IEEE 802.3ba 40 and 100 Gigabit Ethernet Architecture

Ilango Ganga, Intel

IEEE P802.3ba task force Chief Editor
Outline

• IEEE 802.3ba 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
IEEE 802.3ba 40 and 100 GbE overview

- Addresses the needs of computing, network aggregation and core networking applications
- Common architecture for both 40 Gb/s and 100 Gb/s Ethernet
- Uses IEEE 802.3 Ethernet MAC and frame format
- The architecture is flexible and scalable
- Leverages existing 10 Gb/s technology where possible
- Defines physical layer technologies for backplane, copper cable assembly and optical fiber medium
- IEEE 802.3ba standard was ratified in Jun 2010
## IEEE 802.3ba 40G/100G summary

### Physical Layer Specifications

<table>
<thead>
<tr>
<th>Port Type</th>
<th>Description</th>
<th>40 GbE</th>
<th>100 GbE</th>
<th>Solution Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>40GBASE-KR4</td>
<td>Up to at least 1m backplane</td>
<td>✓</td>
<td></td>
<td>4 x 10 Gb/s</td>
</tr>
<tr>
<td>40GBASE-CR4</td>
<td>Up to at least 7m cu (twin-ax)</td>
<td>✓</td>
<td>✓</td>
<td>4 x 10 Gb/s</td>
</tr>
<tr>
<td>100GBASE-CR10</td>
<td>cable</td>
<td>✓</td>
<td>✓</td>
<td>10 x 10 Gb/s</td>
</tr>
<tr>
<td>40GBASE-SR4</td>
<td>Up to at least 100m OM3 MMF</td>
<td>✓</td>
<td>✓</td>
<td>4 x 10 Gb/s</td>
</tr>
<tr>
<td>100GBASE-SR10</td>
<td>(150m OM4 MMF)</td>
<td>✓</td>
<td>✓</td>
<td>10 x 10 Gb/s</td>
</tr>
<tr>
<td>40GBASE-LR4</td>
<td>Up to at least 10km SMF</td>
<td>✓</td>
<td></td>
<td>4 x 10 Gb/s</td>
</tr>
<tr>
<td>100GBASE-LR4</td>
<td>Up to at least 10km SMF</td>
<td></td>
<td>✓</td>
<td>4 x 25 Gb/s</td>
</tr>
<tr>
<td>100GBASE-ER4</td>
<td>Up to at least 40km SMF</td>
<td></td>
<td>✓</td>
<td>4 x 25 Gb/s</td>
</tr>
</tbody>
</table>
40 and 100 GbE layer model

IEEE ComSoc SCV
Santa Clara, CA, Oct 2010

Note: 1. KR4, CR4 & CR10 may use optional FEC
2. XLGMII and CGMII are optional interfaces
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

- **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

- XLGMII\(^1\) (intra-chip)
  - Logical, data/control, clock, no electrical specification

- 40GBASE-R PCS
  - 64B/66B encoding
  - Lane distribution and alignment

- XLAUI\(^1\) (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

- XLPP\(^3\) (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

- **CGMII**<sup>1</sup> (intra-chip)
  - Logical, data/control, clock, no electrical specification
- **100GBASE-R PCS**
  - 64B/66B encoding
  - Lane distribution and alignment
- **CAUI**<sup>1</sup> (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
- **CPPI**<sup>3</sup> (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

- 100 GbE architecture diagram with CPPI
  - CPPI\(^3\) 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 1/64\(^{th}\) 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)

- 40G MAC
- PCS
- PMA (4:4)
- PMA (4:4)
- PMD (x4)
- Optical Module

**MAC integrated with XLPPI**

- 40G MAC
- PCS
- PMA (4:4)
- PMA (4:4)
- PMD (x4)
- Optical Module

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

- 40G MAC
- PCS
- PMA (4:4)
- PMA (4:4)
- FEC
- PMD (x4)
- Optical Module

**Integrated MAC & PHY**

- 40G MAC
- PCS
- FEC
- PMA (4:4)
- PMD (x4)

- Copper or backplane medium
- 4 lanes
100 GbE implementation examples

100GBASE-SR10
(10 lanes)

100GBASE-LR/ER4
(4 WDM lanes)

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

100GBASE-CR10
(integrated PHY)
(10 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 to adapt to current & future needs
• Leverages existing 10 Gb/s technology where possible
• IEEE 802.3ba standard was ratified in Jun 2010
• Future standards related to IEEE Std 802.3ba
  – IEEE P802.3bg task force is developing a std for 40 Gb/s serial single mode fiber PMD
  – 100 Gb/s backplane and copper cable assemblies Call For Interest scheduled for Nov’10