POTENTIAL ISSUES TO DISCUSS IN SA BALLOT

IEEE P802.3dm Task Force
Ragnar Jonsson – Infineon / South Hill Signal
November 2026



Supporters

Adee Ran - Cisco Systems, Inc.

Introduction

- We should expect and welcome comments on delay requirements in Clause 191.14, similar to comments #77 and #78 against D2.1
- There have been offline email discussions on the subject, initiated after the Lisbon meeting
- This presentation reviews key points related to the delay specification
  - Why have delay requirements?
  - How asymmetry changes the delay problem
  - Options available to the Task Force and their trade-offs

Goal

Have the analysis ready so the Task Force can resolve ballot comments efficiently.

Why is delay specified at all?

Two key reasons for delay limits on products:

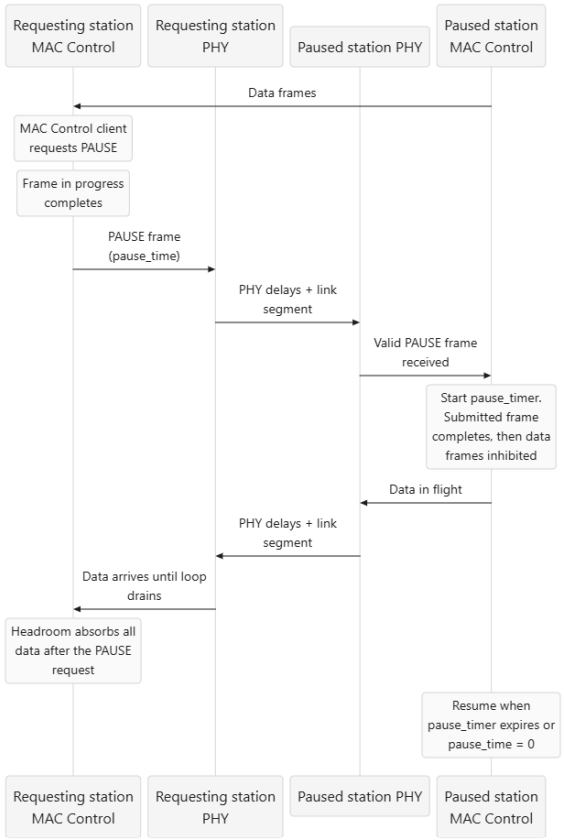
	1. Control signaling latency	2. PAUSE buffer sizing
Purpose	Bound delay of control signals over the link	Bound buffer size needed in higher layers for PAUSE
Interoperability impact	Not significant	**Yes**: sets required PAUSE buffering in the link partner
Where it belongs	Product datasheets	**The standard**

- IEEE 802.3 focuses on interoperability, so only PAUSE-related limits are captured in the standard
  - Normative limit = **worst-case upper bound** for buffer sizing, not a performance target
- Products can be expected to meet additional, tighter delay requirements
  - Driven by customer control-signaling needs
  - Specified and differentiated in datasheets

$$\text{Buffer} \approx R_{data} \times T_{round\text{-}trip} + \text{frame terms}$$

How PAUSE works

- MAC Control PAUSE (Clause 31, Annex 31B) inhibits transmission of data frames for a specified period of time
- Operation:
  1. A MAC Control client (e.g. buffer reaching a threshold) requests PAUSE; MAC Control sends a **PAUSE frame**
  2. The PAUSE frame carries `pause_time`, in pause_quanta (512 bit times)
  3. The receiving station starts its `pause_timer`; a frame already submitted to the MAC is not interrupted, then data frames are inhibited
  4. Transmission resumes when `pause_timer` expires, or a PAUSE frame with `pause_time` = 0 is received
- Data keeps arriving after the PAUSE is requested:
  - Frame in progress at the requesting station
  - **PAUSE response time** of the paused station (31B.3.7: at the MDI for symmetric PHYs; for asymmetric PHYs, MAC part at the XGMII + PHY maximum delays)
  - Frame already submitted at the paused station (not interrupted)
  - **Delay of the link segment** (left to the system designer)
- The requesting station needs **buffer headroom** for all of this
  - Annex 31B expresses the overrun as **max_overrun** = constant + frame_length



> **Note**
>
> The PHY maximum delays are the part of the headroom that the PHY specification controls

PAUSE in camera applications

HS direction (video, camera → ECU)

- The camera generates video continuously; it cannot simply stop the source
- If the camera is paused, it must either:
  - store the video in buffers (costly), or
  - drop video data
- Cameras are less likely than ECUs to have large data buffers
- PAUSE would only **move the dropping of video data from the ECU to the camera**
- So PAUSE is unlikely to be used on the HS link in 802.3dm camera applications

LS direction (control, ECU → camera)

- The camera might send PAUSE to the ECU to throttle control data
- Buffer impact is on the camera side and relatively small (100M data)

> **Tip**
>
> PAUSE support still needs a defined delay bound for interoperability, but this is not the primary use case

Symmetric PHYs: the familiar model

- Same rate in both directions
- Transmit and receive delay per PHY are specified in bit times (BT) or pause_quanta
- One unit of time maps cleanly to both directions
- The two link partners have identical specifications, so a per-PHY limit is enough
- PHY delay limits and the per-speed Annex 31B.3.7 values were written to be consistent with each other

Asymmetric PHYs: what is new

- Two different rates on the same link
  - High-speed path: 2.5G / 5G / 10G
  - Low-speed path: 100M
- Two different PHYs: **PHY_S** (e.g. camera) and **PHY_D** (e.g. ECU)
- Consequences:
  - Bit time and pause_quanta have no single meaning across both directions
  - The PAUSE loop crosses **both** directions and **both** PHYs

> **Important**
>
> The symmetric-PHY method does not carry over directly. A different approach is needed.

Why not pause_quanta for asymmetric PHYs?

- 1 pause_quantum = 512 bit times, so its duration **depends on the data rate**

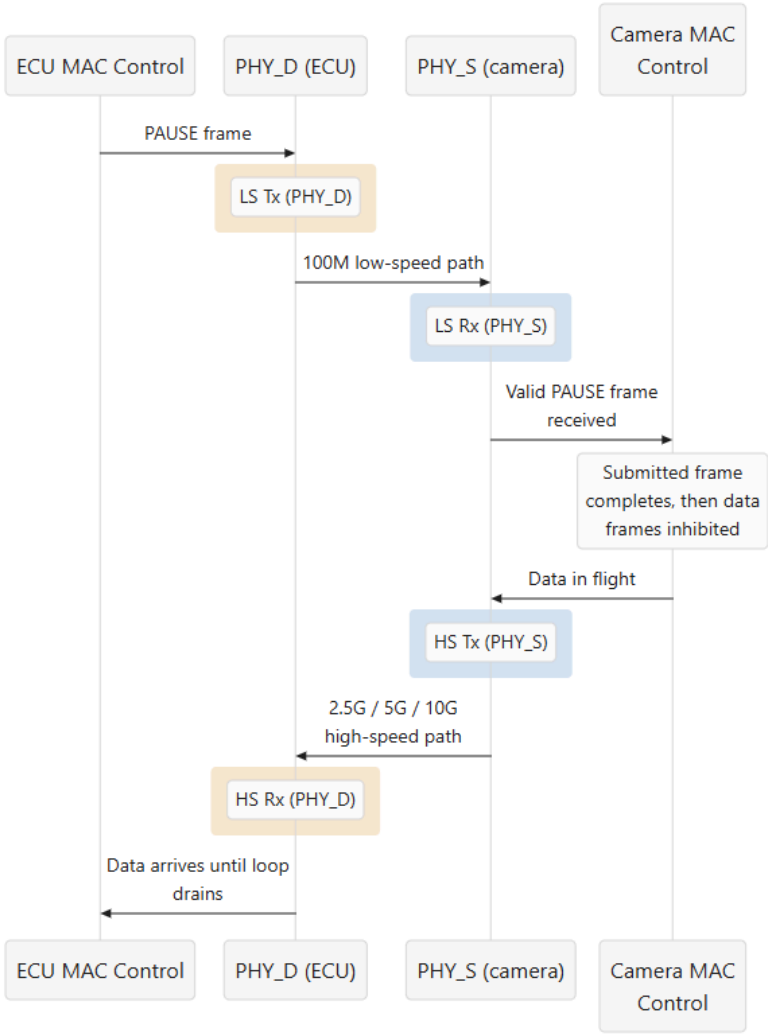
Path	Rate	1 pause_quantum
LS	100 Mb/s	5120 ns
HS	2.5 Gb/s	204.8 ns
HS	5 Gb/s	102.4 ns
HS	10 Gb/s	51.2 ns

- The delay loop includes both HS and LS paths, so there is **no single pause_quantum for the link**
  - LS quanta are very coarse: rounding up can add up to 5 µs of unnecessary buffering at HS rates
- pause_quanta has **no native meaning inside the PHY**
  - It is the unit of `pause_time` in MAC Control; the PHY has no need for it as a requirement
  - Annex 31B (D2.3) uses pause_quanta only for the MAC response, at the XGMII, with the quantum set by the paused station's transmit rate

> **Important**
>
> PHY maximum delays in Clause 191.14 are specified in **nanoseconds**, as in Clause 192

The PAUSE loop in an asymmetric link

- Per-PHY **maximum delay** (PHY share of the PAUSE response time, at the MDI):
  - PHY_S maximum delay = LS Rx + HS Tx
  - PHY_D maximum delay = HS Rx + LS Tx
- **Round-trip delay** = PHY_S + PHY_D maximum delays (link segment delay negligible)
- Reverse case (camera pauses ECU) has the same loop:
  - HS Tx + HS Rx + LS Tx + LS Rx = PHY_S + PHY_D
- **One pair of per-PHY maximum delays covers PAUSE in both directions**



How delay limits are defined in Clause 192

- Delay specified as **maximum delay** limits per PHY, in ns
- Single limit for PHY_S and single limit for PHY_D, **independent of rate**
  - See Table 192–27

Operating mode	Maximum delay [ns]
PHY_S	5120
PHY_D	11776
Round trip	**16896**

- Note that pause_quanta has been dropped from the delay requirements

Annex 31B for asymmetric PHYs (D2.3)

- Base 31B.3.7: PAUSE response time per operating speed, in pause_quanta (mostly at the MDI)
- D2.3 extends Annex 31B for the asymmetric PHYs of Clause 192:
  - 31B.2 `pause_time`: bit time = **transmit MAC bit time at the receiving station**
  - 31B.3.7: a limit per rate combination, measured at the **XGMII** (MAC response only)
  - PICS: MIIea–MIIib, TIM6a–TIM10b

Paused station	31B.3.7 limit (at XGMII)	Equivalent
PHY_S (HS transmit)	4 / 8 / 12 / 16 quanta at 2.5 / 5 / 7.5 / 10G	819 ns at every rate
PHY_D (100M transmit)	1 quantum	5120 ns

- Total PAUSE response at the MDI = **PHY maximum delay** (PHY clause, ns) + **MAC response** (Annex 31B, XGMII)
  - Clear split between PHY and MAC, and the quantum is unambiguous

Extending Annex 31B to Clause 191

- 191.14 cites Annex 31B as the reason for its delay constraints
- The D2.3 Annex 31B updates currently reference **Clause 192 only**
  - 31B.2 `pause_time` definition
  - 31B.3.7 paragraphs per rate combination
  - PICS rows
- The XGMII limits cover the MAC response only, so they do not depend on the PHY
  - The same approach, and likely the same values, can apply to Clause 191 at 2.5 / 5 / 10G
  - The Clause 191 PHY maximum delays (in ns) complete the total response time

> **Opportunity**
>
> Extend the existing D2.3 Annex 31B approach to Clause 191: a small change that keeps both asymmetric PHY clauses aligned

Options for Clause 191

	Option	Normative limits	Notes
A	Per transmit path (current draft)	HS path, LS path (one-way)	Not a limit on each PHY: Tx and Rx of different PHYs both contribute; split is informative in Annex 191A
B	Maximum delay per PHY (TDD style)	PHY_S, PHY_D	Simplest; Annex 191A could stay informative, or be removed later
C	Component limits	LS Tx, LS Rx, HS Tx, HS Rx	Each split needs margin; less trade-off freedom
D	Combined ("6-number")	4 components + PHY_S, PHY_D	Component caps need headroom; per-PHY maximum delay enforces boundary

- Current draft (Option A):
  - A one-way path delay is a system metric, not a compliance requirement for a single PHY
  - Informative split in Annex 191A was added to address this
  - Ballot comments (e.g. #77, #78) indicate there is room to improve this
- Common to all options:
  - Specify in **ns**, not pause_quanta
  - Compliance by design analysis/simulation, not silicon measurement
- Note on Option B: total PAUSE response = PHY Rx + PHY Tx (PHY clause) + MAC response (Annex 31B, at the XGMII)
  - A per-PHY maximum delay (Rx + Tx combined) slots directly into that sum

Splitting limits costs margin

- A per-PHY maximum delay lets designers trade delay between Tx and Rx
- Each split removes that freedom
  - Each component limit needs its own margin
  - So the sum of component limits is greater than an equivalent per-PHY maximum delay
- 802.3dm aims to optimize the relative cost of the camera link
  - Overly tight splits may constrain implementation choices

> **Tip**
>
> Fewer normative numbers → tighter overall bound for similar implementation freedom

Example baseline (for discussion)

Path	Interleave	Tx [ns]	Rx [ns]	Sum [ns]
2.5G HS	1x	680	3830	4510
5G HS	1x	340	1910	2250
5G HS	2x	450	2520	2970
10G HS	1x	170	960	1130
10G HS	2x	220	1260	1480
10G HS	4x	330	1870	2200
100M LS	—	1280	5120	6400

Rate	Interleave	PHY_S [ns]	PHY_D [ns]	Round trip [ns]
2.5G	1x	5800	5110	10910
5G	2x	5570	3800	9370
10G	4x	5450	3150	8600
Single limit (Option B)	all	**5800**	**5110**	**10910**

PHY_S = LS Rx + HS Tx PHY_D = HS Rx + LS Tx

- Highlighted cells: worst-case interleave rows used in the PHY totals; other rows shown for reference
- Maximum allowed values for budgeting PAUSE buffering (products may differentiate with lower delay)
- Less than the Clause 192 round-trip delay (16.9 µs)

Limit granularity: trade-offs

Single limit for all rates and interleaving depths (TDD style)

- ✅ Simplest specification and compliance statement
- ✅ One number per PHY for system integrators
- ⚠️ Bound is set by the slowest rate (2.5G)

Per-rate limits (worst-case interleaving)

- ✅ Smaller HS buffers at higher rates
- ⚠️ More table entries

Per-rate and per-interleaving limits

- ✅ Small additional HS buffer saving, only if deeper interleaving is never used
- ⚠️ Largest table; systems that can use deeper interleaving must size for it anyway

Observation: buffer ≈ rate × delay. Impact quantified on the next two slides.

Limit granularity: PAUSE on the HS direction

Worst case, only if PAUSE is used on the HS direction (unlikely). Buffer ≈ HS rate × round-trip limit (MAC response time and frame terms excluded), using the example baseline

Rate	Interleave	Single limit (all rates, all depths)	Per rate (worst-case depth)	Per rate & depth
2.5G	1x	27.3 kbit	27.3 kbit	27.3 kbit
5G	1x	54.6 kbit	46.9 kbit	43.3 kbit
5G	2x	54.6 kbit	46.9 kbit	46.9 kbit
10G	1x	109.1 kbit	86.0 kbit	75.3 kbit
10G	2x	109.1 kbit	86.0 kbit	78.8 kbit
10G	4x	109.1 kbit	86.0 kbit	86.0 kbit
Normative numbers		**1 per PHY**	**3 per PHY**	**6 per PHY**
10G worst-case buffer		**109.1 kbit (13.6 kB)**	**86.0 kbit (10.8 kB)**	**75.3–86.0 kbit**

- Maximum difference between cases: about **34 kbit (≈ 4 kB)** at 10G
- Single → per rate: about 21% smaller buffer at 10G (about 14% at 5G)
- Per rate → per rate & depth: up to about 12% smaller, only if deeper interleaving is never used
- Absolute sizes are a few kB in every case

Limit granularity: PAUSE on the 100M direction

Buffer ≈ 100 Mb/s × round-trip limit (MAC response time and frame terms excluded), using the example baseline

HS rate	Interleave	Single limit (all rates, all depths)	Per rate (worst-case depth)	Per rate & depth
2.5G	1x	1091 bit (136 B)	1091 bit (136 B)	1091 bit (136 B)
5G	1x	1091 bit (136 B)	937 bit (117 B)	865 bit (108 B)
5G	2x	1091 bit (136 B)	937 bit (117 B)	937 bit (117 B)
10G	1x	1091 bit (136 B)	860 bit (108 B)	753 bit (94 B)
10G	2x	1091 bit (136 B)	860 bit (108 B)	788 bit (99 B)
10G	4x	1091 bit (136 B)	860 bit (108 B)	860 bit (108 B)

- Largest difference between any two cases: about **42 bytes**
  - Less than one minimum-size Ethernet frame (64 B)
  - Small compared with the frame terms (up to 1518 B per maximum-size frame)
- For the likely camera use case, finer limit granularity gives **no meaningful buffer saving**

> **Tip**
>
> A single limit for all rates and interleaving depths is sufficient when PAUSE is used on the 100M direction

Summary

- Delay limits in the standard exist for **PAUSE buffer sizing**; signaling latency belongs in datasheets
- Asymmetry: pause_quanta is not a suitable unit for PHY delays, so specify them in **ns**
- Annex 31B: D2.3 approach (MAC limit at the XGMII, quantum from the paused station's transmit rate) covers Clause 192; extend it to Clause 191
- TDD (Clause 192) precedent: single maximum delay per PHY, which covers PAUSE in both directions
- Fewer splits → less margin needed and more implementation freedom
- PAUSE is unlikely on the HS (video) direction in camera applications
- Limit granularity has a small effect on buffer size (≈ 4 kB HS worst case, ≈ 42 B on 100M)

Key points

- The delay specifications in Clause 191.14 are there for **PAUSE operation**
- **HS direction:** PAUSE is very unlikely to be used
  - If it is, the maximum difference in ECU buffer size is about **34 kbit** (≈ 4 kB): small
- **LS direction:** if PAUSE is used
  - The difference in camera buffer size is about **42 bytes**: very small
- Delay limits are needed for interoperability, but there is **no need to overthink them**
- Observations relevant to the choice of option:
  - The PAR scope targets camera links: "optimized for automotive end-node camera links for operation up to 10 Gb/s in one direction and with a lower data rate in the other direction"
  - Total PAUSE response = per-PHY maximum delay (Rx + Tx) + Annex 31B MAC limit
- The D2.3 Annex 31B updates for asymmetric PHYs should be extended to Clause 191

> **Tip**
>
> Choose a simple, robust specification that the Task Force can agree on quickly

Discussion

- Option B (single PHY_S / PHY_D maximum delay, in ns) as the basis for comment resolution?
- With Option B, is Annex 191A still needed (keep as informative, or remove later)?
- Extend the D2.3 Annex 31B updates (31B.2, 31B.3.7, PICS) to Clause 191?
- Next steps: prepare baseline text for Clause 191.14 / Annex 191A

Call to Action

We should start working on these items now, with the goal of resolving them by the January meeting

THANK YOU!

