FEC and Auto-Neg proposal for 40/100G copper cable assembly

Ilango Ganga
Intel Corporation

May 13, 2008

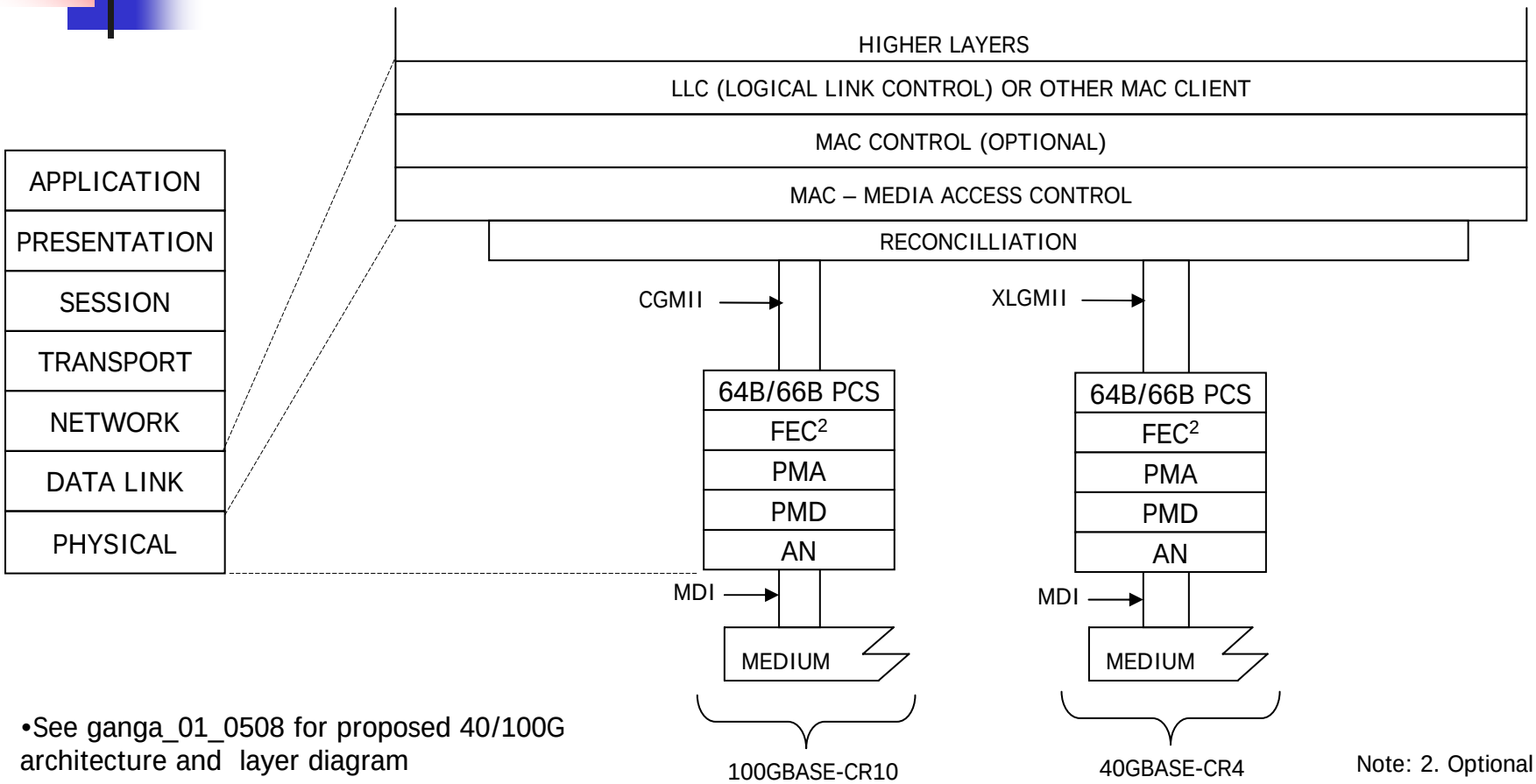
Supporters

- Andre Szczepanek, Texas Instruments
- Joel Goergen, Force 10

Key messages

- Proposal to adopt FEC and Auto-Negotiation for 40/100G copper cable assembly
  - Leverage Clause 74 FEC as an optional sublayer for 40GBASE-CR4 and 100GBASE-CR10
  - Clause 73 Auto-Neg for negotiating FEC, and for providing backward compatibility with 10G

Proposed 40/100G Copper cable assembly layer model



- See ganga_01_0508 for proposed 40/100G architecture and layer diagram
- See diminico_01_0508 for 40/100G copper cable assembly proposal

FEC for 40/100G copper cable assembly

- Leverage 10GBASE-R FEC to specify 40GBASE-CR4 and 100GBASE-CR10 PHYs
  - Clause 74 FEC provides 2 to 2.5dB coding gain
  - Provides additional link budget, better margin
  - Increases BER performance on copper cable assemblies
  - Define as optional sublayer for CR4 and CR10
- Clause 74 FEC overview
  - (2112,2080) Binary burst error correction code
  - Can correct burst errors up to 11 bits
  - Light weight code, low overhead/latency/power
  - No increase in signaling rate or decrease in data rate
  - See http://www.ieee802.org/802_tutorials/july06/10GBASE-KR_FEC_Tutorial_1407.pdf for further details
- Negotiate FEC capability through Auto-Negotiation

Forward Error Correction

- Reuse 10GBASE-R FEC specification (Clause 74)
  - Enumerate the FEC encode and decode functions for 4 lane and 10 lane operation
    - Each lane is encoded and decoded independently
    - The coding is performed on a virtual lane basis
      - 4 in case of 40G
      - 20 incase of 100G
      - See gustlin_01_0508, 40/100G PCS proposal, for further detail on VLs
  - Change to accommodate FEC sync for 4 and 10 (20 VLs) lanes
    - Same state diagram for FEC block lock
    - Report Global Sync achieved for the link only if all lanes are locked
  - Commonality with 40G backplane solution
    - Reuse the management register format

FEC MDIO variable mapping

Table 74–2—MDIO/FEC variable mapping

MDIO variable	PMA/PMD register name	Register/bit number	FEC variable
10GBASE-R FEC ability	10GBASE-R FEC ability register	1.170.0	FEC_ability
10GBASE-R FEC Error Indication ability	10GBASE-R FEC ability register	1.170.1	FEC_Error_Indication_ability
FEC Enable	10GBASE-R FEC control register	1.171.0	FEC_Enable
FEC Enable Error Indication	10GBASE-R FEC control register	1.171.1	FEC_Enable_Error_to_PCS
FEC corrected blocks	10GBASE-R FEC corrected blocks counter register	1.172, 1.173	FEC_corrected_blocks_counter
FEC uncorrected blocks	10GBASE-R FEC uncorrected blocks counter register	1.174, 1.175	FEC_uncorrected_blocks_counter

- Enumerate the following counters for 4 and 10 (20 VLs) lanes
  - FEC_corrected_blocks_counter
  - FEC_uncorrected blocks_counter
  - Possibly use indexed addressing to conserve MDIO address space

Auto-Neg for copper cable assembly

- Adopt Clause 73 Auto-Neg for 40GBASE-CR4 and 100GBASE-CR10
  - Use Auto-Neg to Negotiate FEC capability
  - Auto-Neg allows backward compatibility with legacy 10G CX4 PHYs
- 40GBASE-CR4 PHY when used with a CX4 style connector allows backward compatibility with legacy 10G
  - Allows implementation of dual speed 10/40G systems
- Clause 73 provides parallel detection function for compatibility with legacy PHYs that do not support Auto-Negotiation
  - New 40G PHY can use Parallel Detection for auto-detection of legacy CX4 devices
  - No impact to 10GBASE-CX4 devices
- Auto-Neg could also allow forward compatibility with future 100G devices
  - For example when 40G (4x10G) and 100G (4x25G) share the same connector solution
- Other uses of Auto-Neg
  - Can act as digital signal detect, negotiate Pause ability
  - Allows link partners to startup with a known state

Clause 73 Auto-Neg overview

- Clause 73 defines Auto-Negotiation for backplane Ethernet
  - Provides a means to exchange information between link partners to automatically configure both devices to take maximum advantage of their abilities
  - AN is used to negotiate PHY types, negotiates FEC ability, Pause ability and also acts as a digital signal detect
- AN uses DME signaling with 48-bit base pages to exchange link partner abilities
  - Slower speed (312.5 MBaud) for reliable data transfer
  - 600 to 1200 mVp-p
  - Lane 0 of the MDI is used for Auto-Negotiation, of single or multi-lane PHYs
- AN base page contains
  - Technology Ability field to indicate PHY types
  - FEC capability
  - Pause ability
  - Remote fault

Auto-Neg changes for CR4 & CR10

- Proposed changes for 40GBASE-KR4 and 100GBASE-CR10
  - Add Technology Ability bits from the reserved space to indicate
    - 40GBASE-CR4 ability
    - 100GBASE-CR10 ability
  - Reuse AN management registers
  - No change to negotiate FEC ability
    - FEC when selected to be enabled on all lanes
    - FEC is enabled when both sides advertise FEC ability and at least one side requests to enable FEC
  - No change to Pause ability and Remote Fault bits
  - Parallel detection function to detect legacy 10GBASE-CX4 PHYs

Summary

- Adopt FEC for 40/100G copper cable assembly to increase link budget and BER performance
  - Leverage Clause 74 FEC as an optional sublayer for 40GBASE-CR4 and 100GBASE-CR10 PHYs
- Adopt Auto-Neg for 40/100G copper cable assembly to negotiate FEC ability, and to allow backward compatibility with 10G CX4 systems
  - Use Clause 73 Auto-Neg protocol and signaling