IEEE P802.3ba: Architecture Overview

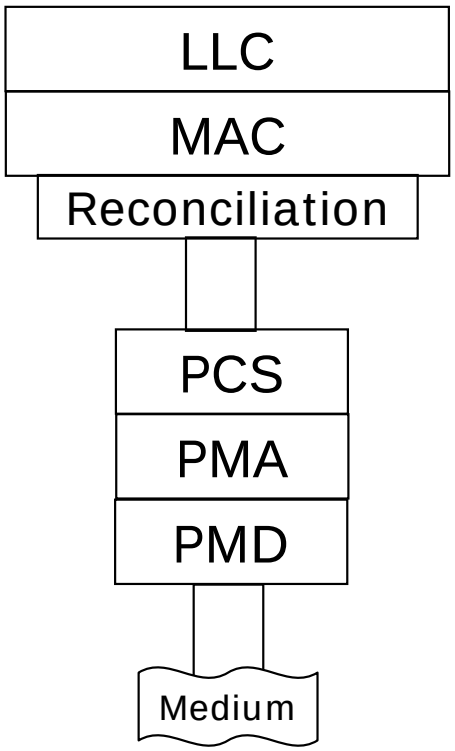
David Law
3COM

John D'Ambrosia
Force10 Networks

Time Synchronization Study Group

- Problem Statement per March CFI:
  - The IEEE supports two related time synchronization standards:
    - IEEE-1588
    - 802.1AS
  - Both need the same facilities from 802.3:
    - Notification of “start of frame” actually being transmitted at a well known point in the physical medium (e.g. the MDI)
    - Notification of “start of frame” being received at the same point
    - Some kind of reporting of the accuracy of the notification
- The IEEE P802.3ba architecture (Draft 2.0) is different than previous generations of Ethernet and needs consideration

Overview of Architecture

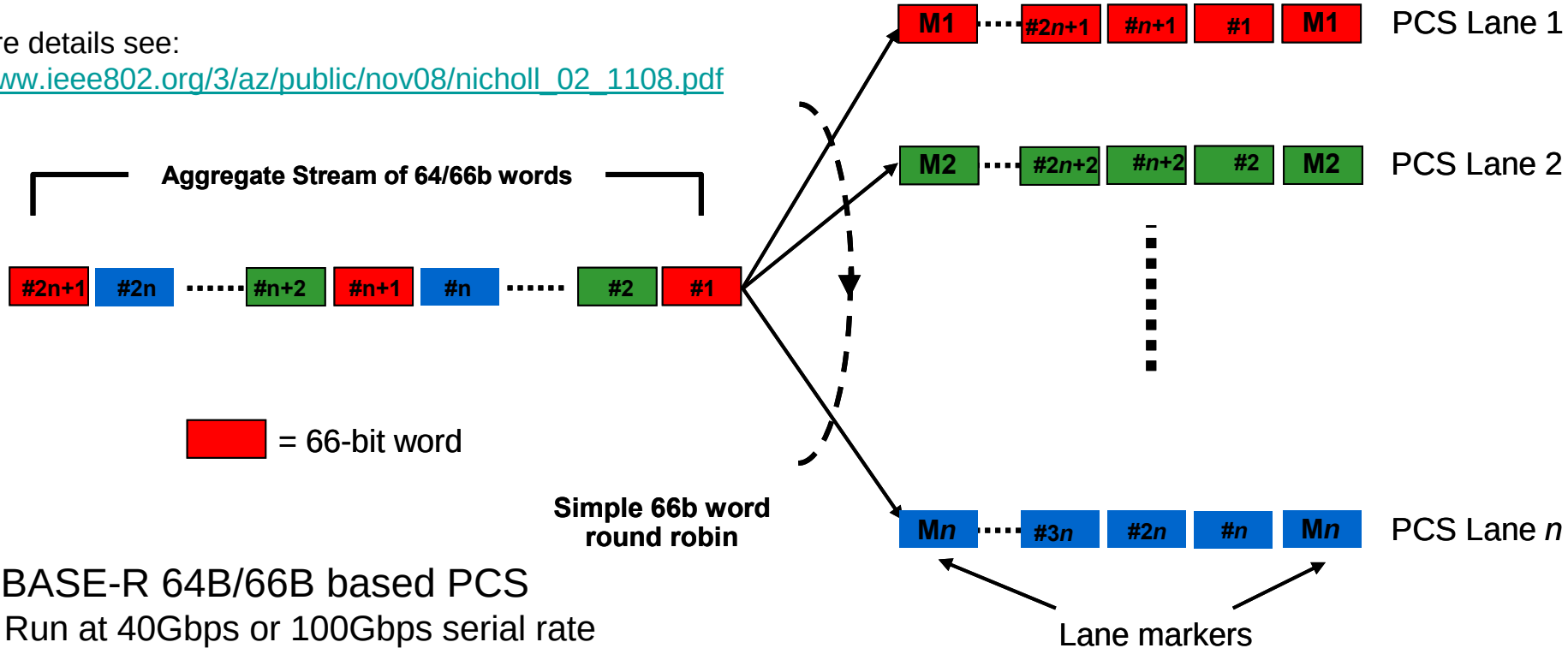


Generalized LAN
CSMA/CD Layers

- Consistent with previous Ethernet rates, extension to 40Gb/s & 100Gb/s data rates
  - Frame format; Services; Management attributes
- Media Access Control (MAC)
  - No changes to the MAC operation
- Physical Coding Sublayer (PCS)
- Physical Medium Attachment Sublayer (PMA)
- Physical Medium Dependent Sublayer (PMD)
- Interface Definitions
- Provide appropriate support for OTN

Physical Coding Sublayer (PCS)

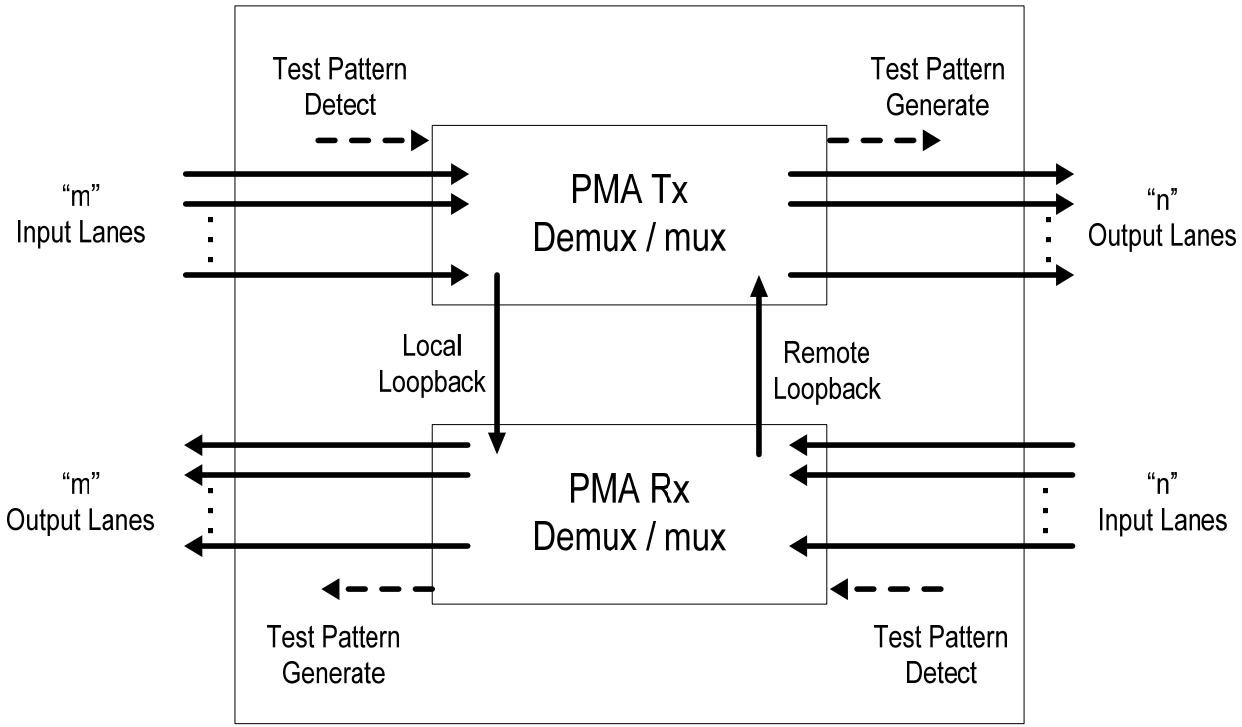
For more details see:
http://www.ieee802.org/3/az/public/nov08/nicholl_02_1108.pdf



- 10GBASE-R 64B/66B based PCS
  - Run at 40Gbps or 100Gbps serial rate
  - Includes 66 bit block encoding and scrambling
- Multi-Lane Distribution
  - Data is distributed across “n” PCS lanes 66 bit blocks at a time
  - 40GbE uses 4 PCS Lanes, 100GbE uses 20 PCS Lanes
  - Round robin distribution
  - Periodic alignment blocks added to each PCS lane for Rx PCS deskew
- Alignment and static skew compensation is done in the Rx PCS only

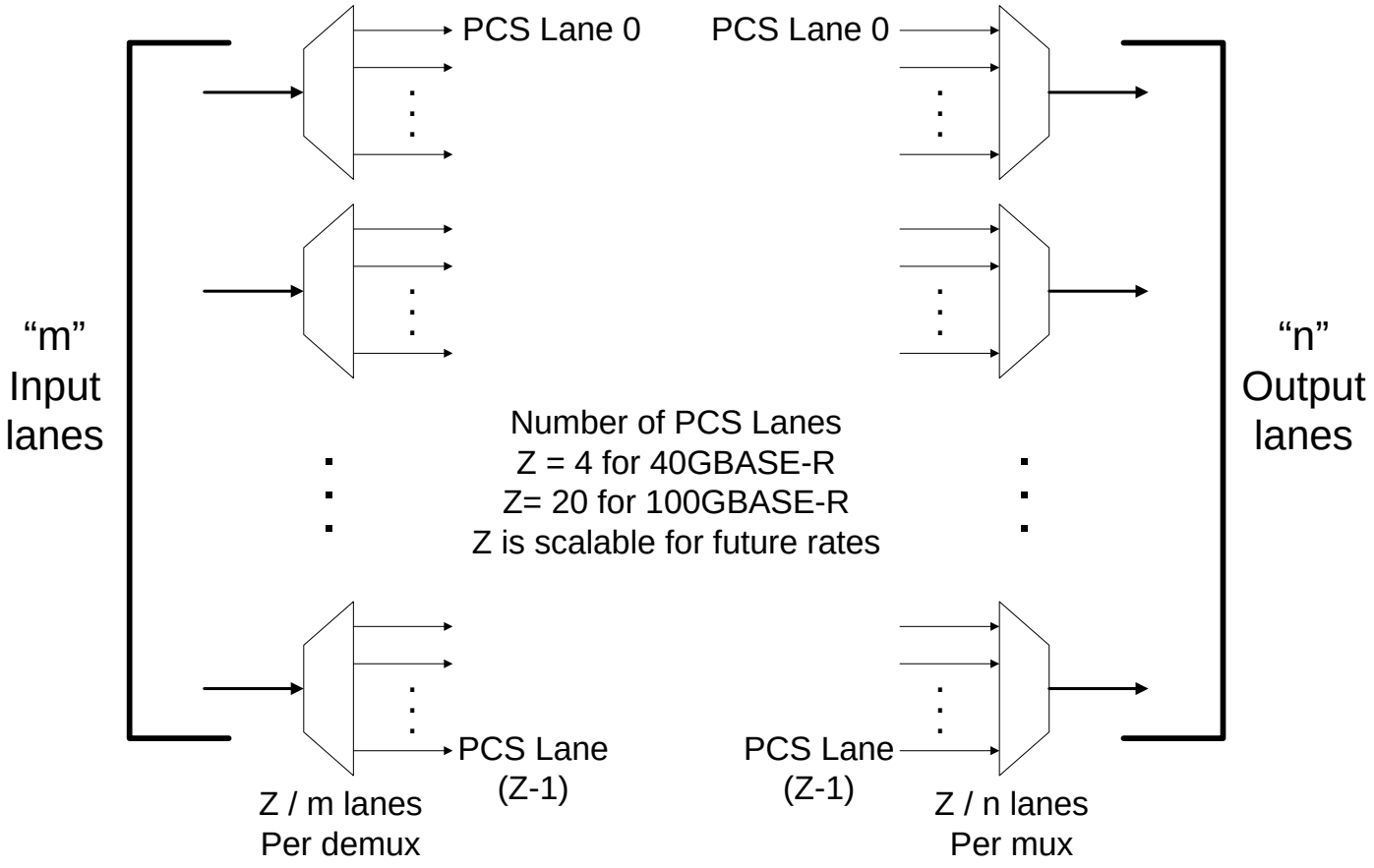
Source: D’Ambrosia, Law, Nowell, “40 Gigabit Ethernet and 100 Gigabit Ethernet Technology Overview,” Ethernet Alliance White Paper, http://www.ethernetalliance.org/images/40G_100G_Tech_overview(2).pdf. , November 2008.

Physical Medium Attachment (PMA)



- Parameterized PMA Sub-layer" Fundamental building block of IEEE P802.3ba Architecture
- Simple bit muxing
- Required for implementing retimed physical interfaces, XLAUI / CAUI (One PMA sub-layer adjacent to each end)
- There must be 1 instance of a PMA sub-layer
- There can be up to 4 instances of PMA sub-layers
- Optional Modes
  - Local loopback
  - Remote loopback
  - PMA Test Patterns
    - PRBS 2^9
    - PRBS 2^31
    - Square Wave

PMA Demux / Mux Functionality

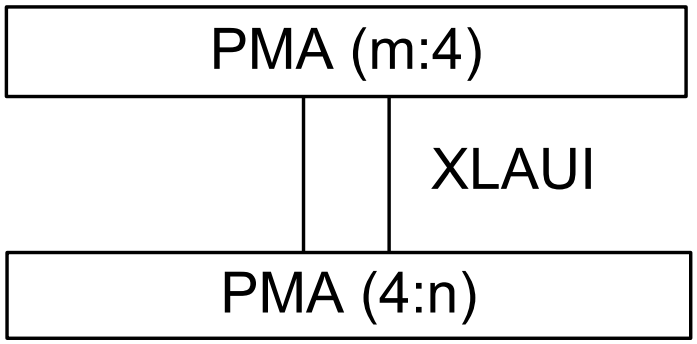


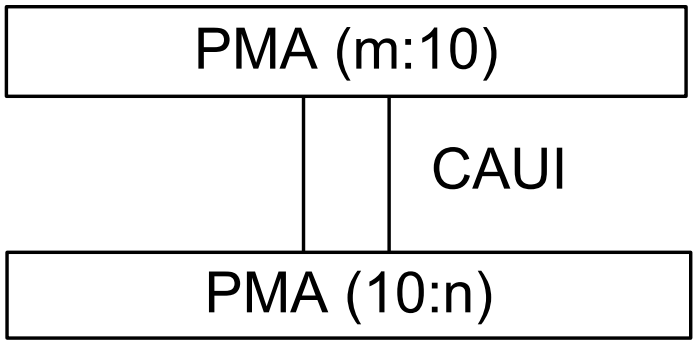
- Enables changing of lane number and rate per lane for multiple physical layer specifications
- For example 100GbE:
  - 10 x10 Gb/s
  - 5 x 20 Gb/s
  - 4 x 25 Gb/s
  - 2 x 50 Gb/s
  - 1 x 100 Gb/s

Shows PMA demux / mux functionality in one direction only

XLAUI / CAUI

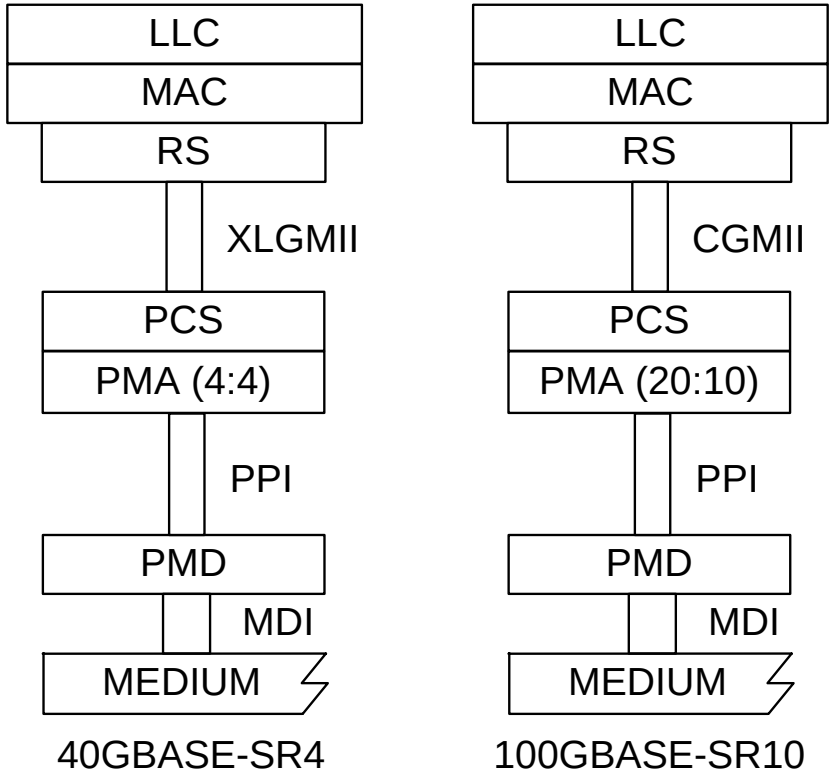
- An optional retimed interface that exists between two PMA sublayers
- XLAUI:
  - 40 GbE
  - 4 lanes x 10.3125 Gb/s
- CAUI
  - 100 GbE
  - 10 lanes x 10.3125 Gb/s
- Applications
  - Chip-to-chip
  - Chip-to-module
- Can have up to 4 instances of PMA sublayers (2 instances of a physical interface)





Parallel Physical Interface

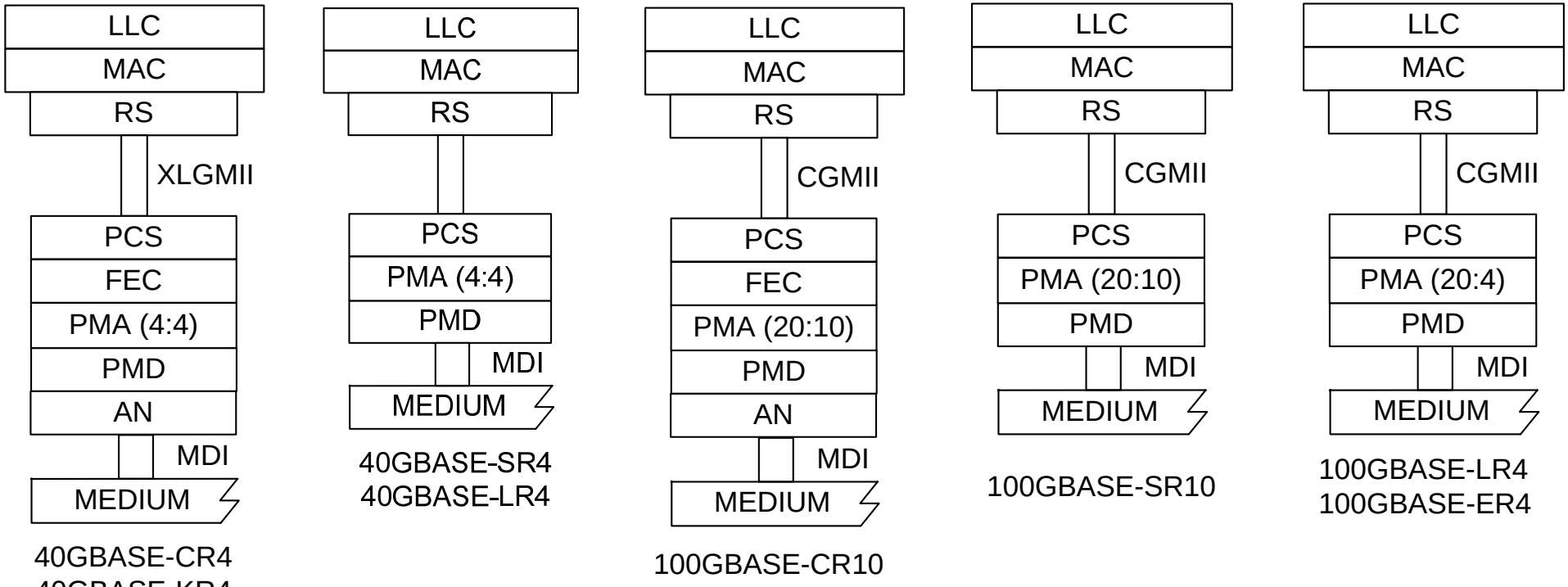
- An optional non-retimed physical instantiation of the PMD Service interface
- Optimized for –SR (≥ 100m OM3 MMF) PHYs
- 40 GbE
  - 4 lanes x 10.3125 Gb/s
- 100 GbE
  - 10 lanes x 10.3125 Gb/s



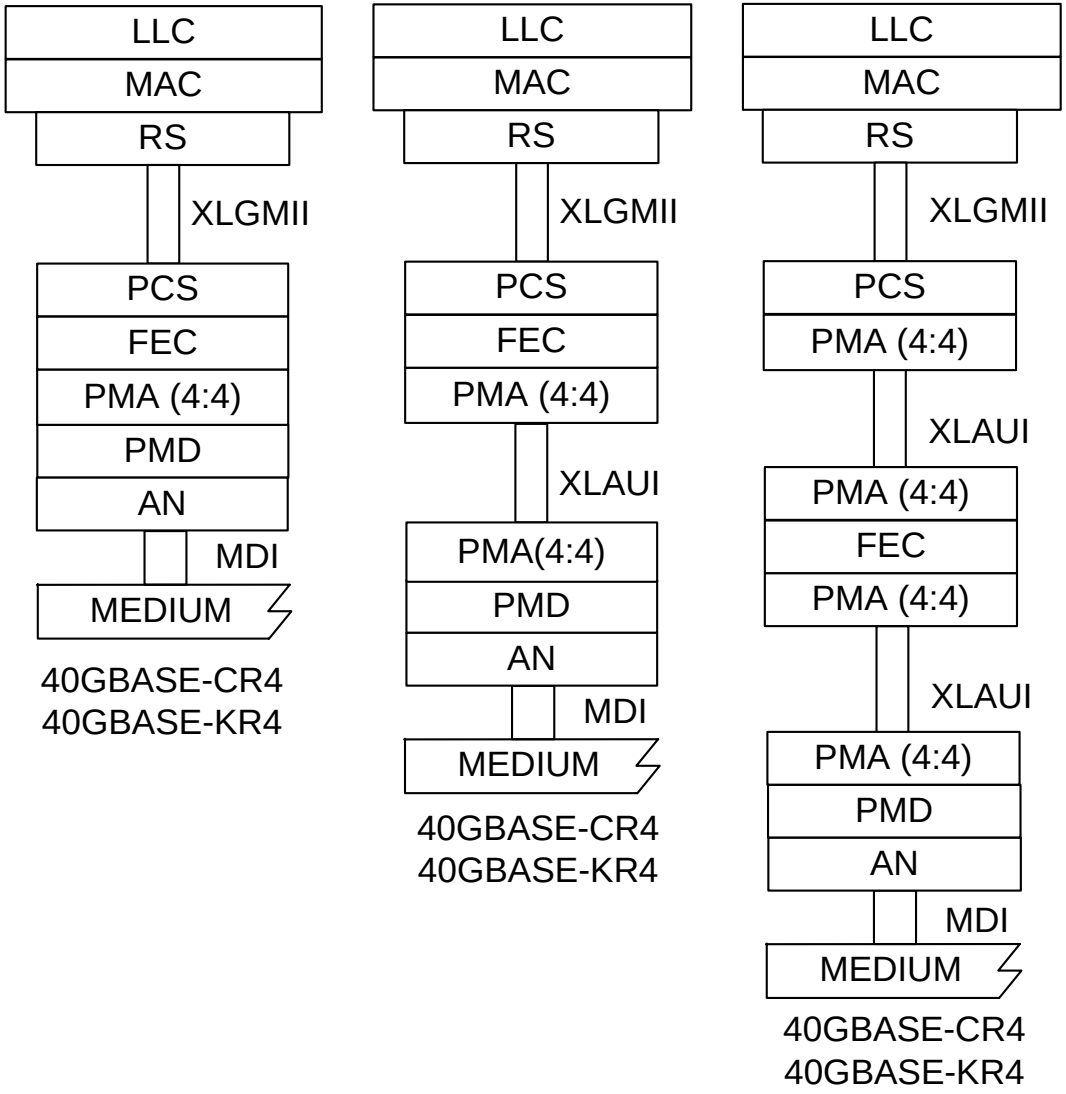
Summary: Physical Layer Specifications

Port Type	Description	40GbE	100GbE	Solution Space
40GBASE-KR4	At least 1m backplane	√		4 x 10 Gb/s
40GBASE-CR4 100GBASE-CR10	At least 10m cu (twin-ax) cable	√	√	“n” x 10 Gb/s
40GBASE-SR4 100GBASE-SR10	At least 100m OM3 MMF	√	√	“n” x 10 Gb/s
40GBASE-LR4	At least 10km SMF	√		4 x 10 Gb/s
100GBASE-LR4	At least 10km SMF		√	4 x 25 Gb/s
100GBASE-ER4	At least 40km SMF		√	4 x 25 Gb/s

The Different Basic Architectures

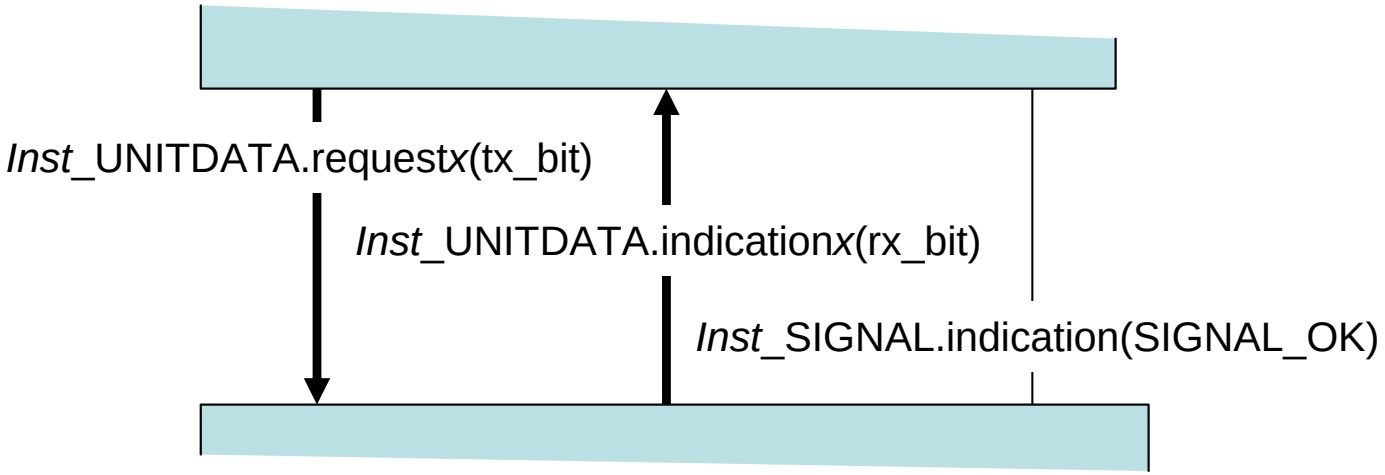


A Flexible Architecture
Example – 40GBASE-(C/K)R4



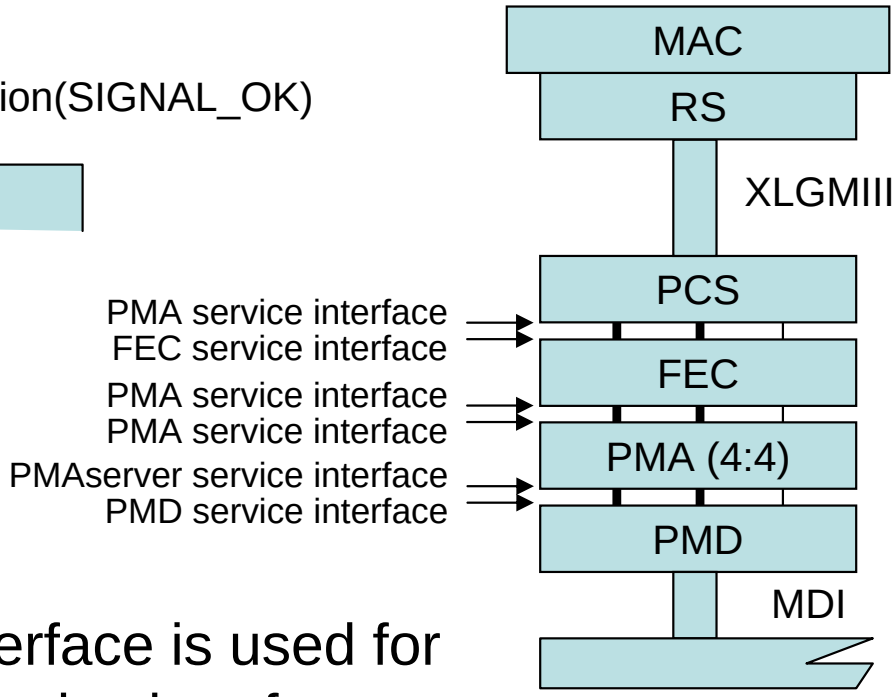
- Fixed locations:
  - PCS Sub-layer
  - PMD Sub-layer
- Flexible architecture to enable multiple implementations
  - Location of PMA Sub-layer(s)
  - Location of FEC Sub-layer
  - Optional Physical Instantiations)
- Key Building Block – the same abstract service interface enables the use of the same optional physical instantiation

Generic multi-lane abstract service interface



x = 0 through *p*-1

Inst = PMA for PMA service interface
PMAserver for PMA service interface
FEC for FEC interface
PMD for PMD interface



40GBASE-KR

The same multi-lane abstract service interface is used for the PMA, PMA server, FEC and PMD service interfaces supporting a flexible architecture with optional FEC and multiple PMA sublayers

Conclusions

- Initial discussions for Time Synchronization have focused on Notification of “start of frame” being transmitted / received at a well known point in the physical medium (e.g. the MDI)
- In the IEEE P802.3ba architecture the time from the MAC to the MDI is implementation dependent
- Location of time-stamping must account for potential for retiming (PCS) that can occur in a IEEE P802.3ba stack
- Passing signals up the stack?
  - One single service interface would need modification
  - Thru XLAUI / CAUI (and possible multiple instances)?
  - Thru PPI?