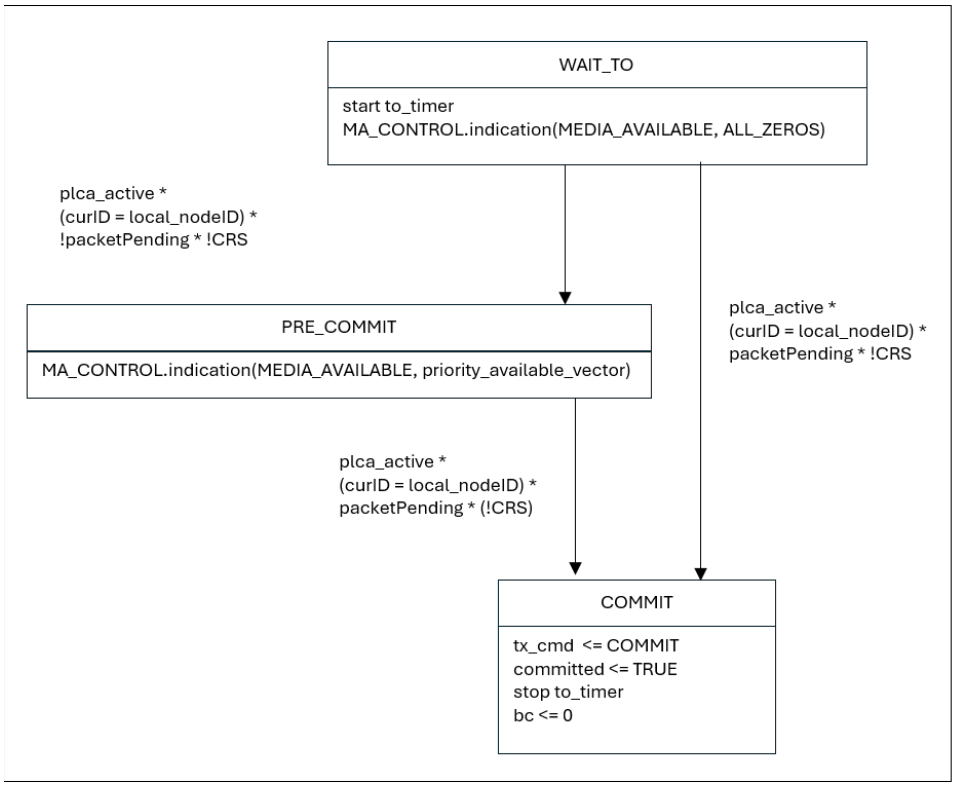
The area of interest in the PLCA Control state diagram is highlighted below:



MEDIA_AVAILABLE Signaling

Change the PLCA Control state diagram as follows and as shown below:

- Generate a null `MEDIA_AVAILABLE` indication in WAIT_TO
- Add a transition from `WAIT_TO->PRE_COMMIT` when it's the local node's TO.
- Add a PRE_COMMIT state where PLCA generates a non-null `MEDIA_AVAILABLE` indication
- Add a transition from `PRE_COMMIT->COMMIT` with the same conditions as `WAIT_TO->PRE_COMMIT`
- Add `PRE_COMMIT->EARLY_RECEIVE` and `PRE_COMMIT->YIELD` transitions with the same conditions as `WAIT_TO->EARLY_RECEIVE` and `WAIT_TO->YIELD`.



The mechanism for generating the `priority_available_vector` value used in the state diagram is beyond the scope of this standard.

PLCA - Delay between TO and Transmission

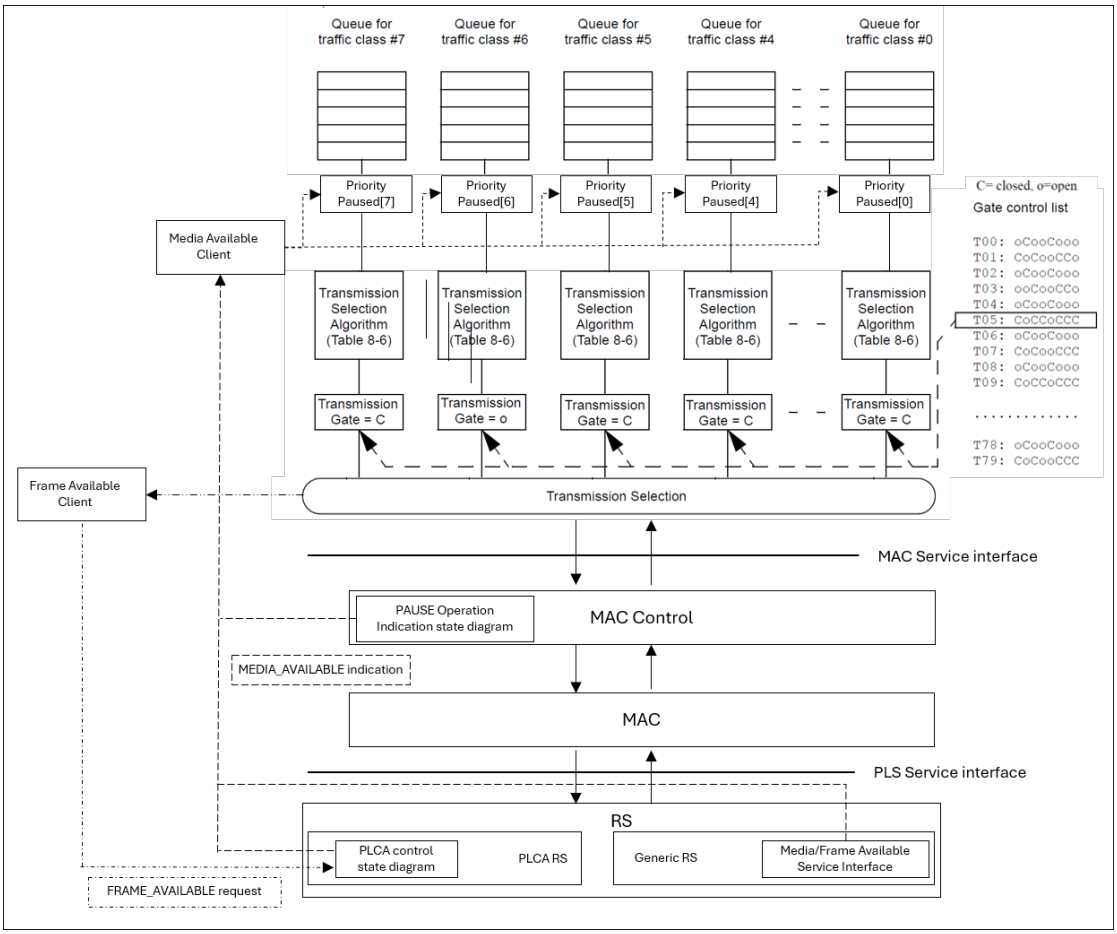
PLCA performance is sensitive to any delay between a TO being claimed and the start of packet transmission. Everything in the path from the PLCA Control State diagram -> Transmission Selection -> PLS Service interface takes "zero time" assuming no deference or collisions, e.g., the MAC assembles the frame and then passes the data bit by bit using the synchronous TransmitBit() procedure[2]. TransmitBit() maps to the PLS_DATA.request() primitive[3],[4]. The PLCA RS inserts these bits into a 4 bit transmit vector[5] to pass to the MI and sets a flag[6] after 4 bits to signal the PLCA Data state diagram that a packet is available. The Data state diagram sets a flag[7] to signal the Control state diagram it has data ready to go. The only delay specified in the standard is the 4 bit times used in the to fill the 4-bit transmit vector.

FRAME_AVAILABLE Signaling

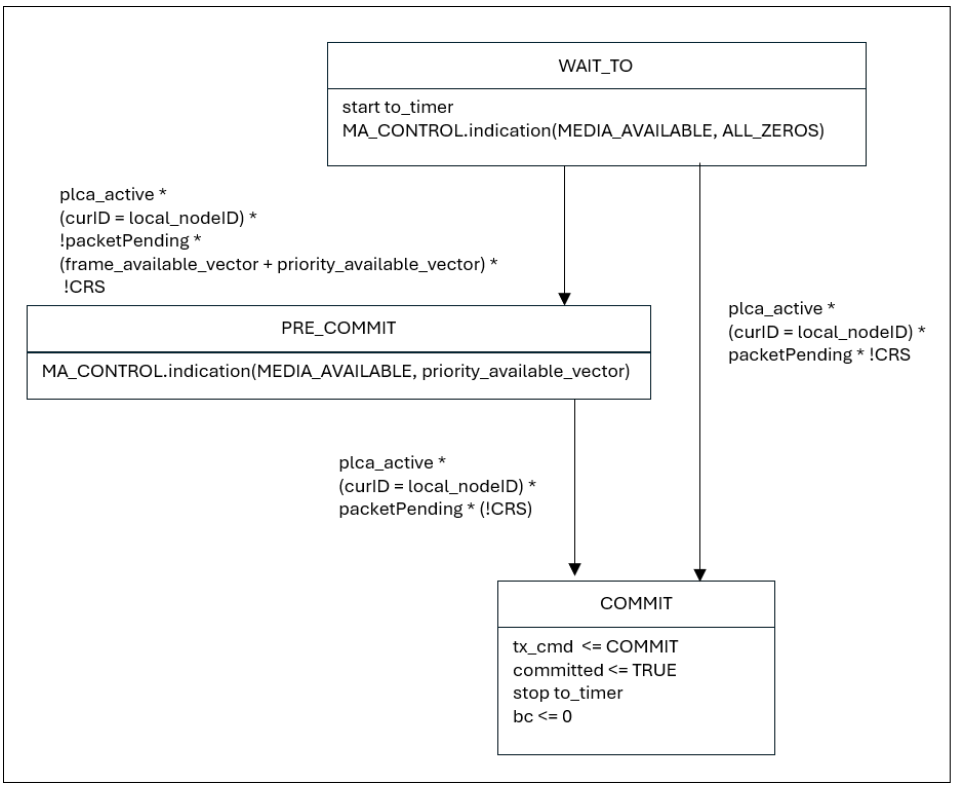
PLCA uses the MAC starting transmission (`packetPending=TRUE`) to decide to claim a TO. PLCA will generate `MEDIA_AVAILABLE` when deciding whether to claim the current TO. Investigation is required to determine the latency from generating `MEDIA_AVAILABLE` to frame data being passed from the MAC to PLCA(setting `packetPending=TRUE`). If this latency would significantly affect performance, it could be addressed using "Frame Available" signaling from transmission selection to PLCA as follows:

1. Define a new `MA_CONTROL.request FRAME_AVAILABLE` opcode which notifies the RS, PLCA in this case, of current frame transmission eligibility using an 8 bit `frame_available_vector` parameter.
2. Define new "Frame Availability" functionality in 802.1Q which generates the `FRAME_AVAILABLE` request when transmission selection state changes
3. Define a new clause 31 annex to process `FRAME_AVAILABLE` requests. Use the `frame_available_vector` parameter to update a new PLCA Control state diagram variable.

[2] 802.3-2022 4.3.3 Services required from the Physical Layer
[3] 802.3-2022 6.3.1.1 PLS_DATA.request
[4] 148.4.2.1 Mapping of PLS_DATA.request
[5] plca_txd<3:0>
[6] plca_txen
[7] packetPending


Queue for traffic class #7
Queue for traffic class #6
Queue for traffic class #5
Queue for traffic class #4
Queue for traffic class #0
Priority Paused[7]
Priority Paused[6]
Priority Paused[5]
Priority Paused[4]
Priority Paused[0]
C= closed, o=open
Gate control list
T00: oCooCooo
T01: CoCooCCo
T02: oCooCooo
T03: ooCooCCo
T04: oCooCooo
T05: CoCCoCCC
T06: oCooCooo
T07: CoCooCCC
T08: oCooCooo
T09: CoCCoCCC
T78: oCooCooo
T79: CoCooCCC
Media Available Client
Transmission Selection Algorithm (Table 8-6)
Transmission Gate = C
Transmission Gate = o
Frame Available Client
Transmission Selection
MAC Service interface
PAUSE Operation Indication state diagram
MAC Control
MEDIA_AVAILABLE indication
MAC
PLS Service interface
RS
PLCA control state diagram
PLCA RS
Generic RS
Media/Frame Available Service Interface
FRAME_AVAILABLE request

Change the PLCA Control state diagram from the figure above by adding a test of (`frame_available_vector * priority_available_vector)` to the `WAIT_TO->PRE_COMMIT` transition as shown below:

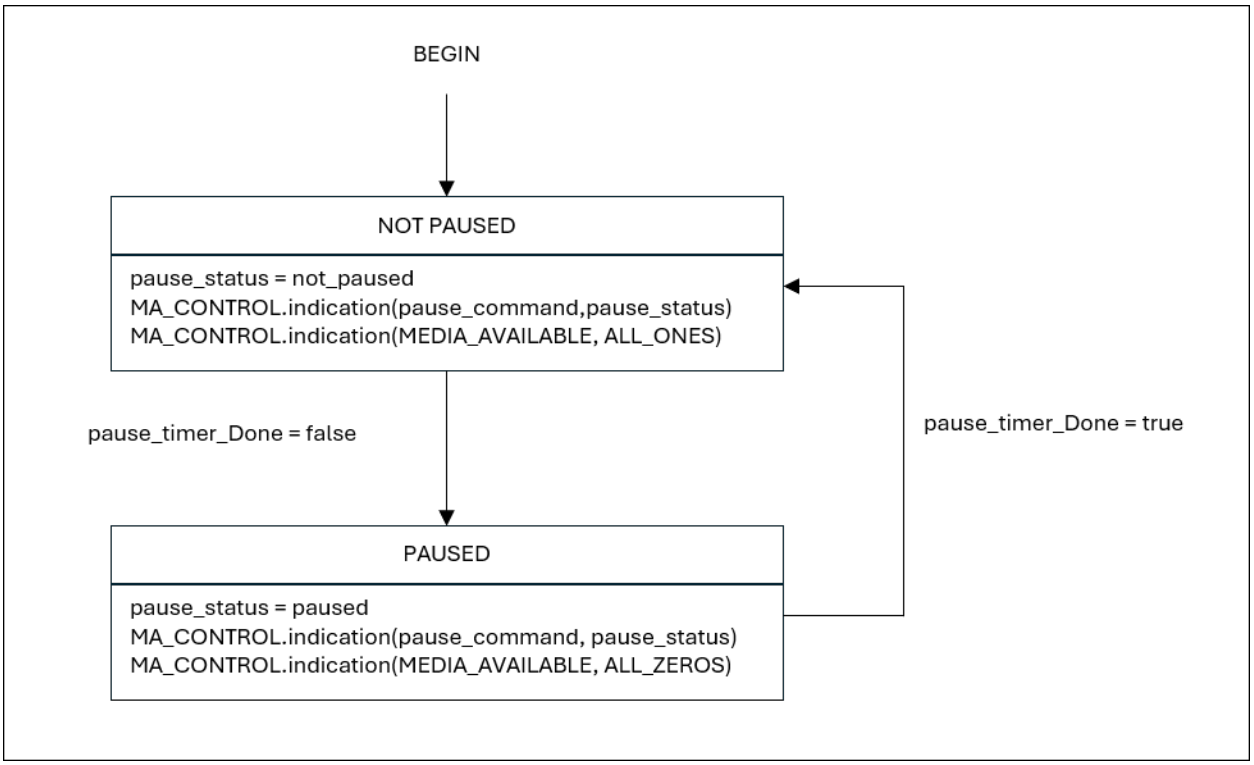


Half Duplex

To support other half-duplex media (e.g. 100BASE-T, 10BASE-T1S not running PLCA), define a new Reconciliation Sublayer extension analogous to clause 90 TSSI. The new RS uses the generic `PLS_CARRIER.indication` to generate `MEDIA_AVAILABLE` indications using `ALL_ZEROS` and `ALL_ONES`.

Full Duplex PAUSE

Deferring transmission selection for full-duplex media using MAC Control PAUSE to when the media is available (i.e., not paused) can be addressed by modifuying the Indication state diagram (31B.3.6.4) to generate `MEDIA_AVAILABLE` indications using `ALL_ZEROS` and `ALL_ONES`.



END OF DOCUMENT