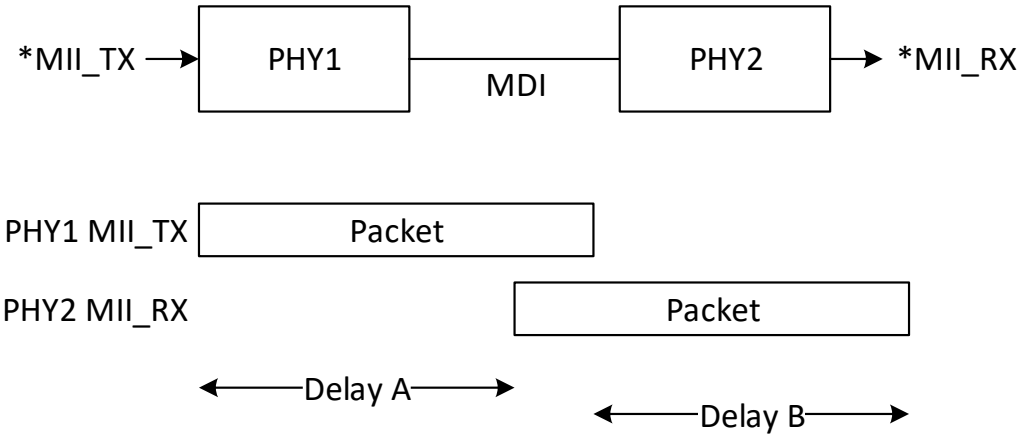
Upstream PHY Latency

William Lo

September 18, 2024

802.3 PHY Latency Reference Points

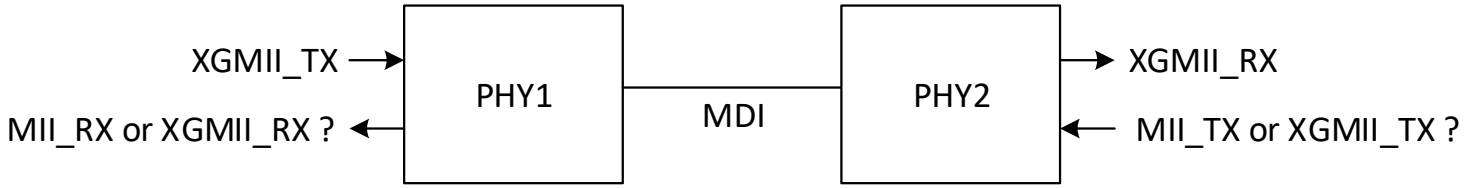
- From *MII_TX of PHY1 to MDI to *MII_RX of PHY2
- Latency bounded mostly for flow control round trip
- Examples
  - 100BASE-T1 MII Clause 96.10
  - 1000BASE-T1 GMII Clause 97.10
  - 2.5/5/10GBASE-T1 XGMII Clause 149.10
- Latency is invariant between start and end of packet
  - Delay A = Delay B



Unbalanced Link

- High Speed – Use XGMII as reference point
- Low Speed – Which to use as reference? MII or XGMII
  - Choice is available only if underlying PHY can burst faster than 100Mb/s
  - Doesn’t make sense to use XGMII if underlying PHY cannot burst faster than 100Mb/s

- Recommend deferring choice of MII vs XGMII in slow direction
  - Pick the one that naturally fits the selected modulation scheme



MII Updates

- 802.3dg adopted ordered sets for MII
- https://www.ieee802.org/3/dg/public/May_2024/Lo_3dg_01a_0724.pdf (slide 3)

- Can send Local and Remote Fault Indication

- So choice using MII vs XGMII on upstream should not be restricted by fault indication

Algorithm Latency

- Discussed at length in https://www.ieee802.org/3/bp/public/jul14/Lo_3bp_01a_0714.pdf
- Example applied in slides 4 to 8 of https://grouper.ieee.org/groups/802/3/dg/public/May_2022/Lo_3dg_01_011024.pdf

- Algorithm latency is the minimum theoretical latency
- Implementation latency assumed to be 0 in current discussion as this is vendor dependent

- Analysis that follows consider the following for algorithm delay
  - Encoder delay
  - FEC underflow prevention delay
  - FEC Frame aggregation delay
  - Data gating delays (i.e. EEE wake up, TDD Burst, rate conversion, etc)
- Other algorithm delay elements are minimal and assume to be 0 to simplify analysis

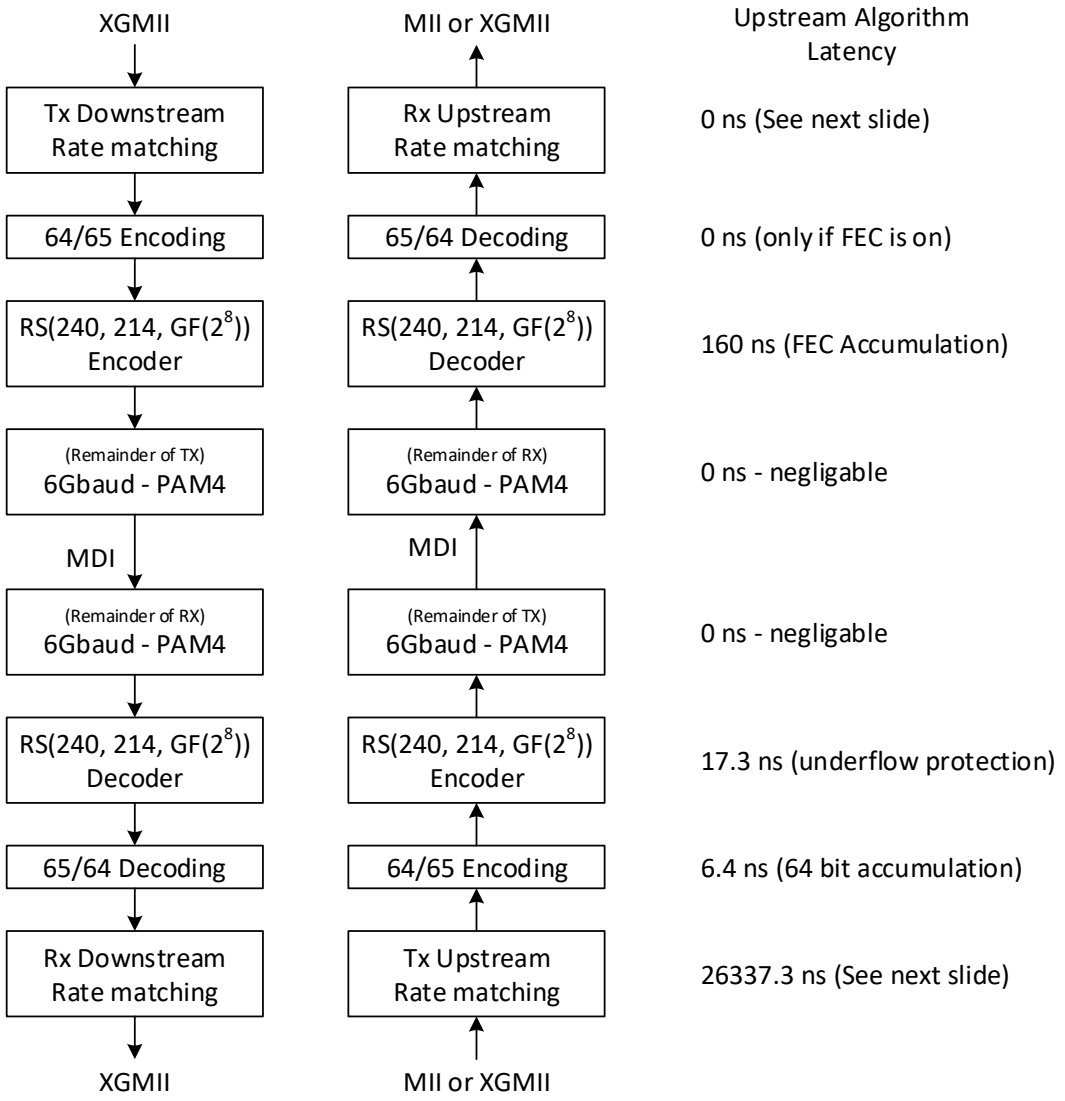
Where in Packet to Measure Delay

- Delay is invariant when measured at same point of packet
  - Doesn't really matter where as long as it is the same point in the packet
  - Side note: PHY delay invariance between packets is taken as a given otherwise causes problem for PTP.
- Realistically, cannot use data in packet until final CRC seen
  - Discard packet with bad CRC

- Analysis that follows measures latency from end of packet to end of packet
- Assume 8 bytes of preamble followed by 64 byte packet
- Assume first 32 byte in first segment and second 32 bytes in second segment
  - Segment is a TDD burst
  - Segment is 2 separate FEC frames or superframes
- Assumes going from sleep to wake in case of EEE

ASA-MLE (TDD) Analysis

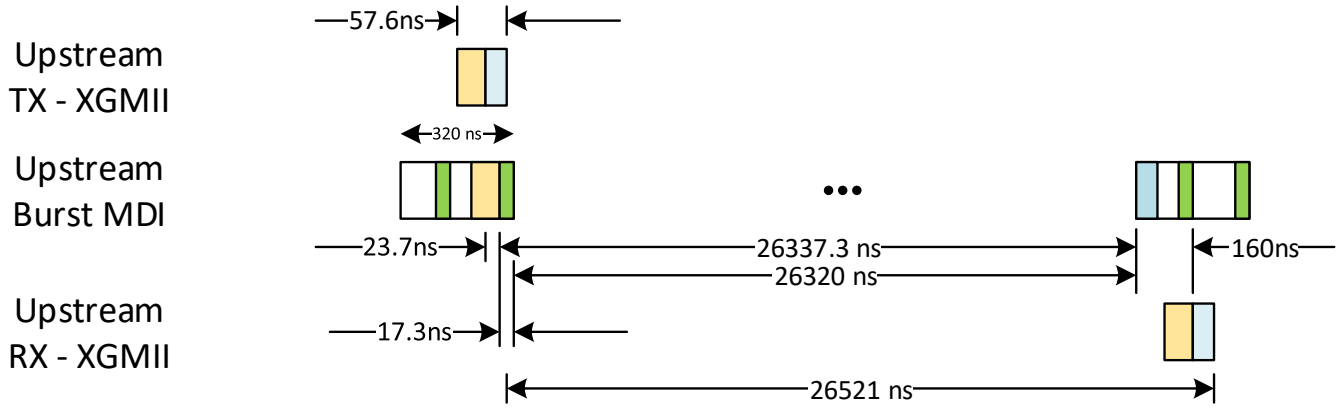
- 6.0 Gbaud PAM4 high speed links in both directions
- Rate matching throttles and schedules up stream and down stream bursts



ASA-MLE (TDD) Analysis – Using XGMII

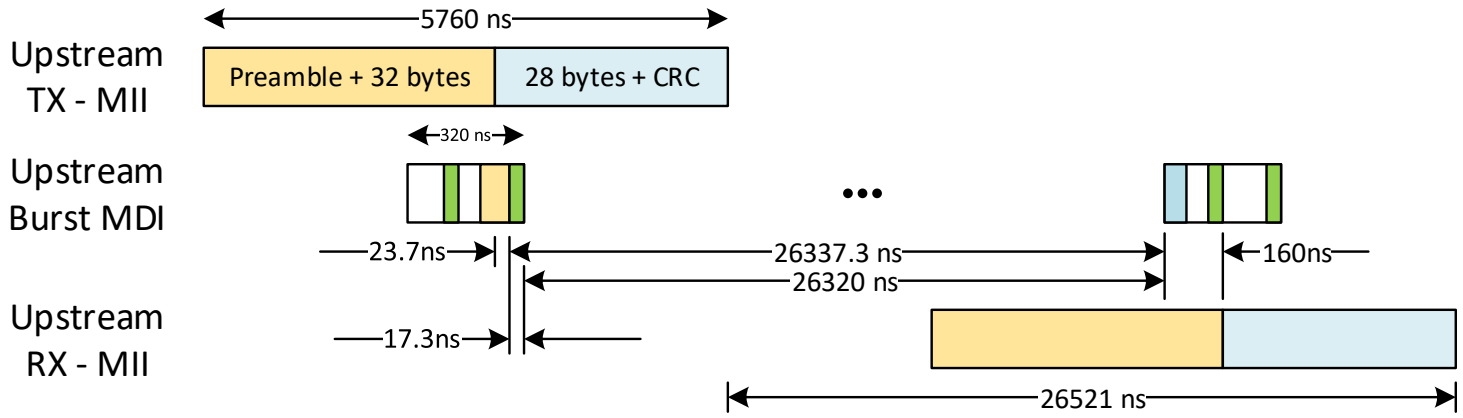
- Latency through PHY
  - 6.4 ns encoder accumulation
  - 17.3 ns transmit FEC underflow protection
  - 17.3 ns transmit FEC parity insertion
  - 26320 ns rate matching delay
  - 160 ns receive FEC accumulation

- Total Upstream PHY Algorithm Latency
  - 26521.0 ns
- Total Downstream PHY Algorithm Latency
  - Replace 26320 ns with 720 ns
  - 921.0 ns



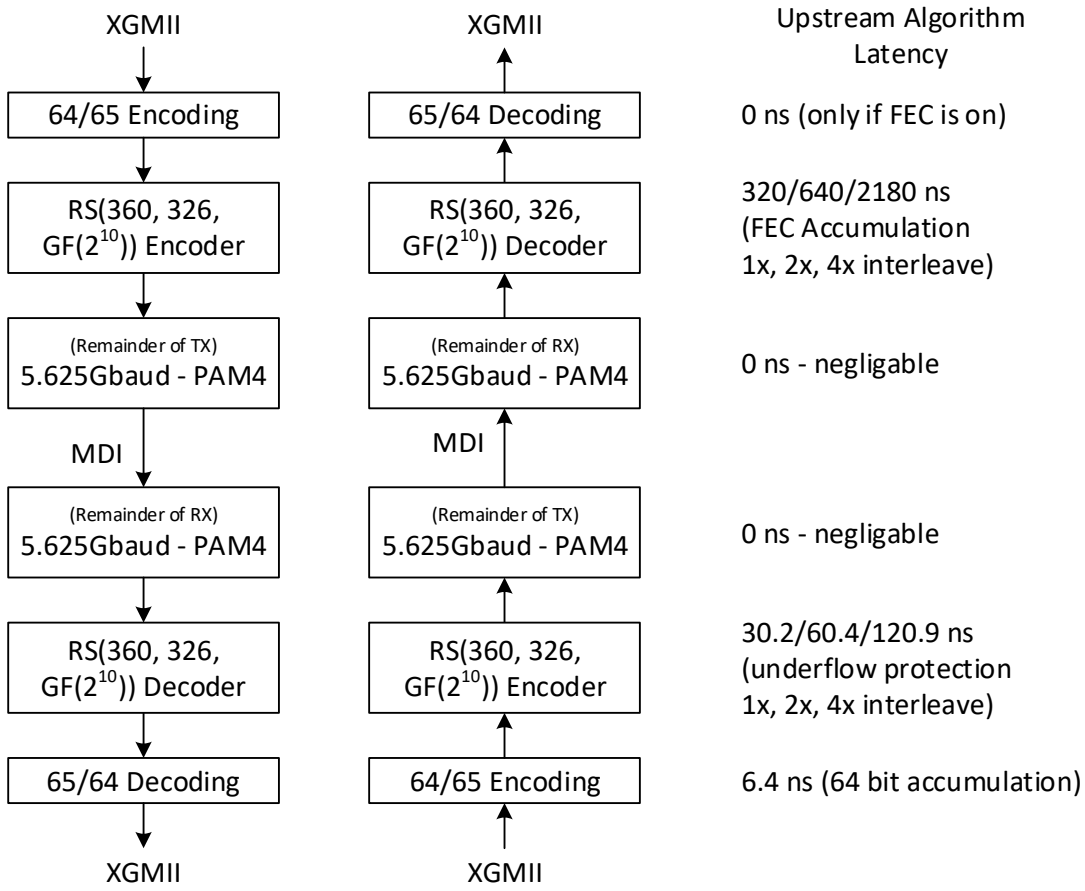
ASA-MLE (TDD) Analysis – Using MII

- Latency through PHY
  - End of packet to end of packet analysis does not change.
  - Same result as calculating with XGMII
- Total Upstream PHY Algorithm Latency
  - 26521.0 ns



10GBASE-T1 Analysis – No EEE

- 5.625 Gbaud PAM4 high speed links in both directions
- Upstream never shuts down
- MAC limits upstream rate but can send at any time

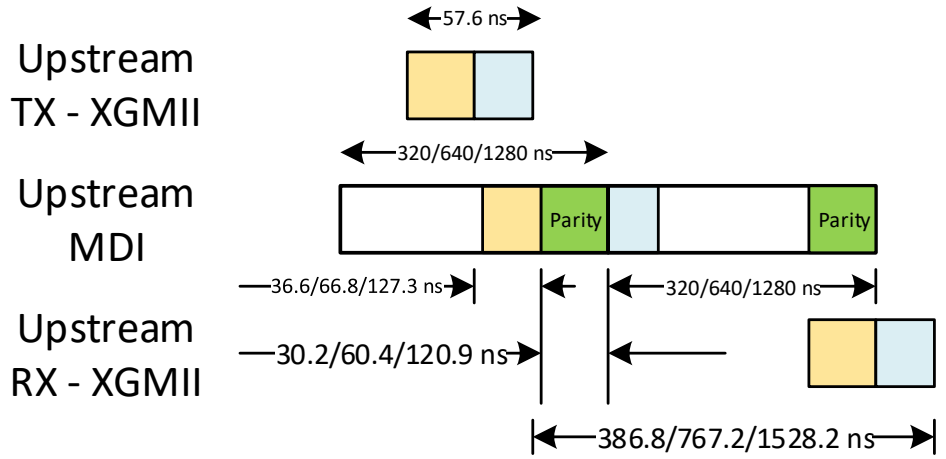


10GBASE-T1 Analysis – No EEE

- Latency through PHY
  - 6.4 ns encoder accumulation
  - 30.2, 60.4, 120.9 ns transmit FEC underflow protection. Depends on interleave depth of 1x, 2x, 4x
  - 30.2, 60.4, 120.9 ns transmit FEC parity insertion
  - 320, 640, 1280 ns receive FEC accumulation Depends on interleave depth of 1x, 2x, 4x

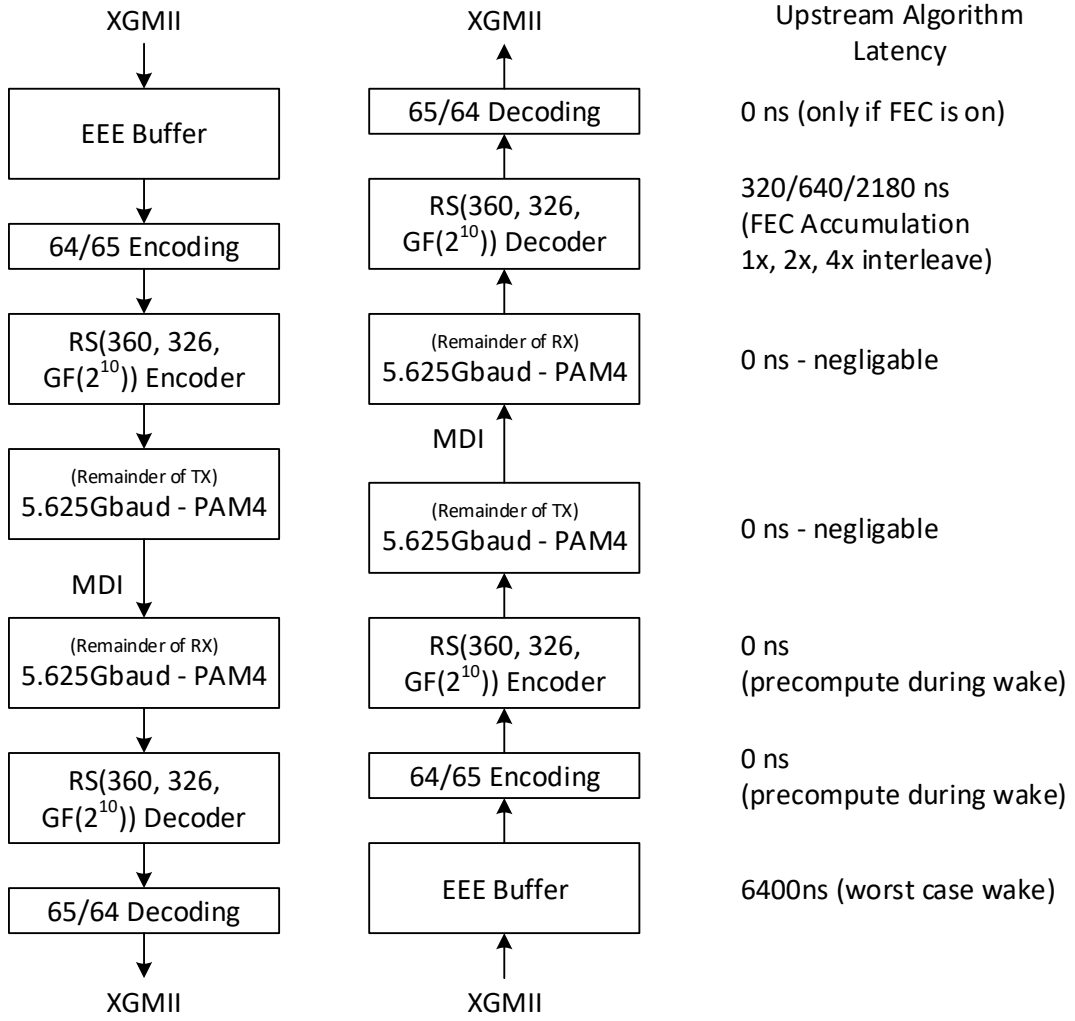
- Total Upstream PHY Algorithm Latency
  - 386.8, 767.2, 1528.2 ns

- Total Downstream PHY Algorithm Latency
  - 386.8, 767.2, 1528.2 ns



10GBASE-T1 Analysis – With EEE

- 5.625 Gbaud PAM4 high speed links in both directions
- Upstream shuts down when not used
- Buffer needed and delay incurred to hold packets to wake PHY

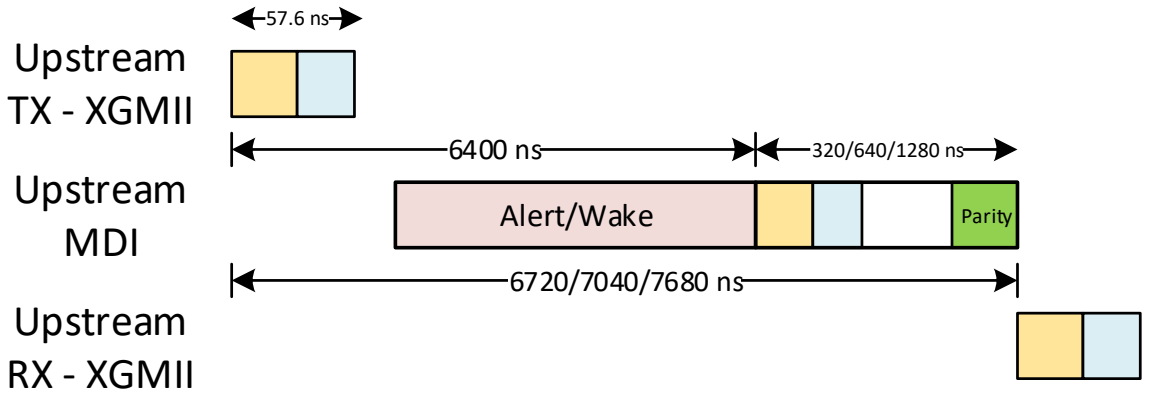


10GBASE-T1 Analysis – With EEE

- Latency through PHY
  - 6400ns worst case normal EEE wakeup
  - Transmitter data can be computed while waiting for wake up
  - 320, 640, 1280 ns receive FEC accumulation Depends on interleave depth of 1x, 2x, 4x

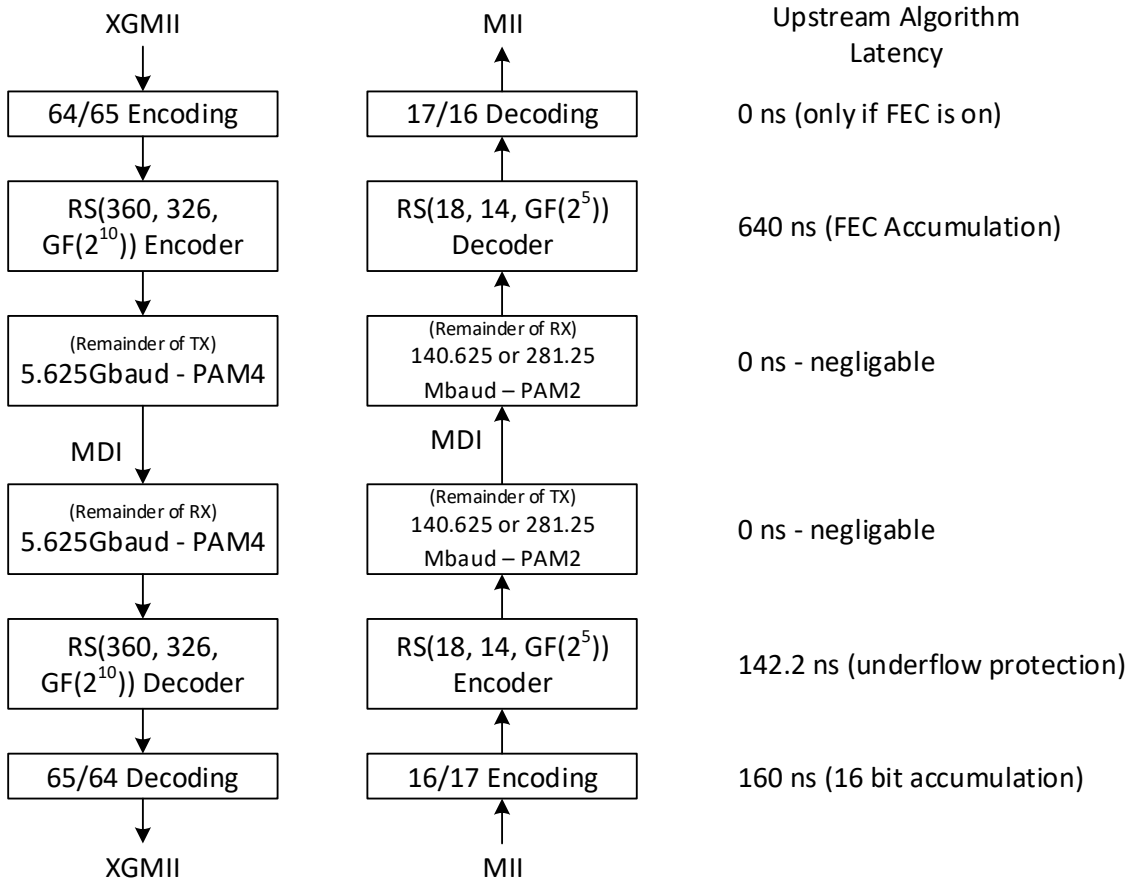
- Total Upstream PHY Algorithm Latency
  - 6720, 7040, 7680 ns

- Total Downstream PHY Algorithm Latency
  - Normal operation (no LPI)
  - 386.8, 767.2, 1528.2 ns



Lo_3dm_02_0924.pdf Proposal

- 140Mbaud or 281Mbaud PAM2 upstream
- 5.625 Gbaud PAM4 downstream

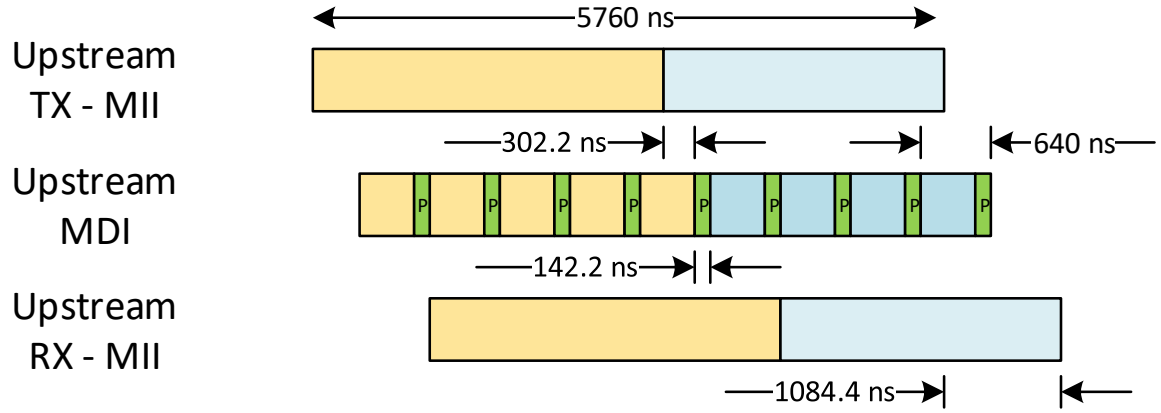


Lo_3dm_02_0924.pdf Analysis

- Latency through PHY
  - 160 ns encoder accumulation
  - 142.2 ns transmit FEC underflow protection.
  - 142.2 ns transmit FEC parity insertion
  - 640 ns receive FEC accumulation

- Total Upstream PHY Algorithm Latency
  - 1084.4 ns

- Total Downstream PHY Algorithm Latency
  - Same as to 10GBASE-T1
  - 386.8, 767.2, 1528.2 ns



Summary

		ASA-MLE (10G_M)	Lo_3dm_02_0924.pdf	10GBASE-T1 - No EEE			10GBASE-T1 - With EEE		
				1X	2X	4X	1X	2X	4X
Up Stream	Algorithm Latency (ns)	26512.0	1084.4	386.8	767.2	1528.2	6720.0	7040.0	7680.0
	Transmitter Speed (MBaud)	6000	140.625 or 281.25	5625					
	PAM	4	2	4					
	Reed Solomon	RS(240, 214, GF(2^8))	RS(18, 14, GF(2^5))	RS(360, 326, GF(2^10))					
	Max Burst Protection (ns)	8.7	71.1	15.1	30.2	60.4	15.1	30.2	60.4
	MII Type	MII or XGMII	MII	XGMII					
Down Stream	Algorithm Latency (ns)	921.0	386.8 / 767.2 / 1528.2	386.8	767.2	1528.2	386.8	767.2	1528.2
	Transmitter Speed (MBaud)	6000	5625						
	PAM	4							
	Reed Solomon	RS(240, 214, GF(2^8)	RS(360, 326, GF(2^10)						
	Max Burst Protection (ns)	8.7	15.1 / 30.2 / 60.4	15.1	30.2	60.4	15.1	30.2	60.4
	MII Type	XGMII							

THANK YOU