SCALING THE MAC AND XGMII FOR 2.5/5GBASE-T

Howard Frazier
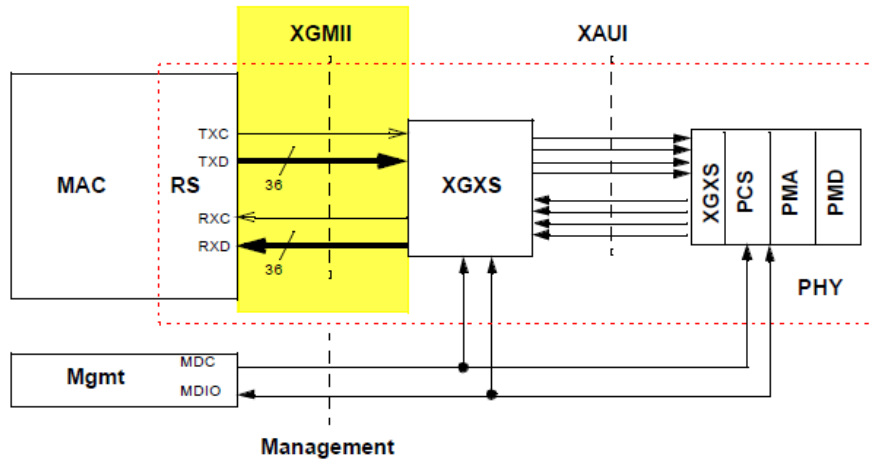
IEEE 802.3bz Task Force

Pittsburg, PA

OUTLINE

- **A blast from the past**

- **Motivation**

- **Layering**

- **XGMII features and benefits**

- **Scaling the XGMII**

- **Scaling the full duplex MAC**

- **Alternatives**

- **Conclusions**

Interface Locations

XGMII   XAUI

MAC   RS   XGXS   XGXS PCS PMA PMD

TXC
TXD
36
RXC
RXD
36

PHY

Mgmt   MDC
MDIO

Management

IEEE 802.3ae
10 Gigabit Ethernet

http://www.ieee802.org/3/ae/public/jul00/frazier_1_0700.pdf

- **PHYs defined in IEEE Std 802.3 are described in terms of two primary interfaces**
  - Medium dependent interface (MDI)
  - Media independent interface (MII)
- **While the MDI is visible to the user, and usually tightly specified, the MII is often used simply as a convenient way to partition the physical layer specifications from the datalink layer specifications**
- **The MII provides a reference model for conveying data and management information**
- **The MII (which ever version) might not be implemented, but the PHY is described as though it were implemented**
- **A logical specification for an MII is an essential part of any IEEE 802.3 PHY**
- **Implementations may use an industry standard derivative of the MII (e.g. SGMII, XFI)**
  - The IEEE 802.3 Ethernet Working Group has resisted writing a standard for such interfaces

LAYERING

XGMII FEATURES AND BENEFITS
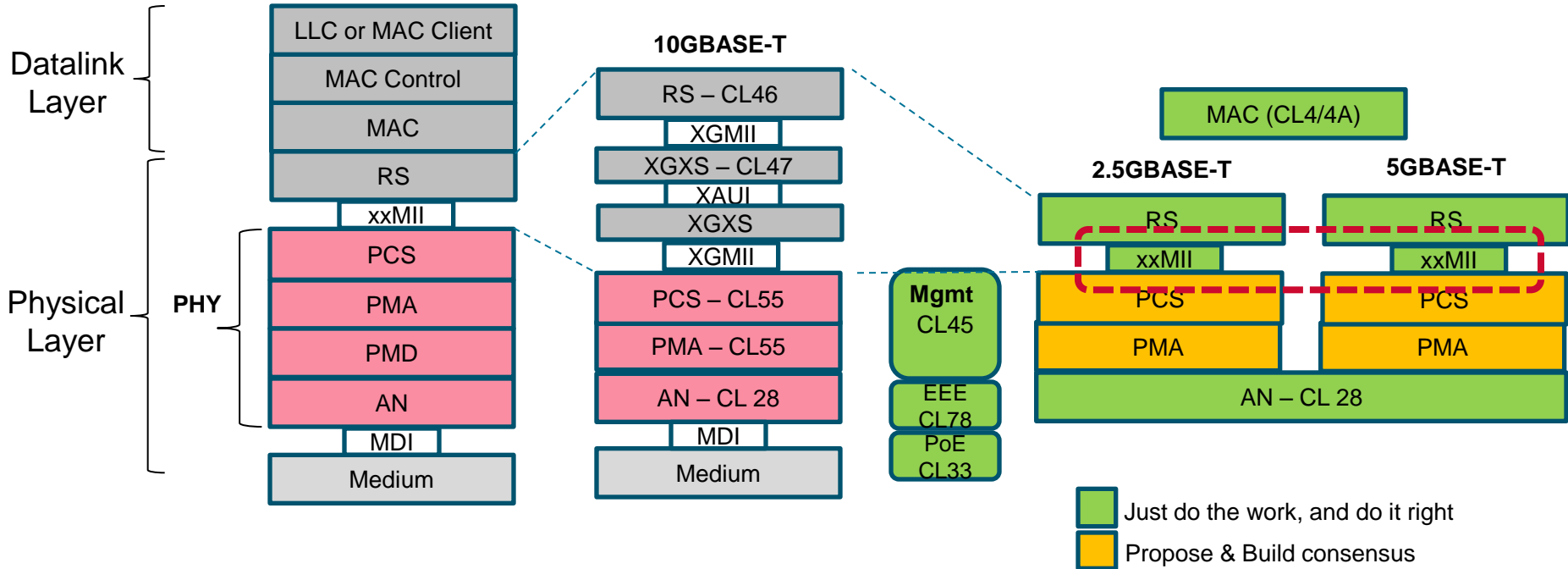
- **XGMII was defined in Clause 46 as part of IEEE Std 802.3ae-2002**
  - Baseline proposal adopted in July, 2000
  - Stable, mature, well understood

- **XGMII represented a significant evolution from the MII and GMII**
  - DDR signaling
  - 32 bit data paths
  - Embedded delimiters
  - Clause 45 MDIO/MDC electricals and register addressing

- **XGMII uses source synchronous timing**
  - Amenable to frequency scaling

- **XGMII supports EEE and 802.1AS/802.3bf time synchronization**

- **The XGMII specification is largely speed independent**

- **The changes needed to accommodate 2.5G and 5G data rates are very modest**

- **Examples follow**

46.3.1.1 TX_CLK (~~10 Gb/s~~ transmit clock)

TX_CLK is a continuous clock ~~used for operation at 10 Gb/s~~. TX_CLK provides the timing reference for the transfer of the TXC<3:0> and TXD<31:0> signals from the RS to the PHY. The values of TXC<3:0> and TXD<31:0> shall be sampled by the PHY on both the rising edge and falling edge of TX_CLK. TX_CLK is sourced by the RS.

The TX_CLK frequency shall be ~~156.25 MHz ±0.01%,~~ one sixty-fourth of the MAC transmit data rate **±0.01%** .

NOTE—For EEE capability, TX_CLK may be halted according to 46.3.1.5.

▪ **Make a similar change in 46.3.2.1 RX_CLK**

46.1.3 Rate(s) of operation

The XGMII supports ~~only the 10 Gb/s~~ **MAC data rates of 10 Gb/s, 5 Gb/s and 2.5Gb/s as defined within this clause. A compliant device may implement any subset of these rates.**

Operation at 10 Mb/s and 100 Mb/s is supported by the MII defined in Clause 22 and operation at 1000 Mb/s by the GMII defined in Clause 35.

~~PHYs that provide an XGMII shall support the 10 Gb/s MAC data rate. 10GBASE-X and 10GBASE-R PHYs operate at a 10 Gb/s data rate.~~

- **Suggest leaving the setup and hold time parameters in Figure 46-16 as is**
- **The t_{pwmin} parameter might be restated as 40% of the clock period**
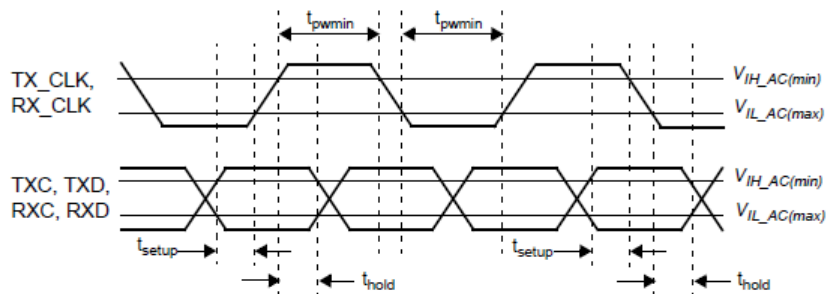
Symbol	Driver	Receiver	Units
t_{setup}	960	480	ps
t_{hold}	960	480	ps
t_{pwmin}	2.5	—	ns

Figure 46–16—TX_CLK and RX_CLK timing parameters

- **The Full duplex MAC as specified in Clause 4 and Annex 4A is inherently scalable**
  - Thank you, IEEE Std 802.3ae for making it so
  - Thank you, IEEE Std 802.3ah for the simplified full duplex MAC in Annex 4A

- **Propose using only the Annex 4A simplified full duplex MAC**

- **The parameters for the full duplex MAC are specified in 4A.4.2 in a rate independent fashion**

4A.4.2 MAC parameters

The parameter values shown in Table 4A–2 shall be used.

Table 4A–2—Full duplex MAC parameter values

Parameters	Values
interPacketGap	96 bits
maxBasicFrameSize	1518 octets
maxEnvelopeFrameSize	2000 octets
minFrameSize	512 bits (64 octets)

- **Except for one pesky detail:**
    - NOTE 3—For 10 Gb/s operation, the spacing between two packets, from the last bit of the FCS field of the first packet to the first bit of the Preamble of the second packet, can have a minimum value of 40 BT (bit times), as measured at the XGMII receive signals at the DTE. This interpacket gap shrinkage may be caused by variable network delays and clock tolerances.

- **Suggest changing this to read:**
    - NOTE 3—For <u>2.5 Gb/s, 5 Gb/s and</u> 10 Gb/s operation, the spacing between two packets, from the last bit of the FCS field of the first packet to the first bit of the Preamble of the second packet, can have a minimum value of 40 BT (bit times), as measured at the XGMII receive signals at the DTE. This interpacket gap shrinkage may be caused by variable network delays and clock tolerances.

- **Scale the GMII from 1 G to 2.5G and 5G**
  - Anybody worried about specifying a 625 MHz single-ended interface?
  - Implies re-using Clause 22 MDIO/MDC electricals and register addressing
- **Adopt and scale a de facto industry standard interface such as SGMII, QSGMII, XFI**
  - Perhaps this is best left outside of IEEE 802.3
- **Create a new interface**
  - Volunteers?
- **No logical/electrical interface specification**
  - Use service primitives instead
- **Create a new clause similar to Clause 106**
  - This might have made some sense for 25G Ethernet, but is not necessary for 2.5G/5G because the timing parameters for the XGMII work just fine at the lower rates

CONCLUSIONS

- **The XGMII specification is well understood and stable**

- **The industry knows how to create serial variants**

- **The XGMII specification can be scaled for 2.5G and 5G operation with modest changes to Clause 46**

- **The Clause 45 MDIO/MDC register addressing scheme is much preferred over the Clause 22 scheme**

- **The simplified full duplex MAC is already specified in a rate independent fashion in Annex 4A**
  - Except for one pesky detail

Gracias!