



Considerations for 40G Backplane Ethernet PHY

Ilango Ganga, Intel

Sep 12, 2007



Outline

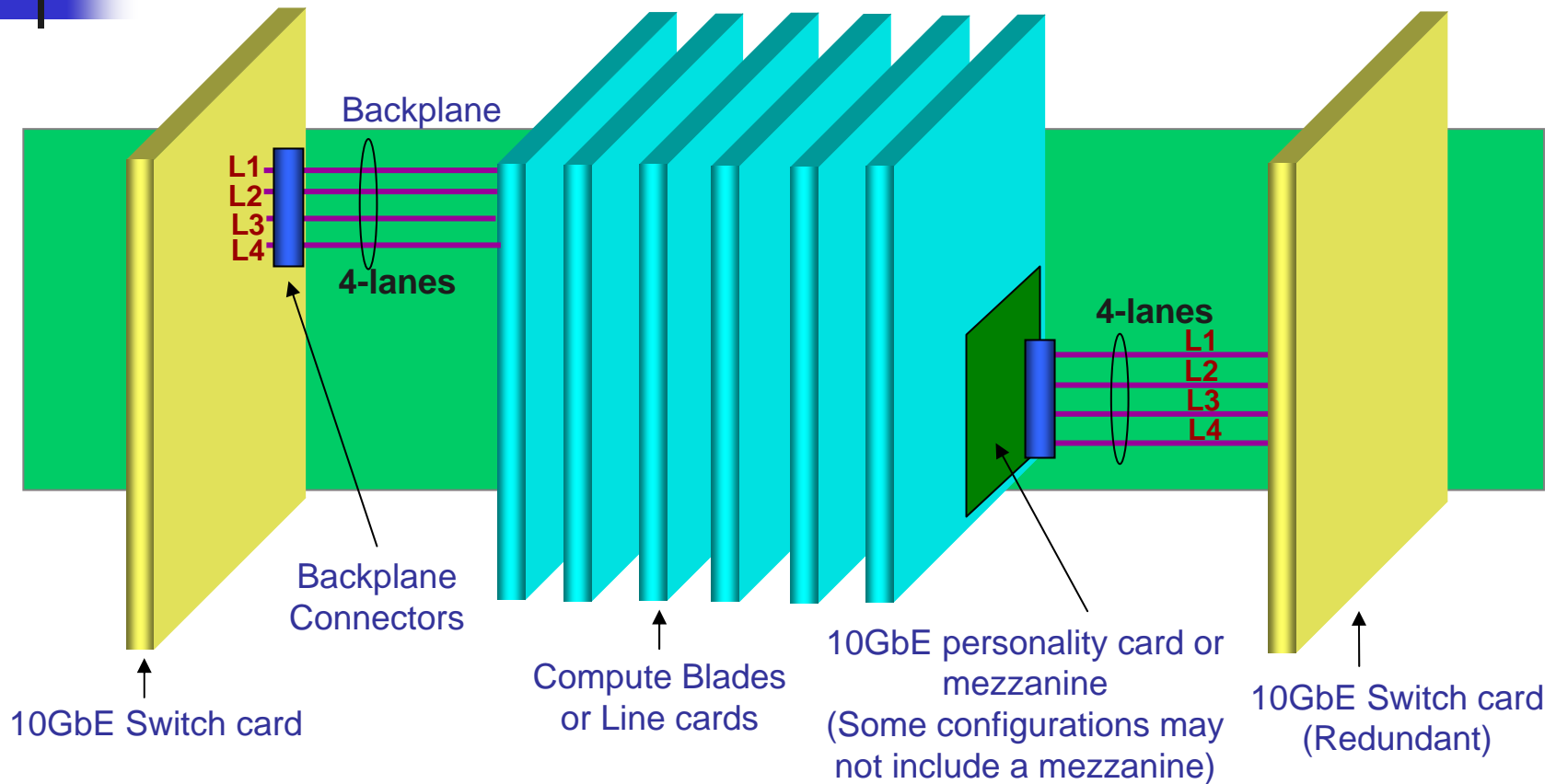
- IEEE 802.3ap Backplane Ethernet overview and typical backplane system
- Architectural positioning of Backplane Ethernet
- Considerations for 40GbE backplane PHY
- Proposed 40GbE layer stack
- Considerations for PCS layer, Startup protocol, FEC and Auto-Negotiation
- Summary & Next steps



Backplane Ethernet overview

- IEEE Std 802.3ap-2007 Backplane Ethernet defines 3 PHY types
 - 1000BASE-KX : 1-lane 1 Gb/s PHY (Clause 70)
 - 10GBASE-KX4: 4-lane 10Gb/s PHY (Clause 71)
 - 10GBASE-KR : 1-lane 10Gb/s PHY (Clause 72)
- Forward Error Correction (FEC) for 10GBASE-KR (Clause 74)
- Auto-negotiation (Clause 73)
 - Auto-Neg between 3 PHY types (AN is mandatory to implement)
 - Parallel detection for legacy PHY support
 - Automatic speed detection of legacy 1G/10G backplane SERDES devices
 - Negotiate FEC capability
- Clause 45 MDIO interface for management
- Channel
 - Controlled impedance (100 Ohm) traces on a PCB with 2 connectors and total length up to at least 1m.
 - Channel model is informative as defined in Annex 69B
- Support a BER of 10^{-12} or better, FEC to increase link budget

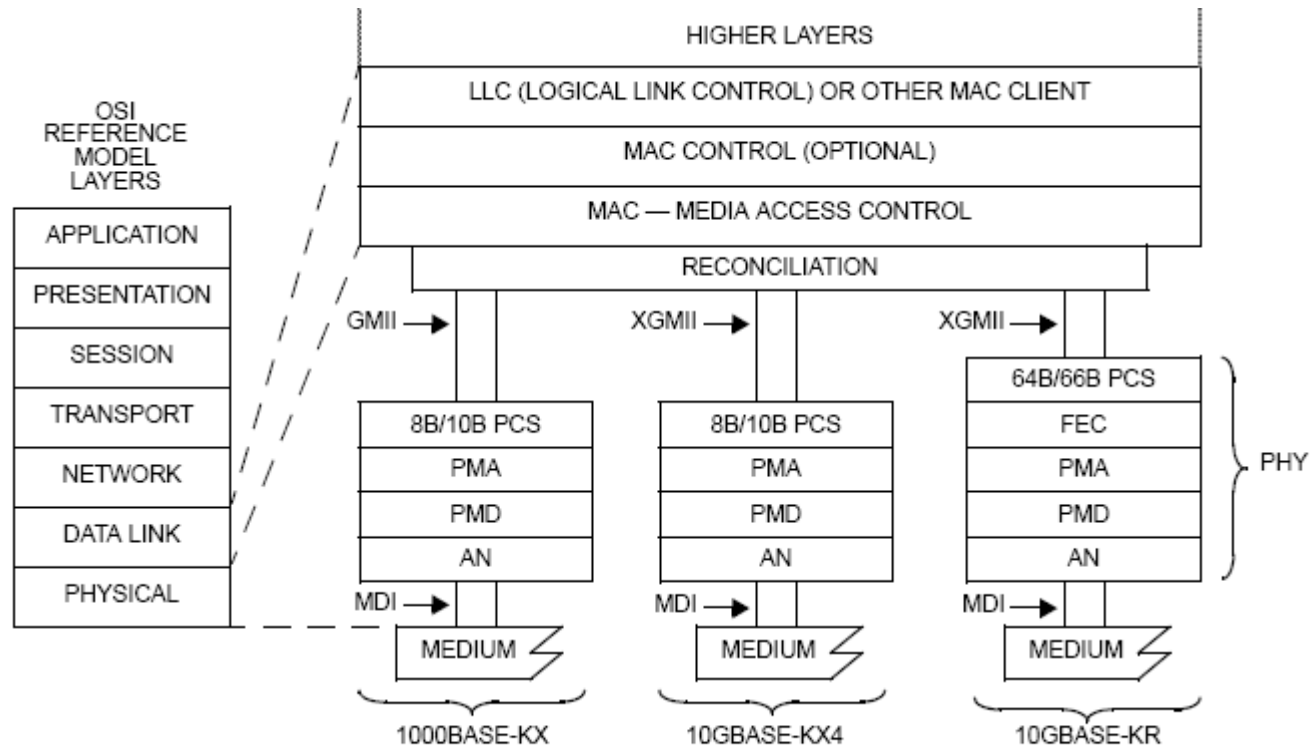
Typical backplane system illustration



Note: The switch cards are shown at the chassis edge for simplicity.

In real systems there could be multiple fabrics located at the center, edge, or rear of the chassis

Architectural Positioning of BPE



AN = AUTO-NEGOTIATION
 GMII = GIGABIT MEDIA INDEPENDENT INTERFACE
 MDI = MEDIUM DEPENDENT INTERFACE
 PCS = PHYSICAL CODING SUBLAYER
 FEC = FORWARD ERROR CORRECTION

PHY = PHYSICAL LAYER DEVICE
 PMA = PHYSICAL MEDIUM ATTACHMENT
 PMD = PHYSICAL MEDIUM DEPENDENT
 XGMII = 10 GIGABIT MEDIA INDEPENDENT INTERFACE

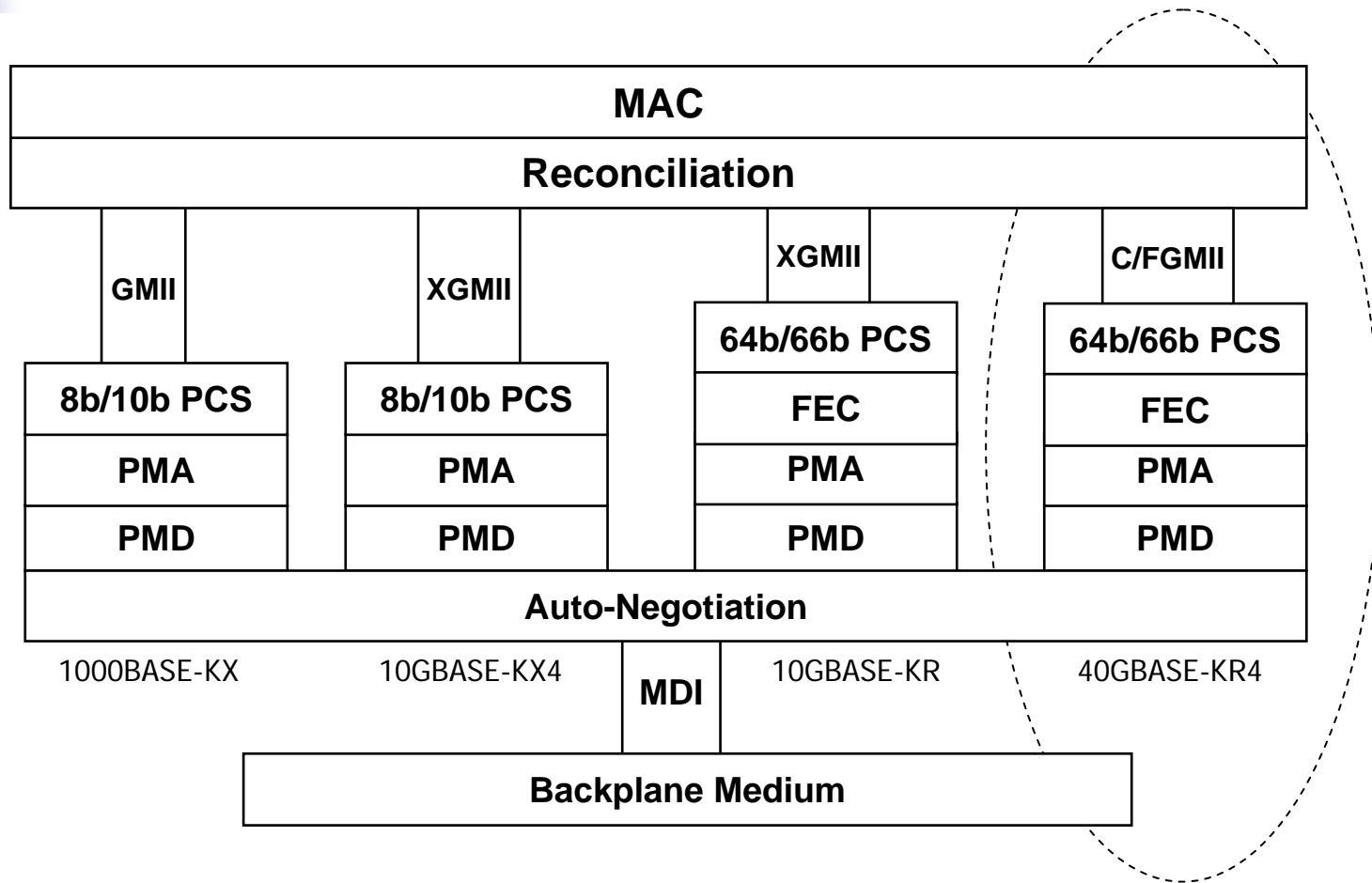
Figure 69-1—Architectural positioning of Backplane Ethernet



Considerations for 40G BPE PHY

- To be architecturally consistent with the Backplane Ethernet layer stack illustrated in Clause 69
- To interface to a 4-lane backplane medium with interconnect characteristics recommended in IEEE Std 802.3ap (Annex 69B)
 - Most generation 2 blade systems are built with 4-lanes (10Gbaud KR ready)
- Leverage 10GBASE-KR technology/specifications (Clause 72) to define 40GBASE-KR PHY:
 - 64b/66b PCS
 - Startup protocol with suitable extensions to run over 4 lanes
 - Signaling speed 10.3125Gbd (per lane)
 - Electrical characteristics
 - Test methodology and procedures
- Optional FEC sublayer
 - PCS to interface to optional FEC sublayer consistent with Clause 74 specification
- Compatible with Backplane Ethernet Auto-Neg (Clause 73)
 - Enhancement to indicate 40GbE ability

Proposed 40GbE BP PHY layer stack





PCS layer for 40GbE backplane PHY

- Use 64b/66b PCS to leverage 10GBASE-KR & FEC
 - Suitable stripping/mapping to run over 4 lanes
 - Signaling rate of 10.3125 Gbaud per lane
- PCS and PMA to interface to an optional FEC sublayer consistent with 10GBASE-KR PHY
- 4 lane PCS/PMA interface
 - Direct mapping of PMD lanes and PCS/PMA interface
- Choose lane encoding schemes to be compatible with Clause 74 FEC functions
 - Encoding, striping and scrambling mechanisms not to interfere with FEC functions



Startup protocol

- 10GBASE-KR startup protocol
 - Facilitates the receiver to tune the transmit equalizer for optimum performance over the BP interconnect
 - Uses a 548 octet training frame with DME signaling (1/4 the rate of KR signaling)
 - Training starts after Auto-Neg completion; indicates “Link Ready” after successful training completion
- Extend state machine to include training of 4 lanes
 - Should the 4 lanes be trained independently?
 - Indicate `signal_detect<=TRUE` when all 4 lanes are trained
 - Or should all the 4 lanes be trained together?
 - Single state machine to control the startup protocol
 - Training control channel on Lane 1 only?
 - Control channel to be extended to include Coefficient exchange for all lanes; same or different?

BP Auto-Neg: 40GbE ability

- IEEE Std 802.3ap defines Auto-Negotiation for backplane Ethernet PHYs
 - AN uses DME signaling with 48-bit base pages to exchange link partner abilities
 - AN is mandatory for 10GBASE-KR backplane PHY, negotiates FEC ability
 - 40GbE BP PHY: Ability to negotiate with other 802.3ap backplane PHYs
- Proposal to add a bit to indicate 40GbE ability
 - Consider to use A3 bit in TA field (currently reserved for future technology)
 - No changes to backplane AN protocol
 - No changes to management register format

Table 73-4—Technology Ability field encoding

Bit	Technology
A0	1000BASE-KX
A1	10GBASE-KX4
A2	10GBASE-KR
A3 through A24	Reserved for future technology
A3	40GBASE-KX4
A4 through A24	Reserved for future technology



Summary & Next Steps

Summary

- Define 40GbE backplane PHY to be architecturally consistent with IEEE Std 802.3ap layer stack
- Leverage 10GBASE-KR technology/specs to define 40GBASE-KR PHY
- Interface to 4 lane backplane medium to take advantage of 802.3ap KR ready blade systems in deployment

Next Steps

- PCS proposals and interface definitions to accommodate backplane Ethernet architecture (including FEC and AN)
- Proposals to enhance startup protocol to run over 4 lanes
- Simulations with 4 lane 10GBASE-KR signaling (40GBASE-KR)