XL/CGMII and RS Proposal

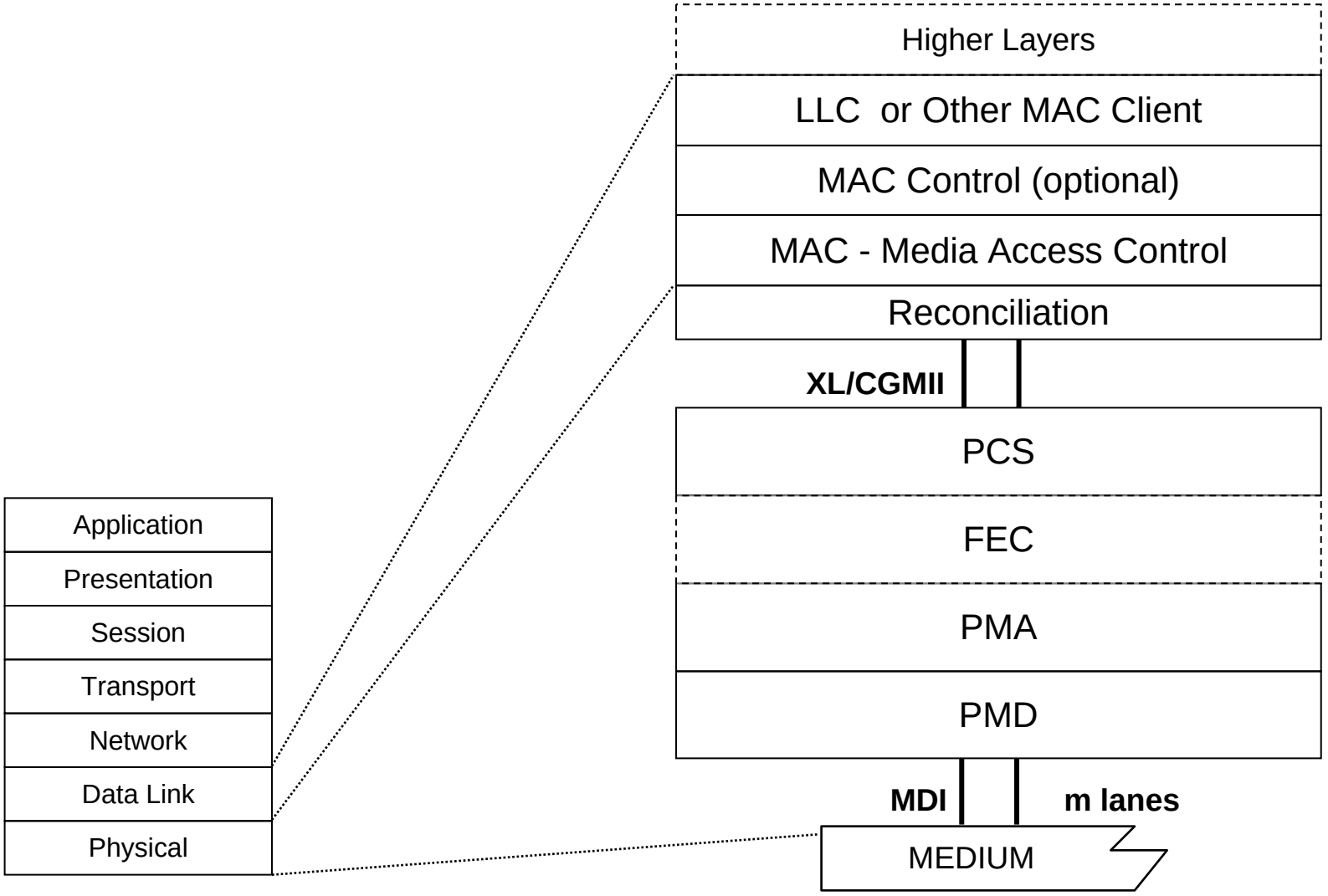
Mark Gustlin - Cisco

IEEE 802.3ba May 2008 Munich

Contributors and Supporters

- Steve Trowbridge - Alcatel-Lucent
- Brad Booth – AMCC
- Dimitrios Giannakopoulos - AMCC
- Piers Dawe – Avago
- Howard Frazier - Broadcom
- Arthur Marris - Cadence Design Systems
- Gary Nicholl – Cisco
- Med Belhadj – Cortina Systems
- Chris Cole - Finisar
- Subi Krishnamurthy - Force10 Networks
- Aris Wong – Foundry
- Shashi Patel - Foundry
- Ryan Latchman - Gennum
- Shinji Nishimura - Hitachi
- Hidehiro Toyoda – Hitachi
- John Jaeger - Infinera
- Andy Moorwood - Infinera
- Thananya Baldwin – Ixia
- Jerry Pepper – Ixia
- Faisal Dada - JDSU
- Norbert Folkens – JDSU
- Jack Jewell – JDSU
- Jeffery J. Maki - Juniper Networks
- David Ofelt – Juniper Networks
- Adam Healey - LSI
- Avigdor Segal – Marvell
- Martin White - Marvell
- Pete Anslow – Nortel
- Song Shang - SMI
- Shimon Muller – Sun
- Farhad Shafai – Sarance
- Andre Szczepanek – TI
- Frank Chang - Vitesse

40GE/100GE Architecture



XL/CGMII Interface

- Why define it?

  Electrically it won't see the light of day

  Some want it for RTL to RTL connections within devices

- The interface is naturally scaled based on speed targets of an implementation

  FPGAs run slower, ASICs faster, next generation ASICs even faster…

- Define it as a logical interface only
  - Service primitives (function calls, pseudo code) +
  - Signals, code-points, syntax, sequences, true/false

XL/CGMII Interface

- Leverage XGMII, but make it 8 lanes instead of 4
- Preserve use of encoded rather than discrete delimiters

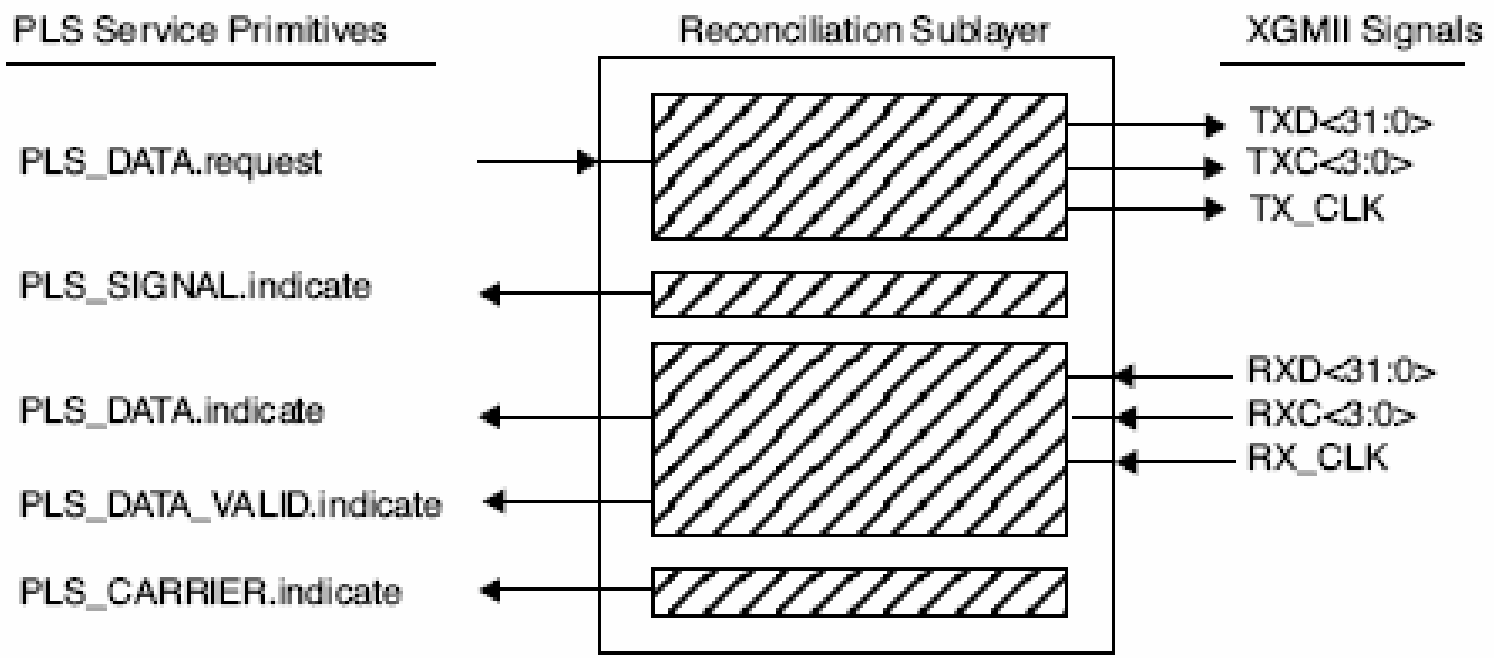
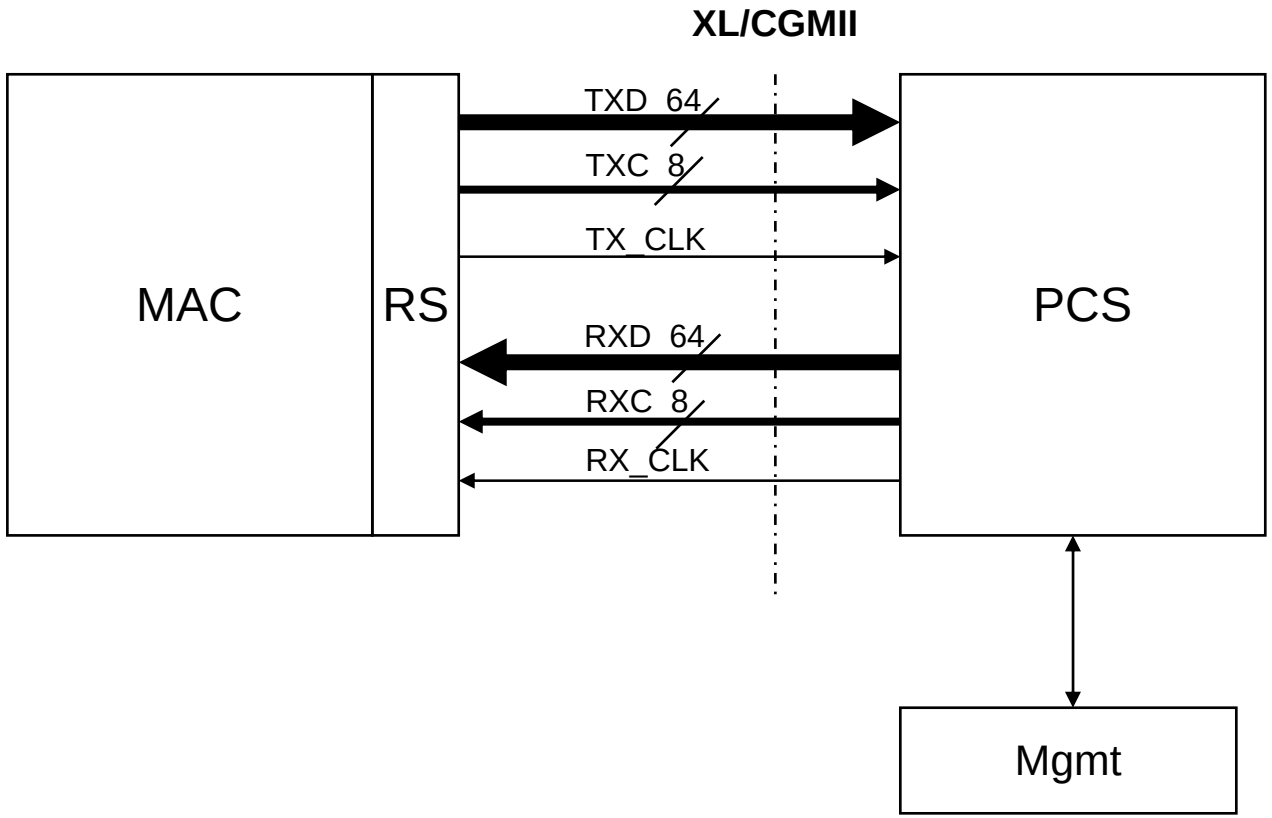


Figure 46–2—Reconciliation Sublayer (RS) inputs and outputs

From 802.3ae

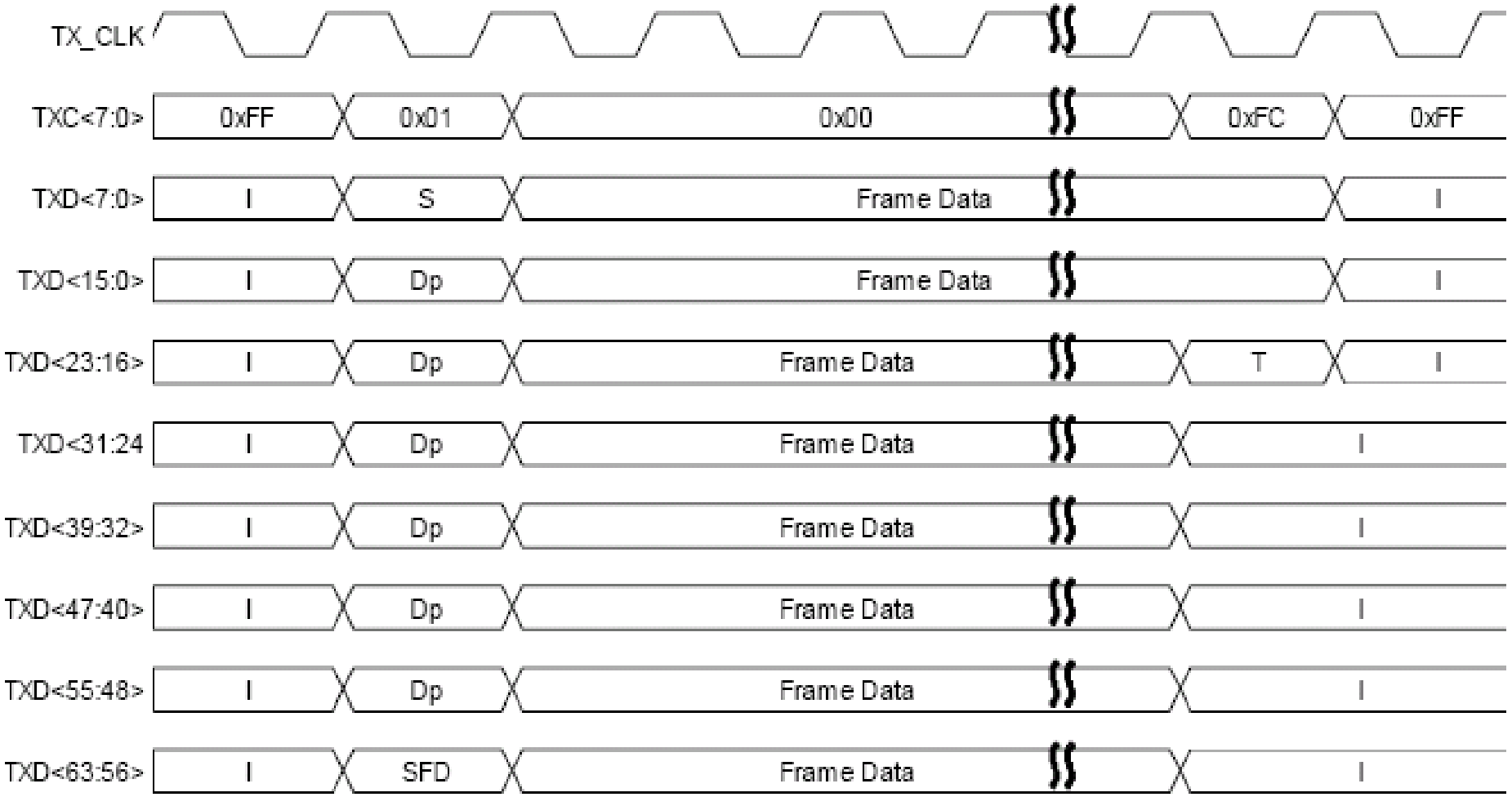
XL/CGMII Interface

- Leverage XGMII, but make it 8 lanes instead of 4
- CLK = 625MHz for 40GE, 1.5625GHz for 100GE
- Clock may be scaled down in frequency by increasing the width from 8 lanes to 16, 24, 32 etc.



XL/CGMII Interface

RX Diagram is Identical



I: Idle control character, S: Start control character, Dp: preamble Data octet, T: Terminate control character, SFD: Start of Frame Delimiter

XL/CGMII Interface

Same encoding as XGMII (for both tx and rx):

Table 46–3—Permissible encodings of TXC and TXD

TXC	TXD	Description	PLS_DATA.request parameter
0	00 through FF	Normal data transmission	ZERO, ONE (eight bits)
1	00 through 06	Reserved	—
1	07	Idle	No applicable parameter (Normal inter-frame)
1	08 through 9B	Reserved	—
1	9C	Sequence (only valid in lane 0)	No applicable parameter (Inter-frame status signal)
1	9D through FA	Reserved	—
1	FB	Start (only valid in lane 0)	No applicable parameter, replaces first eight ZERO, ONE of a frame (preamble octet)
1	FC	Reserved	—
1	FD	Terminate	DATA_COMPLETE
1	FE	Transmit error propagation	No applicable parameter
1	FF	Reserved	—

NOTE—Values in TXD column are in hexadecimal, most significant bit to least significant bit (i.e., <7:0>).

8B vs. 4B alignment

- We could keep the legacy 4B alignment even with the new 8B wide bus
- Or we could go to 8B alignment
  - Only start packets in lane 0
  - Significant gate savings for 100GE, especially in FPGAs
  - Deficit counter goes from 0-7 for 8B alignment (vs. 0-3 for 4)
  - Doubles the buffering required for clock compensation when compared to 4B alignment
- Recommended to go with 8B alignment
- If interface is to be scaled down in frequency (and up in width), packet starts are still on 8B boundaries (lane 0, 8, 16 etc).

IPG Rules for 8B Alignment

- A MAC implementation may be designed to always insert additional idle characters to align the start of preamble on an eight byte boundary.

  Note that this will reduce the effective data rate for certain packet sizes separated with minimum inter-frame spacing.

- Alternatively, the RS may maintain the effective data rate by sometimes inserting and sometimes deleting idle characters to align the Start control character.

  When using this method the RS must maintain a Deficit Idle Count that represents the cumulative count of idle characters deleted or inserted. The counter is incremented for each idle character deleted, decremented for each idle character inserted, and the decision of whether to insert or delete idle characters is constrained by bounding the counter to a minimum value of zero and maximum value of seven.

Summary

- Simple logical interface based on XGMII
- Extended to 8 Bytes
- Naturally scales up and down in width and frequency
- Packet Starts on 8 Byte boundaries