

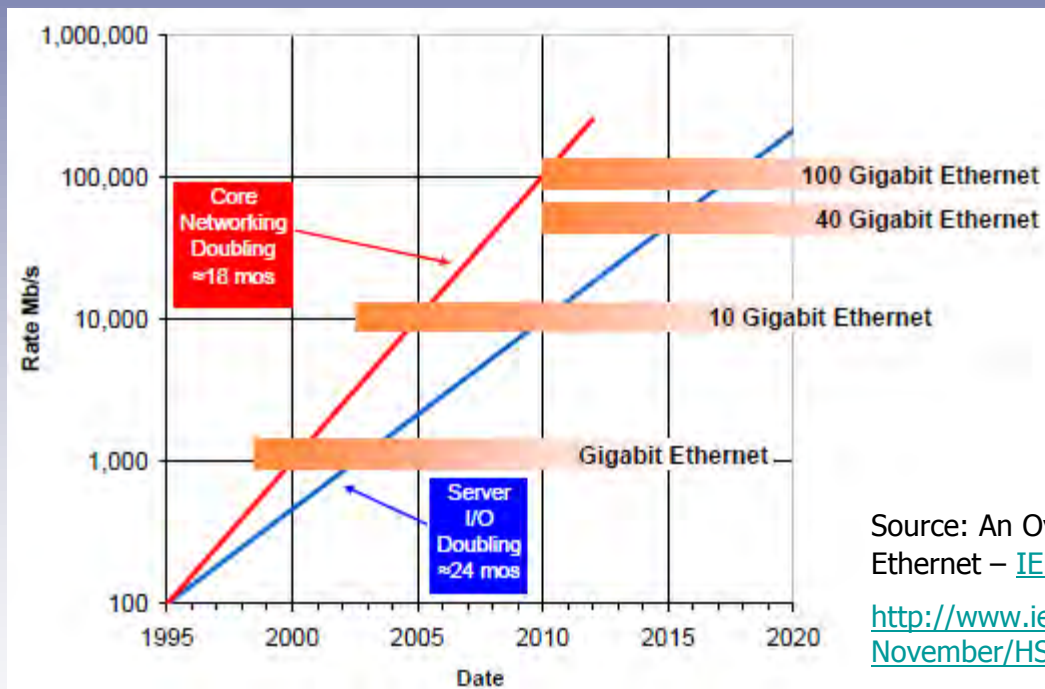
IEEE P802.3ba 40 and 100 Gigabit Ethernet Architecture

Ilango Ganga, Intel

- 40 GbE and 100 GbE
- IEEE P802.3ba overview
- 40 and 100 Gb/s Ethernet layer diagram
- 40 and 100 Gb/s sublayers
- 40 and 100 Gb/s architecture
- Compatibility interfaces
- 40 and 100 Gb/s implementation examples
- Summary

40 GbE and 100 GbE Computing and Networking

- 40 GbE to serve the compute/server BW and server traffic aggregation needs
- 100 GbE to serve the Network core and network aggregation needs



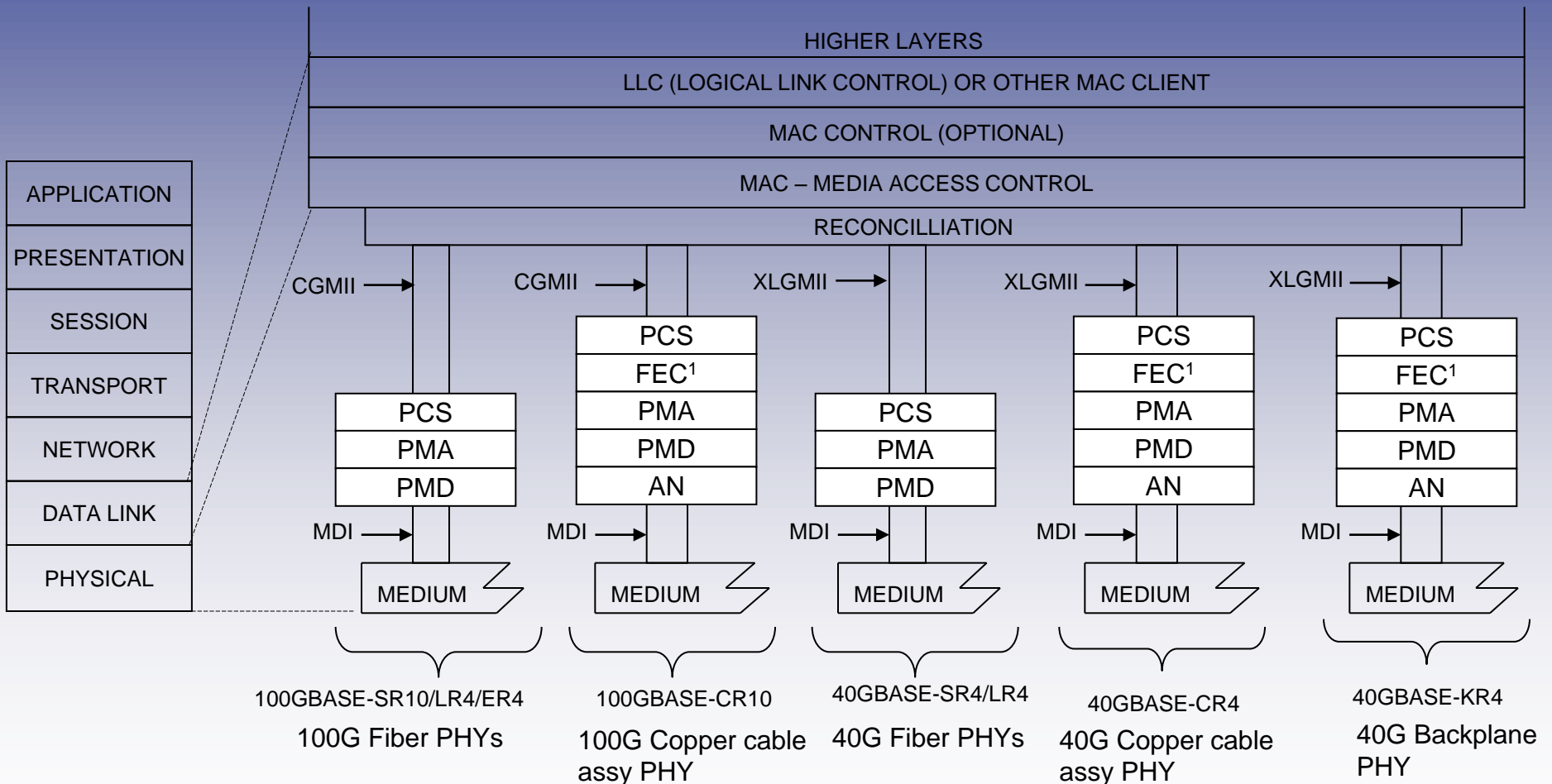
Source: An Overview: Next Generation of Ethernet – [IEEE 802 HSSG Tutorial 1107](http://www.ieee802.org/802_tutorials/07-November/HSSG_Tutorial_1107.zip)

http://www.ieee802.org/802_tutorials/07-November/HSSG_Tutorial_1107.zip

P802.3ba 40 and 100 GbE overview

- Common architecture for both 40 Gb/s and 100 Gb/s
- 40 and 100 GbE uses IEEE 802.3 Ethernet MAC and frame format
- Supports full duplex operation only
- Supports a BER of 10^{-12} or better
- 40 Gb/s PHYs
 - 40GBASE-KR4: 40 Gb/s backplane PHY (over 4 lanes) with a reach of up to at least 1 m
 - 40GBASE-CR4: 40 Gb/s shielded copper cable assy PHY (over 4 lanes) with a reach of up to at least 7 m
 - 40GBASE-SR4: 40 Gb/s multimode fiber PHY (over 4 lanes) with a reach of up to at least 100 m
 - 40GBASE-LR4: 40 Gb/s single-mode fiber PHY (over 4 WDM lanes) with a reach of up to at least 10 Km
- 100 Gb/s PHYs
 - 100GBASE-CR10: 100 Gb/s shielded copper cable assy PHY (over 10 lanes) with a reach of up to at least 7 m
 - 100GBASE-SR10: 100 Gb/s multimode fiber PHY (over 10 lanes) with a reach of up to at least 100m
 - 100GBASE-LR4: 100 Gb/s single-mode fiber PHY (over 4 WDM lanes) with a reach of up to at least 10 Km
 - 100GBASE-ER4: 100 Gb/s single-mode fiber PHY (over 4 WDM lanes) with a reach of up to at least 40 Km

40 and 100 GbE layer model



40 and 100 GbE sublayers

- MAC
 - Same as IEEE 802.3 MAC specified in Clause 4 / Annex 4A
 - Data Encapsulation, Ethernet framing, addressing, error detection (e.g. CRC)
- RS (Reconciliation sublayer)
 - The RS converts the MAC serial data stream to the parallel data paths of XLGMII (40 Gb/s) or CGMII (100 Gb/s)
 - Provides alignment at the beginning frame, while maintaining total MAC transmit IPG
- 40GBASE-R and 100GBASE-R PCS (Physical Coding sublayer)
 - Encodes 64 bit data & 8 bit control of XLGMII or CGMII to 66 bit code groups for communication with 40GBASE-R and 100GBASE-R PMA (64B/66B encoding)
 - Distributes data to multiple lanes, provides lane alignment and deskew
 - Management interface to control and report status
- Forward Error Correction sublayer
 - Optional sublayer for 40GBASE-R and 100GBASE-R to improve the BER performance of copper and backplane PHYs
 - Uses the same FEC functions as defined in Clause 74
 - Operates on a per PCS lane basis at a rate of 10.3125 GBd for 40G and 5.15625 GBd for 100G

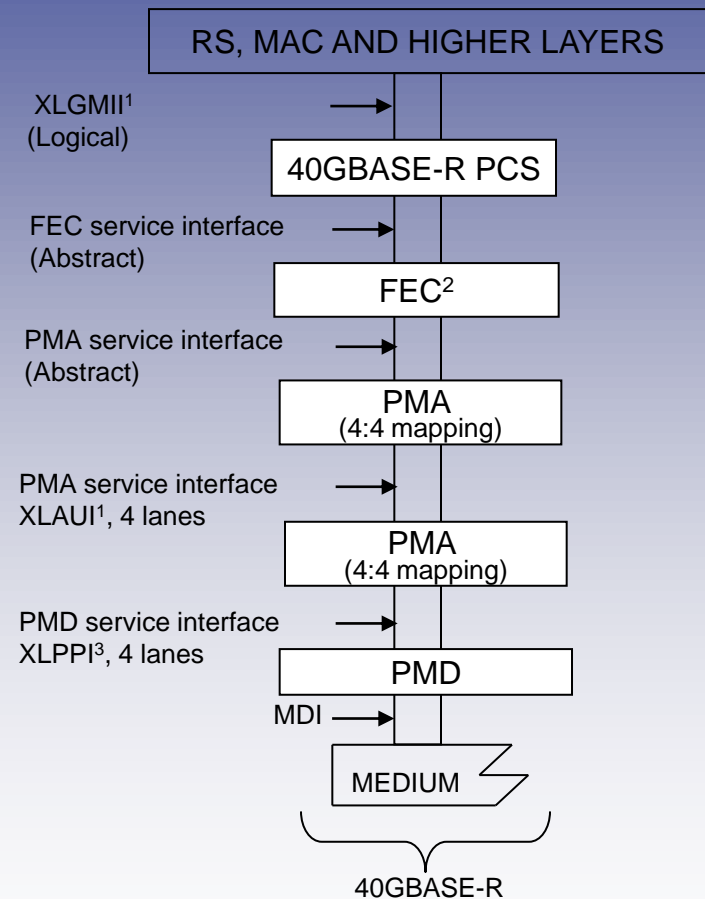
40 and 100 GbE sublayers ..2

- 40GBASE-R and 100GBASE-R PMA (Physical Medium Attachment)
 - Adapts PCS to a range of PMDs
 - Provides bit level multiplexing or mapping from n lane to m lanes
 - Provides clock and data recovery
 - Provides optional loopback and test pattern generation/checking functions
- 40GBASE-R and 100GBASE-R PMD (Physical Medium Dependent)
 - Interfaces to various transmission medium (e.g., backplane, copper or optical fiber medium)
 - Transmission/reception of data streams to/from the underlying medium
 - Provides signal detect and fault function to detect fault conditions
 - 40G PMDs: 40GBASE-KR4, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4
 - 100G PMDs: 100GBASE-CR10, 100GBASE-SR10, 100GBASE-LR4, 100GBASE-ER4
- Auto-Negotiation
 - Clause 73 Auto-Negotiation is used for copper and backplane PHYs to detect the capabilities of the link partners and configure the link to appropriate mode
 - Allows FEC capability negotiation, and provides parallel detection capability to detect legacy PHYs
- Management interface
 - Uses the optional MDIO/MDC management data interface specified in Clause 45 for management of 40 and 100 Gigabit Physical layer devices

40 GbE architecture

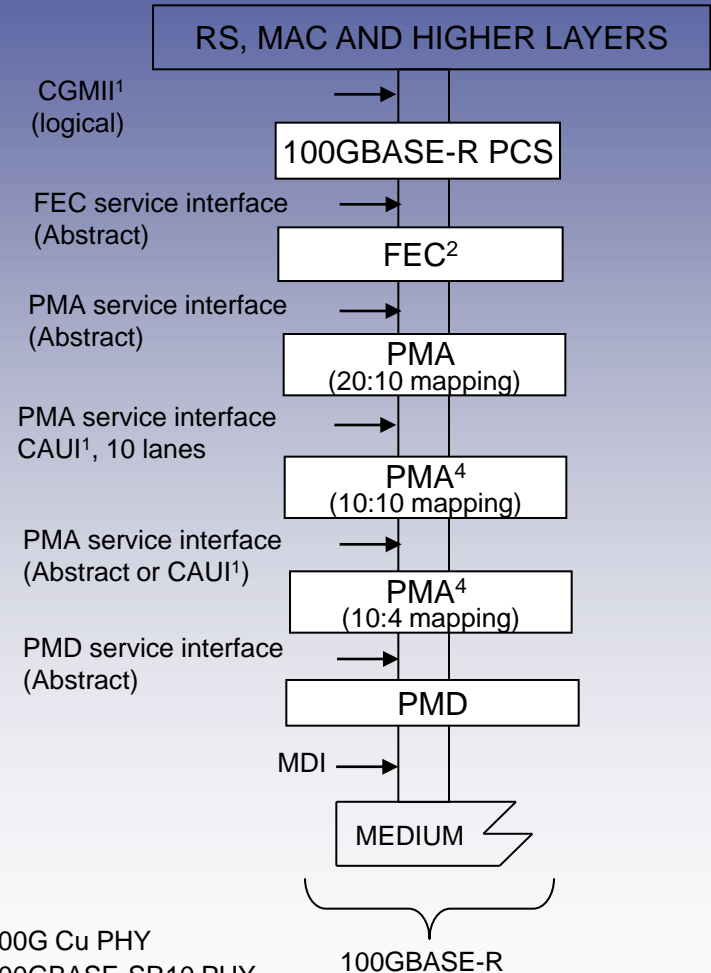
- XLGMII¹ (intra-chip)
 - Logical, data/control, clock, no electrical specification
- 40GBASE-R PCS
 - 64B/66B encoding
 - Lane distribution and alignment
- XLAUI¹ (chip-to-chip or chip-to-module)¹
 - 10.3125 GBaud electrical interface
 - 4 lanes
 - Physical instantiation of PMA service interface
- FEC service interface
 - Abstract
- PMA service interface
 - Abstract, can be physically instantiated as XLAUI electrical interface
- XLPPPI³ (chip-to-module)
 - 10.3125 GBaud electrical interface
 - 4 lanes, optional for use with non retimed 40GBASE-SR4/LR4 optical PHY modules
- PMD service interface
 - Logical

Note: 1. Optional
 2. Optional for 40G Cu & backplane PHYs
 3. Optional for 40G optical PHYs



100 GbE architecture

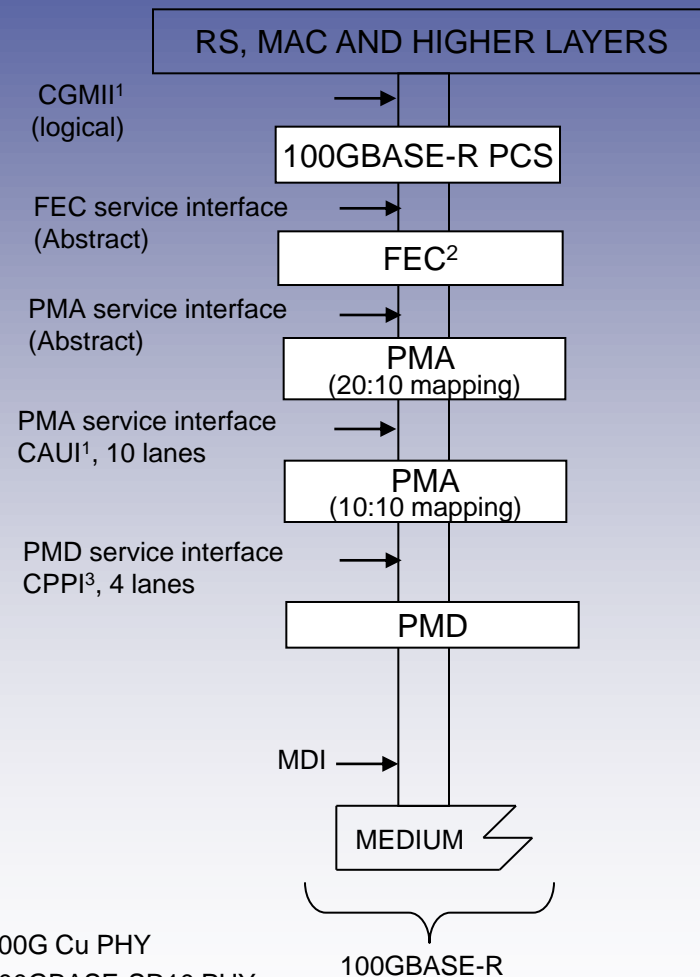
- CGMII¹ (intra-chip)
 - Logical, data/control, clock, no electrical specification
- 100GBASE-R PCS
 - 64B/66B encoding
 - Lane distribution and alignment
- CAUI¹ (chip-to-chip or chip-to-module)
 - 10.3125 GBaud electrical interface
 - 10 lanes
 - Physical instantiation of PMA service interface
- FEC service interface
 - Abstract
- PMA service interface
 - Abstract, can be physically instantiated as CAUI electrical interface
- CPPI³ (chip-to-module)
 - 10.3125 GBaud electrical interface
 - 10 lanes, for use with non retimed 100GBASE-SR10 optical modules
- PMD service interface
 - Logical



Note: 1. Optional
 2. Optional for 100G Cu PHY
 3. Optional for 100GBASE-SR10 PHY
 4. Conditional based on PHY type

100GbE architecture ..2

- 100 GbE architecture diagram with CPPI
 - CPPI³ is physical instantiation of PMD service interface for 100GBASE-SR10 PHYs



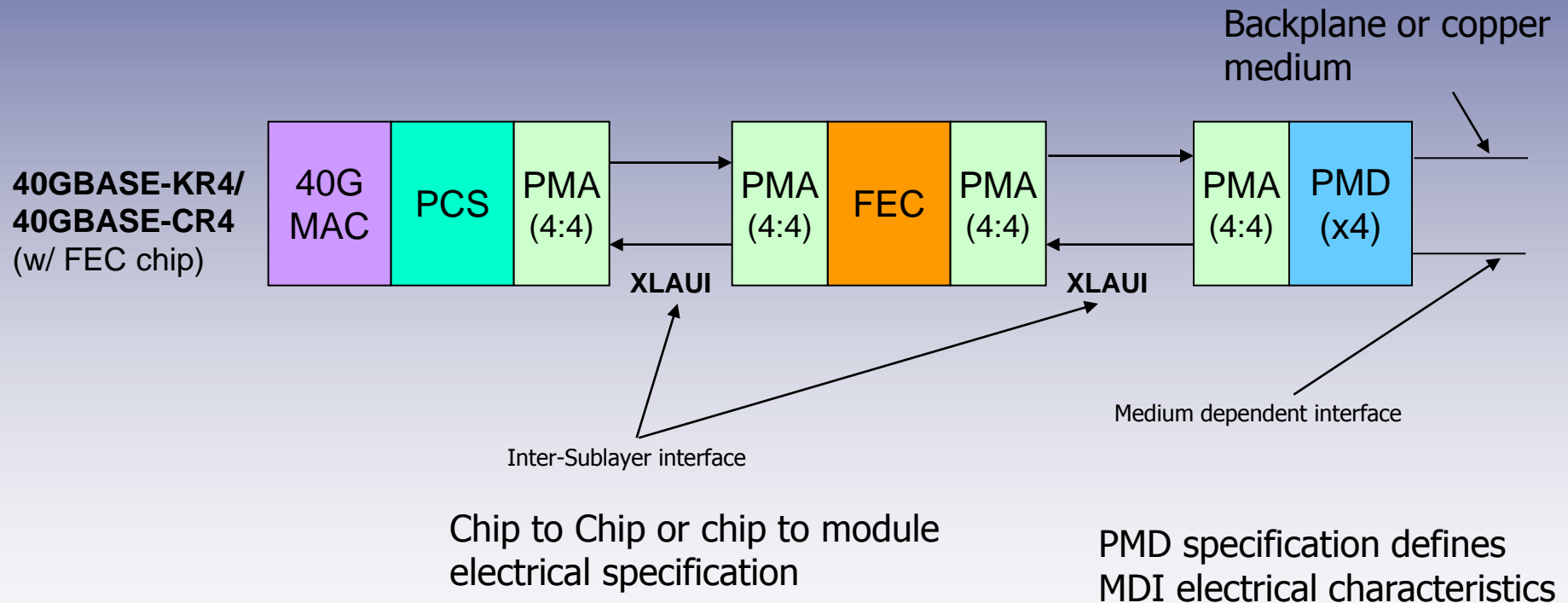
- Note: 1. Optional
 2. Optional for 100G Cu PHY
 3. Optional for 100GBASE-SR10 PHY

Optional compatibility interfaces

- XLGMII and CGMII (40 Gigabit and 100 Gigabit Media Independent Interface)
 - Interface between MAC and PHY layers for intra-chip connectivity
 - Logical definition, data width, control, clock frequency, no electrical or mechanical specifications
 - Independent 64 bit transmit and receive data paths, 8 Tx and Rx control signals
 - Clock is 1/64th of MAC data rate
 - Provides a point of interoperability for multi vendor MAC and PHY implementations
- XLAUI and CAUI (40 Gigabit and 100 Gigabit attachment unit interface)
 - Interface between MAC & PHY layers for chip-to-chip or chip-to-module connectivity
 - Common electrical specification for XLAUI and CAUI
 - 10.3125 GBaud per lane differential signaling
 - 4 lanes in each direction for XLAUI (40 Gb/s) and 10 lanes in each direction for CAUI (100 Gb/s)
- XLPPI and CPPI (40 Gigabit and 100Gigabit parallel physical interface)
 - Chip-to-module interface for use with non retimed optical modules for 40GBASE-SR4, 40GBASE-LR4 and 100GBASE-SR10 PMDs
 - XLPPI is physical instantiation of PMD service interace for 40GBASE-SR4/LR4 PHYs
 - CPPI is physical instantiation of PMD service interace for 100GBASE-SR10 PHYs
 - 10.3125 GBaud per lane differential signaling
 - 4 lanes in each direction for XLPPI and 10 lanes in each direction for CPPI

Electrical interfaces

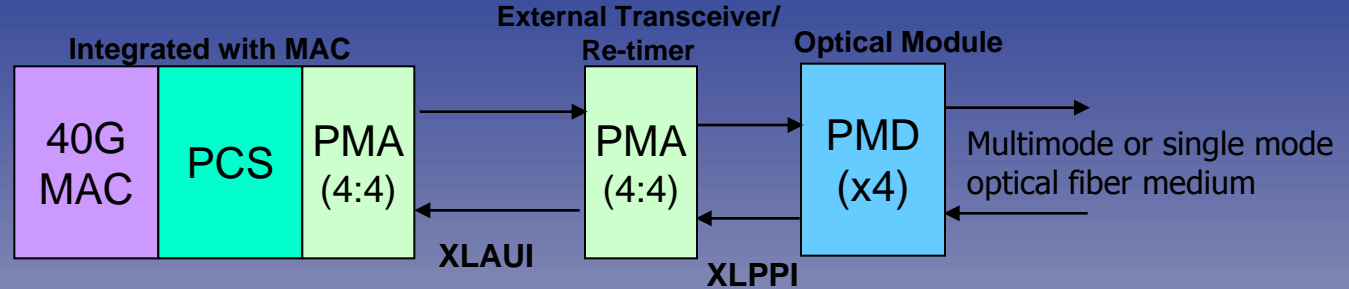
- Illustration of Inter-sublayer interface and Medium dependent interface



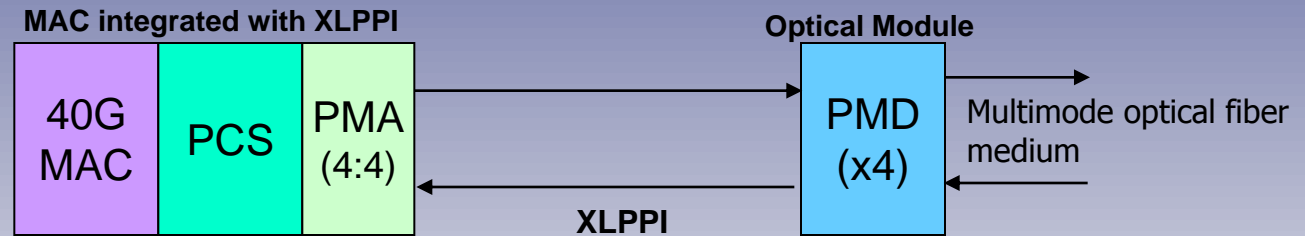
- XLAUI/CAUI and MDI have different electrical characteristics

40 GbE implementation examples

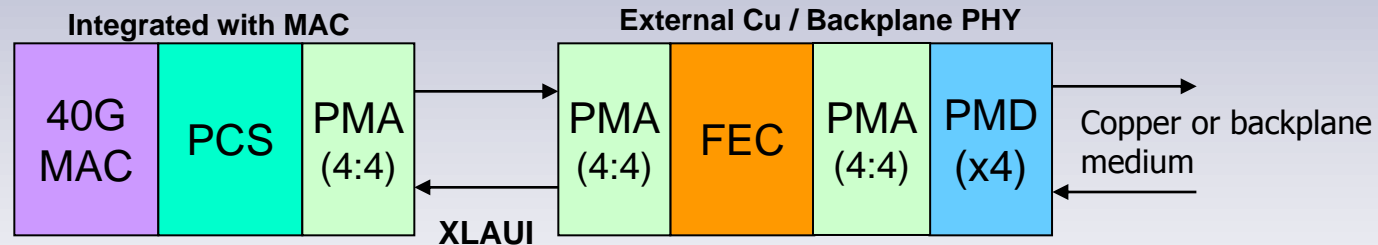
40GBASE-SR4 or LR4
(4 lanes or 4 WDM lanes)



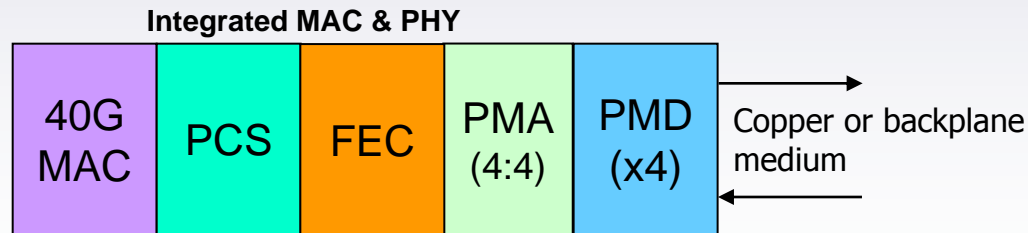
40GBASE-SR4 or LR4
(4 lanes or 4 WDM lanes)



40GBASE-CR4 or KR4
(w/ external PHY chip)

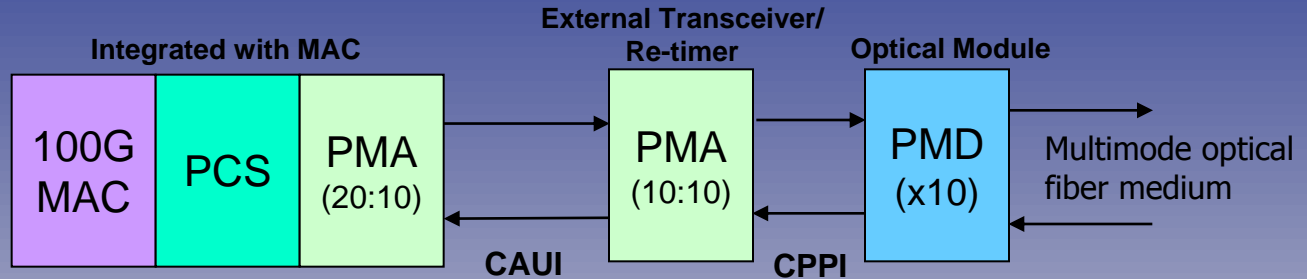


40GBASE-CR4 or KR4
(integrated PHY)

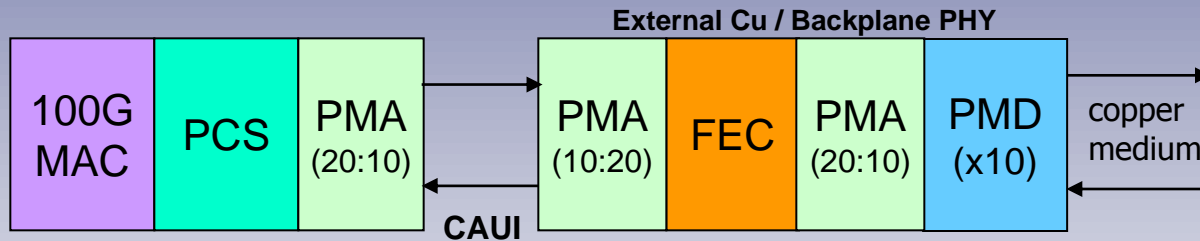


100 GbE implementation examples

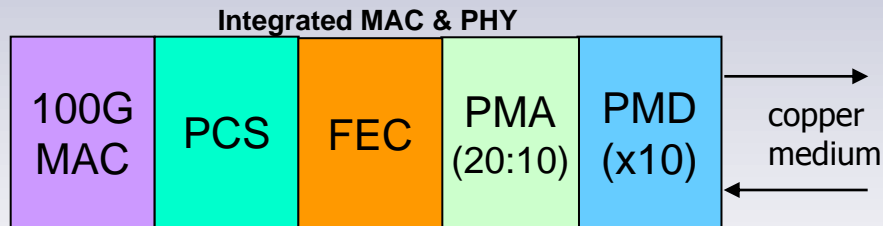
100GBASE-SR10
(10 lanes)



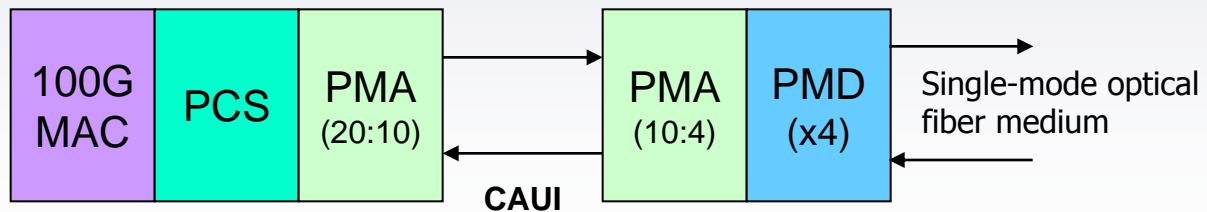
100GBASE-CR10
(w/ external PHY w/FEC)



100GBASE-CR10
(integrated PHY)



100GBASE-LR/ER4
(4 WDM lanes)



Summary

- 40 Gb/s and 100 Gb/s Ethernet use a common architecture
- Addresses the needs of computing, network aggregation and core networking applications
- The architecture is flexible and scalable
- Leverages existing 10 Gb/s technology where possible
- IEEE P802.3ba draft 3.1 is currently in sponsor recirculation ballot
- The target date for P802.3ba standard ratification is Jun 2010
- Future standards related to P802.3ba – An 802.3 study group is working on a PAR for 40 Gb/s serial single mode fiber PMD