Ethernet Training Plan - L2

Knowledge of Specifications

- Download IEEE 802.3 Specification
- Understand state machines for transmit, receive, auto negotiation, block lock.
- Clause 35 GMII interface
- Clause 36 1000BASE-X PCS/PMA
- Clause 37 1000BASE-X AN
- Clause 47 XGMII Interface
- Clause 49 10G PCS 10GBASER
- Clause 73 Backplane AN
- Clause 81 XLGMII and CMII (40 and 100Gb/s) Interface
- Clause 82 PCS 40G/100GBASE-R

Ajaya Dahal - AJ

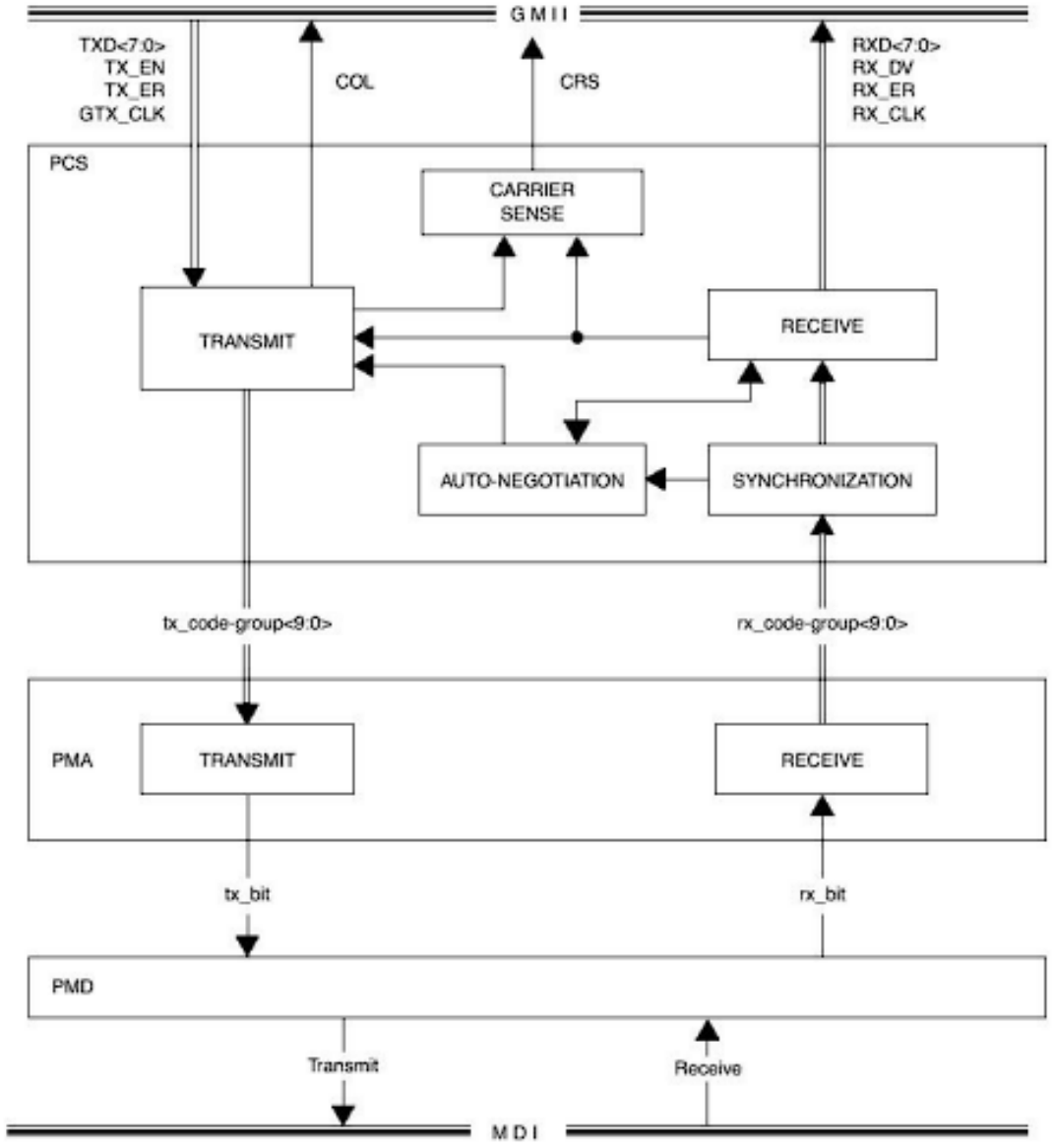


Figure 36–2—Functional block diagram

State diagram variable	MII register	MDIO register
mr_adv_ability[16:1]	4.15:0 Auto-Negotiation advertisement register	7.16.15:0 AN advertisement register
mr_autoneg_complete	1.5 Auto-Negotiation Complete	7.1.5 Auto-Negotiation Complete
mr_autoneg_enable	0.12 Auto-Negotiation Enable	7.0.12 Auto-Negotiation Enable
mr_lp_adv_ability[16:1]	For Base Page: 5.15:0 Auto-Negotiation link partner ability register For Next Page(s): If 6.6=1 and 6.5=1 then 8.15:0 is Auto-Negotiation link partner Received Next Page register If 6.6=1 and 6.5= 0 then 5.15:0 is Auto-Negotiation link partner ability register If 6.6=0 then 8.15:0 or 5.15:0 is Auto-Negotiation link partner Next Page ability register	7.19.15:0 AN LP Base Page ability register
mr_lp_autoneg_able	6.0 Link Partner Auto-Negotiation able	
mr_lp_np_able	6.3 Link Partner Next Page able	
mr_main_reset	0.15 Reset	7.0.15 Reset
mr_next_page_loaded	Set on write to Auto-Negotiation Next Page Transmit register; cleared by Arbitration state diagram	
mr_np_able	6.2 Next Page able	

flp_link_good
Indicates that Auto-Negotiation has completed. Values: false; negotiation is in progress (default). true; negotiation is complete, forcing the Transmit and Receive functions to IDLE.

State machine for transmit

State diagram variable	MII register	MDIO register
mr_np_tx[16:1]	7.15:0 Auto-Negotiation Next Page Transmit Register	7.22.15:0 AN XNP transmit register
mr_np_tx[32:17]	Extended Next Pages not supported by MII register interface	7.23.15:0 Unformatted Code Field 1
mr_np_tx[48:33]	Extended Next Pages not supported by MII register interface	7.24.15:0 Unformatted Code Field 2
mr_page_rx	6.1 Page Received	7.1.6 Page Received
mr_parallel_detection_fault	6.4 Parallel Detection Fault	
mr_restart_negotiation	0.9 Auto-Negotiation Restart	7.0.9 Auto-Negotiation Restart
set if Auto-Negotiation is available	1.3 Auto-Negotiation Ability	7.1.3 Auto-Negotiation Ability

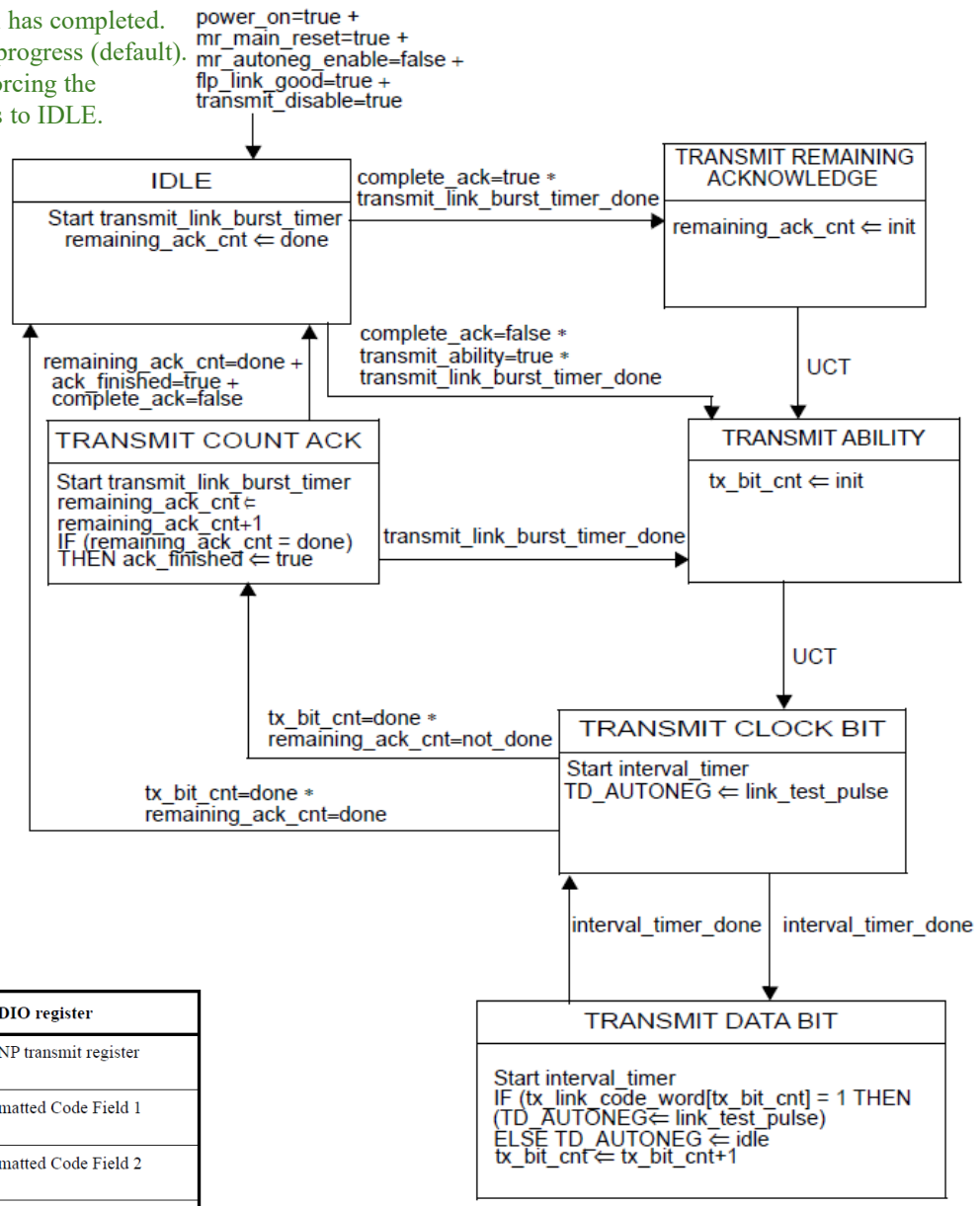


Figure 28–16—Transmit state diagram

transmit_link_burst_timer
Timer for the separation of a transmitted FLP Burst from the next FLP Burst. The transmit_link_burst_timer shall expire 5.7 ms to 22.3 ms after the last transmitted link pulse in an FLP Burst when Extended Next Pages are not supported. When Extended Next Pages are supported, the timer shall expire 5.7 ms to 6.8 ms after the last transmitted link pulse when transmitting 16-bit pages, and shall expire 1.3 ms to 3.1 ms after the last transmitted pulse when
transmitting 48-bit pages.

remaining_ack_cnt
A counter that may take on integer values from 0 to 8. The number of additional link codewords with the Acknowledge Bit set to logic one to be sent to ensure that the Link Partner receives the acknowledgment.
Values: not_done; positive integers between 0 and 5 inclusive.
done; positive integers 6 to 8 inclusive (default).
init; counter is reset to zero.

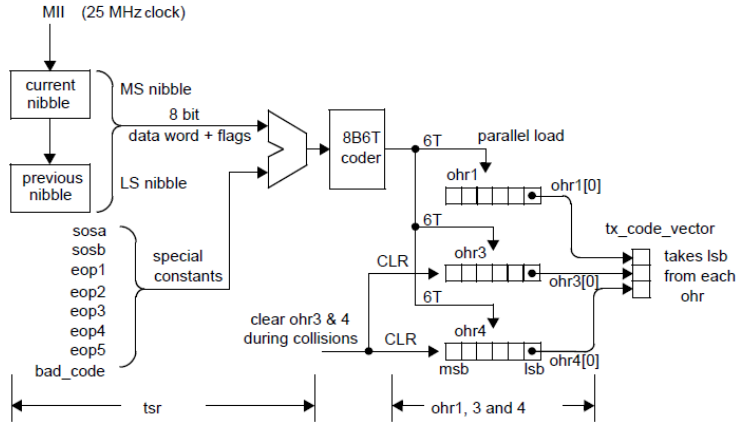
complete_ack
Controls the counting of transmitted link codewords that have their Acknowledge bit set.
Values: false; transmitted link codewords with the Acknowledge bit set are not counted
(default).
true; transmitted link codewords with the Acknowledge bit set are counted.

tx_bit_cnt
A counter that may take on integer values from 1 to 50. This counter is used to keep a count of data bits sent within a DME page. When this variable reaches 50, all data bits have been sent.
Values: not_done; 1 to 49 inclusive.
done; 50.
init; counter is initialized to 1.

State machine for transmit - Variables

All 6T codewords are sent leftmost ternary symbol first.

sosa	A constant that encodes to:	[1 −1 1 −1 1 −1].
sosb	A constant that encodes to:	[1 −1 1 −1 −1 1].
eop1	A constant that encodes to:	[1 1 1 1 1 1].
eop2	A constant that encodes to:	[1 1 1 1 −1 −1].
eop3	A constant that encodes to:	[1 1 −1 −1 0 0].
eop4	A constant that encodes to:	[−1 −1 −1 −1 −1 −1].
eop5	A constant that encodes to:	[−1 −1 0 0 0 0].
bad_code	A constant that encodes to:	[−1 −1 −1 1 1 1].
zero_code	A constant that encodes to:	[0 0 0 0 0 0].



Special constants used by TSR	
start of packet	sosa, sosb
end of packet	eop1-5
TX_ER = 1	bad_code
TX_EN = 0	zero_code

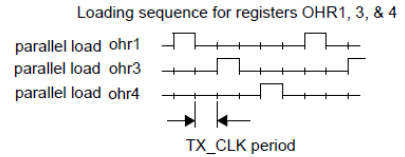


Figure 23–7—PCS Transmit reference diagram

tw1_timer
A continuous free-running timer.
Values: The condition tw1_timer_done goes true when the timer expires.
Restart when: Immediately after expiration (restarting the timer resets condition tw1_timer_done).
Duration: 40 ns nominal.
TX_CLK shall be generated synchronous to tw1_timer (see tolerance required for TX_CLK in 23.5.1.2.10).
On every occurrence of tw1_timer_done, the state diagram advances by one block. The message PMA_UNITDATA.request is issued concurrent with tw1_timer_done.

Register tsr may take on any of the nine constant values listed below (sosa through eop5, bad_code, and zero_code). These values are used to describe the functional operation of the coding process.

State machine for PCS transmit

23.2.4.5 PCS state diagrams

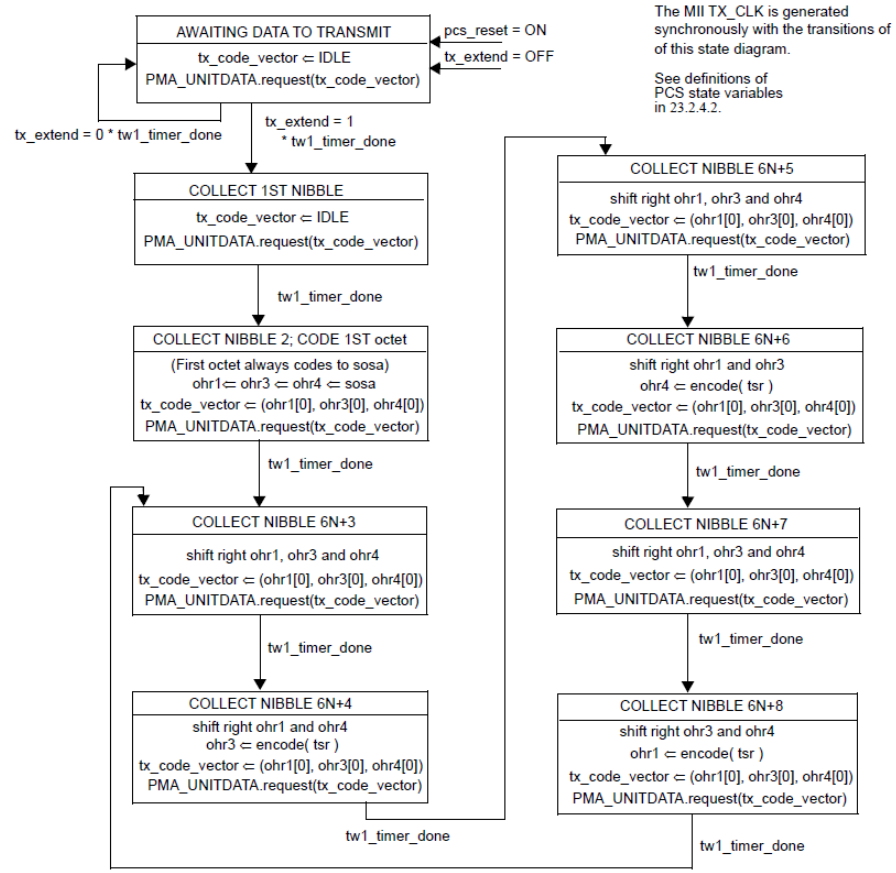


Figure 23–8—PCS Transmit state diagram

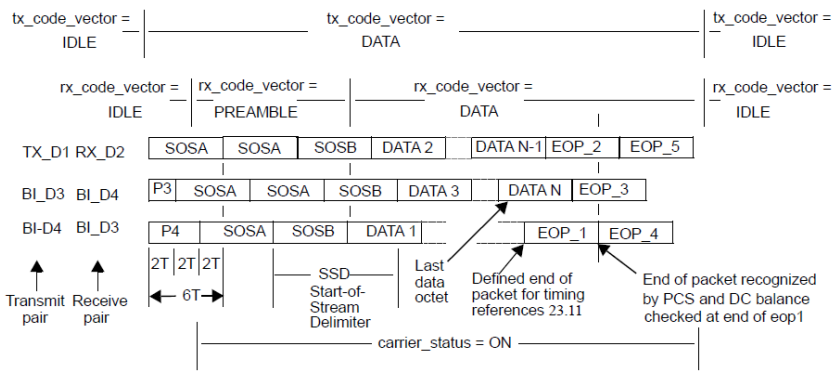


Figure 23–6—PCS sublayer to PMA sublayer frame structure

PMA_UNITDATA.request (tx_symb_vector)
A signal sent to PMA Transmit indicating that a vector of two quinary symbols is available in tx_symb_vector.

flp_test_max_timer
Timer for the maximum time between two link pulses within an FLP Burst. This timer is used in conjunction with the flp_test_min_timer to detect whether the Link Partner is transmitting FLP Bursts. The flp_test_max_timer shall expire 165 µs to 185 µs from the last link pulse.

The basis for all of auto-negotiation's functionality is the Fast Link Pulse (FLP) burst. An FLP burst is simply a sequence of 10BASE-T Normal Link Pulses (NLPs, also known as Link Test Pulses in 10BASE-T world) that come together to form a message, or “word”. Each FLP is composed of 33 pulse positions, with the 17 odd numbered positions corresponding to clock pulses and the 16 even numbered positions corresponding to data pulses. The time between pulse positions is 62.5µs +/- 7µs, and therefore 125µs +/- 14µs between each clock pulse. All clock positions are required to contain a link pulse. However, data positions are not. If there is a link pulse present in a data position, it is representative of a logic one, whereas the lack of a link pulse is representative of a logic zero. The amount of time between FLP bursts is 16ms +/- 8ms, which corresponds to the time between consecutive link test pulses produced by a 10BASE-T device. This was done to allow a fixed speed 10BASE-T device to see FLP bursts and, rather than croaking, remain in the LINK TEST PASS state

State machine for receive

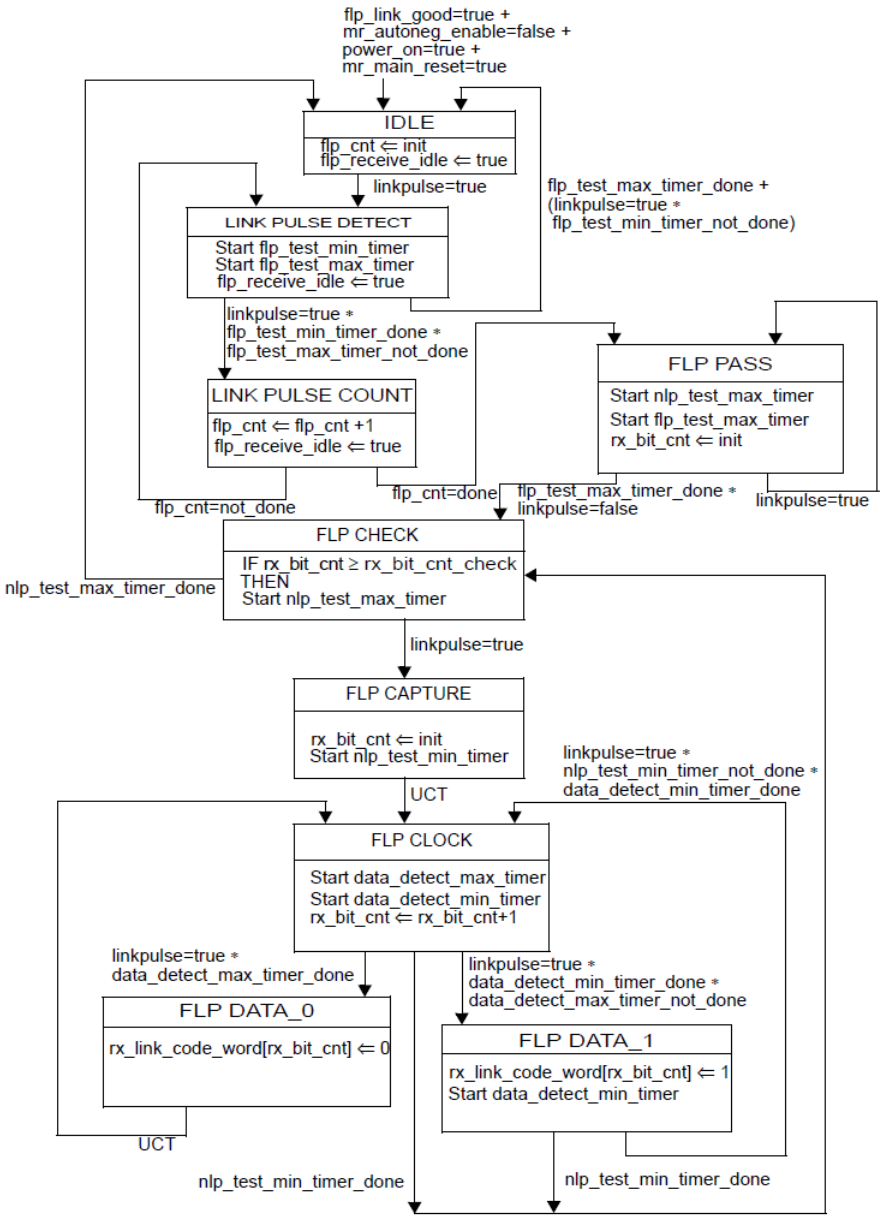


Figure 28–17—Receive state diagram

PMA_UNITDATA.indication (rx_symb_vector)
A signal sent by PMA Receive indicating that a vector of two quinary symbols is available in rx_symb_vector.

All 6T codewords are sent leftmost ternary symbol first.

sosa	A constant that encodes to:	[1 −1 1 −1 1 −1].
sosb	A constant that encodes to:	[1 −1 1 −1 −1 1].
eop1	A constant that encodes to:	[1 1 1 1 1 1].
eop2	A constant that encodes to:	[1 1 1 1 −1 −1].
eop3	A constant that encodes to:	[1 1 −1 −1 0 0].
eop4	A constant that encodes to:	[−1 −1 −1 −1 −1 −1].
eop5	A constant that encodes to:	[−1 −1 0 0 0 0].
bad_code	A constant that encodes to:	[−1 −1 −1 1 1 1].
zero_code	A constant that encodes to:	[0 0 0 0 0 0].

State machine for PCS receive

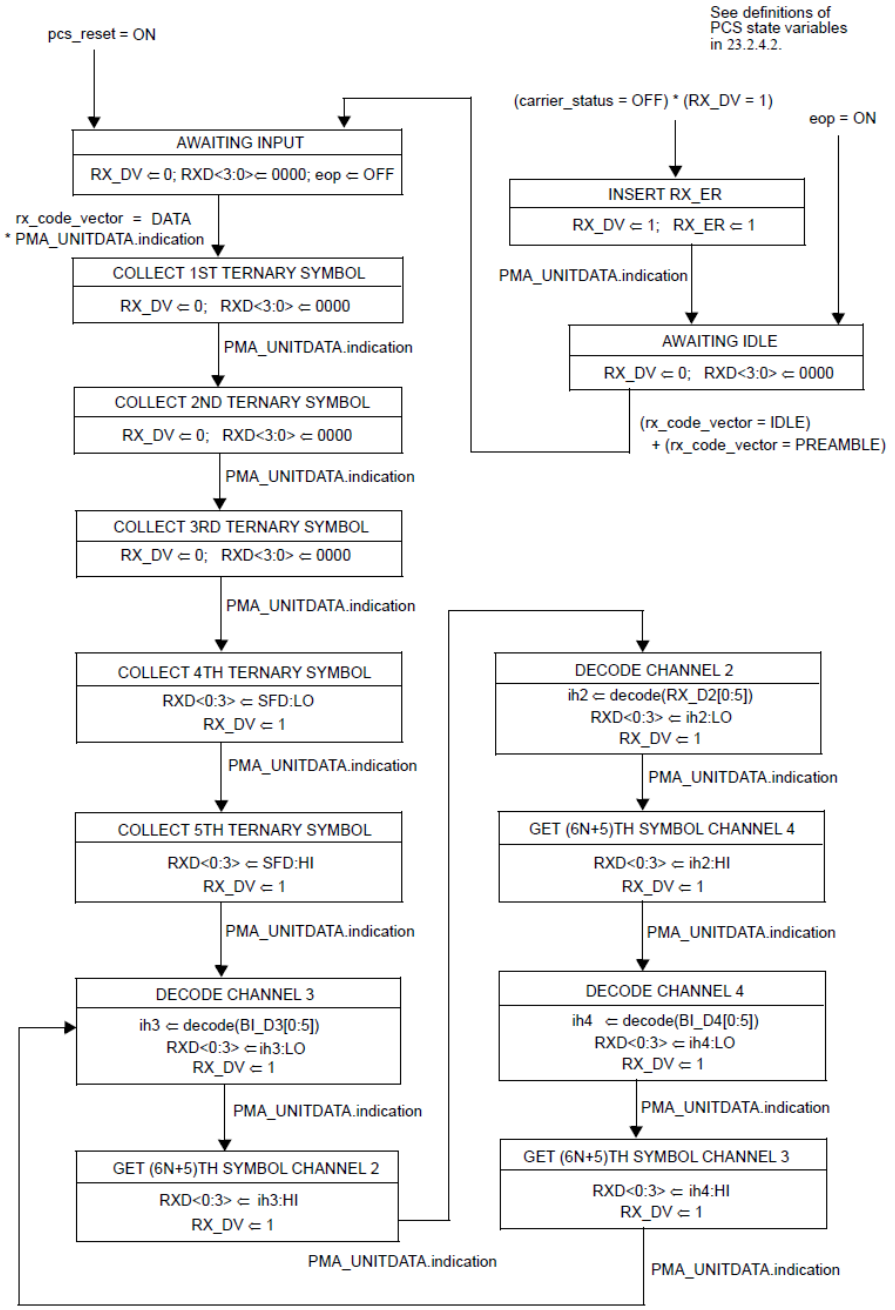
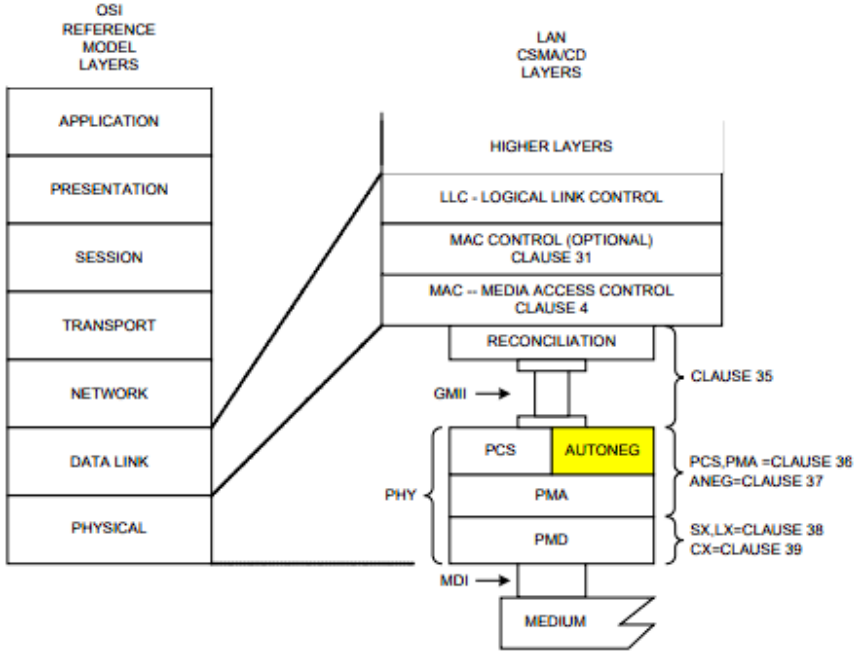


Figure 23–9—PCS Receive state diagram

Auto-Negotiation

AN is located inside the PCS layer

- AN is inside the PCS, which means that the device already needs to have established synchronization with the link partner at Gig speed (1.25 Gigabits/second on the line) before any information can be transferred
- AN has control of what is sent while xmit = CONFIGURATION and IDLE
- The PCS controls what is sent when xmit = DATA
- PCS supplies link status information to the autonegotiation layer



1000BASE-SX
1000BASE-LX
1000BASE-CX

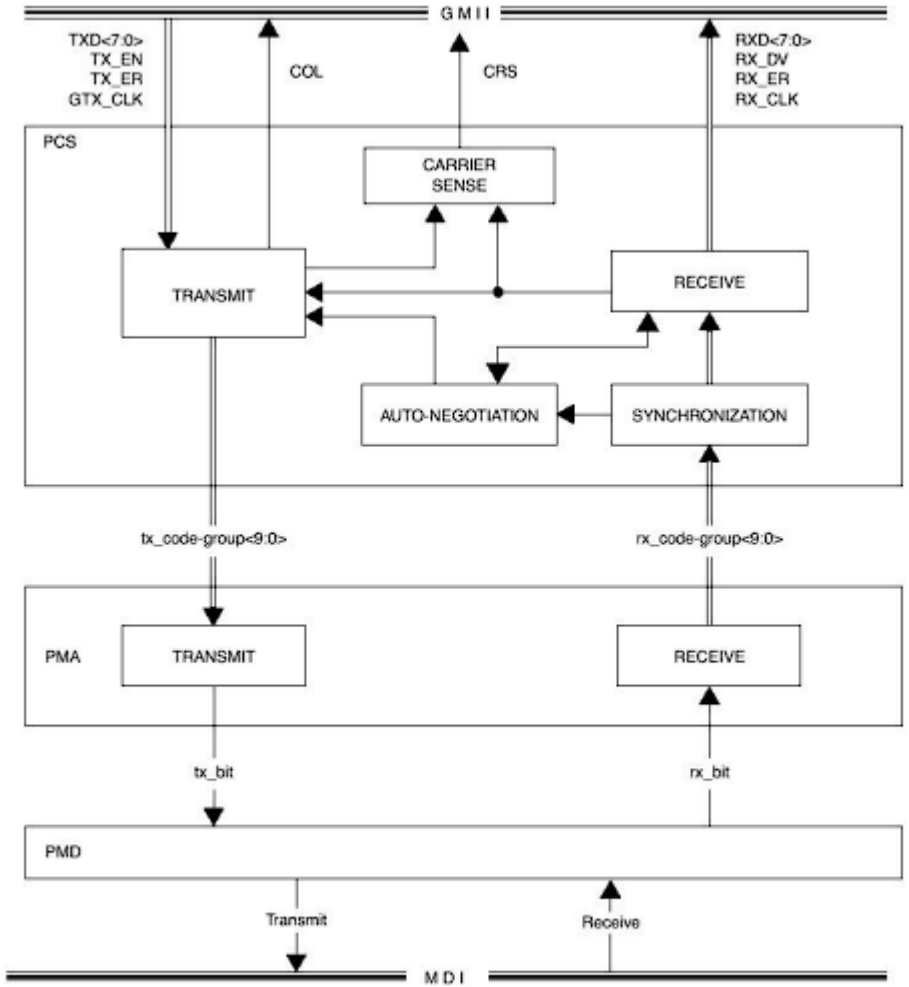
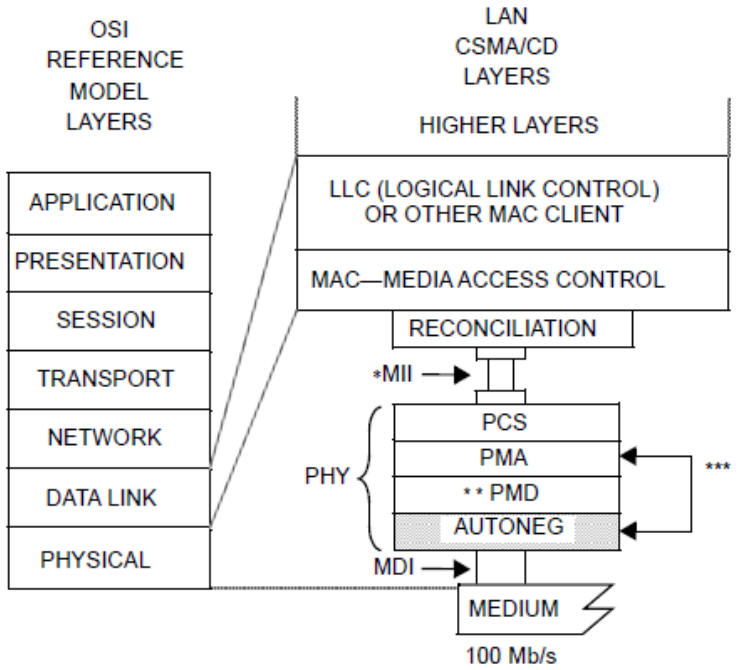


Figure 36–2—Functional block diagram



MDI = MEDIUM DEPENDENT INTERFACE
MII = MEDIA INDEPENDENT INTERFACE
AUTONEG = AUTO-NEGOTIATION

PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PHY = PHYSICAL LAYER DEVICE
PMD = PHYSICAL MEDIUM DEPENDENT

* MII is optional for 10 Mb/s DTEs and for 100 Mb/s systems and is not specified for 1 Mb/s systems.
** PMD is specified for 100BASE-X only; 100BASE-T4 does not use this layer.
*** AUTONEG communicates with the PMA sublayer through the PMA service interface messages PMA_LINK.request and PMA_LINK.indication.

Figure 28–2—Location of Auto-Negotiation function within the ISO/IEC OSI reference model

Auto-Negotiation cont'd

- Auto-Negotiation is responsible for ensuring that the optimal link is resolved between two link partners
- Information is exchanged in a 16-bit word called a configuration register (Config_Reg), which is part of a /C/ Ordered Set
- Duplex and Pause modes are considered when determining the optimal link
- Auto-Negotiation is also responsible for exchanging remote fault and any additional information

Ordered Sets

- An ordered set is defined as a Special (K) code group followed by one or more Data (D) code groups
- There are four ordered sets that are used in Auto-Negotiation
- Two are Idle (/I/) Ordered Sets and two are Configuration (/C/) Ordered Sets

Idle Ordered Sets

- The two possible Idle Ordered Sets are known as I1 and I2
  - I1 is /K28.5/D5.6/
  - I2 is /K28.5/D16.2/
- I1 is used to flip Running Disparity, while I2 is used to maintain Running Disparity

/C/ Ordered Sets

- The two possible /C/ ordered sets are known as C1 and C2
  - C1 is /K28.5/D21.5/Config_Reg
  - C2 is /K28.5/D2.2/Config_Reg
- The Config_Reg is made of two additional Data code groups

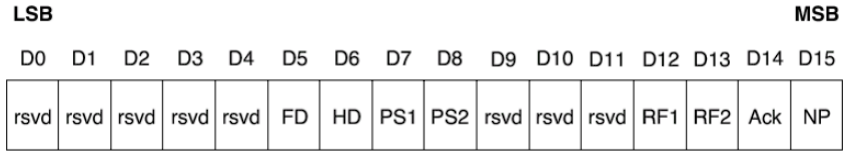
Encoding	Hex	Symbol	Name	Description
K28.5	BC	COM	Comma	Used for Lane and Link initialization and management
K27.7	FB	STP	Start TLP	Marks the start of a Transaction Layer Packet
K28.2	5C	SDP	Start DLLP	Marks the start of a Data Link Layer Packet
K29.7	FD	END	End	Marks the end of a Transaction Layer or Data Link Layer Packet
K30.7	FE	EDB	End bad	Marks the end of a nullified TLP
K23.7	F7	PAD	Pad	Used in Framing and Link Width and Lane ordering negotiations
K28.0	1C	SKP	Skip	Used for compensating for different bit rates for two communicating Ports
K28.1	3C	FTS	Fast Training Sequence	Used within an Ordered Set to exit from L0s to L0
K28.3	7C	IDL	Idle	Used in the Electrical Idle Ordered Set (EIOS)
K28.4	9C	-	-	Reserved
K28.6	DC	-	-	Reserved
K28.7	FC	EIE	Electrical Idle Exit	Reserved in 2.5 GT/s Used in the Electrical Idle Exit Ordered Set (EIEOS) and sent prior to sending FTS at speeds other than 2.5 GT/s

Ref: https://www.iol.unh.edu/sites/default/files/knowledgebase/ge/Clause_37_Auto-Negotiation.pdf

Auto-Negotiation cont'd

Configuration Register

- The Config_Reg is a 16-bit word that contains the information to be transferred
- The Config_Reg can be used to create several types of pages, such as base pages and next pages .The base page looks like:



Duplexes

- The supported duplexes are advertised in bits 5 and 6 of the base page
- There are two possible duplexes
  - Full Duplex – Bit 5 – information is passed bidirectionally
  - Half Duplex – Bit 6 – information can only be passed one way at a time

Pause Modes

- The supported pause modes are advertised in bits 7 and 8 of the base page
- There are two possible pause modes
  - Symmetric pause - Bit 7 – The device can support pause frames flowing in both directions
  - Asymmetric pause – Bit 8 – The device can only support pause frames flowing in one direction

Remote Fault

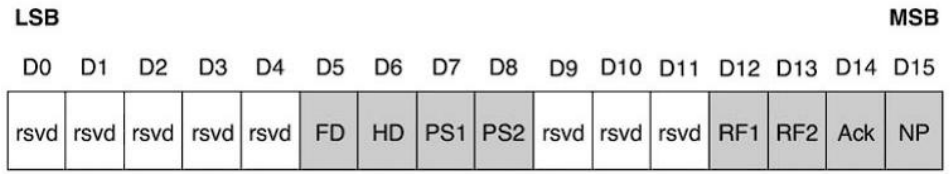
- The remote fault bits are passed in bits 12 and 13 of the base page
- Remote Fault passes link failure information from one device to another

Acknowledge

- The Acknowledge bit is bit 14 of the base page
- The Acknowledge bit lets the link partner know that the device has received the link partner's config_reg
- There is an ability_match function that must be true before the Acknowledge bit can be set to one.

Next Page

- Next page is a method of transferring additional information during the auto-negotiation process
- The Next Page bit is bit 15 of the base page
- The Next Page bit can only be set to one if the device supports next page exchange and wants to send a next page



Ref: https://www.iol.unh.edu/sites/default/files/knowledgebase/ge/Clause_37_Auto-Negotiation.pdf

Physical Layer link signaling for Auto-Negotiation on twisted pair

The functional reference diagram (Figure 28–15) provides a generic example, illustrated with initial PMA implementations and showing the mechanism for expansion. New PMAs are documented in Annex 28D.

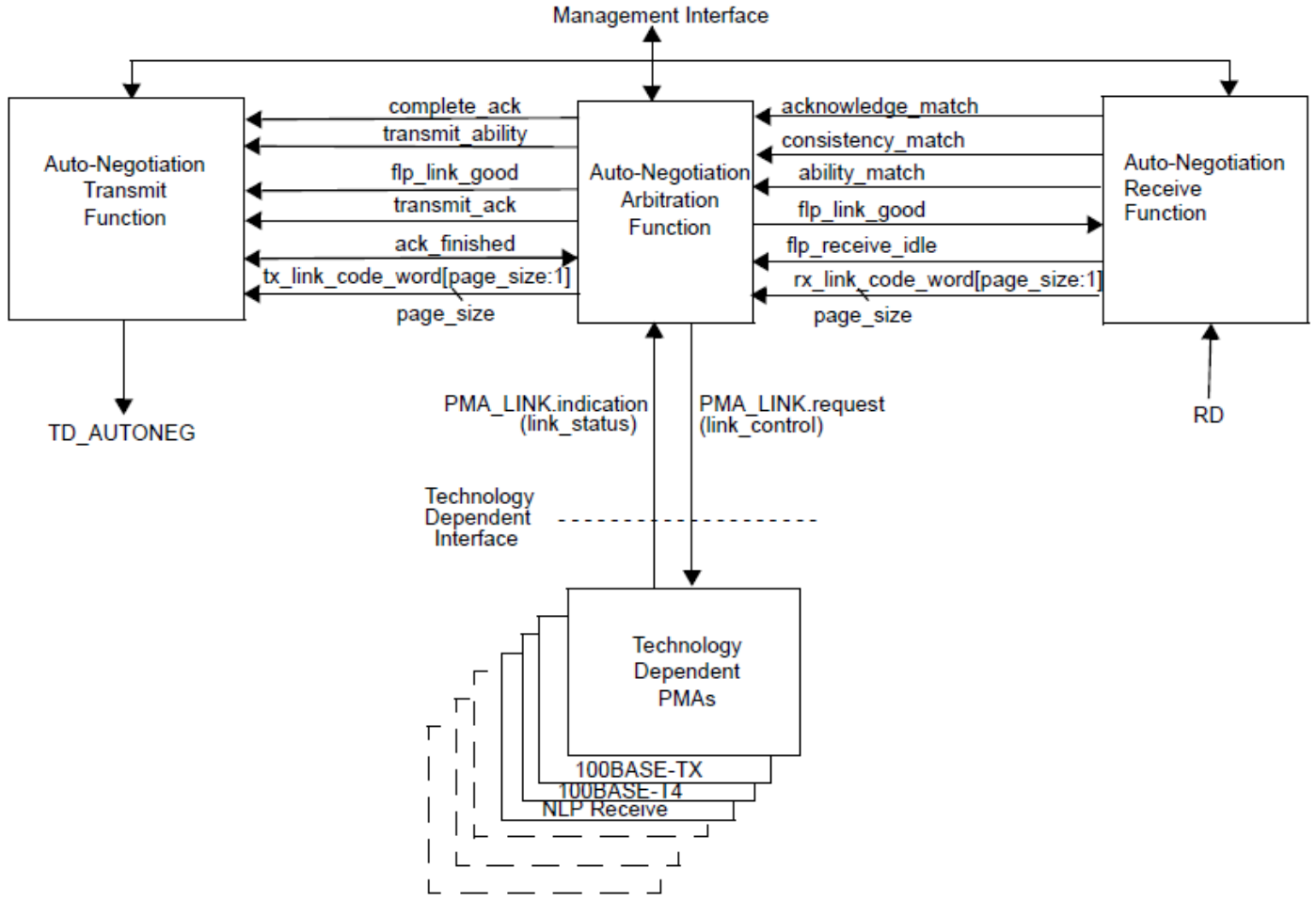


Figure 28–15—Functional reference diagram

- The four key functions in auto-negotiation are ability_match, acknowledge_match, consistency_match, and Idle_match
- The only timer for clause 37 auto-negotiation is link_timer

Ability_Match

- Ability_match is responsible for ensuring that the device receives a reliable config_reg from the link partner
- Ability_match = TRUE when the device has received three consecutive and consistent (**ignoring the ACK bit**) config_regs from the link partner
- The device sets its ACK bit to one when ability_match = TRUE

Acknowledge_Match

• Acknowledge_match is similar to ability_match, except that the device must receive 3 consecutive and consistent config_regs **with the ACK bit set to one**

Consistency_Match

- Consistency_match ensures that the link partner transmitted the same abilities before and after it set its ACK bit
- Consistency_match compares the abilities that were used to set ability_match = TRUE to the abilities that were used to set acknowledge_match = TRUE
- All bits are examined **except for the ACK bit**

Idle_Match

- Idle_match is also similar to ability_match, except that it counts the Idle codes received
- The device cannot establish a link until it has successfully received 3 Idle codes from its link partner

State machine for auto negotiation

Link Timer

- It is defined as 10 ms with a tolerance of +10 ms
- It is used throughout the autonegotiation process to help verify that the link partner has enough time to see everything that happens before the device transitions to the next state

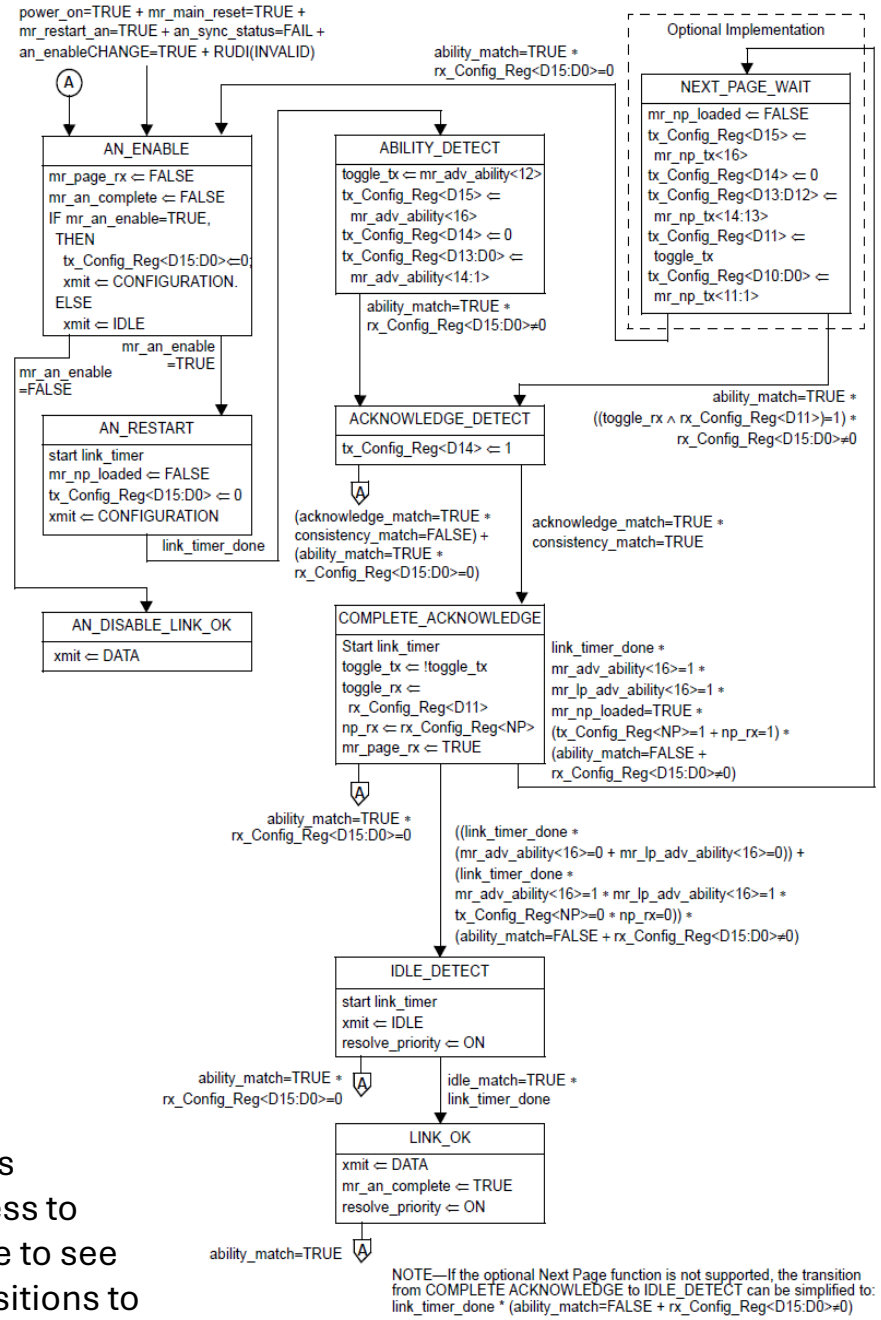
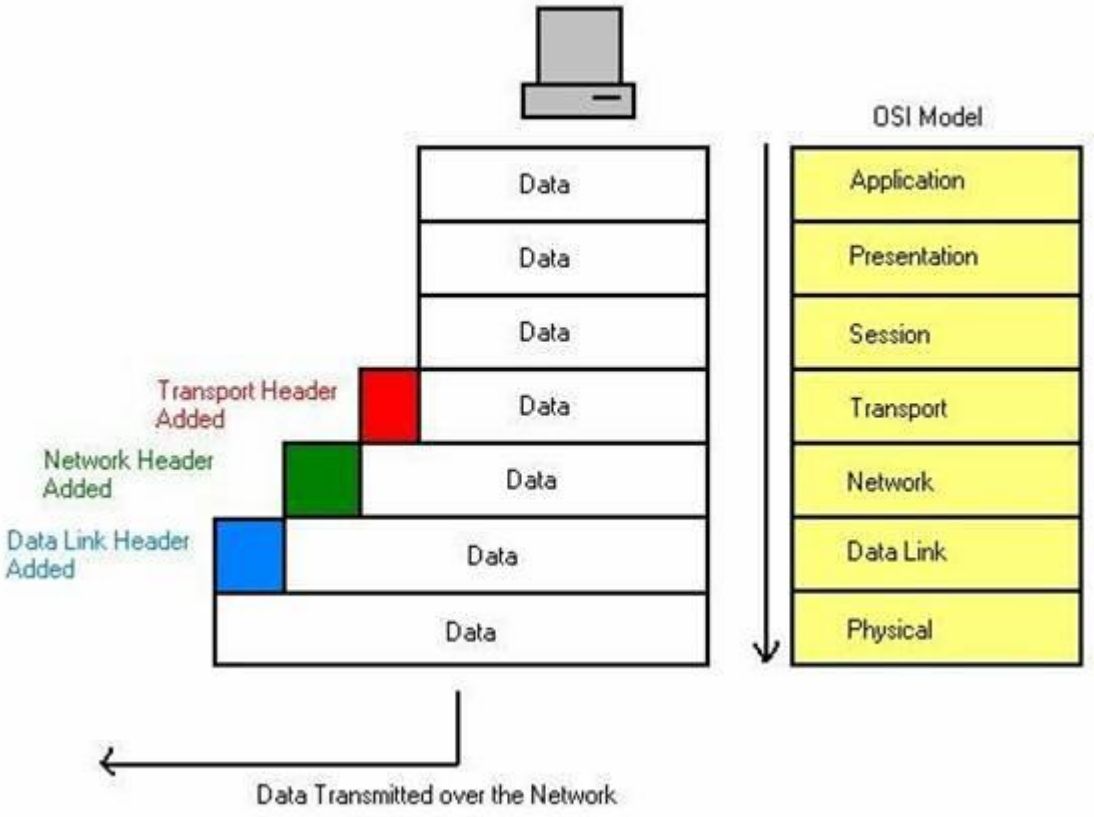
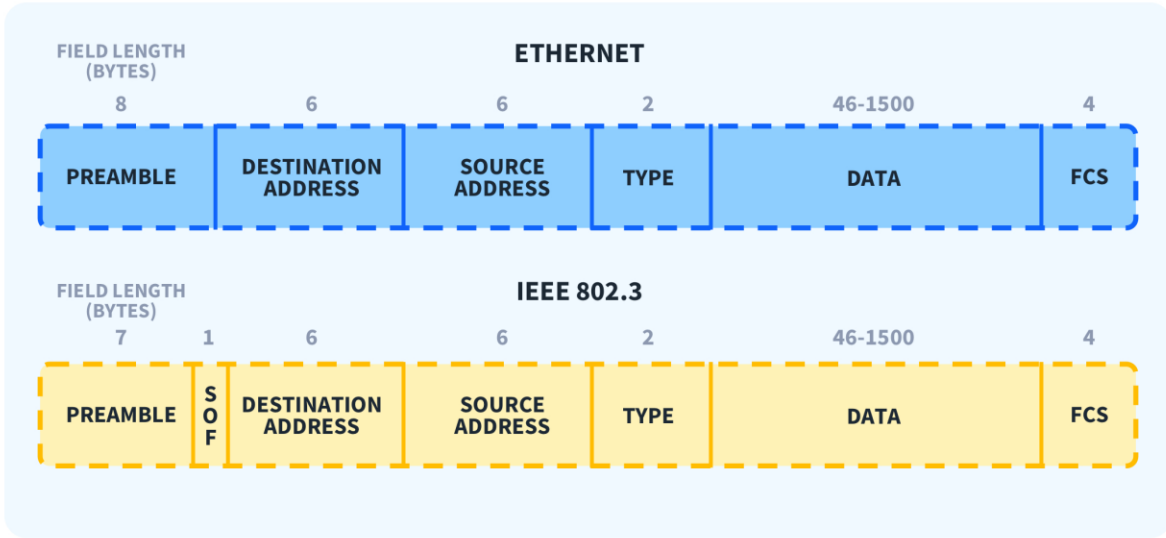


Figure 37–6—Auto-Negotiation state diagram

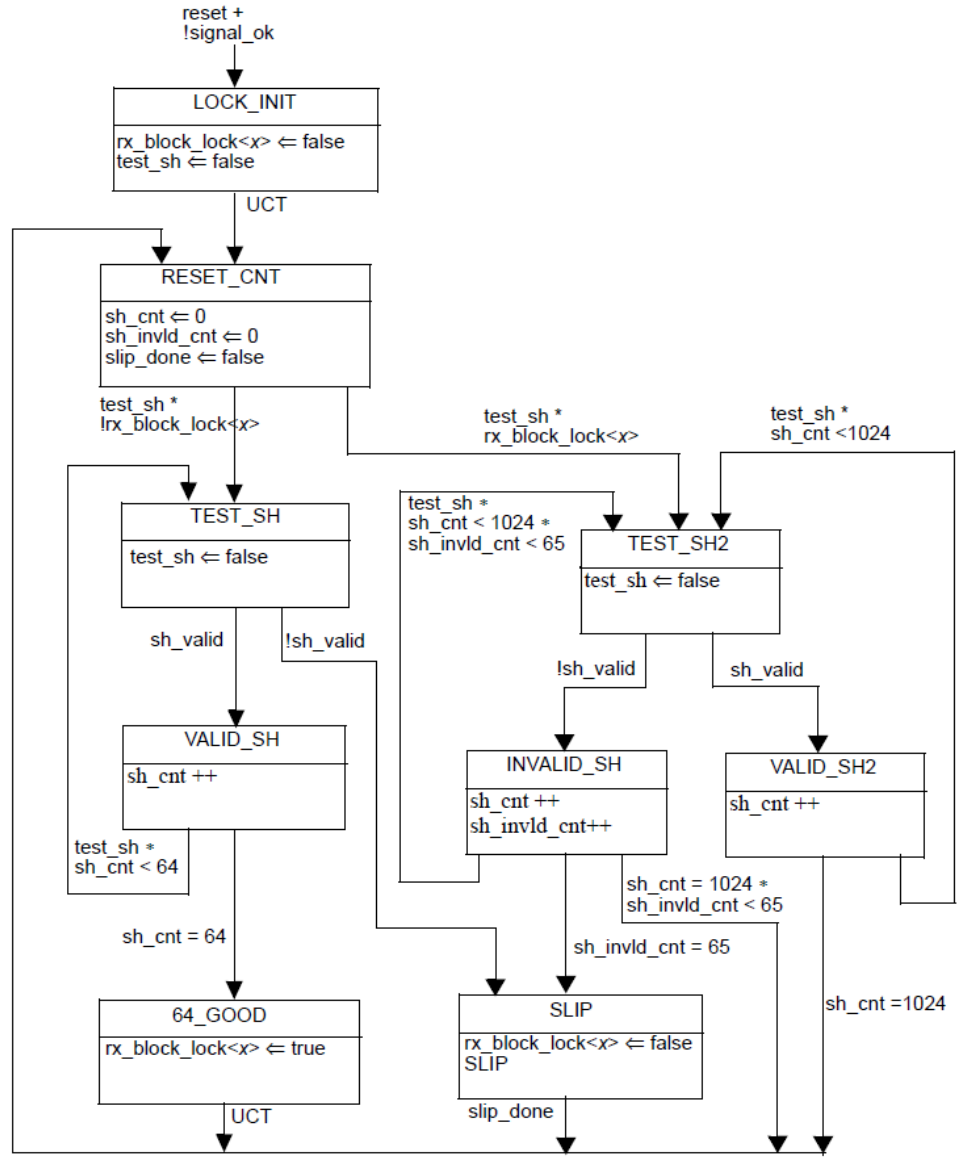
Where is the AN config_reg in ethernet frame?





State machines for block lock

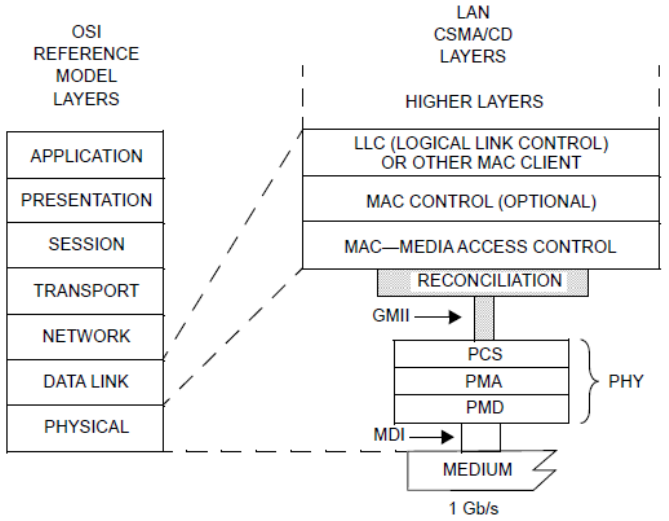
A block lock process operates independently on each lane. Each block lock process looks for 64 valid sync headers in a row to declare lock. A valid sync header is either a 01 or a 10. Once in lock, the lock process looks for 65 invalid sync headers within a 1024 sync window to declare out of lock. An invalid sync header is a 11 or 00.
Once block lock is achieved on a lane, then the alignment marker process starts.



NOTE— rx_block_lock<x> refers to the received lane x of the service interface, where x = 0:3 (for 40GBASE-R) or 0:19 (for 100GBASE-R)

Figure 82–12—Block lock state diagram

Clause 35 GMII Interface



GMII = GIGABIT MEDIA INDEPENDENT INTERFACE
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 35–1—GMII relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

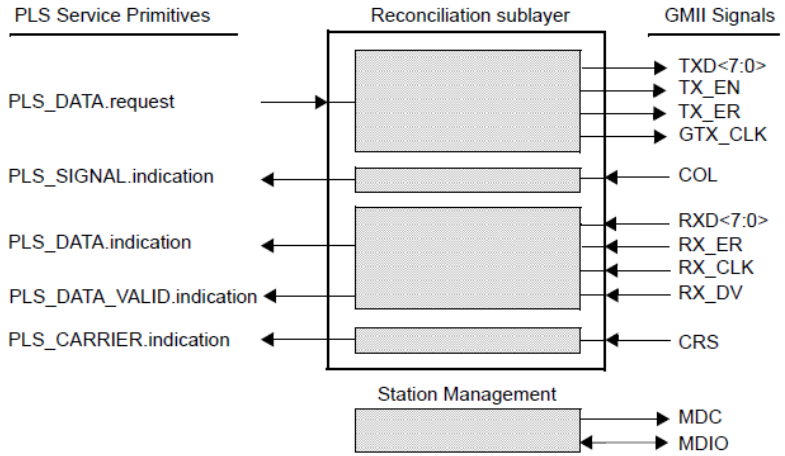
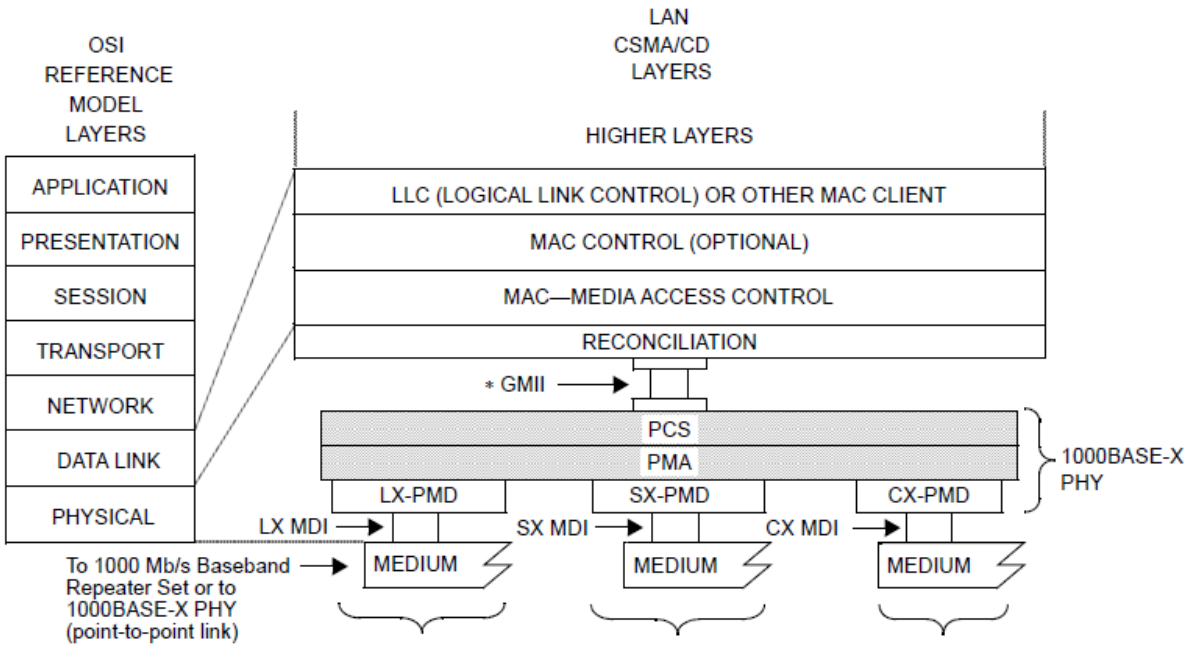


Figure 35–2—Reconciliation Sublayer (RS) inputs and outputs and STA connections to GMII

Clause 36 1000BaseX



MDI=MEDIUM DEPENDENT INTERFACE
GMII=GIGABIT MEDIA INDEPENDENT INTERFACE
PCS=PHYSICAL CODING SUBLAYER
PMA=PHYSICAL MEDIUM ATTACHMENT
PHY=PHYSICAL LAYER DEVICE
PMD=PHYSICAL MEDIUM DEPENDENT
LX-PMD=PMD FOR FIBER—LONG WAVELENGTH, Clause 38
SX-PMD=PMD FOR FIBER—SHORT WAVELENGTH, Clause 38
CX-PMD=PMD FOR 150 Ω BALANCED COPPER CABLING, Clause 39

NOTE—The PMD sublayers are mutually independent.
* GMII is optional.

Figure 36–1—Relationship of 1000BASE-X and the PMDs

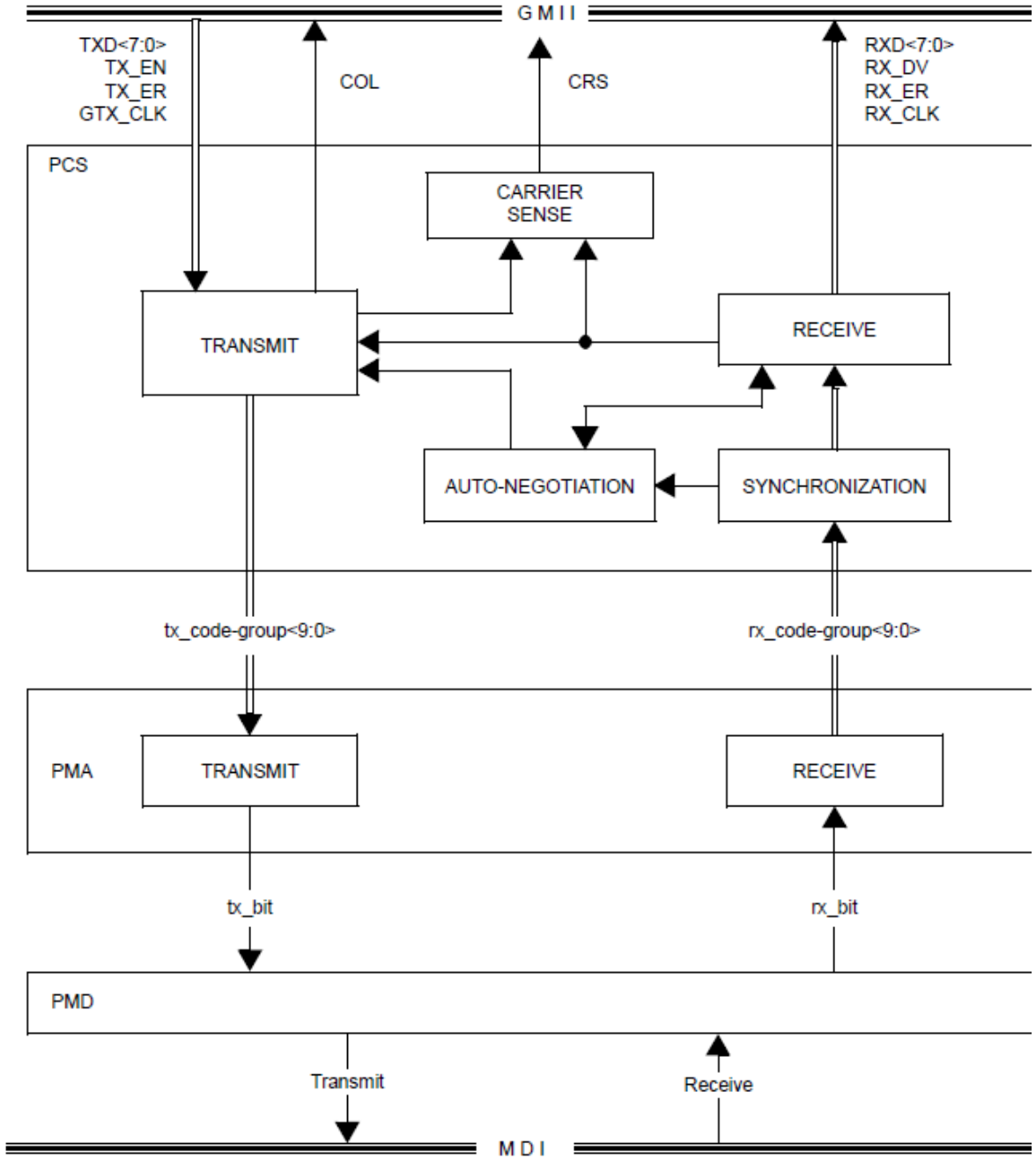


Figure 36–2—Functional block diagram

Clause 47 XGMII Interface

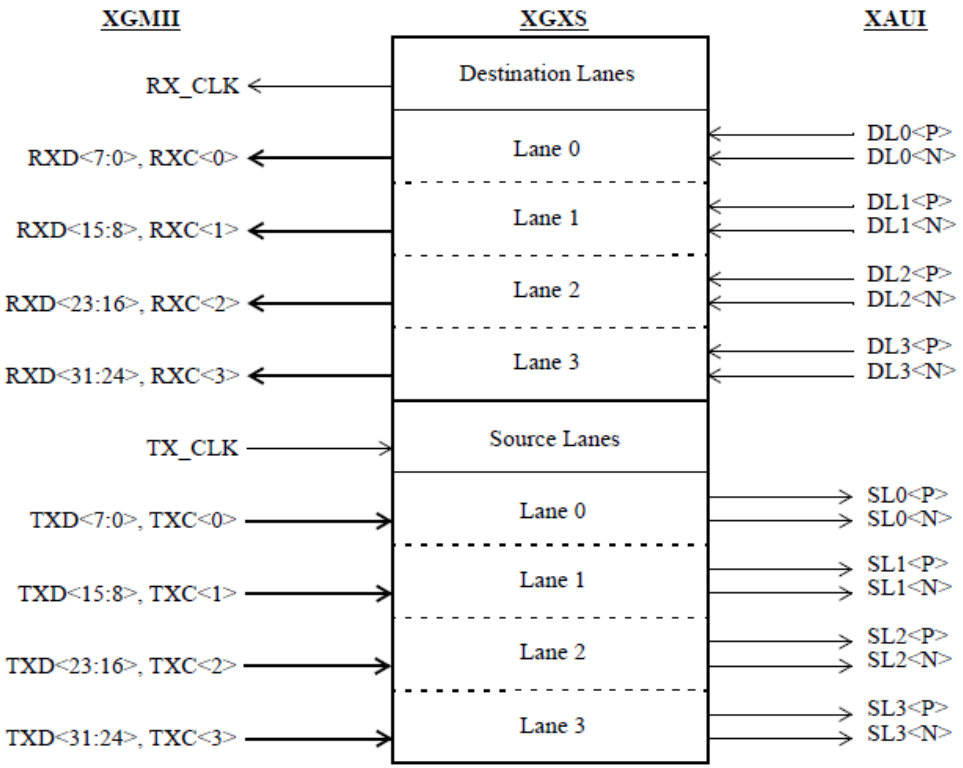
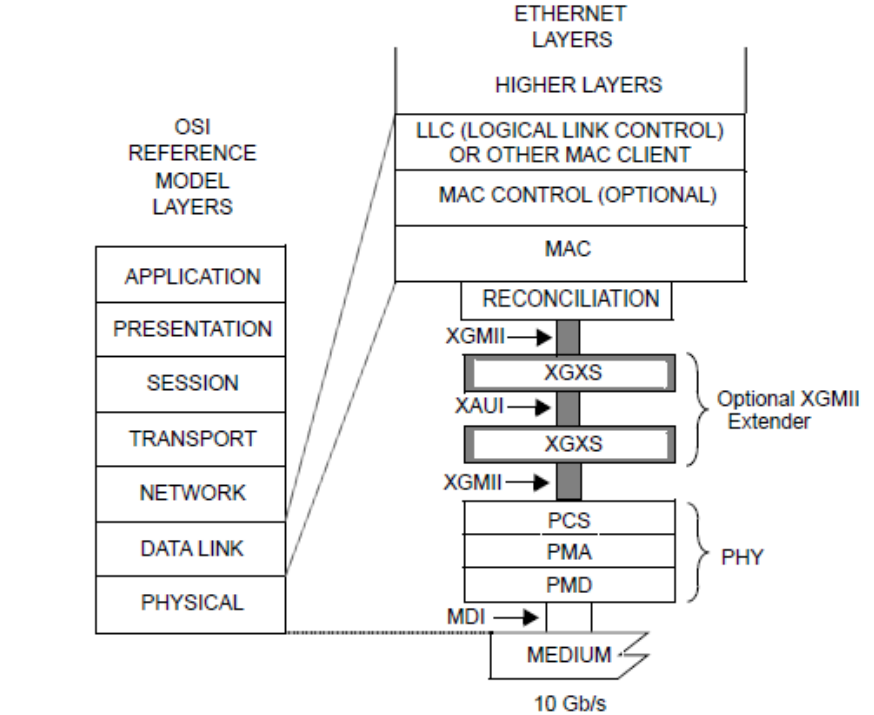


Figure 47–2—XGXS inputs and outputs

This clause defines the functional and electrical characteristics for the optional XGMII Extender Sublayer (XGXS) and 10 Gigabit Attachment Unit Interface (XAUI). Figure 47–1 shows the relationships of the XGMII, XGMII Extender, XGXS, and XAUI.



MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XAUI = 10 GIGABIT ATTACHMENT UNIT INTERFACE
XGMII = 10 GIGABIT MEDIA INDEPENDENT INTERFACE
XGXS = XGMII EXTENDER SUBLAYER

Figure 47–1—XAUI and XGXS relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model

Figure 48–1 depicts the relationships among the 10GBASE-X sublayers (shown shaded), the IEEE 802.3 MAC and RS, and the IEEE 802.2 LLC.

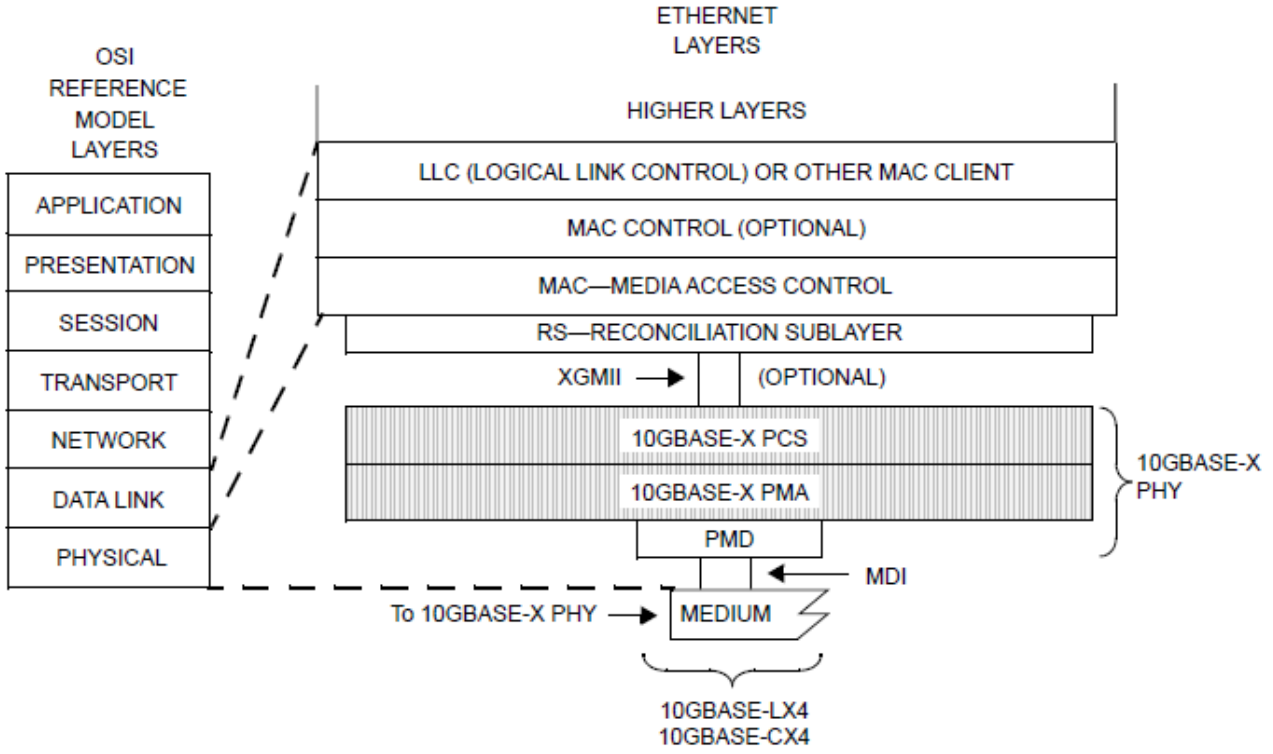
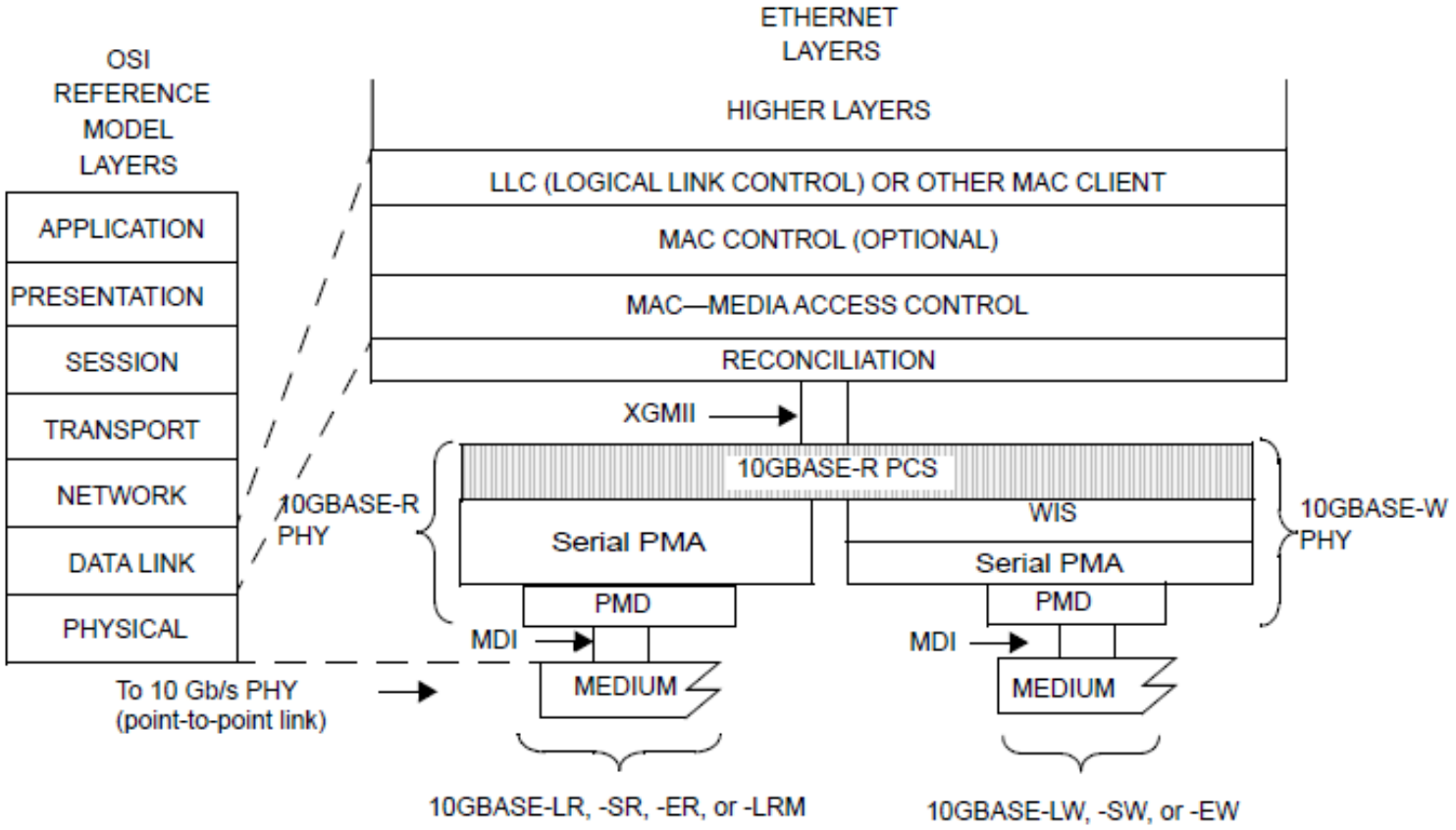


Figure 48–1—10GBASE-X PCS and PMA relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE Ethernet Model

Clause 49 10G PCS



MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
WIS = WAN INTERFACE SUBLAYER
XGMII = 10 GIGABIT MEDIA INDEPENDENT INTERFACE

NOTE—The PMD sublayers are mutually independent.

PMD TYPES:
Medium:
E = PMD FOR FIBER—1550 nm WAVELENGTH
L = PMD FOR FIBER—1310 nm WAVELENGTH
S = PMD FOR FIBER—850 nm WAVELENGTH
M = PMD WITH DISPERSION COMPENSATION FOR MULTI-MODE FIBER
Encoding:
R = 64B/66B ENCODED WITHOUT WIS
W = 64B/66B ENCODED WITH WIS

Figure 49–1—10GBASE-R PCS relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 Ethernet model

49.1.6 Functional block diagram

Figure 49–4 provides a functional block diagram of the 10GBASE-R PHY.

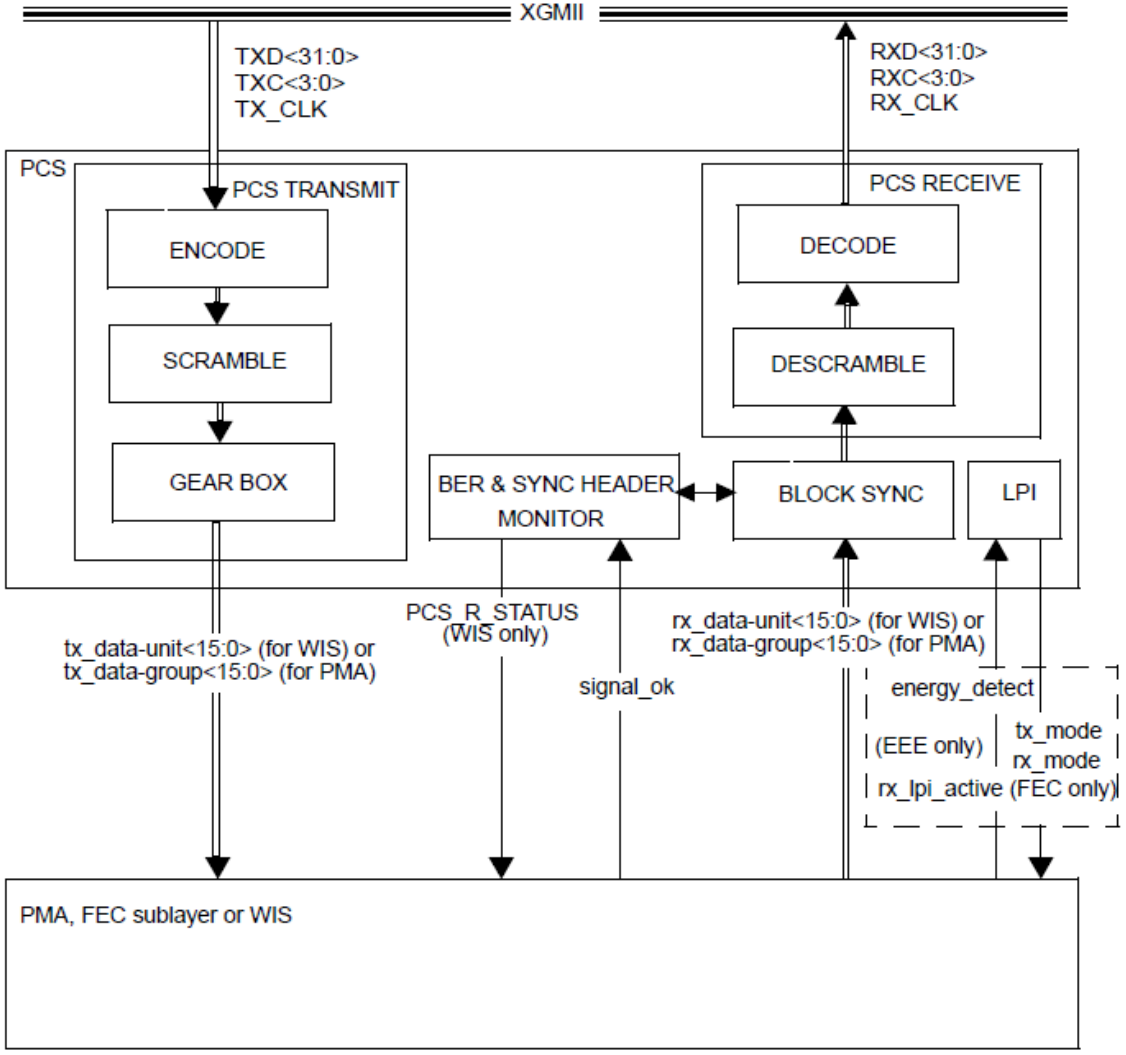


Figure 49–4—Functional block diagram

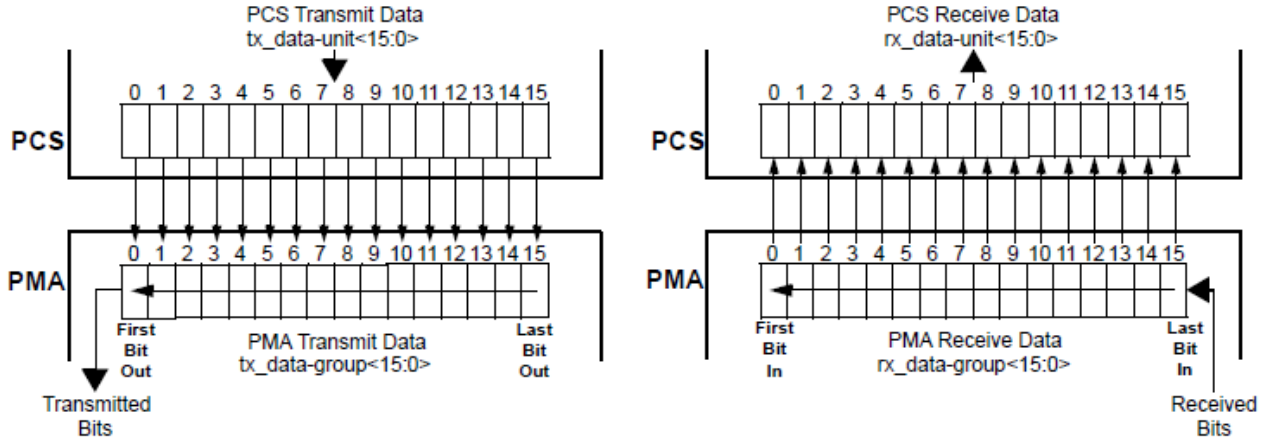


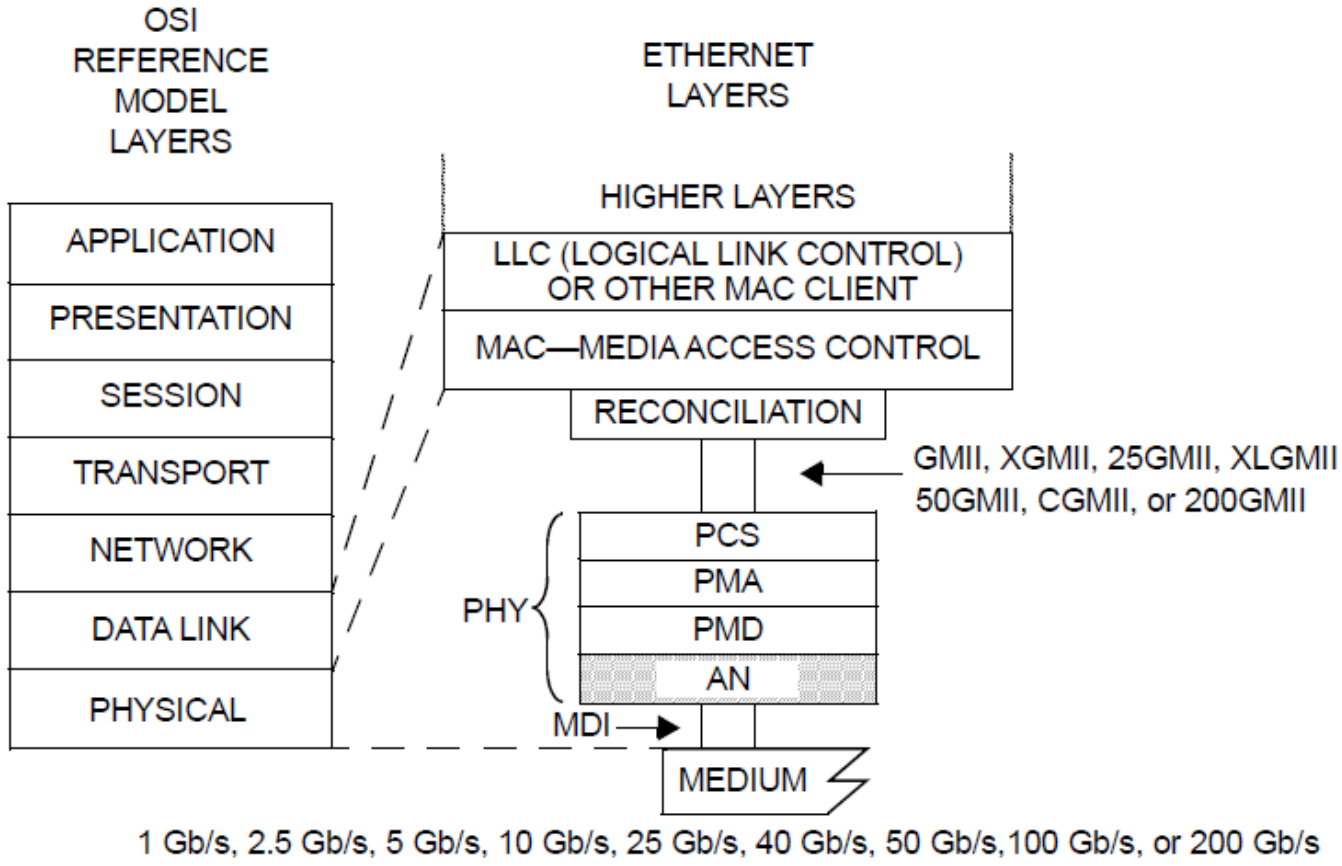
Figure 49–2—10GBASE-R PHY transmission order

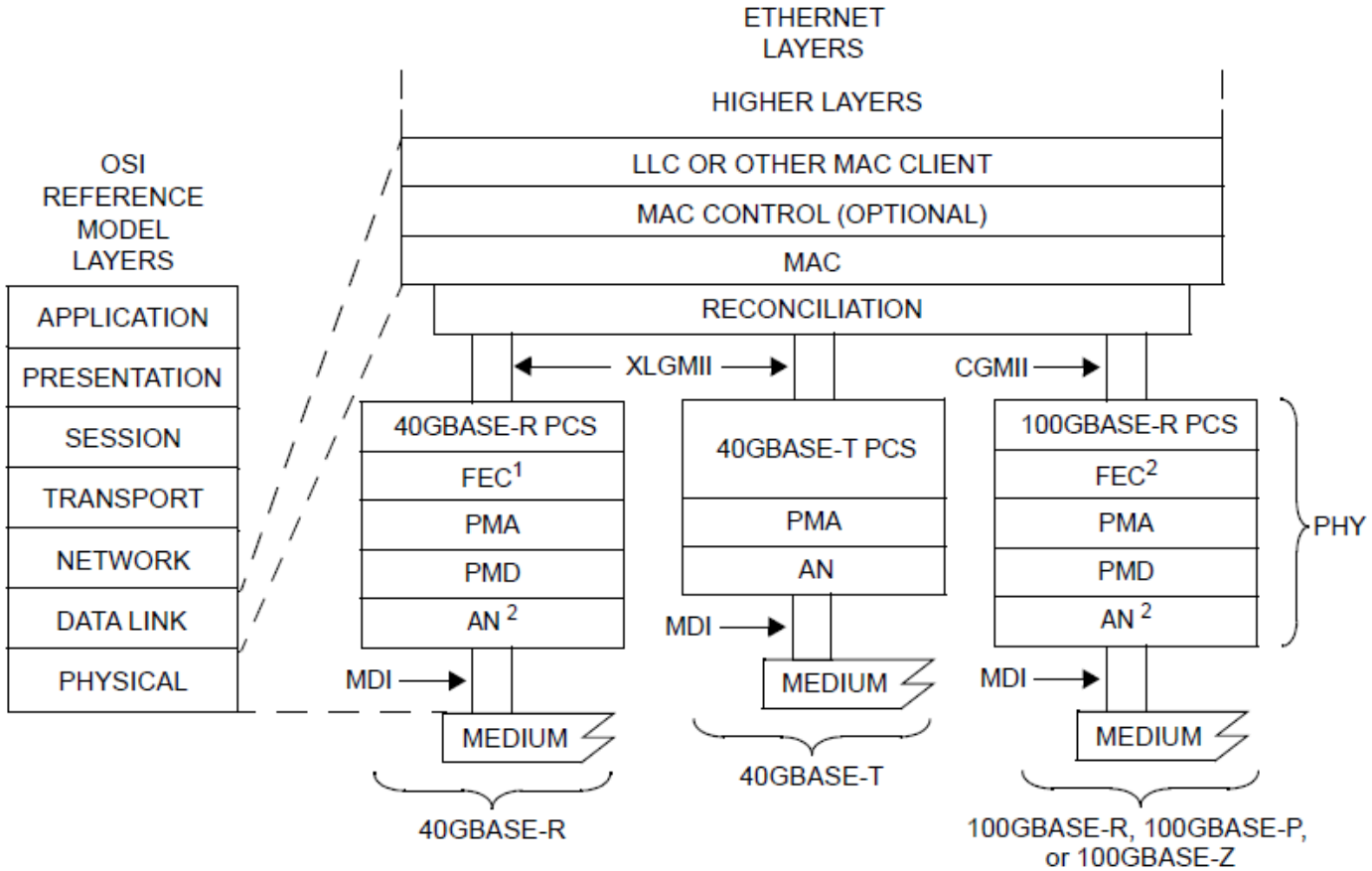
Clause 73 Auto-Negotiation

While implementation of Auto-Negotiation is mandatory for Backplane Ethernet PHYs, the use of Auto-Negotiation is optional. Parallel detection shall be provided for legacy devices that do not support Auto-Negotiation.

Auto-Negotiation is performed using differential Manchester encoding (DME) pages. DME provides a DC-balanced signal. DME does not add packet or upper layer overhead to the network devices.

It is recommended that a device that has negotiated 1000BASE-KX operation through this clause not perform Clause 37 Auto-Negotiation. A device that performs Clause 37 Auto-Negotiation after having negotiated 1000BASE-KX operation through Clause 73 Auto-Negotiation will not interoperate with a device that does not perform Clause 37 Auto-Negotiation. Therefore, a device that intends to enable Clause 37 Auto-Negotiation after Clause 73 Auto-Negotiation has completed shall ensure through an implementation-specific mechanism that the link partner supports Clause 37 Auto-Negotiation and intends to enable it. If Clause 37 Auto-Negotiation is performed after Clause 73 Auto-Negotiation, then the advertised abilities used in the Clause 37 Auto-Negotiation shall match those advertised abilities used in the Clause 73 Auto-Negotiation.





AN = AUTO-NEGOTIATION
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
LLC = LOGICAL LINK CONTROL
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

NOTE 1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE
NOTE 2—CONDITIONAL BASED ON PHY TYPE

Figure 80–1—Architectural positioning of 40 Gigabit and 100 Gigabit Ethernet

Table 80–2—Nomenclature and clause correlation (40GBASE)

Nomenclature	Clause[a]																	
	28	73	74	78	81		82	83	83A	83B	84	85	86	86A	87		89	113
	Auto-Negotiation	Auto-Negotiation	BASE-R FEC	EEE	RS	XLGMII	40GBASE-R PCS	40GBASE-R PMA	XLAUI C2C	XLAUI C2M	40GBASE-KR4 PMD	40GBASE-CR4 PMD	40GBASE-SR4 PMD	XLPPI	40GBASE-LR4 PMD	40GBASE-ER4 PMD	40GBASE-FR PMD	40GBASE-T PCS/PMA
40GBASE-KR4		M	O	O	M	O	M	M	O		M							
40GBASE-CR4		M	O	O	M	O	M	M	O			M						
40GBASE-T	M			O	M	O												M
40GBASE-SR4				O	M	O	M	M	O	O			M	O				
40GBASE-FR				O	M	O	M	M	O	O							M	
40GBASE-LR4				O	M	O	M	M	O	O				O	M			
40GBASE-ER4				O	M	O	M	M	O	O						M		

[a]O = Optional, M = Mandatory.

Table 80–3—Nomenclature and clause correlation (100GBASE copper)

Nomenclature	Clause[a]																		
	73	74	78	81		82	83	83A	83D	85	91	92	93	94	135	135D	135F	136	137
	Auto-Negotiation	BASE-R FEC	EEE	RS	CGMII	100GBASE-R PCS	100GBASE-R PMA	CAUI-10 C2C	CAUI-4 C2C	100GBASE-CR10 PMD	RS-FEC	100GBASE-CR4 PMD	100GBASE-KR4 PMD	100GBASE-KP4 PMD	100GBASE-P PMA	100GAUI-4 C2C	100GAUI-2 C2C	100GBASE-CR2	100GBASE-KR2
100GBASE-KR4	M		O	M	O	M	M	O	O		M		M						
100GBASE-KP4	M		O	M	O	M	O	O	O		M			M					
100GBASE-KR2	M		O	M	O	M		O	O		M				M	O	O		M
100GBASE-CR2	M		O	M	O	M		O	O		M				M	O	O	M	
100GBASE-CR4	M		O	M	O	M	M	O	O		M	M							
100GBASE-CR10	M	O	O	M	O	M	M	O	O	M									

[a]O = Optional, M = Mandatory.

Table 80–4—Nomenclature and clause correlation (100GBASE-R optical)

Nomenclature	Clause[a]														
	78	81		82	83	83A	83B	83D	83E	86	86A	88		91	95
	EEE	RS	CGMII	100GBASE-R PCS	100GBASE-R PMA	CAUI-10 C2C	CAUI-10 C2M	CAUI-4 C2C	CAUI-4 C2M	100GBASE-SR10 PMD	CPPI	100GBASE-LR4 PMD	100GBASE-ER4 PMD	RS-FEC	100GBASE-SR4 PMD
100GBASE-SR4	O	M	O	M	M	O	O	O	O					M	M
100GBASE-SR10	O	M	O	M	M	O	O	O	O	M	O				
100GBASE-LR4	O	M	O	M	M	O	O	O	O			M			
100GBASE-ER4	O	M	O	M	M	O	O	O	O				M		

[a]O = Optional, M = Mandatory.

100GBASE-SR2	O	M	O	M	O	O	O	O	O	M	M	O	O	O	O	M			
100GBASE-DR	O	M	O	M	O	O	O	O	O	M	M	O	O	O	O		M		
100GBASE-FR1	O	M	O	M	O	O	O	O	O	M	M	O	O	O	O			M	
100GBASE-LR1	O	M	O	M	O	O	O	O	O	M	M	O	O	O	O				M

[a]O = Optional, M = Mandatory.

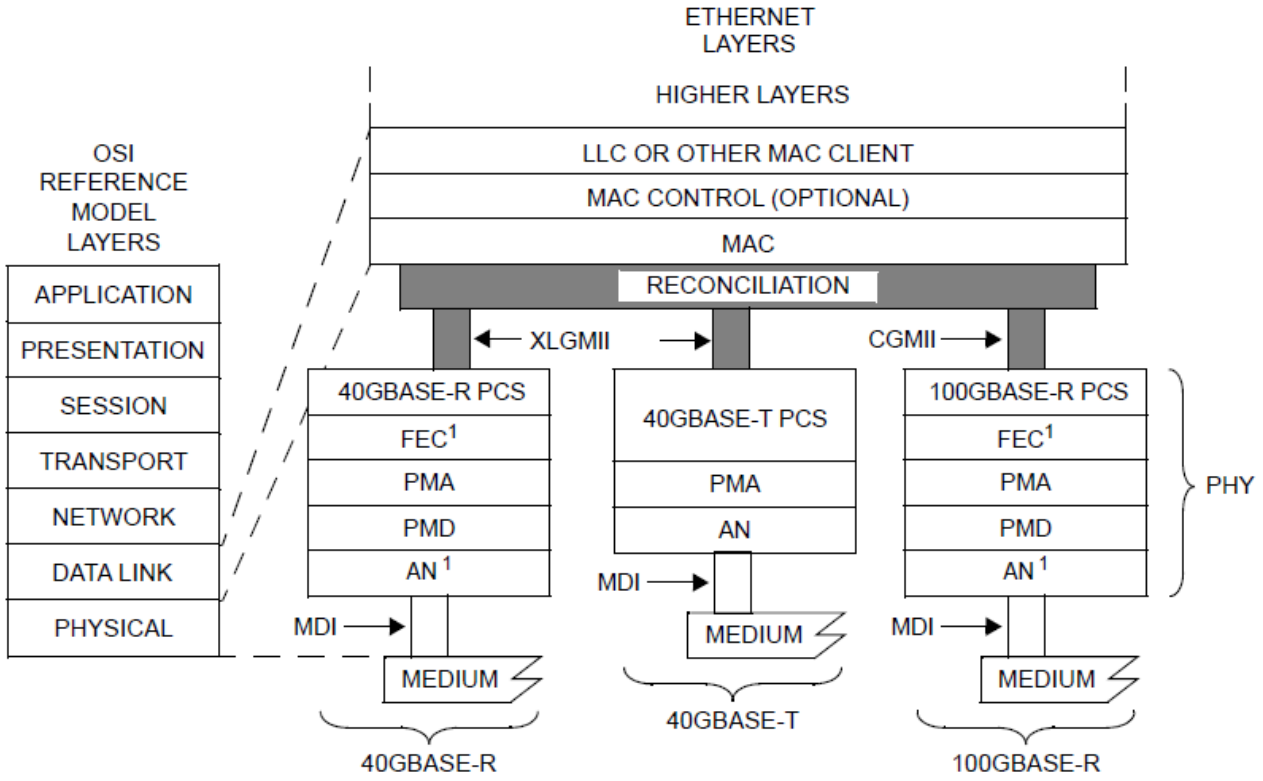
Clause 81 XLGMII and CMII (40 and 100Gb/s) Interface

The XLGMII/CGMII has the following characteristics:

a) The XLGMII supports a speed of 40 Gb/s.
b) The CGMII supports a speed of 100 Gb/s.
c) Data and delimiters are synchronous to a clock reference.
d) It provides independent 64-bit wide transmit and receive data paths.
e) It supports full duplex operation only.

The 40GBASE-T PHY service interface is the XLGMII.

The 100GBASE-T PHY service interface is the CGMII.



AN = AUTO-NEGOTIATION
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
LLC = LOGICAL LINK CONTROL
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

NOTE 1—CONDITIONAL BASED ON PHY TYPE

Figure 81–1—RS and MII relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model

Clause 81 XLGMII and CMII (40 and 100Gb/s) Interface

Summary

The following are the major concepts of the XLGMII/CGMII:
a) The XLGMII/CGMII is functionally similar to other media independent interfaces that have been defined for lower speeds, as they all define an interface allowing independent development of MAC and PHY logic.
b) The RS converts between the MAC serial data stream and the parallel data paths of the XLGMII/CGMII.
c) The RS maps the signal set provided at the XLGMII/CGMII to the PLS service primitives provided at the MAC.
d) Each direction of data transfer is independent and serviced by data, control, and clock signals.
e) The RS generates continuous data or control characters on the transmit path and expects continuous data or control characters on the receive path.
f) The RS participates in link fault detection and reporting by monitoring the receive path for status reports that indicate an unreliable link, and generating status reports on the transmit path to report detected link faults to the DTE on the remote end of the connecting link.
g) The XLGMII and CGMII may also support Low Power Idle (LPI) signaling for PHY types supporting Energy Efficient Ethernet (EEE) (see Clause 78).

Clause 81 XLGMII and CMII (40 and 100Gb/s) Interface

The XLGMII and CGMII, which, when implemented as a logical interconnection port between the MAC sublayer and the Physical Layer device (PHY), uses a 64-bit wide data path

81.1.6 XLGMII/CGMII structure

The XLGMII/CGMII is composed of independent transmit and receive paths. Each direction uses 64 data signals (TXD<63:0> and RXD<63:0>), 8 control signals (TXC<7:0> and RXC<7:0>), and a clock (TX_CLK and RX_CLK). Figure 81–2 depicts a schematic view of the RS inputs and outputs.

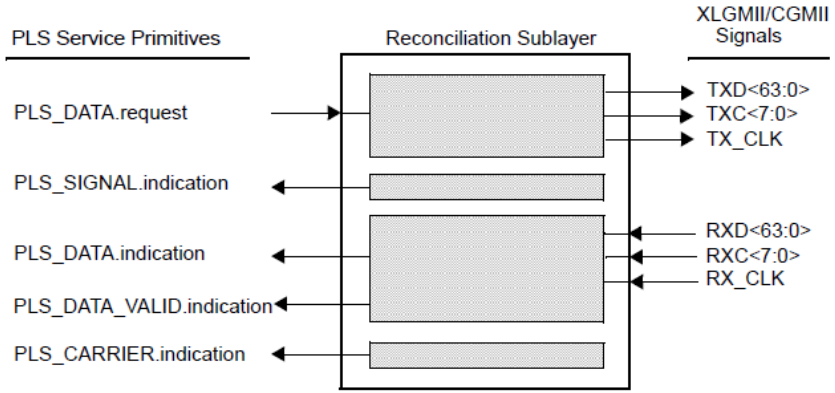


Figure 81–2—Reconciliation Sublayer (RS) inputs and outputs

The 64 TXD and 8 TXC signals shall be organized into eight data lanes, as shall the 64 RXD and 8 RXC signals (see Table 81–2). The eight lanes in each direction share a common clock, TX_CLK for transmit and RX_CLK for receive. The eight lanes are used in round-robin sequence to carry an octet stream. On transmit, each eight PLS_DATA.request transactions represent an octet transmitted by the MAC.

Clause 82 PCS 40G/100GBASE-R

The 40GBASE-R and 100GBASE-R PCSs provide all services required by the XLGMII/CGMII, including
the following:
a) Encoding (decoding) of eight XLGMII/CGMII data octets to (from) 66-bit blocks (64B/66B).
b) Transferring encoded data to (from) the PMA.
c) Compensation for any rate differences caused by the insertion or deletion of alignment markers or
due to any rate difference between the XLMII/CGMII and PMA through the insertion or deletion of
idle control characters.
d) Determining when a functional link has been established and informing the management entity via
the MDIO when the PHY is ready for use.

The PCS shall provide transmit test-pattern mode for the scrambled idle pattern (see 82.2.11), and shall
provide receive test-pattern mode for the scrambled idle pattern (see 82.2.11). Test-pattern mode is activated separately for transmit and receive. The PCS shall support transmit test-pattern mode and receive
test-pattern mode operating simultaneously so as to support loopback testing.

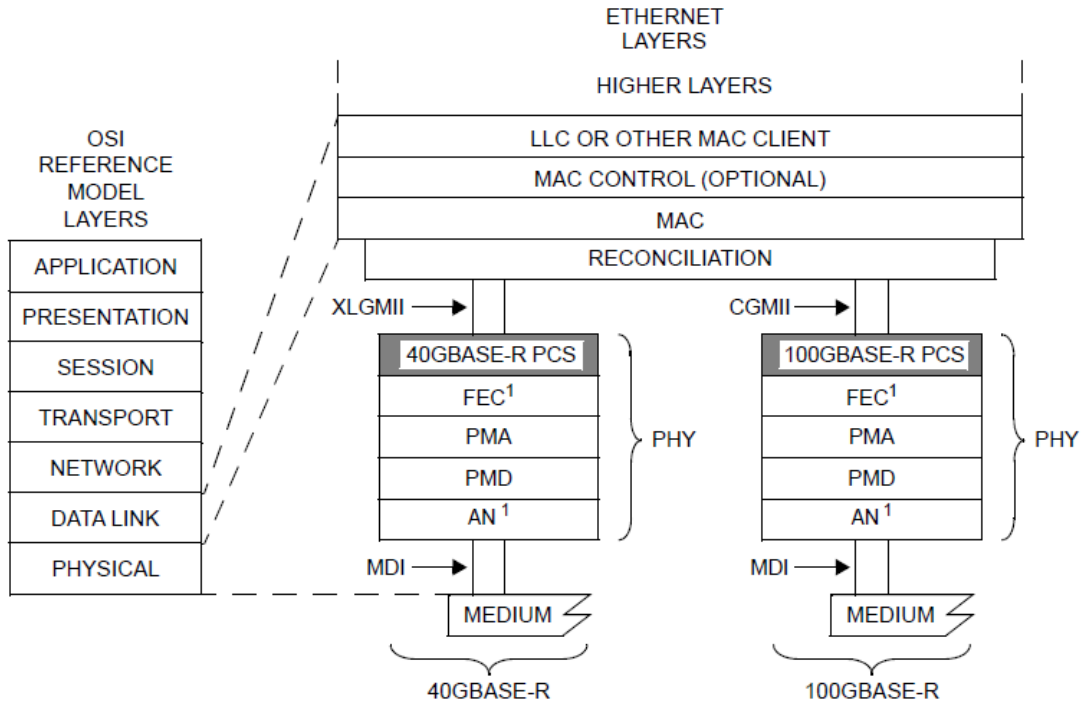


Figure 82–1—40GBASE-R and 100GBASE-R PCS relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 Ethernet model