

800GbE Logic Skew

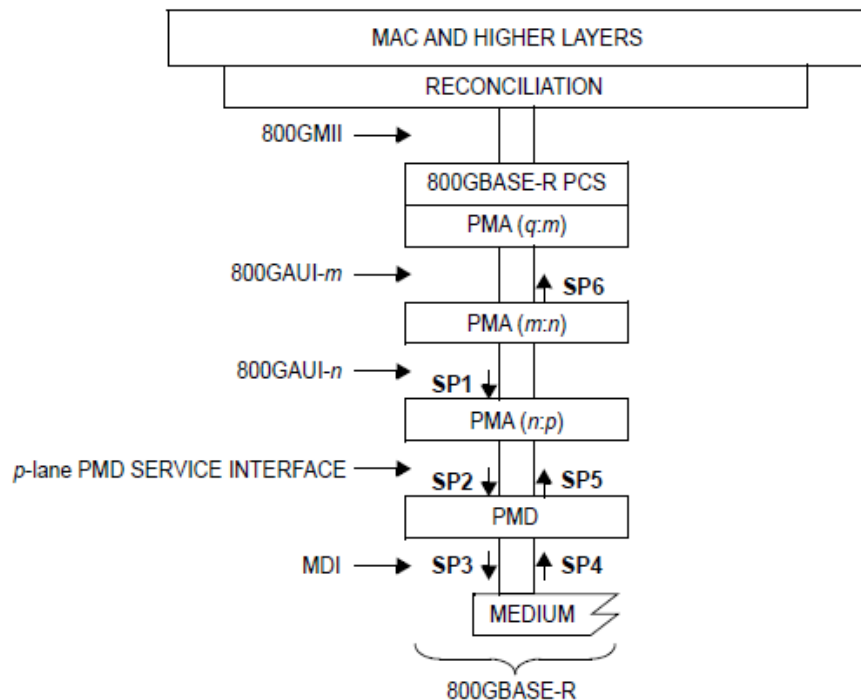
IEEE P802.3df Logic ad hoc
February 21, 2023

Mark Gustlin – Cisco

Introduction

- We have had the same skew budget since 100GbE days, as a comment from Adeo Ran pointed out against draft 1.1
- My view is that keeping that same budget is not a big concern since most designs are port groups and they must support the super set of speeds and skew budgets, but there is likely room for reductions
 - A reduced skew budget could have a small benefit for some single port or single speed designs
- Skew budgets are forever, and will break a link if we get it wrong, so we need to be careful if we reduce the skew budget
- This takes a new look at the total logic skew needed for 800GbE

Current Budget



800GAUI-n = 800 Gb/s ATTACHMENT UNIT INTERFACE
 800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
 MAC = MEDIA ACCESS CONTROL
 MDI = MEDIUM DEPENDENT INTERFACE
 PCS = PHYSICAL CODING SUBLAYER
 PMA = PHYSICAL MEDIUM ATTACHMENT

PMD = PHYSICAL MEDIUM DEPENDENT
 $q=32$
 $m=8$
 $n=8$
 $p=8$

Figure 169–5—800GBASE-R Skew points for a PHY with multiple 800GAUI-n

Each skew point is additive from the previous one

Table 169–5—Summary of Skew constraints

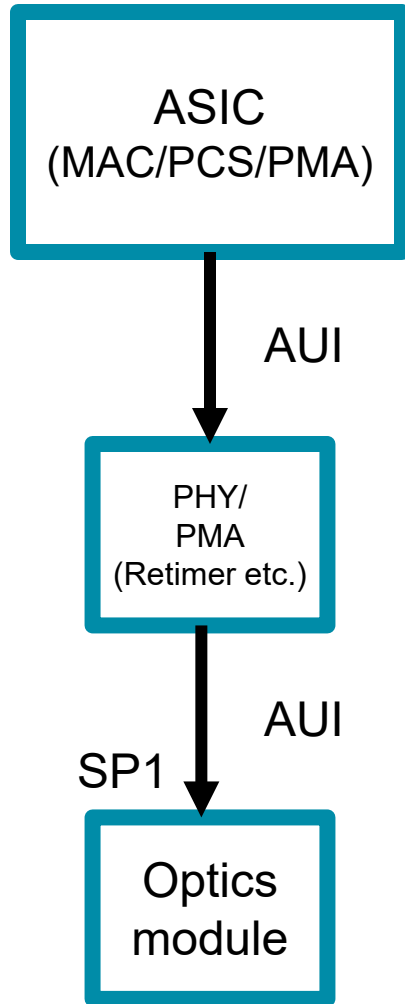
| Skew points | Maximum Skew (ns) ^a | Maximum Skew for 800GBASE-R PCS lane (UI) ^b | Notes ^c |
|----------------|--------------------------------|--|---|
| SP1 | 29 | ≈ 770 | See 173.4.3 |
| SP2 | 43 | ≈ 1142 | See 173.4.3, 124.3.2, 162.6.2, 163.6.2, 167.3.2 |
| SP3 | 54 | ≈ 1434 | See 173.4.3, 124.3.2, 162.6.2, 163.6.2, 167.3.2 |
| SP4 | 134 | ≈ 3559 | See 173.4.3, 124.3.2, 162.6.2, 163.6.2, 167.3.2 |
| SP5 | 145 | ≈ 3852 | See 173.4.3, 124.3.2, 162.6.2, 163.6.2, 167.3.2 |
| SP6 | 160 | ≈ 4250 | See 173.4.3 |
| At PCS receive | 180 | ≈ 4781 | See 172.2.5.1 |

^a The Skew limit includes 1 ns allowance for PCB traces that are associated with the Skew points.

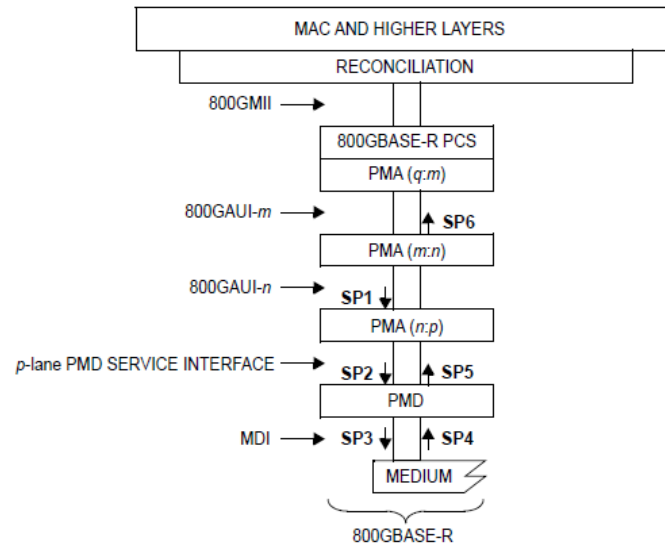
^b The symbol ≈ indicates approximate equivalent of maximum Skew in UI based on 1 UI equals 37.64706 ps at PCS lane signaling rate of 26.5625 GBd.

^c Should there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sublayer clause prevails.

SP1 View



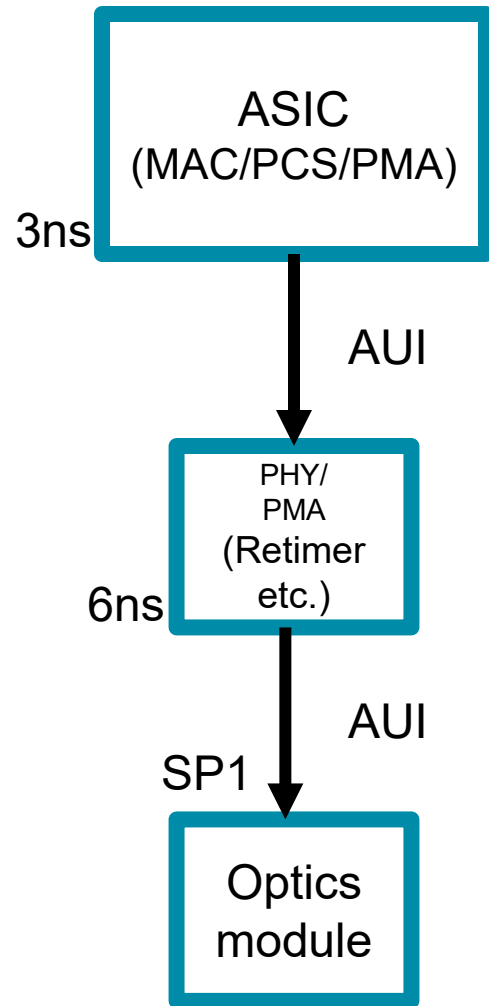
- SP1 has multiple possible components:
 - Skew from a large ASIC
 - Skew from a PHY type device (if used)
 - PCB board skew



800GAUI-n = 800 Gb/s ATTACHMENT UNIT INTERFACE
 800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
 MAC = MEDIA ACCESS CONTROL
 MDI = MEDIUM DEPENDENT INTERFACE
 PCS = PHYSICAL CODING SUBLAYER
 PMA = PHYSICAL MEDIUM ATTACHMENT
 PMD = PHYSICAL MEDIUM DEPENDENT
 q=32
 m=8
 n=8
 p=8

Figure 169-5—800GBASE-R Skew points for a PHY with multiple 800GAUI-n

SP1 View

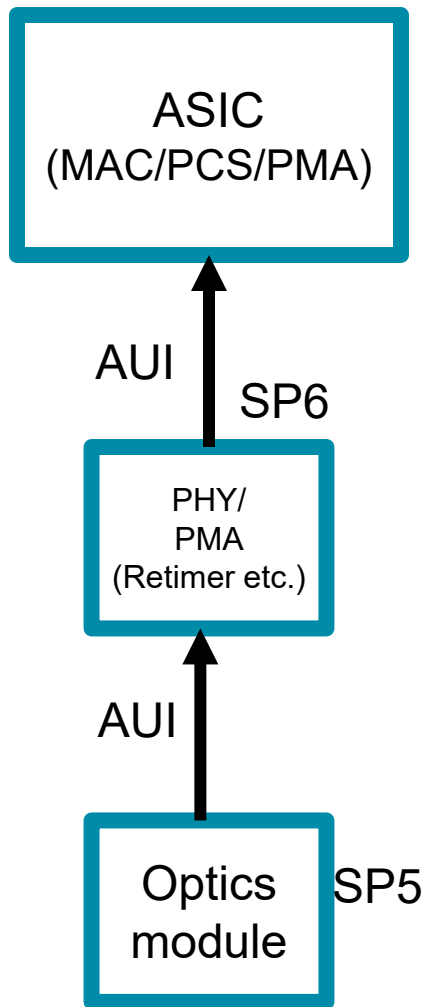


- ASIC contribution could be up to 3ns in the worst case for devices I surveyed
- PHY contribution is up to 6ns for devices I surveyed
 - If the PHY is just a retimer than it adds to the budget
 - If the PHY is a PCS/MAC/MACsec device, then the skew budget is reset at this point
- Board skew contribution
 - 2ns per 12" of PCB, allocate 6" of delta for a port? So assume 1ns?
 - This is pessimistic, but a small adder to the budget
- Total SP1 skew is ~10ns
- Think we should pad it by 5ns, so proposing 15ns total budget
 - Above skew is for a limited number of devices

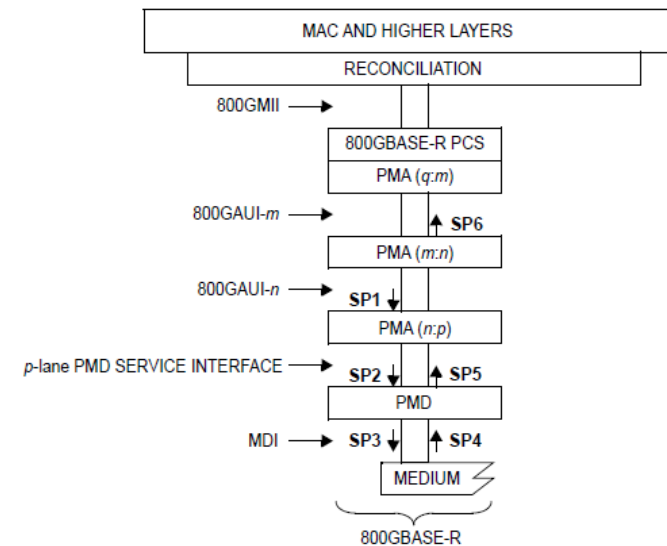
Adee's proposed remedy:

| Skew point | Contributor | Maximum (PCS UI) | Cumulative (PCS UI) | Cumulative (ns) | Reason |
|------------|--------------------------------------|------------------|---------------------|--------------------|------------------------------|
| SP1 | Tx PCS/PMA and possible external PMA | 192 | =770 192 | 29 ≈7.2 | PCS/PMA Tx + PMA Rx + PMA Tx |

SP6 View



- PHY contribution is up to 6ns for devices I surveyed
 - If the PHY is just a retimer than it adds to the budget
 - If the PHY is a PCS/MAC/MACsec device, then the skew budget is reset at this point
- Board skew contribution
 - 2ns per 12" of PCB, allocate 6" of delta for a port?
 - So assume 1ns.
- Unaccounted for skew in the optical module PMA
- Total SP6 skew is ~7ns + optical PMA skew
- Assuming the optical PMA skew is a few ns worst case
- Think we should pad it by 5ns, so proposing 15ns total budget?
 - Above skew is for a limited number of devices
 - Could refine this once we know more about the PMA in the optical module



800GAUI-n = 800 Gb/s ATTACHMENT UNIT INTERFACE
 800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
 MAC = MEDIA ACCESS CONTROL
 MDI = MEDIUM DEPENDENT INTERFACE
 PCS = PHYSICAL CODING SUBLAYER
 PMA = PHYSICAL MEDIUM ATTACHMENT

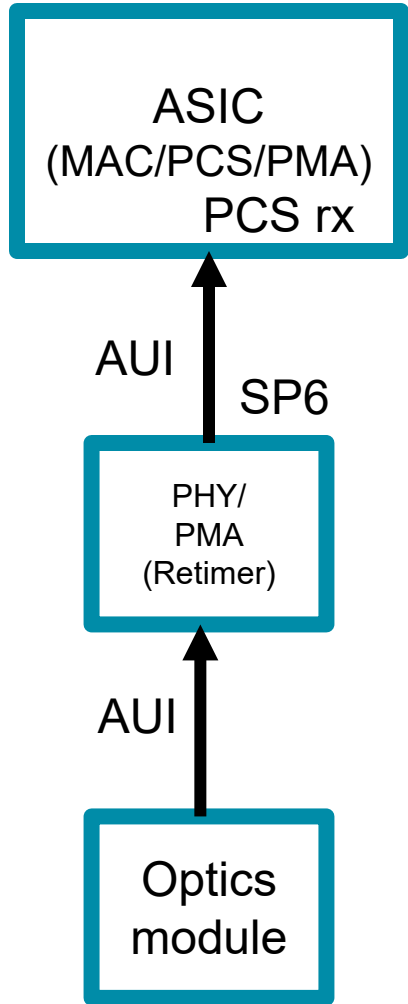
PMD = PHYSICAL MEDIUM DEPENDENT
 q=32
 m=8
 n=8
 p=8

Figure 169-5—800GBASE-R Skew points for a PHY with multiple 800GAUI-n

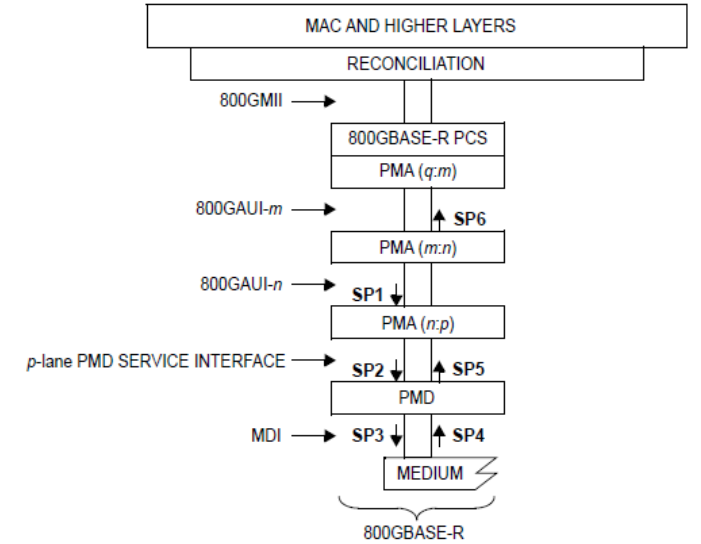
Adee's proposed remedy (9.6ns down from 15ns):

| | | | | | |
|-----|-----------------------------|-----|-----------------------|-----------|-------------------------------|
| SP5 | Module PMD Rx | 128 | ≈3852 1088 | 145 ≈41 | As in PMD Tx |
| SP6 | Module PMA and external PMA | 256 | ≈4250 1344 | 160 ≈50.6 | Two PMAs, each with Rx and Tx |

At PCS RX



- ASIC Internal contribution
 - 3ns, same as SP1 contribution
- Board skew contribution
 - 2ns per 12" of PCB, allocate 6" of delta for a port?
 - So assume 1ns.
- Total PCS skew is 4ns
- Think we should pad it by 4ns, so proposing 8ns total budget
- You could make the case that we don't specify this at all since this contribution is all internal to the ASIC!



800GAUI-n = 800 Gb/s ATTACHMENT UNIT INTERFACE
 800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
 MAC = MEDIA ACCESS CONTROL
 MDI = MEDIA DEPENDENT INTERFACE
 PCS = PHYSICAL CODING SUBLAYER
 PMA = PHYSICAL MEDIUM ATTACHMENT

PMD = PHYSICAL MEDIUM DEPENDENT
 q=32
 m=8
 n=8
 p=8

Figure 169-5—800GBASE-R Skew points for a PHY with multiple 800GAUI-n

Adee's proposed remedy (2.4ns down from 20ns):

| | | | | | |
|-----------|-----------------------------|-----|----------------------|----------------------|-------------------------------|
| SP6 | Module PMA and external PMA | 256 | 4250 1344 | 160 ≈50.6 | Two PMAs, each with Rx and Tx |
| PCS input | Rx PCS/PMA | 64 | 4781 1408 | 180 ≈53 | PMA Rx |

Summary

- I think both SP1 and SP6 should be 15ns each for total skew
 - SP6 can be revisited when we better understand the optics PMA skew budget
- PCS rx skew contribution should be set to 8ns
- We would need to investigate dynamic skew..no recommendations yet

Thanks!