PMA+PMD Delay of an Optical Module

Jeffery Maki

Juniper Networks

Comment #82

Cl **169** *SC* **169.4** *P* **177** *L* **27** # 82

Maki, Jeffery Juniper Networks

Comment Type **T** *Comment Status* **D** *delay (CC)*

The sum of the sublayer delays of 92.16 ns for 800GBASE-R PMA (up to four PMA stages) and 20.48 ns for 800GBASE-VR8/SR8/DR8/DR8-2 PMD is 112.64 ns, which is less than the observed delay of two PMA stages and the PMD. The concern is that these sublayers delays are specified too small in value to be feasible. Excessive delays of about 50% are seen for optical modules (two PMA stages + PMD).

SuggestedRemedy

Increase Delay values for PMA and PMD to align with prevalent implementation.

Proposed Response *Response Status* **W**

PROPOSED REJECT.
The specification of delay for the PMA is rather ambiguous and the delay specified for the PMA and PMD may be smaller than necessary to permit practical implementations. However, a complete proposal with appropriate background and specific changes to the draft is required.
A presentation to address this comment is anticipated.

MAC Control PAUSE – Need for Delay Spec

IEEE Std 802.3-2022, IEEE Standard for Ethernet
SECTION THREE

36.5 Delay constraints

In half duplex mode, proper operation of a CSMA/CD LAN demands that there be an upper bound on the propagation delays through the network. This implies that MAC, PHY, and repeater implementations conform to certain delay minima and maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices. MAC constraints are contained in 35.2.4 and Table 35–5. Topological constraints are contained in Clause 42.

In full duplex mode, predictable operation of the MAC Control PAUSE operation (Clause 31, Annex 31B) also demands that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sublayer, and PHY implementations conform to certain delay maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices.

The reference point for all MDI measurements is the 50% point of the mid-cell transition corresponding to the reference bit, as measured at the MDI.

800G Delay Sublayers

169.3.3 Semantics of inter-sublayer service interface primitives

The semantics of the inter-sublayer service interface primitives for the 800GBASE-R sublayers are described in 116.3.3.1 through 116.3.3.3.

169.4 Delay constraints

Predictable operation of the MAC Control PAUSE operation (Clause 31, Annex 31B) demands that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sublayer, and PHY implementations conform to certain delay maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices. Table 169–4 contains the values of maximum sublayer delay (sum of transmit and receive delays at one end of the link) in bit times as specified in 1.4 and pause_quanta as specified in 31B.2 for 800 Gigabit Ethernet. If a PHY contains an Auto-Negotiation sublayer, the delay of the Auto-Negotiation sublayer is included within the delay of the PMD and medium.

Table 169–4—Sublayer delay constraints (800GBASE)

Sublayer	Maximum (bit time)[a]	Maximum (pause_quanta)[b]	Maximum (ns)	Notes[c]
800G MAC, RS, and MAC Control	196 608	384	245.76	See 170.1.4.
800GBASE-R PCS or 800GXS[d]	640 000	1250	800	See 172.5.
800GBASE-R PMA	73 728	144	92.16	See 173.4.4.
800GBASE-KR8 PMD	32 768	64	40.96	Includes allocation of 14 ns for one direction through backplane medium. See 163.5.
800GBASE-CR8 PMD	32 768	64	40.96	Includes allocation of 14 ns for one direction through cable medium. See 162.5.
800GBASE-VR8 PMD	16 384	32	20.48	Includes 2 m of fiber. See 167.3.1.
800GBASE-SR8 PMD	16 384	32	20.48	Includes 2 m of fiber. See 167.3.1.
800GBASE-DR8 PMD	16 384	32	20.48	Includes 2 m of fiber. See 124.3.1.
800GBASE-DR8-2 PMD	16 384	32	20.48	Includes 2 m of fiber. See 124.3.1.

[a] For 800GBASE-R, 1 bit time (BT) is equal to 1.25 ps. (See 1.4.215 for the definition of bit time.)
[b] For 800GBASE-R, 1 pause_quantum is equal to 640 ps. (See 31B.2 for the definition of pause_quanta.)
[c] Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.
[d] If an implementation includes the 800GMII Extender, the delay associated with the 800GMII Extender includes two 800GXS sublayers.

800G PMA Stages

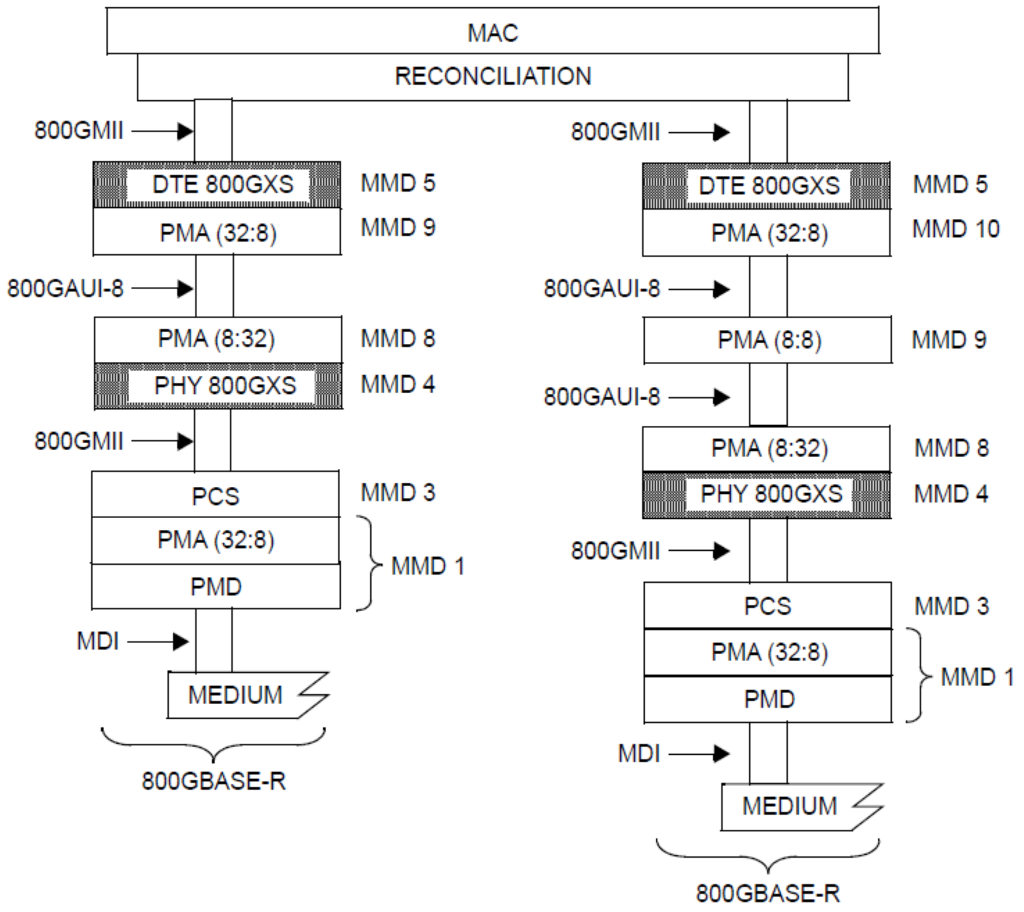
173.4.4 Delay constraints

The maximum cumulative delay contributed by up to four PMA stages in a PHY (sum of transmit and receive delays at one end of the link) shall meet the values specified in Figure 173–1. A description of overall system delay constraints and the definitions for bit-times and pause_quanta can be found in 169.4 and its references.

Table 173–1—Delay constraints

Sublayer	Maximum (bit time)	Maximum (pause_quanta)	Maximum (ns)
800GBASE-R PMA	73 728	144	92.16

3 PMA Stages 4 PMA Stages

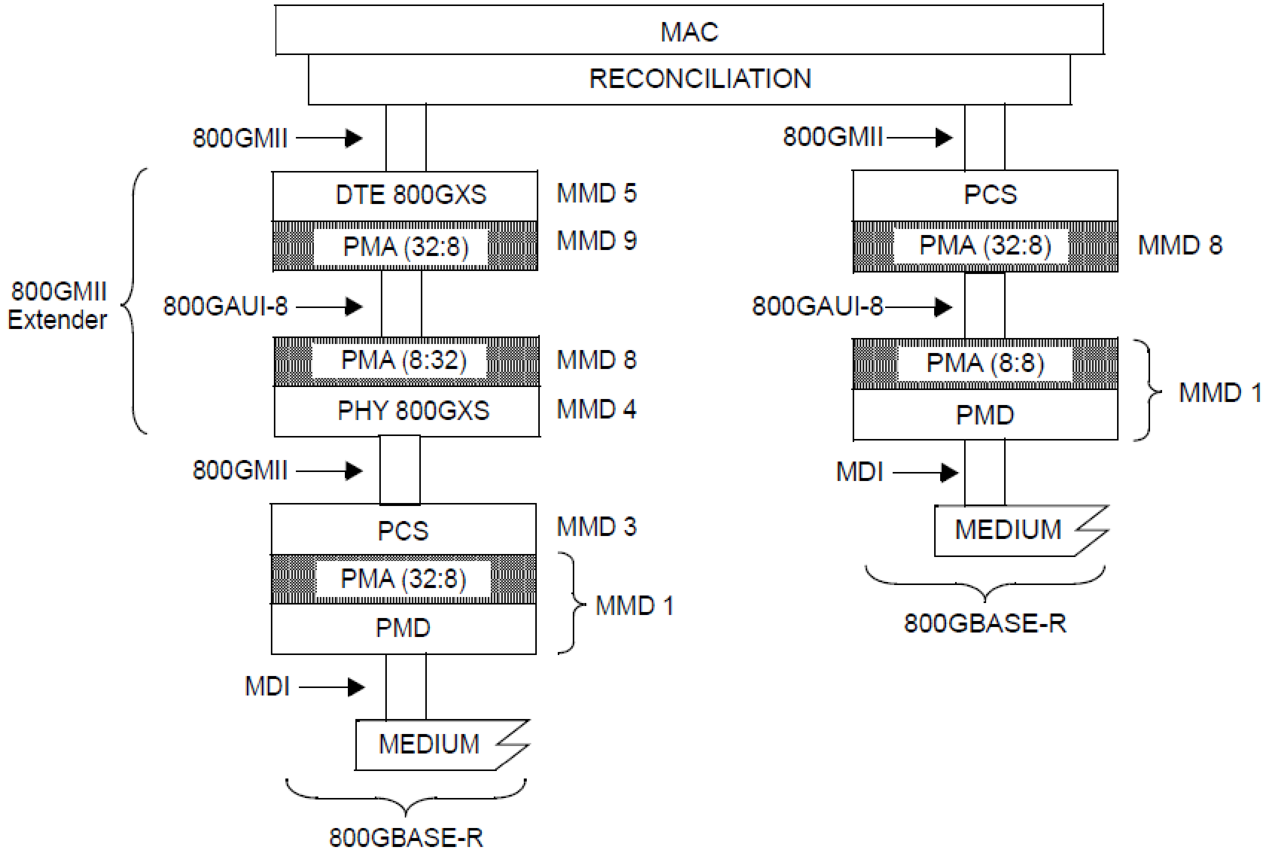


800GAUI = 800 Gb/s ATTACHMENT UNIT INTERFACE
800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
800GXS = 800 Gb/s EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE

MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 171–3—Example 800GBASE-R PMA layering with 800GXS

3 PMA Stages 2 PMA Stages



800GAUI = 800 Gb/s ATTACHMENT UNIT INTERFACE
800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
800GXS = 800 Gb/s EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 173–2—Examples of 800GBASE-R PMA layering

800G Module Delay

800G Module Delay

Total: 43.52 ns

- PMA: 92.16 ns / 4 = 23.04 ns
- Optical PMD: 20.48 ns

169.3.3 Semantics of inter-sublayer service interface primitives

The semantics of the inter-sublayer service interface primitives for the 800GBASE-R sublayers are described in 116.3.3.1 through 116.3.3.3.

169.4 Delay constraints

Predictable operation of the MAC Control PAUSE operation (Clause 31, Annex 31B) demands that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sublayer, and PHY implementations conform to certain delay maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices. Table 169–4 contains the values of maximum sublayer delay (sum of transmit and receive delays at one end of the link) in bit times as specified in 1.4 and pause_quanta as specified in 31B.2 for 800 Gigabit Ethernet. If a PHY contains an Auto-Negotiation sublayer, the delay of the Auto-Negotiation sublayer is included within the delay of the PMD and medium.

Table 169–4—Sublayer delay constraints (800GBASE)

Sublayer	Maximum (bit time)[a]	Maximum (pause_quanta)[b]	Maximum (ns)	Notes[c]
800G MAC, RS, and MAC Control	196 608	384	245.76	See 170.1.4.
800GBASE-R PCS or 800GXS[d]	640 000	1250	800	See 172.5.
800GBASE-R PMA	73 728	144	92.16	See 173.4.4.
800GBASE-KR8 PMD	32 768	64	40.96	Includes allocation of 14 ns for one direction through backplane medium. See 163.5.
800GBASE-CR8 PMD	32 768	64	40.96	Includes allocation of 14 ns for one direction through cable medium. See 162.5.
800GBASE-VR8 PMD	16 384	32	20.48	Includes 2 m of fiber. See 167.3.1.
800GBASE-SR8 PMD	16 384	32	20.48	Includes 2 m of fiber. See 167.3.1.
800GBASE-DR8 PMD	16 384	32	20.48	Includes 2 m of fiber. See 124.3.1.
800GBASE-DR8-2 PMD	16 384	32	20.48	Includes 2 m of fiber. See 124.3.1.

[a] For 800GBASE-R, 1 bit time (BT) is equal to 1.25 ps. (See 1.4.215 for the definition of bit time.)
[b] For 800GBASE-R, 1 pause_quantum is equal to 640 ps. (See 31B.2 for the definition of pause_quanta.)
[c] Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.
[d] If an implementation includes the 800GMII Extender, the delay associated with the 800GMII Extender includes two 800GXS sublayers.

Reported Values

Optical module Delay (PMA+PMD) reported by individuals affiliated with the following companies

- 55 ns by individual(s) from Marvell
- 60 ns by individual(s) from Broadcom

Proposal

- For the PMA Stage that appears next to the PMD, specify the sum PMA+PMD
- Leave a separate budget for the other PMA Stages
- Use a value of **200 ns** providing plenty of implementation variation for all cases where the PMD is optical
  - 800GBASE-VR8
  - 800GBASE-SR8
  - 800GBASE-DR8
  - 800GBASE-DR8-2

Summary

- For the PAUSE application, the key parameter is the total Delay (overall system latency)
- Delay is the sum of the Tx and Rx paths
- Optical module portion of the Delay depends upon the PMA and PMD Sublayers
  - PMD Sublayer is in the module
  - PMA Sublayer is distributed across the module and host
- Define the sum of the PMA and PMD Delay in the module
  - 200 ns